JMIR Medical Education

Impact Factor (2024): 3.2
Volume 6 (2020), Issue 1 ISSN 2369-3762 Editor-in-Chief: Blake J. Lesselroth, MD, MBI, FACP, FAMIA

Contents

Original Papers

Viewpoints

Review

Original Paper

Comparison of Assessment by a Virtual Patient and by Clinician-Educators of Medical Students' History-Taking Skills: Exploratory Descriptive Study

Jean Setrakian[1], MD; Geneviève Gauthier[1], PhD; Linda Bergeron[1], MA; Martine Chamberland[1], MD, PhD; Christina St-Onge[1], PhD

Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Jean Setrakian, MD
Faculté de médecine et des sciences de la santé
Université de Sherbrooke
3001 12ème Avenue Nord
Sherbrooke, QC, J1H 5N4
Canada
Phone: 1 8193461110 ext 74948
Email: jean.setrakian@usherbrooke.ca

Abstract

Background: A virtual patient (VP) can be a useful tool to foster the development of medical history–taking skills without the inherent constraints of the bedside setting. Although VPs hold the promise of contributing to the development of students' skills, documenting and assessing skills acquired through a VP is a challenge.

Objective: We propose a framework for the automated assessment of medical history taking within a VP software and then test this framework by comparing VP scores with the judgment of 10 clinician-educators (CEs).

Methods: We built upon 4 domains of medical history taking to be assessed (breadth, depth, logical sequence, and interviewing technique), adapting these to be implemented into a specific VP environment. A total of 10 CEs watched the screen recordings of 3 students to assess their performance first globally and then for each of the 4 domains.

Results: The scores provided by the VPs were slightly higher but comparable with those given by the CEs for global performance and for depth, logical sequence, and interviewing technique. For breadth, the VP scores were higher for 2 of the 3 students compared with the CE scores.

Conclusions: Findings suggest that the VP assessment gives results akin to those that would be generated by CEs. Developing a model for what constitutes good history-taking performance in specific contexts may provide insights into how CEs generally think about assessment.

(JMIR Med Educ 2020;6(1):e14428) doi:10.2196/14428

KEYWORDS

virtual patients; medical history taking; automated scoring; simulation training; medical education; medical students; educational assessment; computer software; medical history–taking skills; medical history–taking skills assessment

Introduction

Background

Virtual patients (VPs) are increasingly used in health professions education (HPE) [1,2], including the teaching of diagnostic reasoning and interviewing [3]. Despite VPs' positive impact on learning and skill development [4-7], their usefulness and effectiveness as learning tools have been challenged [8,9], and questions have been raised about which competencies students develop through VPs [10] and how VPs align with, and complement, learning outcomes in HPE curricula [1,11]. One main issue is the lack of outcome measures to monitor the impact of VPs on student learning.

Developing and measuring specific learning outcomes is challenging for many reasons, including the inherent variability in the ways to solve complex problems in HPE [12] as well as the impact of developmental and contextual perspectives on skills and competencies [13,14]. Consequently, current outcome measures for VPs have mostly focused on pre-post satisfaction, knowledge, or global correlation with other measures or tests [15], which only provide partial insight into competency development and mastery. More specific and accurate outcome measures are required to explore further and document a VP's potential positive impact on students' learning. One such outcome measure is how well the assessment *by a VP* can reproduce teachers' assessment of students' performance.

Assessment aligned with teachers' judgment could become an integral part of the utilization of VP software (1) by learners for individual practice with feedback by the VP on performance and (2) by teachers as a tool for illustration and evaluation. A VP could be used for assessing reasoning and interviewing skills [16-19] and would be readily acceptable to students [20]. Creating a realistic, credible, and multidimensional VP is challenging [21]. The complexity of assessing integration of reasoning and interviewing skills [22,23] adds further to the challenge.

Our goal was to develop, and implement in a VP, an automated assessment of medical students' history-taking skills and document how this assessment aligns with the perspectives of clinician-educators (CEs).

Assessing Medical History–Taking Skills

Medical history is central to making a correct diagnosis, with real as well as simulated patients [24,25]. Good history-taking requires both skillful diagnostic reasoning and interviewing [26]. Observation by CEs of students obtaining a patient's history at the bedside provides a valuable (and often the only) opportunity for teaching and assessing how these twin skills are integrated [27,28]. Several tools exist and can be used to document parts of the medical history–taking skills such as the Cambridge-Calgary model [29], the History-Taking Rating Scale (HTRS) [30], the Maastricht History-taking and Advice Checklist (MAAS) [31], and the Brown Interviewing Checklist (BIC) [32].

The items assessed in these tools are broad and require human judgment to assess. For example, "picking up clues" is an item included in the section on "gathering information" of the Cambridge-Calgary model. Such items as are covered under "gathering information" may be self-evident for CEs, yet translating them into an automated assessment is complex. For instance, VP software can be set up to assess whether a student picks up verbal clues. To do this, the VP must first be programmed with specific instances in the simulation during which the patient gives a verbal clue that must be picked up. The software can then document the student's *behavior* (did he or she act on the clue?) and use it as *evidence* that he or she did indeed pick up the clue. Picking up verbal clues is one of the many skills that could be programmed in this fashion (ie, instances and assessment of behavior), with this degree of fine granularity into a history-taking VP.

A framework is required in the development of automated assessment by VPs, modeled on how CEs' assess history taking. Then from this framework, specific implementation rules can be programmed into a VP to provide feedback on performance to the learner (formative assessment). Once established, such a framework could eventually lead to the development of additional evaluation tools (summative assessment).

We developed a framework to precisely articulate skills assessed in history taking by breaking down their broad components into operational objective measures. To explore whether such measures can be used in the ways outlined earlier, we tested whether they were comparable with CEs' assessments. Articulating how we assess these skills furthers our knowledge of how we assess history taking at the bedside through tools such as the Calgary-Cambridge model.

The objectives of our study were as follows: (1) to present a framework for assessing medical history–taking skills through VP software and (2) to examine, using this framework, the alignment of VP assessment with that of CEs.

A Framework for Virtual Patient Assessment of Medical History–Taking Skills

Our goal was first to clarify expectations and assumptions about medical history–taking skills, exploring ranges of acceptable performance in the context of medical history taking [33]. Our work thus began by operationalizing expected medical history–taking skills at the clerkship level by identifying the characteristics of a successful performance.

Building on years of experience assessing the bedside skills of students such as those described by the HTRS, the MAAS, the BIC, and the Calgary-Cambridge model and through iterative consultations with colleagues from a Canadian University, the principal investigator (JS) set out to break down the skills into a framework comprising bite-sized specific instances and behaviors that can be automated and thus programmed into a VP. These were classified into 4 domains: breadth of data gathering, depth of data gathering, logical sequence of questions, and interviewing technique. These domains were then adapted to be implemented into a specific VP environment. See Table 1 for the framework's definitions and operationalization for implementation rules.

Table 1. Framework for virtual patient assessment of medical history–taking skills.

Domain	Description	Implementation rules
Breadth		
Breadth of data gathering	Extent of exploration to find all relevant problem areas in the patient's situation	Symptoms identified: as percentage out of total number of relevant symptoms
Depth		
Depth of data gathering	Extent of exploration to find all relevant details about each problem area	Details asked about the symptoms: percentage out of total number of details programmed in the VP[a]
Logical sequence		
Sequence of questions	Logical sequence that reflects thinking through the relevant diagnostic possibilities	Differential scoring for overall order of identification of symptoms and for alternative sequences (see Multimedia Appendix 1)
Interviewing technique		
Component (a): appropriate use of generic questions	Asking for generic details that apply to each and every symptom, such as duration, severity, course, and precipitating factors	Generic questions: percentage out of total questions–>scoring performances using a range established by the CE[b]
Component (b): appropriate use of transitioning statements	Appropriate use of transitioning statements such as "yes," "no," "let me ask you a few questions," and "that's normal". The ideal number varies from encounter to encounter	Opening and follow-up questions, interruptions, yes or no answers, reassurance and transition statements; –>scoring performances using a range established by the CE
Component (c): appropriate flow	Avoidance of jumping from 1 topic to the next without apparent reason, or of leaving some areas not fully explored before moving on to others	Number of times the student passes from 1 category of questions (eg, GI) to another (eg cardiac)–>scoring according to acceptable numbers established by the CE
Component (d): successful handling of KIE[c]	Combination of a number of events or instances that require an understanding of implicit communication rules (clues, misunderstandings, tangential answers, incomplete answer, vague answer, imprecise answer)	Binary scoring of success or failure of events if encountered in any given KIE

[a]VP: virtual patient.

[b]CE: clinician-educator.

[c]KIE: key interview element.

The first 2 domains (ie, breadth and depth) concerned completeness of data gathering. Are all the patient's symptoms obtained, and are they obtained in sufficient detail? During bedside teaching, although CEs are unaware of all the patient's symptoms and the details thereof, they routinely make a judgment of a student's thoroughness. For the VP, we defined breadth as the percentage of the VP's symptoms (eg, dizziness, pallor, fatigue, hematochezia) identified by the students and depth as the percentage of programmed symptom details identified (eg, dizziness for 3 weeks, worse upon standing, first instance, without loss of consciousness).

The third domain, logical sequence of questions, reflected CEs' judgment of students navigating through a differential diagnosis. Although diagnostic reasoning cannot be assessed directly, inferences are made about students' reasoning through the sequence of questioning about symptoms. For example, asking about fever right after finding out about a cough is taken as indirect evidence that the student entertained the possibility of an infectious cause for cough. Without limiting the "right" sequence exclusively to an expert path, the VP assessment was made to attach different scores to various optional sequences of exploring 2, 3, 4, or 5 symptoms to reflect this type of assessment of diagnostic reasoning.

The fourth domain, interviewing technique, comprises 4 components. The first 3 components, use of a combination of generic vs system-specific questions, transition statements, and number of jumps between topics, are described in Table 1. These 3 components could be easily monitored by the VP. As to the fourth component, colleagues who were consulted for the design of the VP pointed out that specific interviewing pitfalls occurring during medical history taking constituted a key component of their assessment of the performance of students: Did they miss a clue, were they thrown off by a tangential answer, or were they able to stay on course and come back to explore the tangent later? We operationalized these elements through key interview elements (KIEs; see Table 1). These elements, based on the common challenges encountered in interviews, were programmed in a sufficiently large number to ensure that each student would encounter on average 3 or 4 instances.

Each of the 4 domains described earlier were implemented into the VP to provide 4 different scores and a global score: virtual patient–breadth score (VP–BS), virtual patient–depth score (VP–DS), virtual patient–logical sequence score (VP–LSS), and virtual patient–interviewing technique score (VP–ITS), as well as a virtual patient–global score (VP–GS). Although the VP was programmed to provide domain scores from its data, the relative importance of score components and thresholds for specific errors were left to be adapted to the educational context of use.

Methods

Design of Study

In this exploratory descriptive study, we articulated and tested a framework for assessing medical history–taking skills with a VP. First, we implemented this framework into a specific VP and then compared global and domain scores assigned by the VP to those assigned by 10 experienced CE participants. The study was approved by our institution's ethics committee.

Participants

A total of 10 CEs, all general internists from a Canadian Department of Medicine, were recruited by convenience sampling. The sample consisted of 6 men and 4 women, with a mean (SD) of 16.5 years (9.2) of medical specialty practice and a mean (SD) of 14.3 years (8.3) of evaluating medical students' history taking. None of the participants had been involved in the elaboration or consultation that had led to the programming of the VP. All participants gave consent to participate in the study.

Materials

Screen Recordings of Student Interviews or Stimuli

Screenshot videos of 3 third-year medical students' interview with the VP in a clinical case of colon cancer were used as stimuli. Students were recruited through convenience sampling. The screenshot videos were created using Camtasia Studio 7, conserving the students' anonymity. A total of 2 students were in the first trimester and 1 in the last trimester of the clerkship of a 4-year medical curriculum. Each student was met with individually, and a consent form was signed that authorized the use of recorded data in the research project.

Each student was first introduced to the software. Each part of the screen interface, as well as the navigation boxes, was explained. The student had 10 min to navigate freely and get familiarized with the software. The student was then invited to take a medical history from the VP just as he or she would do with a real patient at the start of a hospital admission. Within a time limit of 30 min, the student was asked to go at his or her own rhythm, without "racing with the clock." The students readily used the software in all its components, without asking for further explanations. Although the software allowed the students to enter their most likely diagnosis at the end, the screenshot recording was interrupted before they entered their diagnosis, as this was not the focus of the CE's assessment.

Rating Tool

A rating tool was developed for CE by 2 team members (JS and CS). The rating scale mirrored the assessment scheme implemented in the VP with a global performance score and scores for each domain (breadth, depth, logical sequence, and interviewing techniques). Each score was described by 1 question. The CE participants had to provide ratings on descriptive 10-cm visual analog scales with 3 descriptors: 1 at each end labeled "below average" and "above average" and 1 in the middle of the line labeled "average" (see Multimedia Appendix 2), referring in this case to a third-year student's performance. A visual analog scale was decided upon over a percent score to avoid assessors assigning a typical range of marks between 60% and 100%.

Survey on Assessment Practice

A survey was developed by the authors to collect the CE participants' collective assessment practice. More specifically, the survey documented (a) their relative domain weighting (breadth, depth, logical sequence, interviewing technique) for a global score, (b) their weighting of interviewing technique elements (specific instances, use of statements, use of generic questions, number of jumps between topics), and (c) their acceptable and desirable ranges for the (1) use of statements, (2) use of generic questions, and (3) number of jumps between topics.

To help CE participants better understand some of the terms used (eg, "specific instances") and how to express the upper and lower limits, the survey included definitions, examples, and visual aids (see Multimedia Appendix 3).

The Virtual Patient Software

The VP has been developed to provide students with feedback on diagnostic reasoning and interviewing skills during medical history taking.

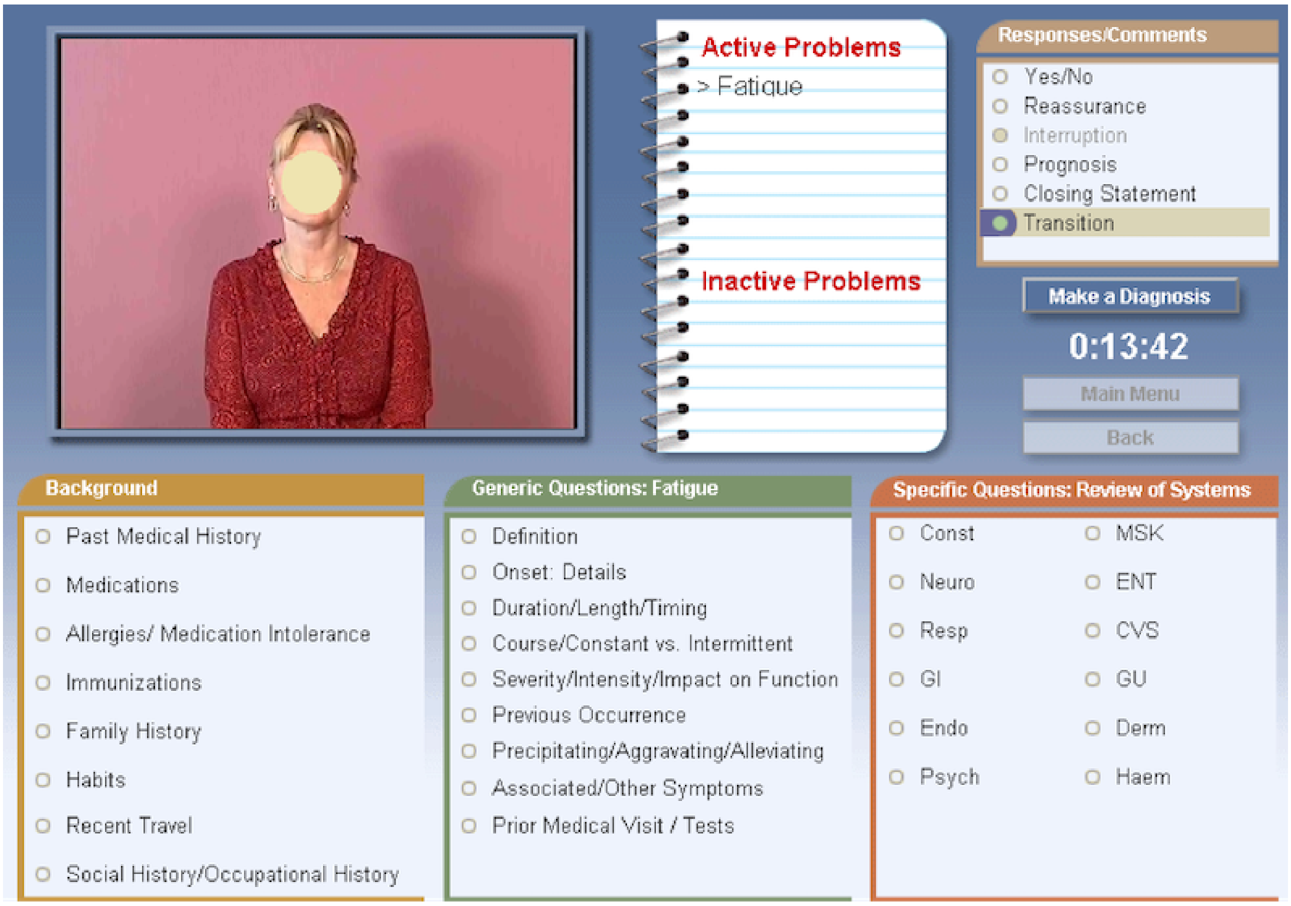
The software entailing three clinical cases including one of colon cancer was developed by the author (JS), a CE who provided the instructional design and content (eg, questions/responses, components of the panels). Instructional design and graphic design support as well as programming in Java were provided by the Instructional Communications Centre, McGill University, Montreal, between 2002 and 2006. The software uses a set of predetermined questions to be used for the interview. Video answers were created to have a set of default responses, provided by an actress, for all the questions available to the medical interviewer. Some reactions of annoyance, irritation, or anxiety were also recorded to keep the interview more realistic. Responses specific to each of the 3 clinical cases were recorded to be substituted to the default questions depending on each clinical situation.

The screen interface consists of various panels (see Figure 1) including the following: (1) a video of the patient, (2) a note pad where the symptoms appear as they are revealed by the patient's answers that the student can then drag and drop between an active and an inactive problem list, (3) three panels of questions (background questions, generic questions, and a review of systems), (4) a responses-and-comments panel, (5) a clock, (6) a box allowing the student to make the diagnosis at the end, and (7) two buttons ("main menu" and "back"), allowing the student to browse.

The questions available in the software (around 500) are divided into the 3 main categories (illustrated by the different panels shown earlier) and the responses-and-comments panel. The first category (background) includes questions on medical history, medications, allergies, immunizations, family history, habits, recent travel, and social history. The second category lists *generic* questions that can be applied to each symptom, ranging from "What happened just before the symptom started?" to "Have you seen a physician for that symptom?". Each question can be applied to each of the patient's symptoms, and the wording of the question changes as the student clicks on a different item on the problem list (on the note pad). The third category consists of a review of systems containing 350 questions. As the student clicks on a system, a list of questions about the chosen system appears. In addition, the student may click on responses or comments that include transition statements, interruption statements, and reassurance statements. A *follow-up* button is available once an answer has been provided by the patient, and allows a choice of 4 follow-up questions: "You need to tell me more about that," "Let me ask you once more," "Pardon?," and "Are you sure?".

As the patient reveals her symptoms or items of her medical history, they appear in the list of "active problems." Items can be moved (drag and drop) between the lists of "active problems" and "inactive problems" at any time. When the student is ready, he or she may click on "make a diagnosis" and choose one or more items among a list of diagnoses.

Figure 1. Virtual patient screen interface.



Procedure

A research assistant met the CE participants for 2 hours. After a brief introduction to the simulation and the project, CE participants had 5 min to navigate for themselves the simulation software, on a different case from the stimulus, to become familiar with the interface and the choice of questions.

The CE participants were then given succinct and nonquantitative definitions of breadth, depth, logical sequence, and interviewing technique, without revealing the corresponding VP operational definitions. After watching each student's screen recording, they were asked to score the student's performance using the *rating tool*. They did not see the student's choice of diagnosis, as the screenshot video was interrupted before, and they were not familiar with the diagnosis of the VP. Afterward, they had to complete the *survey on assessment.*

Analyses

Virtual Patient Scores

The data of the *survey on assessment* were subsequently used to compute the VP-derived scores for these components. For example, to compute the VP–GS, each VP domain score (breadth, depth, logical sequence, and interviewing technique) was multiplied by the mean weight that the CE participants attached to each domain. Furthermore, to compute the VP–ITS, the CE participants' mean suggested cutoffs were used for acceptable and optimal ranges of generic vs specific questions, transitioning statements, and number of jumps between topics.

Clinician-Educator Scores

The response on each participant's visual analog scale was converted to a score out of 100 by measuring with a ruler the position of the respondent's pen mark, with 10 cm representing 100%. For each student, the mean values for global performance, breadth, depth, logical sequence, and interviewing technique scores provided by the 10 CE participants on the assessment grids constituted the CE scores (clinician-educator–global score [CE–GS], clinician-educator–breadth score [CE–BS], clinician-educator–depth score [CE–DS], clinician-educator–logical sequence score [CE–LSS], clinician-educator–interviewing technique score [CE–ITS], respectively).

Results

Students' scores, from VP and CE, are presented in Figure 2. The single line represents the VP software assessment, and the boxplot represents the range of assessment made by the CE participants. There is a boxplot for each of the 3 students' performance for each of the 5 scores. The goal of these descriptive analyses is to explore how the assessment provided by the VP using our framework compares with the gold standard, that is, the assessment provided by the CE. The aim is having a VP score that is within the range of scores that CE have assigned to each student (see Multimedia Appendix 4).

Overall, the scores provided by the VP were slightly higher but comparable with the ones assigned by the CE for the global performance and for the domains of depth, logical sequence, and interviewing technique. For breadth, the VP scores were higher, and they did not fall within the range of the CE scores for student A and C. On interviewing technique, which includes 4 components, only the score for student C from the VP was not within the range of CE scores.

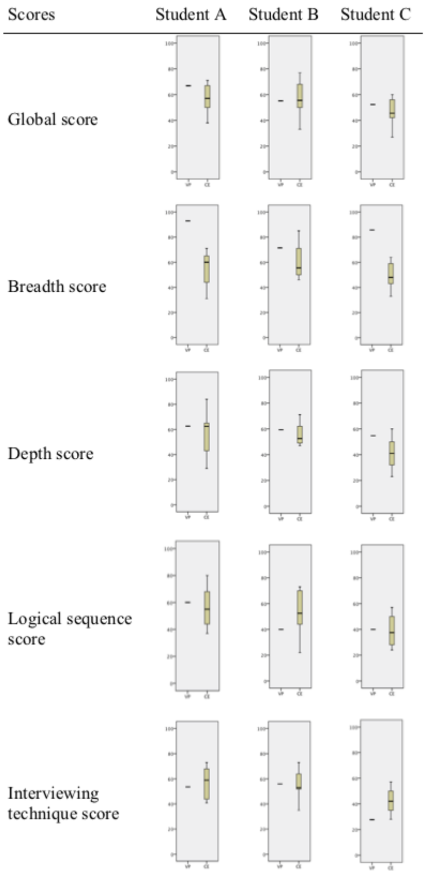


Figure 2. Boxplots displaying the virtual patient and clinician-educator scores for each student and for each score.

Discussion

Comparing Virtual Patient and Clinician-Educators' Scores

We implemented within a specific VP a framework for assessing medical history taking by breaking down broad skills into bite-sized assessment points and then tested the framework against the judgment of 10 CEs. Our findings suggest that through such a framework, assessment by VP can produce scores akin to those generated by a CE. We discuss our results and reflect on the relevance of each domain in terms of the proposed assessment scheme and on its implementation.

An advantage of using an assessment framework embedded in a VP to assess medical history–taking skills is the reliability of the assessment [34]. When referring to reliability as "the consistency of scores across replications of a testing procedure" [34], it is clear that automated assessment can contribute to the reliability of the assessment. Given that reliability is often considered necessary to the validity of the assessment scores, it stands to reason that we wanted to document if our framework embedded in a VP yielded valid assessments. Thus, we compared VP scores with "gold standards," that is, CEs' assessment of the history-taking skills of medical students.

The VP–GS was computed from the 4 domain scores, albeit with relative weighting calibrated according to the survey of CE participants (see Multimedia Appendix 4 for details).The CE–GSs were an appraisal by CE of the students' overall performance and were not derived from the CE's domain scores. CE's global appraisals typically have the gestalt quality of a true expert assessment [35] and represent the gold standard for the VP–GS. Overall, the global scores from the VP fell within the range of the CE–GSs, suggesting that depth, breadth, logical sequence, and interviewing technique are appropriate and sufficient domains to approximate an expert's gestalt assessment, as otherwise VP– and CE–GSs might have differed. Furthermore, when CEs were asked for additional domains that they considered important, they named aspects of the medical interview that could not be seen on a screen recording, such as students' empathy, body language, and tone; skills such as picking up nonverbal clues; and the ability to organize the interview between introduction and conclusion, which they felt, rightly, that the VP did not allow. It bears to be pointed out that some of these aspects of assessing medical history taking fall outside the "gathering information" section of the Calgary-Cambridge model, whereas others such as picking up nonverbal clues are within that section and could have been programmed into a VP (eg, video of the VP fidgeting) but were not addressed by our framework.

The VP scores for breadth, that is, the identification of the full range of the VP's various symptoms, are higher than the breadth scores given by CE for 2 of the 3 students. VP scores were simply the percentage of symptoms identified by the student out of the total number of symptoms programmed. Unlike the VP, CE had no knowledge of the total number of symptoms programmed and made a judgment as to what other symptoms this type of patient might have. There could be two main reasons for this difference between VP– and CE–BSs: (1) the VP may not have been programmed with a sufficiently large number of symptoms to be a realistic representation of this type of patient or (2) CE may have expected a broader range of questions about general symptoms, the so-called "review of systems." We did not identify missing details that should have been programmed into the VP after repeated use of the case with students and consultation with CE, suggesting that rather than the VP having too few symptoms, CEs expect a review of systems as part of any medical history taking. Of note, all 10 CE participants were general internists, who likely incorporate such a generalist approach in their own practice. The review of systems was not taken into account in the VP–BS.

The VP scores for depth, that is, the level of detail about each of the VP's symptoms, are within the range of scores given by CEs. Again, VP scores were simply the percentage of symptom details identified by the student out of the total number of symptom details programmed. The fact that the CE's judgment is aligned with this simple ratio suggests that CEs were able to estimate the details about symptoms that were missed or not missed by the students.

The VP scores for logical sequence, which reflects systematic thinking through the relevant diagnostic possibilities, are well aligned with the range of scores given by the CE. Implementation in a VP was much more complex than that for breadth or depth as it involved assigning different scores to a number of potential sequences of questions relevant to the VP's symptoms. Indeed, this domain required a set of rules that reflected the existence, as for all complex problem solving, of not just one so-called expert path, but of several acceptable paths to reaching the diagnosis. In addition, this domain score, unlike the first 2, could not be improved by the students simply clicking on as many questions as they could, as the scoring depended on sequence of questioning rather than the sheer number of questions asked.

The VP scores for interviewing technique, which is a combination of 4 components (appropriate use of generic questions, transition, flow, and handling of KIEs), are within the ranges of scores by CE for students A and B and slightly less for student C. This other complex measure, which has been calibrated using the ranges suggested by CEs as to the ideal and acceptable limits for the number of jumps between topics, the use of transitioning statements, and the use of generic questions and specific KIEs, seems to provide VP scores that are in the lower range than the corresponding CE scores. The VP scores were binary and may have been too restrictive in their application of CE's suggested ideal and acceptable ranges.

Survey on Assessment Practice

The responses of the 10 CEs to the study survey documented their relative domain weighting (breadth, depth, logical sequence, and interviewing technique) for a global score, their weighting of interviewing technique elements (specific instances, use of statements, use of generic questions, and the number of jumps between topics), and their acceptable and desirable ranges for the use of statements, the use of generic questions, and the number of jumps between topics (see Multimedia Appendix 4). These surveys allowed us to refine the framework at the final step of computing scores from the raw data of the VP. Such an iterative process ensured that an automated assessment reflected the CE's priorities and values in judging student performance.

Reflection on Proposed Outcome Measures

Developing a framework for assessment of history-taking skills to program into a VP and comparing VP scores with CEs' judgment enables us to reflect both on the proposed framework and on its implementation into a specific VP. For example, as we reflect on how the breadth score is underestimated by the VP, we know we are probably missing an element of breadth as defined by CEs, likely a wider-ranging review of systems, as described earlier. We are therefore considering the integration of an additional component of the number of systems (eg, cardiovascular, renal) the student explores through specific questions into the VP's domain score. Similarly, when we are reflecting on the implementation of our framework, we want to review how the ranges of acceptable numbers of generic questions or transition statements are calculated. Instead of applying discrete cutoffs (eg, less than 26.4% is given zero, based on the mean from the CE survey), we would possibly need to try using an incremental cutoff to better reflect CEs' judgment and resulting scores.

Numerous studies related to VPs have centered on their impact on knowledge acquisition and skills [36]. This study focuses on developing an assessment framework aligned with educators' assessment practices. Inviting CEs' perspective [1,37] allows for the creation of VP aligned with CEs' educational objectives, while in turn providing CE with an opportunity to understand better their students' skill development. After implementation, using CEs' judgment to validate and test the assessment framework, as we have done here, further helps improve implementation and alignment with objectives. The ultimate goal is better VP integration into the formal curriculum, and a smooth transition from VP to bedside teaching, as it is clear that no VP could ever replace real interaction with patients. Assessment provided by VP must make sense to all actors in the learning environment, and reflect as faithfully as possible current assessment practices, ultimately to promote genuine improvement in performance.

Limitations

The study's CE vs VP comparison results are preliminary, as they include the use of a single case and limited number of students' performances. Our results need to be tested with other cases and a larger audience in a variety of settings. Medical students at the clerkship level are the intended audience for this specific VP software dealing with diagnostic reasoning and interviewing skills, and the results may not hold true for different levels of students and additional assessments such as communication skills and body language. In addition to the small number of students' performances, their narrow spread represents another limitation. The 3 students did not have extremes of high- and low-quality performance. Using a larger pool of students and selecting specific performances purposefully for validating a broad range of performance would enable us to test better for VP scores' discriminative ability. Also, this VP software is not intended to assess the nonverbal communication skills inherent to the history-taking skills, the focus being more on most of the other aspects of gathering information as part of the medical interview.

Conclusions

We developed a framework for assessment of medical history–taking skills and programmed it into a VP software that aligned with assessment by CEs in our small observational study. Through an iterative process, our study also provided insight into how CEs assess specific domains of medical history taking, allowing us to refine further the scheme programmed into the VP. Our results suggest that some skills that are usually assessed at the bedside can be assessed by software, provided reasoning is judged with flexibility through a range of logical sequences rather than an "expert path" and that broad descriptive terms such as "picks up clues" can be translated into operational, observable behaviors by the student and the VP is then specifically programmed to include situations that call upon the student to demonstrate these skills by engaging in specific behaviors (such as clarification, following up on clues, asking a logical sequence of questions, using open-ended questions) Further steps in this direction, with more diverse VPs and ongoing consultation and exchange with CEs can be expected to result in producing a generation of VPs that are programmed to provide feedback to learners and to assist teachers in their assessment of performance.

Acknowledgments

This project was funded by the *Société des médecins de l'Université de Sherbrooke*, which was not involved in the study or in the writing of the manuscript. We would like to thank Sìlvia Mamede and Remy Rikers for their help in the design of this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Differential scoring for overall order of identification of symptoms and for alternative sequences.
[PDF File (Adobe PDF File), 53 KB - mededu_v6i1e14428_app1.pdf]

Multimedia Appendix 2
Rating tool.
[PDF File (Adobe PDF File), 69 KB - mededu_v6i1e14428_app2.pdf]

Multimedia Appendix 3
Survey on assessment practice.
[PDF File (Adobe PDF File), 157 KB - mededu_v6i1e14428_app3.pdf]

Multimedia Appendix 4

Mean weight for each domain and components and mean limits for acceptable and optimal range for domains and components given by clinician-educator participants.
[PDF File (Adobe PDF File), 52 KB - mededu_v6i1e14428_app4.pdf]

References

1. Berman NB, Durning SJ, Fischer MR, Huwendiek S, Triola MM. The role for virtual patients in the future of medical education. Acad Med 2016 Sep;91(9):1217-1222. [doi: 10.1097/ACM.0000000000001146] [Medline: 26959224]
2. Hege I, Kononowicz AA, Tolks D, Edelbring S, Kuehlmeyer K. A qualitative analysis of virtual patient descriptions in healthcare education based on a systematic literature review. BMC Med Educ 2016 May 13;16:146 [FREE Full text] [doi: 10.1186/s12909-016-0655-8] [Medline: 27177766]
3. Parsons TD, Kenny P, Rizzo AA. Virtual Human Patients for Training of Clinical Interview and Communication Skills. In: Proceedings of the 7th Virtual Reality and Associated Technology Conference with ArtAbilitation. UK: The University of Reading; 2008 Presented at: ICDVRAT'08; September 8-11, 2008; Maia, Portugal.
4. Robinson KE, Allen PJ, Quail M, Beilby J. Virtual patient clinical placements improve student communication competence. Interact Learn Environ 2018:1-11. [doi: 10.1080/10494820.2018.1552869]
5. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in health professions education: a systematic review and meta-analysis. Acad Med 2010 Oct;85(10):1589-1602. [doi: 10.1097/ACM.0b013e3181edfe13] [Medline: 20703150]
6. Bergeron D, Champagne J, Qi W, Dion M, Thériault J, Renaud J. Impact of a student-driven, virtual patient application on objective structured clinical examination performance: observational study. J Med Internet Res 2018 Feb 22;20(2):e60 [FREE Full text] [doi: 10.2196/jmir.7548] [Medline: 29472175]
7. Lehmann R, Bosse HM, Simon A, Nikendei C, Huwendiek S. An innovative blended learning approach using virtual patients as preparation for skills laboratory training: perceptions of students and tutors. BMC Med Educ 2013 Feb 12;13:23 [FREE Full text] [doi: 10.1186/1472-6920-13-23] [Medline: 23402663]
8. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. Med Educ 2012 Jul;46(7):636-647. [doi: 10.1111/j.1365-2923.2012.04243.x] [Medline: 22616789]
9. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. Acad Med 2013 Jun;88(6):872-883. [doi: 10.1097/ACM.0b013e31828ffdcf] [Medline: 23619073]
10. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
11. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. Acad Med 2014 Mar;89(3):387-392 [FREE Full text] [doi: 10.1097/ACM.0000000000000130] [Medline: 24448038]
12. Gauthier G, Lajoie SP. Do expert clinical teachers have a shared understanding of what constitutes a competent reasoning performance in case-based teaching? Instr Sci 2014 Jul;42(4):579-594. [doi: 10.1007/s11251-013-9290-5]
13. Bates J, Ellaway RH. Mapping the dark matter of context: a conceptual scoping review. Med Educ 2016 Aug;50(8):807-816. [doi: 10.1111/medu.13034] [Medline: 27402041]
14. Teunissen PW, Kogan JR, Cate OT, Gruppen LD, Lingard LA. Learning in practice: a valuation of context in time-variable medical training. Acad Med 2018 Mar;93(3S Competency-Based, Time-Variable Education in the Health Professions):S22-S26. [doi: 10.1097/ACM.0000000000002070] [Medline: 29485483]
15. Botezatu M, Hult H, Tessma MK, Fors U. Virtual patient simulation: knowledge gain or knowledge loss? Med Teach 2010;32(7):562-568. [doi: 10.3109/01421590903514630] [Medline: 20653378]
16. Chapman DM, Calhoun JG, van Mondfrans AP, Davis WK. Measuring clinical reasoning competency using a virtual patient model. J Clin Reason Proced Competency 2013 Mar;1:1-9 [FREE Full text]
17. Johnsen K, Raij A, Stevens A, Lind DS, Lok B. The Validity of a Virtual Human Experience for Interpersonal Skills Education. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. United States: ACM; 2007 Presented at: CHI'07; April 28 - May 3, 2007; California, San Jose, USA p. 1049-1058 URL: https://dl.acm.org/doi/10.1145/1240624.1240784
18. Stevens A, Hernandez J, Johnsen K, Dickerson R, Raij A, Harrison C, et al. The use of virtual patients to teach medical students history taking and communication skills. Am J Surg 2006 Jun;191(6):806-811. [doi: 10.1016/j.amjsurg.2006.03.002] [Medline: 16720154]
19. Triola MM, Huwendiek S, Levinson AJ, Cook DA. New directions in e-learning research in health professions education: report of two symposia. Med Teach 2012;34(1):e15-e20. [doi: 10.3109/0142159X.2012.638010] [Medline: 22250691]
20. Gesundheit N, Brutlag P, Youngblood P, Gunning WT, Zary N, Fors U. The use of virtual patients to assess the clinical skills and reasoning of medical students: initial insights on student acceptance. Med Teach 2009 Aug;31(8):739-742. [doi: 10.1080/01421590903126489] [Medline: 19811211]
21. Posel N, Shore BM, Fleiszer D. Virtual patient cases: a qualitative study of the requirements and perceptions of authors. Int J Med Educ 2012;3:175-182. [doi: 10.5116/ijme.5038.a1e1]

22. Cohen DS, Colliver JA, Marcy MS, Fried ED, Swartz MH. Psychometric properties of a standardized-patient checklist and rating-scale form used to assess interpersonal and communication skills. Acad Med 1996 Jan;71(1 Suppl):S87-S89. [doi: 10.1097/00001888-199601000-00052] [Medline: 8546794]
23. Cutrer WB, Sullivan WM, Fleming AE. Educational strategies for improving clinical reasoning. Curr Probl Pediatr Adolesc Health Care 2013 Oct;43(9):248-257. [doi: 10.1016/j.cppeds.2013.07.005] [Medline: 24070582]
24. Peterson MC, Holbrook JH, von Hales D, Smith NL, Staker LV. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. West J Med 1992 Feb;156(2):163-165 [FREE Full text] [Medline: 1536065]
25. Tsukamoto T, Ohira Y, Noda K, Takada T, Ikusaka M. The contribution of the medical history for the diagnosis of simulated cases by medical students. Int J Med Educ 2012;3:78-82. [doi: 10.5116/ijme.4f8a.e48c]
26. Cary J, Kurtz S. Integrating clinical communication with clinical reasoning and the broader medical curriculum. Patient Educ Couns 2013 Sep;92(3):361-365. [doi: 10.1016/j.pec.2013.07.007] [Medline: 23896126]
27. Atkinson K, Ajjawi R, Cooling N. Promoting clinical reasoning in general practice trainees: role of the clinical teacher. Clin Teach 2011 Sep;8(3):176-180. [doi: 10.1111/j.1743-498X.2011.00447.x] [Medline: 21851565]
28. Latta L, Tordoff D, Manning P, Dent J. Enhancing clinical skill development through an Ambulatory Medicine Teaching Programme: an evaluation study. Med Teach 2013 Aug;35(8):648-654. [doi: 10.3109/0142159X.2013.801553] [Medline: 23758182]
29. Kurtz S, Silverman J, Benson J, Draper J. Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. Acad Med 2003 Aug;78(8):802-809. [doi: 10.1097/00001888-200308000-00011] [Medline: 12915371]
30. Evans BJ, Sweet B, Coman GJ. Behavioural assessment of the effectiveness of a communication programme for medical students. Med Educ 1993 Jul;27(4):344-350. [doi: 10.1111/j.1365-2923.1993.tb00279.x] [Medline: 8412875]
31. Kraan HF, Crijnen AA, de Vries MW, Zuidweg J, Imbos T, van der Vleuten CP. To what extent are medical interviewing skills teachable? Med Teach 1990;12(3-4):315-328. [doi: 10.3109/01421599009006637] [Medline: 2095449]
32. Novack DH, Volk G, Drossman DA, Lipkin M. Medical interviewing and interpersonal skills teaching in US medical schools. Progress, problems, and promise. J Am Med Assoc 1993 Apr 28;269(16):2101-2105. [Medline: 8468764]
33. Durning SJ, Artino AR, Schuwirth L, van der Vleuten C. Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. Acad Med 2013 Apr;88(4):442-448. [doi: 10.1097/ACM.0b013e3182851b5b] [Medline: 23425980]
34. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC, USA: Amer Educational Research Assn; 2014.
35. Hodges B. Assessment in the post-psychometric era: learning to love the subjective and collective. Med Teach 2013 Jul;35(7):564-568. [doi: 10.3109/0142159X.2013.789134] [Medline: 23631408]
36. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, et al. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jul 2;21(7):e14676 [FREE Full text] [doi: 10.2196/14676] [Medline: 31267981]
37. Gauthier G. Using Teaching Analytics to Inform Assessment Practices in Technology Mediated Problem Solving Tasks. In: Proceeding of the 3rd Conference on Learning Analytics and Knowledge. 2013 Presented at: LAK'13; April 8-12, 2013; Leuven, Belgium URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.415.5386&rep=rep1&type=pdf

Abbreviations

BIC: Brown Interviewing Checklist
CE: clinician-educator
CE–BS: clinician-educator–breadth score
CE–DS: clinician-educator–depth score
CE–GS: clinician-educator–global score
CE–ITS: clinician-educator–interviewing technique score
CE–LSS: clinician-educator–logical sequence score
HPE: Health Professions Education
HTRS: History-Taking Rating Scale
KIE: key interview element
MAAS: Maastricht History-taking and Advice Checklist
VP: virtual patient
VP–BS: virtual patient–breadth score
VP–DS: virtual patient–depth score
VP–GS: virtual patient–global score
VP–ITS: virtual patient–interviewing technique score
VP–LSS: virtual patient–logical sequence score

Edited by G Eysenbach; submitted 17.04.19; peer-reviewed by S Edelbring, I Hege, L Woodham, M Gaman; comments to author 01.10.19; revised version received 25.11.19; accepted 16.12.19; published 12.03.20.

Please cite as:
Setrakian J, Gauthier G, Bergeron L, Chamberland M, St-Onge C
Comparison of Assessment by a Virtual Patient and by Clinician-Educators of Medical Students' History-Taking Skills: Exploratory Descriptive Study
JMIR Med Educ 2020;6(1):e14428
URL: http://mededu.jmir.org/2020/1/e14428/
doi:10.2196/14428
PMID:32163036

Viewpoint

Email Use Reconsidered in Health Professions Education: Viewpoint

Jennie C De Gagne[1], PhD, DNP, RN, NPD-BC, CNE, ANEF, FAAN; Yesol Yang[1], MSN, NP-C; Sharron Rushton[1], DNP, MS, RN, CCM, CNE; Paula D Koppel[1], MS, RNC, GNP-BC, AHN-BC, NBC-HWC; Katherine Hall[2], MSN, RN, ONC

[1]Duke University School of Nursing, Durham, NC, United States

[2]West Virginia University School of Nursing, Morgantown, WV, United States

Corresponding Author:
Yesol Yang, MSN, NP-C
Duke University School of Nursing
307 Trent Drive, DUMC 3322
Durham, NC, 27710
United States
Phone: 1 713 927 1698
Fax: 1 919 681 8899
Email: yesol.yang@duke.edu

Abstract

Email has become a popular means of communication in the past 40 years, with more than 200 billion emails sent each day worldwide. When used appropriately, email can be an effective and useful form of correspondence, although improper practices, such as email incivility, can present challenges. Email is ubiquitous in education and health care, where it is used for student-to-teacher, provider-to-provider, and patient-to-provider communications, but not all students, faculty members, and health professionals are skilled in its use. This paper examines the challenges and opportunities posed by email communication in health professions education and reveals important deficiencies in training, as well as steps that can be taken by health professions educators to address them. Recommendations are offered to help health professions educators develop approaches for teaching email professionalism.

(JMIR Med Educ 2020;6(1):e19300) doi:10.2196/19300

KEYWORDS

communication; electronic mail; professionalism; faculty; health occupations; health occupations students

Introduction

Given the increasing complexity of the health care system, health professions educators must ensure that future clinicians are prepared to use effective communication intraprofessionally, interprofessionally, and with patients and their caregivers within and across health care settings. Communication skills are foundational competencies in education and patient care [1,2], and health care communication is occurring more frequently in an electronic manner [3]. Although email is ubiquitous in education and health care, its pervasiveness does not ensure that students, faculty members, or health professionals are skillful in its use. In fact, this review of email use within health care and educational settings reveals important training deficiencies and the need for specific steps to be taken by health professions educators. It is imperative that health professionals have the ability to use and select electronic technologies appropriately [1] in order to foster communication and civility among teams in the health care sector.

The birth of email can be traced back to the staff of the Massachusetts Institute of Technology who used electronic notes to communicate on multiuser computers in the 1960s [4]. In 1972, Ray Tomlinson, a computer engineer contracted by the US Defense Department, sent the first electronic message over the earliest form of the internet, the ARPAnet [4]. Ever since email transitioned from technical exchanges among elite programmers to mass communication, researchers have been studying the use of email in higher education institutions [5]. Early studies from the late 1980s focused on the utilization of email as a research tool, user perception and adoption of email in instructional settings, and the effects of email communication on users [5]. In 1997, it was estimated that 17.5 million adults in the United States used the internet for medical information, and by the late 1990s, physicians were beginning to use email for consulting, obtaining laboratory information, following up on patient outcomes, reviewing and disseminating research, and communicating with patients [6]. Today, there are an estimated 5.2 billion registered email accounts globally [7], and they send an estimated 220 billion emails per day [8]. Additionally, 72% of internet users now state that they search the internet for health information [9], and 1%-10% of patients utilize email to communicate with their physicians between appointments [10].

The private, corporate, health care, and higher education sectors incorporate email as a foundational mode of modern communication [11,12]. As such, email has become ubiquitous in higher education and has greatly improved the networking and collaborating capabilities of faculty, staff, and students [13,14]. Email is the means of communication preferred by students and faculty owing to its affordability, accessibility, and ability to send accompanying files [15]. Although the benefits of email include simplicity and speediness of communication, its use can involve unwanted outcomes such as uncivil or inconsiderate behavior. For example, hostile and antagonistic email messages containing aggressive comments, insults, and personal attacks have been frequently reported [16]. Students often take for granted the instant access they have to faculty and take up a considerable amount of faculty members' work time by asking for information that has been posted [14].

Cyberincivility is defined as communicative behavior against social norms that is exhibited in computer-mediated interactions, such as those involving email and text messages, or on online social networking sites [17]. Because health professional students who demonstrate cyberincivility in school appear to continue the same behavior after they complete their education [18-20], prelicensure education on email netiquette is especially important. In spite of the need, training in email netiquette is not occurring consistently and is not having consistent results in interprofessional discipline training programs. De Gagne et al [16] noted that only half of nursing students reported receiving information on netiquette, with only 6% being aware of the Nursing Council of State Boards of Nursing guidelines on social media. In addition, Oakley et al [20] found that computer-mediated communication guidelines and some training for dental students did not result in adequate outcomes. A study by Barnhart et al [21] reported that the inclusion of training involving the professional use of email in the curriculum for family medicine residents led to some improvements in communication practices, but unwanted behaviors continued.

Given the increased dependence on email in health care and health professions education, and the risk of undesirable outcomes associated with ineffective email communication, it is imperative that health professions educators prepare their students to engage in appropriate and efficient email netiquette. Our viewpoint paper considers the theoretical foundations of netiquette and cybercivility, as well as relevant literature reviews; its purpose is to promote a culture of cybercivility in health professions education in order to foster responsible and effective use of email in the academic and clinical settings.

Theoretical Foundations of Netiquette and Cybercivility

Netiquette

Netiquette, or internet etiquette, encourages the use of good manners when communicating in cyberspace [22], thereby promoting users to become better cyber citizens. The core roles of netiquette are to provide ethical and moral concepts of right or wrong, as well as operational guidelines [22] for civil behaviors in the online community. Several theoretical or conceptual frameworks have been posited as the underpinning mechanisms and dynamics behind these social phenomena.

Politeness Theory

The politeness theory attempts to explain why people do not always express themselves clearly, directly, or efficiently [23]. According to this theory, people are motivated either by positive faces (ie, a desire to be approved by or connected to others) or negative faces (ie, a desire for disconnection with others or to remain independent) [23]. In order to maintain one's own positive or negative face, an individual must be socially supportive of others' needs or faces. When a person feels intimidated by factors, such as disagreement, criticism, disapproval, and skepticism during social interactions, they can respond with a face-threatening act; variations of this protective mechanism include responding (1) without politeness, (2) with positive politeness, (3) with negative politeness, or (4) indirectly or off-record [23]. Although concern has been expressed that the politeness theory may not account for cultural differences in perception or expectations of politeness, it provides an overall foundation for understanding acts of good manners and civility in linguistic and social structures [23,24].

Social Information Processing Theory

The social information processing (SIP) theory, which was coined by Joseph Walther in 1992 [25], explains how people connect and develop relationships in computer-mediated environments without nonverbal signals [26]. Although it is often believed that insufficient verbal cues make it difficult for people to form inferences about others, SIP theorists posit that people can effectively and intimately build a relationship in cyberspace in the absence of face-to-face interactions [26]. In computer-mediated environments where communication is mainly text based (eg, emails, chat rooms, and instant messaging), people can develop favorable impressions of others by seeking out cues in the messages and by choosing words to counteract the lack of nonverbal cues [26]. From SIP perspectives, the characteristics and rate of information exchanged in cyberspace differ from those in face-to-face environments, which may explain why and how uncivil email communications occur and are perceived. The main challenge to the SIP theory is that people often engage in "hybrid" relationships (neither strictly online or offline) [26]. An understanding of how the dynamics of online and offline communication complement each other could advance the development of email civility strategies and other communication techniques to address challenges in cyberspace.

Awareness to Action Educational Framework

In the current digital age, where there is no defining line between public and private space, the private life of a professional can impact their professional image [27]; similarly, an individual cannot separate how they portray themselves in cyberspace from how their character is perceived [27]. The Awareness to Action (A2A) framework encompasses an assessment (proactive) and a decision-making (reactive) tool to facilitate awareness of what is acceptable or unacceptable and appropriate or inappropriate in online communication, and to help individuals make informed decisions regarding online behavior [28]. The three components of the A2A framework (clarity, context, and confirmability, or the three Cs) require an explicit guideline or policy for application to incidents or events [28]. The three Cs should be considered in sequence, and if clarity is not breached, context and confirmability do not need to be assessed [28]. The main question for clarity is as follows: "Does the behavior explicitly breach policy or guidelines?" The question for context is as follows: "Can you explain or describe the context of the situation or when and where it occurred?" and the question for confirmability is as follows: "Can you confirm the consequences and the outcome?" [28]. The A2A framework can be useful for self-efficacy in promoting email civility as it (1) facilitates the reflection of online behaviors and (2) helps to set norms, consensus, consistency, and evidence for decisions in regard to cybercivility.

Email Use in an Academic Setting

Academic Cyberincivility

Email provides a number of benefits to faculty and students in the academic setting. Although academics in the late 1980s were reluctant to adopt this new method of communication [29], email has now replaced other modes of communication in higher education [14]. Email is used in traditional and web-based learning environments to facilitate class activities, enable mentoring and collaboration, and disseminate course information and assignments [30-32]. Email is also incorporated into educational environments to facilitate learning and engagement [33]. When used as a pedagogical tool, it allows the instructor to facilitate the dissemination of information and to support conversations that would not normally take place during a class session [13]. Moreover, emails sent by course instructors help to motivate students toward successful learning outcomes [34]. The instantaneous and continuous nature of email permits increased interaction between faculty and students, which is crucial for increasing the quality of education and facilitating an effective learning environment [30]. The asynchronous feature of email supports the careful construction of questions and responses by allowing each party to consider their message before sending [34]. Such a delay can benefit shy or reluctant students by (1) removing the competitive nature of classroom discussions, (2) providing time to reflect on the topic, and (3) allowing students to develop a response that demonstrates a higher level of critical and reflective thinking [34], as well as their communication skills and professionalism [13].

Although there are many positive benefits to email in an academic setting, this form of communication can present challenges related to workload and compromised relationships. Both professionals and academics in higher education are overwhelmed by the number of emails they receive and the pressure to respond to the emails immediately [11]. One study found that associate professors and full professors received an average of 84 emails per working day [14]. When faculty receive overly casual messages from students, they may view the senders as less credible and their messages as poor in quality [13]. Instructors may be less likely to comply with a student's request after receiving a causal message [13]. Additionally, the relationships between faculty and students and the resulting learning outcomes are at risk of degradation when inappropriate and misinterpreted messages are exchanged [13]. Uncivil emails from students can lead faculty members to have unpleasant feelings toward them and a decreased willingness to collaborate with them [13]; thus, the impact of incivility on relationships is an important rationale for teaching civil behavior in an online environment.

Email Use Policies in an Academic Setting

Cybercivility is an important component in our increasingly prevalent online interactions and impacts learning in online educational platforms; however, the scope and availability of cybercivility guidelines in US schools of health professions are limited. Email guidelines have been identified for selected schools of health professionals, including a dental school [35] and medical school [36]. A study by De Gagne et al [16] explored the prevalence and composition of cybercivility policies or guidelines regarding email correspondence in US graduate nursing schools (n=230). Only 8% (n=19) of these nursing schools had guidelines for email use. Additionally, best practices for netiquette were found in 84% (n=16) of email guidelines, and 63% (n=12) outlined behaviors to be avoided or reduced. Protocols encouraged a "cooling off" period for emotional and disruptive emails (n=4, 21%) and recommended caution relative to privacy concerns and the potential for miscommunication due to the nonverbal email format (n=9, 47%). Out of 19 guidelines, 9 (47%) specified the consequences for violations. The most common themes found in the guidelines were professionalism, confidentiality/privacy, and forbidden behaviors [16]. These components were also endorsed in previous studies [22,37]. The presence of guidelines for email use can have an impact on the professional and ethical behaviors that are essential for student-faculty relationships [16]; however, guidelines must be accessible and embedded into the curriculum to ensure both awareness and understanding by faculty and students of the professional and ethical behaviors necessary when using email communication [16].

Email Use in a Clinical Setting

Intra- and Interprofessional Collaboration

Since the 1990s, health professionals have used email to communicate among colleagues and to schedule meetings [12]. Today, health professionals use email when collaborating and obtaining consults from other professionals because it conveniently enables the dissemination of information, enhances effective communication, and may facilitate patient care [38]. A survey of oncology physicians found that all respondents had used email to communicate with colleagues, including 78% (n=650) who had received results via email [39]. Furthermore, email was found to facilitate communication between inpatient and outpatient settings and was identified as the preferred method of communication among primary care providers [40]. In 2012, a Cochrane systematic review was conducted to determine the effect of email on clinical care. The systematic review identified a single randomized controlled study showing that emails to physicians positively impacted their osteoporosis guideline adherence [41]. In 2015, the authors conducted another systematic review on the same topic but did not identify any new studies [12].

Email communication among providers has been shown to increase the speed and reliability of communication within an interdisciplinary intensive care team, resulting in improved patient outcomes [42]. One study, which looked at the content of emails exchanged between physicians and nurses, found that the majority of emails were of a nonurgent information-sharing nature and that more than 40% did not require any response [43]. In a study of patients with advanced heart failure and ventricular-assist devices, physicians and pharmacists with established connections used email as an adjunct to face-to-face communication for medication management, enabling the initiation and titration of medication therapy [44]. Similarly, a study of smartphone and email use in the clinical setting revealed that the majority of participants felt that email improved their ability to receive a direct and immediate response from other health care providers [45]; however, the study also reported a potential decrease in interprofessional relationships and an increase in uncivil behaviors by trainees who frequently attended to the device [45]. The use of email among professions can entail similar risks. Information inappropriately or incorrectly shared among health care professionals can result in privacy and confidentiality breaches, as well as medical errors [38]. Email incivility in the intraprofessional setting has not been studied extensively, but Resendes et al [37] noted that unprofessional email communication among health providers can induce negative perceptions of the sender and a delay in response time.

Patient Email Use

Email communication provides a valuable tool for provider-patient interactions when used appropriately and in a secure manner. In fact, many patients prefer communicating with their health care provider via email [20] because it expands opportunities for consultation, treatment, and patient care [38]. Email has been described as environmentally and economically friendly, as well as efficient because it quickly connects the individuals providing and receiving care [46]. In a previous report, 80% of oncology physicians surveyed had communicated with a patient via email [39]. The two most common topics of emails identified by primary care physicians were answering patient questions (82% of respondents; n=219) and changing appointments (72% of respondents; n=192) [47]. Additionally, the use of email can improve the management of chronic diseases and continuity of care because it enables patients to disclose sensitive or embarrassing issues that they might have difficulty discussing face-to-face [48]. Research on the impact of physician email communication has generally been positive with both patients and providers noting convenience and improved quality of care [49-51], although at least one study [52] found that patients preferred telephone or direct communication over email on their military health secure messaging system. In 2018, Wagg et al [53] reviewed 31 studies involving computer-mediated communication (eight of these studies focused on email) and found that 81% (n=26) demonstrated a positive impact on patients. The outcomes noted in this review were increases in access to health care providers, enhancements in communication between patients and providers, improvements in meeting the informational needs of patients, increases in patient empowerment, and improvements in blood sugar control among diabetic patients who received supportive emails. Although the benefits have been noted, physician use of email communication with patients is low compared to both patient and provider willingness (6%-19% vs 70%) [54].

The impact of email on patient-provider communication has been generally positive, but there are concerns about its use. Patients have expressed concerns about whether physicians actually receive their emails, and if so, how quickly [49]. Additionally, socioeconomic indicators of patients have been identified as barriers to email use [49,51]. Makarem and Antoun [55] described much lower use of email communication between patients and physicians in developing countries, as well as differences in how patients and physicians view the use of email and its importance. Physicians using email identified workload, lack of reimbursement for time responding to emails, and inappropriate use of email by some patients as barriers to its effective role [49-51].

The ease and speed of email communication can result in unprofessional and miscommunicated messages [46]. The negligent construction of email messages (eg, no subject lines, no proper salutations, excessive lingo usage, and slang) can negatively impact the professional rapport physicians must maintain with their patients [38]. Missing or inappropriate information provided by patients may cause confusion or a delay in treatment [48]. Furthermore, research demonstrates that the content and tone of emails between patients and providers is generally task-oriented and focuses on nonurgent health-related issues, although some content relates to emotional needs and relationship building [49,56]. While Hogan et al [56] found tone and content to be generally positive, the need for patient-centered improvements and proactive communication by providers was noted. Patients and physicians have shared concerns about confidentiality and security [49-51]; however, Mold et al [51] found no harm or privacy violations in a systematic review of 17 studies.

Organizational Use of Email

Email has the potential to impact several aspects of organizations. First, health care organizations use email to conduct routine business tasks and value it for its ability to easily share documents, facilitate collaboration and workflow, and hold workers accountable [57]. Second, email facilitates organization information sharing both intraorganizationally and extraorganizationally. Within the organization, email has been used to disseminate evidence-based practice information [58]. A case study by Medland [59] demonstrated that email can be used as a leadership tool to promote connection, build competency, and increase coworkers' sense of being valued. Outside the organization, the Veteran Administration conducted a survey to better understand how veterans would like to receive information in the event of a natural disaster, and veterans ranked email messages as one of the top three helpful communication modes for those less than 64 years old [60]. Third, email has been utilized by organizations as part of quality improvement efforts. An organization embedded email reminders in electronic health records to improve admission medication reconciliation by resident physicians [61]. Email consults based on a template have also been implemented for professionals ranging from primary care providers to specialists within a national health system, showing a decrease in wait times and reduced cost [62]. Fourth, the use of email has been studied as a means of improving processes from a research perspective. Another hospital developed an automated system to send emails to physicians for tests pending at patient discharge to improve follow-up care [63].

Email Policies in Health Care

While professional standards have been established for professional communication in the health care sector [46], as well as behavioral health provider standards or guidelines [64], lack of training or guidelines on electronic or email communication for health care professionals has been clearly documented [16,20,46]. It has been suggested that an understanding of how to use email does not necessarily ensure appropriate professional communication [20,65,66]. Accordingly, studies on email communication across health care disciplines, including nursing [16,17], medicine [27,37,38,46,67,68], and mental health [69], have called for the development of policies and guidelines to enhance this form of communication. Malka et al [38] and Railey et al [46] found a lack of formal guidelines for email use by physicians. Guidelines for patient-physician email and text message exchanges have been recently published by the American Medical Association [70] to address such issues, including (1) establishing a turnaround time for messages, (2) retaining copies of email communications with patients, and (3) refraining from sending angry, sarcastic, harshly critical, and libelous references to third parties [70]. In addition to professional organizations, some health care systems have created policies to guide electronic communication among their employees [71,72].

Recommendations

Evidence-Based Educational Strategies

We concur with the interventional strategies for integrating cybercivility into the following areas of health professions education proposed by De Gagne et al [73]: (1) ethical knowledge and skills, (2) curriculum development and content delivery, and (3) praxis. Their study demonstrated that students in health professions lack knowledge on e-professionalism and would benefit from online resources that facilitate reflective discussions. They recommend the following: (1) integrating cybercivility and digital communication into course curriculum to facilitate formal assessment and evaluation of these learning objectives; (2) evaluating content to ensure it is accessible, feasible, and effective; (3) incorporating writing and reflective exercises to help uncover and make visible any unwritten or unintended "hidden curriculum;" (4) providing training at both individual and organizational levels; and (5) promoting partnerships and faculty development through cybercivility training, including interprofessional training for currently employed health care professionals [56].

Existing Models

Faculty can use models that have already been developed in their teaching procedures. Railey et al [46] created the SURE model with key questions for health care providers and students to use when composing an email. In this model, S is related to checking spelling and syntax, and including a subject and signature; U is related to urgency and an unprofessional tone; R is related to reviewing for content and confirming a recipient; and E is related to emotions and ethical concerns [46]. The previously described A2A framework by Ryan [28] is another potential model. Regardless of the models or teaching approaches employed, curriculum for both professional development and health professions education should include the content outlined in the next four subsections of this paper.

Basic Email Etiquette

Poorly constructed or uncivil email communications, including those that leave out relevant information (eg, subject line or name) [14], disregard spelling and grammar errors, insert inappropriate abbreviations and slang, or use informal and impolite tones [13,74], can damage the credibility of the sender and cause the reader to underestimate the sender's competency [13]. While working to understand the unfavorability of selected characteristics in professional emails, researchers identified nonwhite background color, hard to read fonts, and lack of a subheading as the top three unfavorable characteristics that make recipients less likely to reply [37]. Using the "reply all" button when it is not absolutely necessary can also create frustration for those to whom the topic does not apply. Some examples of statements that should never be sent as "reply all" are as follows: (1) Congrats! (2) Thank you! (3) I agree, (4) Please remove me from this mailing list, (5) LOL, and (6) Please stop reply all to this thread [75].

Ethical, Legal, and Professional Implications

The importance of writing emails that reflect and communicate professional and ethical values must be emphasized in educational curriculum and practice guidelines. The effects of inappropriate communication on student-faculty relationships, patient-provider relationships, and the reputation of professional sectors within the larger global community are far reaching. Faculty, staff, and students should be aware of how email communication is impacted by federal, state, and local laws; in particular, students and clinicians should be aware of what types of patient information must be excluded from emails in order to comply with the Health Insurance Portability and Accountability Act. Faculty should carefully consider any potential violations of the Family Educational Rights and Privacy Act when communicating with and about students via email.

Clear Guidelines for Expectations and Repercussions for Infractions

Cain and Romanelli [27] recommended honor codes and professional socialization as potential techniques to improve professional email communication. Formal guidelines should include clear direction on e-professionalism, cybercivility, and cyberincivility, as well as specific behaviors to be avoided, expectations for conduct, and consequences for inappropriate communication [16]. It is important to define and address the consequences of uncivil behaviors and to provide educational resources for both those enacting uncivil behaviors and their victims [73]. The development of a common set of guidelines and standards across disciplines could be an important collaborative opportunity pioneered by interprofessional health educators.

Research

Additional research is needed to provide an increased understanding of email incivility in the clinical and academic settings. Current faculty and professional development approaches for educating students and other professionals on email netiquette should be evaluated in order to identify strategies that enhance learning, encourage behavioral change, and enforce guidelines. Further research is needed to explore the use of email in interprofessional communication, as well as the development of conceptual models, according to the theories discussed in this paper, to support curriculum inclusion and guideline development [23,25]. Finally, additional studies are needed to (1) understand the effective use of email communication between health care providers and patients and (2) identify the types of information and interactions that are most effective for specific populations.

Conclusions

There is agreement in the literature that email incivility has a negative impact on student-faculty, faculty-faculty, and patient-provider relationships; however, interventions to address this problem in health professions education have not been well documented. This paper discussed current knowledge from a review of the literature on theoretical foundations of cybercivility and made recommendations for strengthening email netiquette. Fostering e-professionalism requires the cultivation of not only knowledge and skill but also ethical and moral reasoning. Given the importance of web-based learning platforms and digital communication, the need for effective strategies for educators is paramount.

Acknowledgments

This work was supported by a 2018-2019 Duke Academy for Health Professions Education and Academic Development Supporting Health Professions Educators grant awarded to JCD. The authors would like to thank the Duke University Compact for Open Access Publishing Equity program for its support on the open access publication of this manuscript.

Conflicts of Interest

None declared.

References

1. Interprofessional Education Collaborative. 2016. Core Competencies for Interprofessional Collaborative Practice: 2016 Update URL: https://nebula.wsimg.com/2f68a39520b03336b41038c370497473?AccessKeyId=DC06780E69ED19E2B3A5&disposition=0&alloworigin=1 [accessed 2020-03-20]
2. Schmitt M, Blue A, Aschenbrener CA, Viggiano TR. Core competencies for interprofessional collaborative practice: reforming health care by transforming health professionals' education. Acad Med 2011 Nov;86(11):1351. [doi: 10.1097/ACM.0b013e3182308e39] [Medline: 22030650]
3. Norfolk ER, Glazier J, Hamilton E, Maddux FW. Electronic communication and text messaging among nephrology providers. Am J Kidney Dis 2013 Oct;62(4):656-658. [doi: 10.1053/j.ajkd.2013.04.024] [Medline: 23820067]
4. Briddock D. Shibboleth Authentication Request. Tech Origins Email URL: http://www.libproxy.wvu.edu/login?url=https://search-proquest-com.www.libproxy.wvu.edu/docview/1516515275?accountid [accessed 2020-01-03]
5. Tao L, Reinking D. What Research Reveals about Email in Education. 1996 Nov 03 Presented at: Annual Meeting of the College Reading Association; November 1996; Charleston, SC.
6. Moyer CA, Stern DT, Katz SJ, Fendrick AM. "We got mail": electronic communication between physicians and patients. Am J Manag Care 1999 Dec;5(12):1513-1522 [FREE Full text] [Medline: 11066618]
7. The Radacati Group. 2014. Email Statistics Report, 2014-2018 URL: http://www.radicati.com/wp/wp-content/uploads/2014/01/Email-Statistics-Report-2014-2018-Executive-Summary.pdf [accessed 2020-05-05]
8. Internet Live Stats. Internet Users URL: http://www.internetlivestats.com/internet-users/ [accessed 2020-05-05]
9. Fox S, Duggan M. Pew Research Center. Health Online 2013 URL: https://www.pewresearch.org/internet/2013/01/15/health-online-2013/ [accessed 2020-05-05]
10. Welch W, Mathew M, Welch R, McShane B. Email as an Encumbrance to Physicianpatient Communication. Cureus 2019;11(1):e3816. [doi: 10.7759/cureus.3816]

11. Chase NM, Clegg B. Effect of email utilization on higher education professionals. Int J Technol Hum Interact 2011;7(4):31-45. [doi: 10.4018/jthi.2011100103]
12. Goyder C, Atherton H, Car M, Heneghan CJ, Car J. Email for clinical communication between healthcare professionals. Cochrane Database Syst Rev 2015 Feb 20(2):CD007979. [doi: 10.1002/14651858.CD007979.pub3] [Medline: 25698124]
13. Kim D, Yoon HB, Yoo D, Lee S, Jung H, Kim SJ, et al. Etiquette for medical students' email communication with faculty members: a single-institution study. BMC Med Educ 2016 Apr 27;16:129 [FREE Full text] [doi: 10.1186/s12909-016-0628-y] [Medline: 27121179]
14. Pignata S, Lushington K, Sloan J, Buchanan F. Employees' perceptions of email communication, volume and management strategies in an Australian university. Journal of Higher Education Policy and Management 2015 Mar 26;37(2):159-171. [doi: 10.1080/1360080x.2015.1019121]
15. Willemse JJ, Bozalek V. Exploration of the affordances of mobile devices in integrating theory and clinical practice in an undergraduate nursing programme. Curationis 2015;38(2):e1-e10 [FREE Full text] [doi: 10.4102/curationis.v38i2.1510] [Medline: 26852423]
16. De Gagne JC, Yamane SS, Conklin JL, Chang J, Kang HS. Social media use and cybercivility guidelines in U.S. nursing schools: A review of websites. J Prof Nurs 2018;34(1):35-41. [doi: 10.1016/j.profnurs.2017.07.006] [Medline: 29406136]
17. De Gagne JC, Choi M, Ledbetter L, Kang HS, Clark CM. An Integrative Review of Cybercivility in Health Professions Education. Nurse Educ 2016;41(5):239-245. [doi: 10.1097/NNE.0000000000000264] [Medline: 27022683]
18. Klainberg MB, McCrink A, Eckardt P, Schecter R, Bongiorno A, Sedhom L. Perspectives on academic misconduct: implications for education and practice. J N Y State Nurses Assoc 2014;44(1):11-21. [Medline: 26021131]
19. Woodworth JA. Promotion of Nursing Student Civility in Nursing Education: A Concept Analysis. Nurs Forum 2016 Jul;51(3):196-203. [doi: 10.1111/nuf.12138] [Medline: 26488502]
20. Oakley M, Horvath Z, Weinberg SM, Bhatt J, Spallek H. Comparison of e-mail communication skills among first- and fourth-year dental students. J Dent Educ 2013 Nov;77(11):1413-1424. [Medline: 24192406]
21. Barnhart A, Lausen H, Smith T, Lopp L. Electronic health communication: an educational application for this principle of the Patient-centered Medical Home. Fam Med 2010 May;42(5):322-326 [FREE Full text] [Medline: 20455107]
22. Ross ST. Albion.com. 2011. Netiquette Home Page URL: http://www.albion.com/netiquette [accessed 2020-01-20]
23. Goldsmith D, Normand E. Politeness theory: how we use language to save face. In: Braithwaite DO, Schrodt P, editors. Engaging theories of interpersonal communication: multiple perspectives. Thousand Oaks, CA: Sage; 2014:267-278.
24. Dunn CD. Formal forms or verbal strategies? Politeness theory and Japanese business etiquette training. Journal of Pragmatics 2011 Dec;43(15):3643-3654. [doi: 10.1016/j.pragma.2011.06.003]
25. Walther JB, Burgoon JK. Relational Communication in Computer-Mediated Interaction. Human Comm Res 1992 Sep;19(1):50-88. [doi: 10.1111/j.1468-2958.1992.tb00295.x]
26. Walther JB. Social information processing theory: impressionsrelationship development online. In: Braithwaite DO, Schrodt P, editors. Engaging theories of interpersonal communication: multiple perspectives, 2nd ed. Thousand Oaks, CA: Sage; 2014:417-428.
27. Cain J, Romanelli F. E-professionalism: a new paradigm for a digital age. Currents in Pharmacy Teaching and Learning 2009 Dec;1(2):66-70. [doi: 10.1016/j.cptl.2009.10.001]
28. Ryan GS. E-professionalismnursing education. In: Dyson S, McAllister M, editors. Routledge international handbook of nurse education. London: Routlegde; 2019:269-396.
29. Komsky SH. A Profile of Users of Electronic Mail in a University. Management Communication Quarterly 1991;4(3):310-340. [doi: 10.1177/0893318991004003003]
30. Cung B, Xu D, Eichhorn S. Increasing Interpersonal Interactions in an Online Course: Does Increased Instructor E-mail Activity and a Voluntary In-Person Meeting Time Facilitate Student Learning? OLJ 2018 Sep 01;22(3):193-215. [doi: 10.24059/olj.v22i3.1322]
31. Jacobs K, Doyle N, Ryan C. The nature, perception, and impact of e-mentoring on post-professional occupational therapy doctoral students. Occup Ther Health Care 2015 Apr;29(2):201-213. [doi: 10.3109/07380577.2015.1006752] [Medline: 25821883]
32. Mubuan R, Ebron G. A blended learning approach to teaching writing: using e-mail in the ESL classroom. Asian EFL Journal 2019(100):80-103.
33. Rohrer Vitek CR, Nicholson WT, Schultz C, Caraballo PJ. Evaluation of the use of clinical decision support and online resources for pharmacogenomics education. Pharmacogenomics 2015;16(14):1595-1603. [doi: 10.2217/pgs.15.100] [Medline: 26419532]
34. Lewandowski K. Distance learning by exam as support of traditional method of education. Int J Electron Telecommun 2015;61(3):231-235. [doi: 10.1515/eletel-2015-0030]
35. Arthur A. Dugoni School of Dentistry. Email Guidelines URL: https://www.dental.pacific.edu/departments-and-groups/information-technology/school-technology-resources/email-guidelines [accessed 2020-01-19]
36. The George Washington University. Social Media and Email Policy for School of Medicine and Health Sciences URL: https://tinyurl.com/yb5wfofq [accessed 2020-01-19]

37. Resendes S, Ramanan T, Park A, Petrisor B, Bhandari M. Send it: study of e-mail etiquette and notions from doctors in training. J Surg Educ 2012;69(3):393-403. [doi: 10.1016/j.jsurg.2011.12.002] [Medline: 22483143]
38. Malka ST, Kessler CS, Abraham J, Emmet TW, Wilbur L. Professional e-mail communication among health care providers: proposing evidence-based guidelines. Acad Med 2015 Jan;90(1):25-29. [doi: 10.1097/ACM.0000000000000465] [Medline: 25162617]
39. Cook N, Maganti M, Dobriyal A, Sheinis M, Wei AC, Ringash J, et al. E-Mail Communication Practices and Preferences Among Patients and Providers in a Large Comprehensive Cancer Center. J Oncol Pract 2016 Jul;12(7):676-684. [doi: 10.1200/JOP.2015.008722] [Medline: 27352950]
40. Holleck JL, Gunderson CG, Antony SM, Gupta S, Chang JJ, Merchant N, et al. The "Hand-in" Project: Jump-starting Communication Between Inpatient and Outpatient Providers. South Med J 2017 Nov;110(11):694-698. [doi: 10.14423/SMJ.0000000000000724] [Medline: 29100218]
41. Pappas Y, Atherton H, Sawmynaden P, Car J. Email for clinical communication between healthcare professionals. Cochrane Database Syst Rev 2012 Sep 12(9):CD007979. [doi: 10.1002/14651858.CD007979.pub2] [Medline: 22972116]
42. O'Connor C, Friedrich JO, Scales DC, Adhikari NK. The Use of Wireless E-Mail to Improve Healthcare Team Communication. Journal of the American Medical Informatics Association 2009 Sep 01;16(5):705-713. [doi: 10.1197/jamia.m2299]
43. Smith CN, Quan SD, Morra D, Rossos PG, Khatibi H, Lo V, et al. Understanding interprofessional communication: a content analysis of email communications between doctors and nurses. Appl Clin Inform 2012;3(1):38-51 [FREE Full text] [doi: 10.4338/ACI-2011-11-RA-0067] [Medline: 23616899]
44. Lekura J, Tita C, Lanfear DE, Williams CT, Jennings DL. Assessing the potential of e-mail for communicating drug therapy recommendations to physicians in patients with heart failure and ventricular-assist devices. J Pharm Pract 2014 Oct;27(5):478-480. [doi: 10.1177/0897190013513618] [Medline: 24319078]
45. Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, et al. An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study. J Med Internet Res 2011 Aug 29;13(3):e59 [FREE Full text] [doi: 10.2196/jmir.1655] [Medline: 21875849]
46. Railey K, Tuttle B, Weiss J. Are You SURE You Want to Send That? A Model for Email Professionalism in Medical Education. J Physician Assist Educ 2017 Dec;28(4):205-209. [doi: 10.1097/JPA.0000000000000165] [Medline: 29189651]
47. Dash J, Haller DM, Sommer J, Junod Perron N. Use of email, cell phone and text message between patients and primary-care physicians: cross-sectional study in a French-speaking part of Switzerland. BMC Health Serv Res 2016 Oct 05;16(1):549 [FREE Full text] [doi: 10.1186/s12913-016-1776-9] [Medline: 27716256]
48. Atherton H, Sawmynaden P, Sheikh A, Majeed A, Car J. Email for clinical communication between patients/caregivers and healthcare professionals. Cochrane Database Syst Rev 2012 Nov 14;11:CD007978. [doi: 10.1002/14651858.CD007978.pub2] [Medline: 23152249]
49. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. Patient Educ Couns 2010 Aug;80(2):266-273 [FREE Full text] [doi: 10.1016/j.pec.2009.09.038] [Medline: 19914022]
50. Macdonald GG, Townsend AF, Adam P, Li LC, Kerr S, McDonald M, et al. eHealth Technologies, Multimorbidity, and the Office Visit: Qualitative Interview Study on the Perspectives of Physicians and Nurses. J Med Internet Res 2018 Jan 26;20(1):e31 [FREE Full text] [doi: 10.2196/jmir.8983] [Medline: 29374004]
51. Mold F, de Lusignan S. Patients' Online Access to Their Primary Care Electronic Health Records and Linked Online Services: Implications for Research and Practice. J Pers Med 2015 Dec 04;5(4):452-469 [FREE Full text] [doi: 10.3390/jpm5040452] [Medline: 26690225]
52. Hernandez BF, Morgan BJ, ISh J, Agbator LO, Lindo-Moon S, Stotler FF. Communication preferences and satisfaction of secure messaging among patients and providers in the military healthcare system. Military Medicine 2018;183:e383-e390. [doi: 10.1093/milmed/usy094]
53. Wagg AJ, Callanan MM, Hassett A. The use of computer mediated communication in providing patient support: A review of the research literature. Int J Nurs Stud 2018 Jun;82:68-78. [doi: 10.1016/j.ijnurstu.2018.03.010] [Medline: 29609154]
54. Antoun J. Electronic mail communication between physicians and patients: a review of challenges and opportunities. Fam Pract 2016 Apr;33(2):121-126. [doi: 10.1093/fampra/cmv101] [Medline: 26711957]
55. Makarem NN, Antoun J. Email communication in a developing country: different family physician and patient perspectives. Libyan J Med 2016;11:32679 [FREE Full text] [doi: 10.3402/ljm.v11.32679] [Medline: 27855773]
56. Hogan TP, Luger TM, Volkman JE, Rocheleau M, Mueller N, Barker AM, et al. Patient Centeredness in Electronic Communication: Evaluation of Patient-to-Health Care Team Secure Messaging. J Med Internet Res 2018 Mar 08;20(3):e82 [FREE Full text] [doi: 10.2196/jmir.8801] [Medline: 29519774]
57. Gill B. Harvard Business Review. 2013. Email: not dead, evolving URL: https://hbr.org/2013/06/e-mail-not-dead-evolving [accessed 2020-02-05]
58. Harris C, Garrubba M, Melder A, Voutier C, Waller C, King R, et al. Sustainability in Health care by Allocating Resources Effectively (SHARE) 8: developing, implementing and evaluating an evidence dissemination service in a local healthcare setting. BMC Health Serv Res 2018 Mar 02;18(1):151 [FREE Full text] [doi: 10.1186/s12913-018-2932-1] [Medline: 29499702]

59. Medland J. Leveraging Email to Build Healthy Effective Work Teams. Nurse Leader 2020 Apr. [doi: 10.1016/j.mnl.2020.01.010]
60. Lukowsky LR, Der-Martirosian C, Dobalian A. Seeking Information from the VA During Natural Disasters. Hosp Top 2018;96(4):114-122. [doi: 10.1080/00185868.2018.1521251] [Medline: 30588867]
61. Johnson K, Burkett GS, Nelson D, Chen AR, Matlin C, Garger C, et al. Automated E-mail Reminders Linked to Electronic Health Records to Improve Medication Reconciliation on Admission. Pediatr Qual Saf 2018;3(5):e109 [FREE Full text] [doi: 10.1097/pq9.0000000000000109] [Medline: 30584636]
62. Patterson V, Humphreys J, Henderson M, Crealey G. Email triage is an effective, efficient and safe way of managing new referrals to a neurologist. Qual Saf Health Care 2010 Oct;19(5):e51. [doi: 10.1136/qshc.2009.038513] [Medline: 20511602]
63. Dalal AK, Schnipper JL, Poon EG, Williams DH, Rossi-Roh K, Macleay A, et al. Design and implementation of an automated email notification system for results of tests pending at discharge. J Am Med Inform Assoc 2012;19(4):523-528 [FREE Full text] [doi: 10.1136/amiajnl-2011-000615] [Medline: 22268214]
64. Reamer FG. Evolving standards of care in the age of cybertechnology. Behav Sci Law 2018 Mar;36(2):257-269. [doi: 10.1002/bsl.2336] [Medline: 29659068]
65. Essary AC. The impact of social media and technology on professionalism in medical education. J Physician Assist Educ 2011;22(4):50-53. [doi: 10.1097/01367895-201122040-00009] [Medline: 22308935]
66. Thompson LA, Dawson K, Ferdig R, Black EW, Boyer J, Coutts J, et al. The intersection of online social networking with medical professionalism. J Gen Intern Med 2008 Jul;23(7):954-957 [FREE Full text] [doi: 10.1007/s11606-008-0538-8] [Medline: 18612723]
67. Cain J, Linos E, Chretien KC. Cyberbullying in Academic Medicine: A Framework for Managing Social Media Attacks. Academic Medicine 2019;94(5):626-629. [doi: 10.1097/acm.0000000000002649]
68. Spector ND, Matz PS, Levine LJ, Gargiulo KA, McDonald MB, McGregor RS. e-Professionalism: challenges in the age of information. J Pediatr 2010 Mar;156(3):345-346. [doi: 10.1016/j.jpeds.2009.12.047] [Medline: 20176180]
69. Deen SR, Withers A, Hellerstein DJ. Mental health practitioners' use and attitudes regarding the Internet and social media. J Psychiatr Pract 2013 Nov;19(6):454-463. [doi: 10.1097/01.pra.0000438184.74359.88] [Medline: 24241499]
70. American Medical Association. Guidelines for patient-physician electronic mail and text messaging 2018 URL: https://policysearch.ama-assn.org/policyfinder/detail/Guidelines%20for%20Patient-Physician%20Electronic%20Mail%20and%20Text%20Messaging%20H-478.997?uri=%2FAMADoc%2FHOD.xml-0-4344.xml [accessed 2020-05-05]
71. The Newcastle upon Tyne Hospitals NHS Foundation Trust. 2010. Email and electronic communications policy URL: http://www.newcastle-hospitals.org.uk/downloads/policies/Information%20Technology/EmailPolicy200705.pdf [accessed 2020-05-04]
72. North Bay General Hospital. 2008. Appropriate Use Of Internet and Email URL: http://www.nbrhc.on.ca/mainstreetblog/wp-content/uploads/2014/05/Internet-EmailPolicy.pdf [accessed 2020-02-01]
73. De Gagne JC, Yamane SS, Conklin JL. Evidence-based strategies to create a culture of cybercivility in health professions education. Nurse Educ Today 2016 Oct;45:138-141. [doi: 10.1016/j.nedt.2016.07.007] [Medline: 27472429]
74. Lewin-Jones J, Mason V. Understanding style, language and etiquette in email communication in higher education: a survey. Research in Post-Compulsory Education 2014 Mar 05;19(1):75-90. [doi: 10.1080/13596748.2014.872934]
75. Osman H. Don't reply all: 18 email tactics that help you write better emails and improve communication with your team. Scotts Valley, CA: CreateSpace Independent Publishing; 2015:18-58.

Abbreviations

A2A: Awareness to Action
SIP: social information processing

Edited by G Eysenbach; submitted 12.04.20; peer-reviewed by A Ramazan; comments to author 01.05.20; revised version received 09.05.20; accepted 14.05.20; published 01.06.20.

Please cite as:
De Gagne JC, Yang Y, Rushton S, Koppel PD, Hall K
Email Use Reconsidered in Health Professions Education: Viewpoint
JMIR Med Educ 2020;6(1):e19300
URL: http://mededu.jmir.org/2020/1/e19300/
doi:10.2196/19300
PMID:32478659

Original Paper

Current and Future Trends in Life Sciences Training: Questionnaire Study

William Magagna[1*], MEd; Nicole Wang[2*], MSED; Kyle Peck[2*], PhD

[1]Siemens Healthineers, Newark, DE, United States

[2]Department of Learning Performance System, Pennsylvania State University, University Park, PA, United States

[*]all authors contributed equally

Corresponding Author:
Nicole Wang, MSED
Department of Learning Performance System
Pennsylvania State University
Keller Building
University Park, PA
United States
Phone: 1 2154072119
Email: nwang@psu.edu

Abstract

Background: Every year, the life science field spends billions of dollars on educational activities worldwide. The continuing professional development of employees, especially in this field, encompasses great challenges. Emerging technologies appear to offer opportunity, but relatively little research has been done on the effectiveness of pedagogies and tools that have been used in the life sciences, and even less research has been devoted to understanding the potential power of emerging options that might determine the field's future.

Objective: In collaboration with the Life Sciences Trainers & Educators Network (LTEN), this study investigated the current state of the pedagogies and tools currently adopted by corporate training professionals in the life sciences as well as the professionals' perceptions of the impacts of emerging technologies on training.

Methods: This study adopted a mixed methods approach that included a survey and a follow-up interview. The survey consists of 18 broad questions with 15 subquestions in each of the five specific sectors of the life sciences field. Interviews were conducted by phone and lasted approximately 40 minutes, covering 18 questions designed to follow-up on findings from the survey items.

Results: Both survey and interview results indicated that the professionals were not satisfied with the status quo and that training and education in this field need to change. Most of the techniques and tools currently used have been used for some time. The professionals surveyed were not satisfied with the current techniques and tools and did not find them cost-effective. In addition, the respondents pictured the future of training in this field to be more engaging and effective.

Conclusions: This is the first study in a series designed to better understand education and training in the life sciences on a macro level, in order to build a foundation for progress and evolution of the future landscape. Next steps involve developing strategies for how to extend this vision throughout individual organizations.

(JMIR Med Educ 2020;6(1):e15877) doi:10.2196/15877

KEYWORDS

professional training; training with technologies; life sciences professionals; mixed methods

Introduction

Background

The life sciences have been evolving at a staggering rate in every aspect. According to Kaufman [1], "Medicine has gone through major changes over the last 50 years. Today it is recognized that medical knowledge doubles every 6-8 years, with new medical procedures emerging every day." In gene ontology, there are about 250 ontologies accessible to professionals in the field [2]. In addition, new services, technologies, and applications have emerged through the evolution of life sciences education [3,4]. For instance, studies have suggested the use of artificial intelligence in improving medical imaging [5,6] and automating medical diagnosis and prediction [7,8].

Recent advances in technology have also drastically changed how people teach and learn. As Bonk et al [9] mentioned, "Recent technological developments have converged to dramatically alter conception of teaching and learning process." These advancements offer innovative approaches to teaching and learning professionals within the life sciences field. For instance, Garcia-Pallares et al [10] and Tymcznska [11] discussed approaches in which instructors incorporate a course management system into life sciences professional training. In addition, Mantovani [12], Stansfield et al [13], and Barsom et al [14] addressed the potential to integrate virtual reality (VR) into training to enhance and improve learning experiences. As the rapid development and application of artificial intelligence (AI) continues [15], opportunities for adopting AI into training in the life sciences field will also become increasingly evident.

Although the fields of life sciences, teaching, and learning are moving forward, professional development of people in the life sciences field seems to fall behind. Gorman et al [2] pointed out that current and even future health care professionals are trained using the 100-year-old apprenticeship model, which is "see one, do one, and teach one." In addition to the apprenticeship model, common strategies include lectures and films as the basis of a life sciences professional's training, even though billions of dollars are dedicated to continuing educational activities in the life sciences worldwide [16,17]. When it comes to professional training in the life sciences, studies [18] indicate the similarities of the constant need for learning new knowledge, skills, and attitudes required due to the complexity of the field.

At presently, literature exploring the current state of teaching and learning in the life sciences field is sparse, especially within the professional training realm. Additionally, literature exploring the cost-effectiveness of these emerging approaches for training medical professionals is almost nonexistent. Although there is an underlying assumption that there is a direct relationship between continuing professional development and the performance of recipients, only a few studies have attempted to validate this assumption. According to Bloom et al, [19] and Umble and Cervero [20], continuing professional development can be effective, but its effectiveness varies.

In addition, a plethora of studies [21-24] have explored how emerging technologies can change the learning landscape across sectors. For instance, Dubey and Gunasekaran [25] investigated how AI can impact the transportation sector, and Gavish et al [26] explored how augmented reality (AR) and VR can transform industrial training. However, few studies have explored how emerging technologies can impact professional training and education in the life sciences field.

Objectives

There is a limited amount of research into the educational tools and approaches currently employed within the life sciences. Furthermore, there is a lack of studies investigating how life sciences training might evolve under the influence of emerging technologies and the increasing emphasis on cost-effectiveness. This study aims to understand the current state of teaching and learning in the life sciences, and teaching professionals' perceptions of the impact of new technologies and practices on the field. Specifically, this study will investigate the following:

- What technologies and pedagogies are educational professionals in Life Sciences Trainers & Educators Network (LTEN) member organizations using now?
- Which currently used approaches are most cost-effective and judged as most satisfactory by training and education professionals in the life sciences?
- How do life science training and education professionals think emerging technologies might change current practice in the near future?

Methods

Overview

This study used a mixed methods approach [27] to better understand the current state of teaching and learning in life sciences and training professionals' perceptions of the impacts of emerging technologies. Mixed methods research requires data triangulation from quantitative and qualitative approaches, which strengthens the construct validity of the study [28]. In addition, 57 members from a life sciences education not-for-profit organization, LTEN, participated in the survey, and 9 participants who responded to the survey were interviewed. In compliance with the Pennsylvania State University Institutional Review Board protocols, all participants signed the informed consent release form.

Quantitative Method

The survey consists of 18 broad questions with 15 subquestions in five specific sectors of the life sciences field (sales, clinical, product-related, customer-related, and other). Questions included demographic information along with detailed questions on the use and perceptions of pedagogies and tools.

Qualitative Method

The primary data were collected through semistructured interviews. Interviews were conducted by phone and lasted approximately 40 minutes, covering 18 questions designed to follow-up findings from the survey questions. The interview protocol was designed based on the theoretical framework proposed by Seidmen [29], which consists of three general genres: personal experiences of emerging technology, attitudes toward specific technologies, and future expectations of emerging technologies. Researcher memos also served as a secondary data source [30].

Results

Demographics

Survey respondents represented the diversity of the LTEN membership. Founded in 1971, LTEN has grown to more than 1900 individual members who work in pharmaceuticals, biotech, medical device, and diagnostic companies, and industry partners who support the life sciences training departments. Additionally, LTEN has members across noncommercial disciplines including clinical, manufacturing, compliance, regulatory, quality, and general practice training roles [31]. This study invited 326 active members who are directly involved in the training department of member organizations, to participate in the survey through email. A total of 57 participants completed the survey.

As shown in Table 1, of the 57 participants, 24 (42%) of the respondents were education/training directors, 15 (27%) were corporate executives, 11 (19%) were education/training managers, and 7 (13%) were training developers. The respondents also had diverse responsibilities within their organizations (Table 2), with 23 (40%) working at US commercial-only organizations and 15 (26%) working at an entirely global organization.

Table 1. Respondents' role.

Role	Respondents, n (%)[a]
Education/training developer	7 (13)
Education/training manager	11 (19)
Education/training director	24 (42)
Corporate executive responsible for education and training	15 (27)

[a]Percentages may not add up to 100 due to rounding.

Table 2. Respondents' responsibilities.

Responsibility	Respondents, n (%)
US commercial operations only	23 (40)
An entire US organization	10 (18)
Global organization, but commercial operations only	9 (16)
An entire global organization	15 (26)

Quantitative Results

Current Pedagogies Identified

A set of survey questions asked what pedagogies respondents currently use as their teaching strategies. Once they indicated their pedagogies, the respondents were asked to rank their selected pedagogies in order of importance. Fourteen pedagogies were presented as choices in the survey, and a weighted vote methodology was used to compare the pedagogies most commonly used and those perceived to be most important. A weighted ranking was produced by assigning a rank of 14 points to the item identified as most important, a score of 13 to the second most important, and so forth. The process was repeated for each category of trainee and for each training topic category, and a sum was calculated within each category of trainee and topic. The number of all the ranked items in each topic was used to produce a weighted percentage of the approaches used, which was divided by the total of all scores.

As shown in Table 3, there were 4 pedagogies (instructor-led training, virtual instructor-led training, online readings, and role play activities) that captured about 55% of the weighted importance rankings, with the other 10 pedagogies comprising the other half.

Table 3. Respondent-ranked importance of pedagogies.

Pedagogies	Weighted percentages[a]
Instructor-led training	17
Virtual instructor-led training	10
Role play activities	10
Competency-based learning	9
Case studies	9
Simulations	8
Field-based activities	8
Online readings	8
Collaborative learning	7
Problem-based learning	6
Games	4
Online discussions	2
Project-based learning	2
Other (please indicate)	1

[a]Percentages may not add up to 100 due to rounding.

Additionally, we identified some relationships in the differences between the pedagogies that respondents generally use and those they judged as most important. Figure 1 shows the top 5 ranked pedagogies in terms of use and importance (shown with numbers representing their ranks in each category) in sales across 3 training topics (clinical, product, and skills). Across all 3 topics, instructor-led training (ILT) is perceived to be the most used and most important pedagogy for salespeople, and virtual instructor-led training (VILT) methods are among the top 3 in terms of use but are much lower in terms of importance in product and skills topics.

Figure 1. Weighted comparison between pedagogy use and importance in sales. ILT: instructor-led training; VILT: virtual instructor-led training.

Clinical topics		Product topics		Skills topics	
Weighted percent of approaches used	Weighted importance	Weighted percent of approaches used	Weighted importance	Weighted percent of approaches used	Weighted importance
1. ILT	1. ILT	1. ILT	1. ILT	1. ILT	1.ILT
2. VILT	2. VILT	2. Role play	2. Competency-based	2. Role play	2.Role play
3. Online readings	3. Role play	3. VILT	3. Case studies	3. VILT	3.Simulations
4. Role play	4. Online readings	4. Online readings	4. Simulations	4. Simulations	4.Collaborative learning
5. Case studies	5. Problem-based learning	5. Simulations	5. Collaborative learning	5. Case studies	5.Competency-based

Cost-effectiveness of Current Pedagogies Used

To determine respondents' perceptions of the pedagogies and tools they use, we asked them to indicate a number between 0 and 100 that best indicated their satisfaction with each approach as well as its cost-effectiveness. To minimize the work and time required for respondents, we only asked them to do this for the top 3 pedagogies they had selected. We then averaged these satisfaction ratings and ranked approaches based on these averages. As shown in Table 4, the order of the satisfaction list indicates which approaches practitioners favor. There is a notable discrepancy between their satisfaction with a given approach and its cost-effectiveness.

In terms of satisfaction with the approaches, the most used approaches (ILT and VILT) are not the ones with which respondents are most satisfied. Project-based learning, the option with which most respondents were satisfied, was not among the most used.

In terms of cost-effectiveness, role-playing activities are most highly ranked, but are among the least frequently used. This may be due to the fact that role playing generally involves several trainees and trainers simultaneously, which may present logistical difficulties especially when conducted face-to-face. Interestingly, project-based learning, ranked as the most satisfying approach, is ranked as among the least cost-effective.

Table 4. Comparison between the most satisfying and most cost-effective approaches.

Rank	Most satisfying approach	Most cost-effective approach
1	Project-based learning	Role-play activities
2	Case study	Competency-based learning
3	Online discussion	Instructor-led training
4	Instructor-led training	Field-based learning
5	Problem-based learning	Problem-based learning
6	Other	Simulations
7	Competency-based learning	Other
8	Simulation	Online-reading
9	Virtual instructor	Project-based learning
10	Field-based learning	Case study
11	Collaborative learning	Virtual instructor
12	Role-play activities	Collaborative learning
13	Online reading	Games
14	Games	Online discussion

Now Versus the Future

We also asked respondents to consider a list of 8 technologies and asked how important they felt these technologies are now and how important respondents felt these technologies will be in 5 years, on a scale of 1-10. Respondents predicted a decrease in the importance of course or learning management systems (LMS) and a large increase in the importance of AI (Table 5). They also predicted that webinars and course development systems will decrease slightly in importance, while predicting that simulation creation tools will become the most important approach. A rather significant decline in the use of online games was also projected.

Table 5. Comparison between tools perceived to be most important now and in 5 years.

Rank	Now	Future
1	LMS[a]	Simulation creation tools
2	Webinars (live)	Artificial intelligence
3	Course development systems	Webinars (live)
4	Simulation creation tools	Course development systems
5	Online games	Virtual reality
6	Virtual reality	LMS
7	Augmented reality	Augmented reality
8	Artificial intelligence	Online games

[a]LMS: learning management systems.

Qualitative Results

Qualitative phone interviews consisted of 18 questions related to contextual background and detailed information on both current learning strategies and experiences with technology. Participants' perceptions of the future of training were also queried. After using an open coding approach [30] and thematic analysis [32] of 9 interview transcripts and researcher's memos, the results revealed the emergence of three themes.

The Status Quo

All 9 respondents recognized that current training and education strategies in this field are at a clear risk of being abandoned in favor of rapidly evolving technologies. Most of the current techniques and tools have been used for some time, and the respondents noted that they are not very satisfied with them and that they do not perceive them to be very cost-effective. In particular, 8 of 9 respondents mentioned that the dominant learning tool, the LMS, is not conducive to learning, and the purpose of it is largely for administrative record tracking. One respondent said:

> *Most LMS have not developed with the user experiences in mind. In a global organizational level, LMS is just, "read the pdfs and do the quiz." It's absolutely boring.*

Another respondent mentioned that LMS are generally not "mobile friendly," adding that his sales team cannot have access to the learning material on-demand:

> *They are on the move all the time. It is unrealistic to expect them to have time to sit in the office and go through all the learning information.*

Additionally, respondents raised concerns about measuring learning. Two respondents confessed that they have no valid understanding of whether learners are engaged with the tools and solutions currently employed and whether they are devoting the effort required to benefit in a meaningful way. One respondent said:

> *We only have a self-assessed checklist for our folks to fill out after the training session, but we do not know their progress at all.*

Another respondent raised concern about how to devise valid evaluation metrics to assess his staff members. He provided an example that there are some staff members who got perfect scores on sales training assessments, yet have among the worst sales performance, while other staff members might have poor scores on sales training quizzes but have the highest sales performance in their district.

Respondents acknowledged the reality that making any change involving control and administration of the learning enterprise is challenging. However, respondents also pointed out that if properly leveraged, technologies offer vast opportunities for learning and could demonstrably enhance business productivity, even in the heavily regulated life science market where disruptive change is difficult to enact.

Call for Change

All the respondents identified that there is a clear call to action for change within both the life sciences learning professionals and the companies in which they work. More specifically, respondents indicated the need to provide new learning solutions and leverage technologies for both employees and customers. They reported the belief that this would help drive a successful business and that organizations not taking this approach would be at risk of being left behind. One respondent even joked that she did not want to be "asleep at the wheel."

Throughout the interviews, all the respondents were reflective about their roles and performance as learning leaders and practitioners, coupled with a strong desire to be strategic in their decision-making concerning technology and to be open to change. One participant mentioned:

> *We, as practitioners and leaders alike, face a compelling need to improve in terms of competency, speed, quality, cost and overall return on investment.*

Incorporating new technologies is perceived by these professionals as offering great promise, despite their awareness of the challenges that they realize will inevitably arise as they attempt to change the perceptions of those at the highest levels of their organizations toward embracing new solutions that leverage technology.

Additionally, respondents addressed the need for change to meet the training requirements for different generations and different ways of learning. One respondent said:

> *The old ways of teaching and training will not work on the young generations. Millennials are on their smartphones all the time; they are addicted to the technologies.*

Other respondents indicated that people come from different backgrounds in learning and have different learning styles; therefore, the traditional ways of learning will become obsolete.

The Future of Teaching and Learning

When asked about the future and the roles technology might play in learning, all of the respondents described a future in which training programs and processes are accelerated, impactful, and engaging. One respondent said:

> *Our goal is to create learning experiences, so that when people walk away, they are like, wow, that never happened to me before, I am going to remember that. We are now integrating technologies to involve all the senses to create a new unique learning events to improve impacts and effectiveness.*

They also mentioned seeing solutions evolve in which learners are given more control over the learning, and where learning practitioners shift from content developers to content curators. These predictions were based on observable trends, encompassing where technologies seem to be heading as well as the changing behaviors and product lines from the providers of learning materials and training development tools.

Additionally, respondents identified the potential to increase their impact through enhanced ability to develop social connectivity during learning, and to make learning personally relevant, interesting, engaging, easily accessible, self-driven, and even fun. Expanding on the concept of social connectivity, many commented on the growing nature of learner-centered environments in which learners engage interactive resources to meet their needs, working "on *their* terms" through learning experiences increasingly embedded in the workflow rather than as a separate formal learning event. One respondent said:

> *Learning is socio-cultural, if you limit the level of interactivity, you limit learning. We always need to look for opportunities for the learners to take control of the ability to connect and learn from each other. We need to look for ways to democratize data and have learning occurs [sic] down to the peer to peer level.*

All respondents perceived AI, AR, and VR as having transformative near-term potential. One respondent commented that these emerging technologies are the ones "to take people's knowledge and skills to the next level." At the same time, two respondents indicated concern regarding how to adopt AI in this heavily regulated field.

Discussion

Principal Findings

This study explored the current state of training in the life science professional field through a mixed methods approach. Survey results indicated that instructor-led training is perceived to be the most used and most important pedagogy, and virtual instructor-led methods are among the top 3 in terms of use but are much lower in importance. In terms of satisfaction with the approaches, it is interesting to note that the most used approaches (instructor-led training and virtual instructor-led training) are not the approaches with which respondents were most satisfied. Although they were most satisfied with project-based learning, this approach was among the least cost-effective ones.

From the cost-effectiveness perspective, role-playing activities are most highly ranked, but are among the least frequently used. This may be due to the fact that role-playing generally involves several trainees and trainers simultaneously, which may present logistical difficulties, especially when conducted face-to-face, which may cause busy professionals to dislike the activity.

Looking into the future, both survey and interview results indicated that respondents are not satisfied with the status quo and that teaching and learning in this field need to change. Most of the techniques and tools currently used have been used for a long time, and the professionals are not very satisfied with them and do not find them very cost-effective. Unfortunately, the market and organizations in which these practitioners work are very complex, making change difficult. The interview results indicate that there is an evolution underway, but also highlight the need to get better at incorporating new technologies. This will require changing the perception that training and education are large expenses incurred by the organization without much evidence-based justification regarding effectiveness, and will require design thinking [33] to consider both new approaches and new ways to demonstrate the effectiveness of training efforts based on the contributions training and education make to the providing organizations. Interviewees described a future in which training programs and processes are accelerated, impactful, and engaging, and in which learners are given more control. Emerging technologies such as AI and VR were seen to have increasingly important roles to play, allowing learning to become more interesting, engaging, and perhaps even fun. The vision is of learner-centered environments in which learners engage interactive resources to meet their needs, perhaps in learning experiences that are increasingly embedded in the workflow. These professionals see a bright, exciting future and a challenging path to realize this vision.

Limitations

Two limitations should be considered when interpreting these results. In the quantitative component of the study, 57 participants responded to the survey, which represents 17% of the sample pool of 326 people. However, this result is consistent with a typical noncompensated survey response rate (10%-15%) [34]. Future studies should expand the survey to a larger membership body to increase the number of respondents. In addition, the participants were recruited through a single organization (LTEN); future studies should expand recruitment to encompass a broader spectrum of education and training professionals in the life sciences.

Conclusions

This is the first study in a series designed to better understand education and training in the life sciences on a macro level, in order to build a foundation for progress and evolution of the future landscape. All respondents in this study seemed very aware that rapid and potentially beneficial change is underway, fueled by emerging technologies. They acknowledge that while the pace of its emergence is increasing in less complex contexts, aspects of this particular industry seem likely to inhibit the pace of change. In addition, respondents also acknowledge that the adoption of emerging technologies is impeded by the absence of data demonstrating a compelling return on investment, and insufficient time and resources. As important as this perspective is, understanding is a necessary but not sufficient first step. Next steps involve developing strategies for how to extend this vision throughout the individual organizations. Beyond this, we will need to determine how to expose the existing dissatisfaction with traditional, ineffective ways of operating, and create realistic "first steps" in the desired direction. Then, leaders in the field must gather data, make modifications, adjust, and document the effects. The process is not unlike the development of the organization's core products, from research and design to operational practices. The process will be challenging, particularly in the highly regulated, relatively conservative life sciences market. This set of challenges might become less daunting through projects like this initial study, and the extended conversations it may generate among those ready to act as pioneers.

Acknowledgments

This research is supported by LTEN. LTEN is the only global 501(c)(3) nonprofit organization specializing in meeting the needs of life sciences learning professionals.

Conflicts of Interest

None declared.

References

1. Kaufmann C. Computers in surgical education and the operating room. Annales Chirurgiae et Gynaecologiae 2001:141-146.

2. Gorman PJ, Meier AH, Rawn C, Krummel TM. The future of medical education is no longer blood and guts, it is bits and bytes. The American Journal of Surgery 2000 Nov;180(5):353-356. [doi: 10.1016/s0002-9610(00)00514-6]
3. Becker C, Frishman WH, Scurlock C. Telemedicine and Tele-ICU: The Evolution and Differentiation of a New Medical Field. The American Journal of Medicine 2016 Dec;129(12):e333-e334. [doi: 10.1016/j.amjmed.2016.05.045]
4. Khosla S, White R, Medina J, Ouwens M, Emmas C, Koder T, et al. Real world evidence (RWE) – a disruptive innovation or the quiet evolution of medical evidence generation? F1000Res 2018 Aug 29;7:111. [doi: 10.12688/f1000research.13585.2]
5. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018 May 17;18(8):500-510. [doi: 10.1038/s41568-018-0016-5]
6. Langlotz CP, Allen B, Erickson BJ, Kalpathy-Cramer J, Bigelow K, Cook TS, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology 2019 Jun;291(3):781-791. [doi: 10.1148/radiol.2019190613]
7. Park SH, Han K. Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. Radiology 2018 Mar;286(3):800-809. [doi: 10.1148/radiol.2017171920]
8. Shen J, Zhang CJP, Jiang B, Chen J, Song J, Liu Z, et al. Artificial Intelligence Versus Clinicians in Disease Diagnosis: Systematic Review. JMIR Med Inform 2019 Aug 16;7(3):e10010 [FREE Full text] [doi: 10.2196/10010]
9. Bonk CJ, Hara N, Dennen V, Malikowski S, Supplee L. We're in TITLE to Dream: Envisioning a Community of Practice, The Intraplanetary Teacher Learning Exchange. CyberPsychology & Behavior 2000 Feb;3(1):25-39. [doi: 10.1089/109493100316201]
10. García-Pallarés J, Sánchez-Medina L, Carrasco L, Díaz A, Izquierdo M. Endurance and neuromuscular changes in world-class level kayakers during a periodized training cycle. Eur J Appl Physiol 2009 Apr 26;106(4):629-638. [doi: 10.1007/s00421-009-1061-2]
11. Tymczynska M. Integrating in-class and online learning activities in a healthcare interpreting course using Moodle. The Journal of Specialised Translation, 12 2009:148-164. [doi: 10.1057/9781137478412.0005]
12. Mantovani F. VR learning: potential challenges for the use of 3D environments in education training. In: Riva G, Galimberti C, editors. Towards Cyberpsychology. Amsterdam: Ios Press; 2001:207-225.
13. Stansfield S, Shawver D, Sobel A, Prasad M, Tapia L. Design and Implementation of a Virtual Reality System and Its Application to Training Medical First Responders. Presence: Teleoperators and Virtual Environments 2000 Dec;9(6):524-556. [doi: 10.1162/105474600300040376]
14. Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 2016 Feb 23;30(10):4174-4183. [doi: 10.1007/s00464-016-4800-6]
15. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Jun 21;2(4):230-243. [doi: 10.1136/svn-2017-000101]
16. Brown C. A. , Belfield, C. R., & Field, S. J. Cost effectiveness of continuing professional development in health care: a critical review of the evidence.Bmj 2002;324(7338):652-655. [doi: 10.1136/bmj.324.7338.652]
17. Vaughn HT, Rogers JL, Freeman JK. Does Requiring Continuing Education Units for Professional Licensing Renewal Assure Quality Patient Care? The Health Care Manager 2006;25(1):78-84. [doi: 10.1097/00126450-200601000-00011]
18. Peck C, McCall M, McLaren B, Rotem T. Continuing medical education and continuing professional development: international comparisons. BMJ 2000 Mar 12;320(7232):432-435 [FREE Full text] [doi: 10.1136/bmj.320.7232.432] [Medline: 10669451]
19. Bloom G. Blended coaching: Skills and strategies to support principal development. Thousand Oaks, CA: Corwin Press; 2005.
20. Umble KE, Cervero RM. Impact Studies in Continuing Education For Health Professionals. Eval Health Prof 2016 Jun 29;19(2):148-174. [doi: 10.1177/016327879601900202]
21. Lee I, Hsu T, Chen T, Zheng M. The Application of AR Technology to Spatial Skills Learning in Carpentry Training. IJIET 2019;9(1):56-60. [doi: 10.18178/ijiet.2019.9.1.1173]
22. Xu J, Tang Z, Yuan X, Nie Y, Ma Z, Wei X, et al. A VR-based the emergency rescue training system of railway accident. Entertainment Computing 2018 Aug;27:23-31. [doi: 10.1016/j.entcom.2018.03.002]
23. Kang SG, Ryu BJ, Yang KS, Ko YH, Cho S, Kang SH, et al. An Effective Repetitive Training Schedule to Achieve Skill Proficiency Using a Novel Robotic Virtual Reality Simulator. Journal of Surgical Education 2015 May;72(3):369-376. [doi: 10.1016/j.jsurg.2014.06.023]
24. Clifford R. Evaluating the effects of realistic communication disruptions in VR training for aerial firefighting. In: 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). 2018 Sep 05 Presented at: 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); 2018; Würzburg, Germany p. 1-8. [doi: 10.1109/vs-games.2018.8493423]
25. Dubey R, Gunasekaran A. The role of truck driver on sustainable transportation and logistics. Ind and Commercial Training 2015 Apr 07;47(3):127-134. [doi: 10.1108/ict-08-2014-0053]
26. Gavish N, Gutiérrez T, Webel S, Rodríguez J, Peveri M, Bockholt U, et al. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interactive Learning Environments 2013 Jul 18;23(6):778-798. [doi: 10.1080/10494820.2013.815221]

27. Creswell J. A Concise Introduction To Mixed Methods Research. Thousand Oaks, California: Sage Publications, Inc; 2014.
28. Creswell J. Editorial: Mapping the Field of Mixed Methods Research. Journal of Mixed Methods Research 2008 Nov 15;3(2):95-108. [doi: 10.1177/1558689808330883]
29. Seidman I. Interviewing As Qualitative Research: A Guide For Researchers In Education And The Social Sciences. New York: Teachers College Press; 2013.
30. Saldaña J. Ethnotheatre: Research from page to stage. New York: Routledge; 2016.
31. LTEN website. 2018. URL: https://www.l-ten.org/page/LTEN_History [accessed 2018-11-01]
32. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Brown T, Wyatt J. Design Thinking for Social Innovation. Development Outreach 2010 Jul;12(1):29-43. [doi: 10.1596/1020-797x_12_1_29]
34. Conn KM, Mo CH, Sellers LM. When Less Is More in Boosting Survey Response Rates. Social Science Quarterly 2019 Apr 12;100(4):1445-1458. [doi: 10.1111/ssqu.12625]

Abbreviations

AI: artificial intelligence
AR: augmented reality
ILT: instructor-led training
LMS: learning management systems
LTEN: Life Sciences Trainers & Educators Network
VILT: virtual instructor-led training
VR: virtual reality

Edited by G Eysenbach; submitted 14.08.19; peer-reviewed by A Johnson, K Aguirre; comments to author 28.10.19; revised version received 16.12.19; accepted 10.02.20; published 24.04.20.

Please cite as:
Magagna W, Wang N, Peck K
Current and Future Trends in Life Sciences Training: Questionnaire Study
JMIR Med Educ 2020;6(1):e15877
URL: http://mededu.jmir.org/2020/1/e15877/
doi:10.2196/15877
PMID:32329739

Original Paper

A Cardiology Handbook App to Improve Medical Education for Internal Medicine Residents: Development and Usability Study

Asad Torabi[1], MD; Abhishek Khemka[2], MD; Pantila V Bateman[2], MD

[1]Indiana University School of Medicine, Indianapolis, IN, United States

[2]Krannert Institute of Cardiology, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Pantila V Bateman, MD
Krannert Institute of Cardiology
Indiana University School of Medicine
1801 N Senate Blvd, Suite 4000
Indianapolis, IN, 46204
United States
Phone: 1 3172740995
Fax: 1 3179622501
Email: pvanicha@iu.edu

Abstract

Background: At most institutions, internal medicine residents struggle with balancing clinical duties and learning opportunities, particularly during busy cardiology ward rotations. To improve learning experiences for residents, we helped develop a cardiology handbook app to supplement cardiology education.

Objective: The aim of this study was to report the development, implementation, and preliminary impact of the Krannert Cardiology Handbook app on graduate medical education.

Methods: In June 2017, 122 residents at Indiana University were invited to download the digital handbook in the Krannert app. The Krannert app featured a total of 13 chapters written by cardiology fellows and faculty at Indiana University. Residents were surveyed on their self-reported improvement in cardiology knowledge and level of satisfaction after using the Krannert app. Residents were also surveyed regarding their preference for a digital handbook app versus a paper handbook.

Results: Of the 122 residents, 38 trainees (31.1%) participated in survey evaluations. Among all respondents, 31 app users (82%) reported that the app helped improve their cardiology knowledge base. The app had an overall favorable response.

Conclusions: The Krannert app shows promise in augmenting clinical education in cardiology with mobile learning. Future work includes adding new topics, updating the content, and comparing the app to other learning modalities.

(JMIR Med Educ 2020;6(1):e14983) doi:10.2196/14983

KEYWORDS

mobile learning; medical reference app; cardiology; internal medicine

Introduction

Clinical education is a complex process. The need to strengthen the quality of medical education has led to educational innovation and novel instructional strategies such as simulation and mobile technology [1,2]. Medical reference apps are being recognized as an increasingly important asset to improve medical education. In a survey polling over 3000 residents, fellows, and staff in Accreditation Council for Graduate Medical Education (ACGME) programs, the most popular app types requested were those on textbook and reference material, in-training exam material, classification and treatment algorithms, and general medical knowledge [3]. Respondents in the survey also indicated a demand for higher-quality apps.

Despite the growing popularity of medical reference apps, there is insufficient data on the use and effectiveness of smartphones and mobile learning in graduate and undergraduate medical education [4-9]. Currently, most reference apps cover general medical knowledge, with some serving as a calculator. In cardiology, there are limited specialty-specific reference apps available [10].

A recent review demonstrates that published medical education research in cardiology is lacking [11]. In general, there is a paucity of resident-level teaching tools available in cardiovascular education. There are also limited publications covering narrow cardiovascular topics, including the evaluation of chest pain and myopericarditis [12,13], but no comprehensive teaching tools to cover common cardiac diseases. An easily accessible teaching tool in cardiology would be beneficial to residents and supplement their existing cardiology curriculum.

At our institution, residents (of internal medicine, medicine-pediatrics, and preliminary medicine) were polled about their learning experiences during their inpatient cardiology rotation. A review of their free-text responses indicated the need for further learning opportunities. After reviewing the needs assessment, our overarching goal was to modernize teaching in cardiology and improve bedside learning by creating a digital handbook as a practical reference for trainees. Learning objectives were accomplished by residents going through the digital handbook individually at their own pace during the rotation. The aim was to provide trainees with an easily accessible and reliable source to better understand cardiac disease pathology, diagnostics, and management. This intervention was not designed to replace didactics or traditional bedside learning but rather to supplement them. In this paper, we report the development, implementation, and preliminary evaluation of the Krannert Cardiology Handbook app.

Methods

Overview

The Krannert app was piloted in 2017 as a cardiology reference tool for residents beginning their inpatient cardiology months at Indiana University, an urban academic medical center. A total of 122 residents (categorical, medicine-pediatrics, and preliminary) from 2017-2018 were invited to download the app before and during their orientation for an inpatient cardiology rotation. The study was deemed exempt from Institutional Review Board (IRB) review by the Kuali Coeus IRB Office of Research Compliance from Indiana University. Learners were provided with informed consent and no personal information was collected.

Needs Assessment

For curriculum development, we conducted a targeted needs assessment at our own institution. We reviewed free-text responses from an electronic survey, via MedHub, asking residents to describe their learning experiences (ie, lectures, conferences, case discussions, overall quality of faculty teaching, bedside teaching, adequate time for reading and studying) during their clinical rotation. The results of this assessment highlighted difficulties in balancing service and educational duties on a clinically demanding service as seen in other specialties [14]. The residents noted a lack of adequate learning opportunities and specifically requested more teaching in electrocardiographic interpretation and echocardiography. In response to these comments, we created the Krannert app curriculum to help improve educational experience.

Curricular Design

We arranged the digital handbook into 13 chapters, which were written by cardiology faculty members and cardiology fellows from the Krannert Institute of Cardiology at Indiana University. All materials were peer-reviewed by faculty members (ie, senior attendings) who work with medical trainees, including fellows, residents, and medical students. All attendings are board certified in cardiovascular disease and, depending on their specific subspecialty, they may also be board certified in internal medicine, interventional cardiology, electrophysiology, echocardiography, and nuclear cardiology. The cardiology fellows were board certified in internal medicine and in the process of completing their 3-year training in general cardiovascular disease in an ACGME accredited program at Indiana University. The names of the authors were listed in each chapter.

In the beginning, the reading content and specific topics were selected by the authors, who also serve as key clinical educators within the Cardiology Division. Specific topics were chosen based on common diagnoses for patients admitted to the cardiology care unit as well as for patients from the cardiology consultation services (Figure 1). It was also based on findings from our targeted needs assessment. The app content was written in a succinct, outline format to maintain brevity and serve as a quick reference along with basic classification and treatment algorithms. The app as a resource was designed to provide practical knowledge and not as a reference for in-training exam material. We made a point to include a guide on hemodynamics to review the fundamentals of right heart catheterization, valvular and pericardial pathology, and intra-aortic balloon pump because these are common topics that general medicine residents may not commonly encounter on other hospital rotations. Learning material involving ST-elevation myocardial infarction, non–ST-elevation myocardial infarction, cardiogenic shock, and arrhythmias, with potential complications and management, were included to help trainees feel more comfortable managing their acutely ill patients. Content developed by fellows was written from a peer teaching perspective, and many chapters included figures, tables, and images as well as references for deeper understanding of basic concepts (see Table 1 for app objectives and Multimedia Appendix 1 for sample content). After the content was written, app development by The Center for Physician Education at Indiana University Health took approximately 100 hours. The cost was US $10,000.

Figure 1. Table of contents of the Krannert Cardiology Handbook app. Chapters not visualized in the figure are tachyarrhythmia, bradyarrhythmia, syncope, pulmonary hypertension, cardiogenic pharmacology, and introduction to echocardiography.

Table 1. Krannert Cardiology Handbook app objectives.

Objective	Description
Acute coronary syndrome	• List features of unstable angina/non–ST elevation myocardial infarction from ST elevation myocardial infarction • List indications and complications of heart catheterization and intra-aortic balloon pump placement. Recognize appropriate pressure wave forms • Identify complications postmyocardial infarction • List common antianginal medications. List indications and contraindications for thrombolytics and anticoagulants • List indications for ionotropic stimulation for stress testing
Miscellaneous	• Identify workup for syncope • List types of pulmonary hypertension
Cardiac arrhythmias	• Identify electrocardiographic abnormalities and manage supraventricular and ventricular arrhythmias • Recognize and manage bradyarrhythmias
Cardiac failure	• Recognize and manage patients with acute decompensated heart failure including acute pulmonary edema and hypotension; understand use of inotropic and afterload reducing agents • Differentiate cardiogenic shock from other types of shock
Hemodynamics	• Recognize and manage valvular heart disease, including emphasis of physical exam findings indicative of severe valvular disease. Recognize need for surgical correction • Identify basic views of 2D echocardiogram and examples of valvular regurgitation and stenosis • Identify types of pericardial disease

The handbook was released as an app on Apple's App Store [15] in June 2017 and subsequently on Google Play [16] in September 2017; both are available free of charge. The mobile app was introduced to all medicine residents at Indiana University as a single iteration, by an email sent from the internal medicine department. The app could be utilized by trainees on any rotation but was targeted toward residents on the cardiology service. One week prior to starting the cardiology service, residents received an email with orientation materials, which included information regarding the app, advertised as a reference tool.

Survey Design

To evaluate the Krannert app, we gathered data throughout the 2017-2018 academic year with an anonymous self-generated survey. This was an open survey built into the app, completely voluntary, and advertised by the internal medicine department through email. No incentives were offered to fill in the survey. The questions were not randomized, and the target population was residents who used the app. The questionnaire had not been tested before, as this was a pilot study. Respondents were able to review and change answers before submission. We only analyzed completed surveys, and we did not weigh survey items or use propensity scores. We did not use IP (Internet Protocol) checks to evaluate whether responses were unique because learners were allowed to use the app on any device that they wished to use, and we did not have respondents enter any personal information.

Outcome Measures

We were interested in three outcomes: (1) self-reported improvement in cardiology knowledge; (2) resident satisfaction with the handbook; and (3) resident learning preference (ie, the Krannert app versus a traditional paper handbook). Responses were recorded using a 5-point Likert-type scale from 1="strongly disagree" to 5="strongly agree" (Multimedia Appendix 2). We used descriptive statistics to summarize all data in survey item responses. Responses like "agree" or "strongly agree" to a statement were considered positive. We used self-reported learning results to measure the effectiveness of the app in achieving the stated educational objectives.

Results

Participants

Of the 122 residents invited to download the app, 38 (31% completion rate) participated in the survey.

Medical Trainee–Perceived Learning

We used a single survey item from 38 respondents to assess self-reported learning after using the app. The majority of residents (n=31; 82%) reported that they agreed or strongly agreed that the app helped them improve their cardiology knowledge base (Multimedia Appendix 3, item 1).

Medical Trainee Satisfaction

Among all respondents, 90% (n=34) indicated the app was easy to use and 87% (n=33) reported that the app content was delivered in a user-friendly manner. Overall, trainees found the app to be acceptable (n=33; 87%). From all respondents, 76% (n=29) reported that the amount of app content was appropriate and 71% (n=27) reported that the app met their educational needs in cardiology (Multimedia Appendix 3, items 2-6).

Learning Preference

Of the 38 respondents, 74% of medical trainees (n=28) reported a preference for the mobile app over a traditional paper cardiology handbook. Nearly 24% (n=9) were neutral in their response to this survey item (Multimedia Appendix 3, item 7).

Discussion

Principal Findings

We introduced a novel educational app to improve learning opportunities for our trainees on their cardiology rotation and received favorable results. Easy access to the app as a teaching tool appeared to play a key role in improving their learning experience at our institution. Although 74% (28/38) of trainees preferred learning through mobile resources over traditional paper resources, other trainees did not have a strong inclination. It is our opinion that the benefits of a digital handbook include convenience and transportability to help answer clinical questions at the point of need at a cost that is fairly lower than printing a textbook. Learning resources should be readily accessible, and use of a mobile app provides trainees with this benefit. Mobile devices can also be used to customize educational materials. The Krannert app curriculum can be used by a variety of learners, including students in undergraduate medical education, and can serve as a teaching aid for institutions teaching adult cardiology. An educational handbook app may also be beneficial to other medical specialties, in general.

Our relatively positive experiences with a mobile app are comparable to other studies that introduced mobile textbook apps to trainees [17-19]. Hardyman et al [17] complied medical textbooks into an app called iDoc. Using self-reported patient encounters, the house staff in Wales reported an improvement in their efficiency, effectiveness, and timeliness in patient care. Therefore, we believe that the use of medical apps in education will likely continue to be helpful and popular among medical trainees. However, the need is great for high-quality medical apps that are accurate and up-to-date. There is also a growing need to learn more about how medical apps can impact and improve medical education. This is a distinctively unique learning platform in the ever-expanding digital world.

Strengths and Limitations

We had some limitations in our study. Because this is a pilot feasibility study, the sample size is small. We did not design the study to compare results between a control group and an intervention group. We did not assess measurable gains in learner’s knowledge with pre- and postknowledge assessments, which may yield different results than self-reported outcomes. Additionally, we did not compare the results of the app to other learning experiences, such as bedside teaching, lectures, or case conferences. Tracking the number of app downloads by our own residents was limited due to privacy restrictions from Apple and Google. We also did not assess whether there were certain topics that learners felt were ineffective or had difficulties with using the app interface.

Recommendations and Future Research

Further research should study larger populations of residents across multiple institutions and include written practical knowledge assessments. Future work will also include a comparison of the app with other learning modalities or experiences and assess whether learning from the app leads to sustained learning. Future steps in improving our app include adding new topics such as preventive cardiology, congenital heart disease, and cardio-oncology, updating the content, and making the app more interactive. We plan to review the app content annually and use the survey link within the app to directly incorporate resident feedback into the app content and layout design.

Conclusions

The development of a digital handbook app improved medical education for trainees in cardiology and appeared to play a key role in improving their learning experience at our institution. Future steps in improving our app include adding new topics, updating the content, and making the app more interactive. Trainees from other programs may also benefit from having an educational handbook app in cardiology as well as other medical specialties.

Acknowledgments

The authors thank Peng-Sheng Chen, MD, of Krannert Cardiology and J Alex Trimpe of The Center for Physician Education, Indiana University Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the Krannert Cardiology Handbook app.
[DOCX File , 328 KB - mededu_v6i1e14983_app1.docx]

Multimedia Appendix 2
A postcurriculum survey regarding thoughts on the Krannert app and its cardiology curriculum.
[DOCX File , 14 KB - mededu_v6i1e14983_app2.docx]

Multimedia Appendix 3
Survey results for self-reported learning (item 1), trainee satisfaction (items 2-6), and learning preference (item 7) (N=38).

[PNG File , 423 KB - mededu_v6i1e14983_app3.png]

References

1. Gaglani SM, Topol EJ. iMedEd. Academic Medicine 2014;89(9):1207-1209. [doi: 10.1097/acm.0000000000000361]
2. Ludmerer KM, Johns MME. Reforming Graduate Medical Education. JAMA 2005 Sep 07;294(9):1083. [doi: 10.1001/jama.294.9.1083]
3. Franko OI, Tirrell TF. Smartphone App Use Among Medical Providers in ACGME Training Programs. J Med Syst 2011 Nov 4;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7]
4. Masters K, Ellaway RH, Topps D, Archibald D, Hogue RJ. Mobile technologies in medical education: AMEE Guide No. 105. Medical Teacher 2016 Mar 24;38(6):537-549. [doi: 10.3109/0142159x.2016.1141190]
5. Wallace S, Clark M, White J. 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. BMJ Open 2012 Aug 24;2(4):e001099. [doi: 10.1136/bmjopen-2012-001099]
6. Ellaway RH, Fink P, Graves L, Campbell A. Left to their own devices: Medical learners' use of mobile technologies. Medical Teacher 2013 Nov 07;36(2):130-138. [doi: 10.3109/0142159x.2013.849800]
7. O'Connor P, Byrne D, Butt M, Offiah G, Lydon S, Mc Inerney K, et al. Interns and their smartphones: use for clinical practice. Postgrad Med J 2013 Nov 15;90(1060):75-79. [doi: 10.1136/postgradmedj-2013-131930]
8. Webb K, Bullock A, Dimond R, Stacey M. Can a mobile app improve the quality of patient care provided by trainee doctors? Analysis of trainees' case reports. BMJ Open 2016 Sep 21;6(9):e013075. [doi: 10.1136/bmjopen-2016-013075]
9. Short SS, Lin AC, Merianos DJ, Burke RV, Upperman JS. Smartphones, Trainees, and Mobile Education: Implications for Graduate Medical Education. Journal of Graduate Medical Education 2014 Jun;6(2):199-202. [doi: 10.4300/jgme-d-13-00238.1]
10. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. J Med Internet Res 2012 Sep 27;14(5):e128 [FREE Full text] [doi: 10.2196/jmir.1994] [Medline: 23017375]
11. Allred C, Berlacher K, Aggarwal S, Auseon AJ. Mind the Gap: Representation of Medical Education in Cardiology-Related Articles and Journals. Journal of Graduate Medical Education 2016 Jul;8(3):341-345. [doi: 10.4300/jgme-d-15-00285.1]
12. Hariton EL, Lester M, Richards J. An Atypical Cause of Chest Pain: An Octopus Trap. MedEdPORTAL Publications 2014;10 [FREE Full text] [doi: 10.15766/mep_2374-8265.9732]
13. Stair R. Myopericarditis. MedEdPORTAL Publications 2010:1676 [FREE Full text] [doi: 10.15766/mep_2374-8265.1676]
14. Quinn A, Brunett P. Service versus education: finding the right balance: a consensus statement from the Council of Emergency Medicine Residency Directors 2009 Academic Assembly. Acad Emerg Med 2009 Dec;16 Suppl 2:S15-S18 [FREE Full text] [doi: 10.1111/j.1553-2712.2009.00599.x] [Medline: 20053203]
15. App Store - Apple. Krannert Cardiology Handbook URL: https://apps.apple.com/us/app/krannert-cardiology-handbook/id1217564931 [accessed 2020-01-21]
16. Google Play. Krannert Cardiology Handbook URL: https://play.google.com/store/apps/details?id=org.iuhcpe.krannertcardiology&hl=en_US [accessed 2020-01-21]
17. Hardyman W, Bullock A, Brown A, Carter-Ingram S, Stacey M. Mobile technology supporting trainee doctors' workplace learning and patient care: an evaluation. BMC Med Educ 2013 Jan 21;13(1). [doi: 10.1186/1472-6920-13-6]
18. Lau C, Kolli V. App Use in Psychiatric Education: A Medical Student Survey. Acad Psychiatry 2016 Nov 23;41(1):68-70. [doi: 10.1007/s40596-016-0630-z]
19. Sandholzer M, Rurik I, Deutsch T, Frese T. Medical Students' Expectations Towards an Implementation of a Family Medicine Textbook as a Comprehensive App in Germany. J Med Syst 2014 Aug 30;38(10). [doi: 10.1007/s10916-014-0125-y]

Abbreviations

ACGME: Accreditation Council for Graduate Medical Education
IP: Internet Protocol
IRB: Institutional Review Board

Edited by G Eysenbach; submitted 09.06.19; peer-reviewed by M Ponum, O Adebayo, T Muto; comments to author 16.12.19; revised version received 24.01.20; accepted 07.02.20; published 16.04.20.

Please cite as:
Torabi A, Khemka A, Bateman PV
A Cardiology Handbook App to Improve Medical Education for Internal Medicine Residents: Development and Usability Study
JMIR Med Educ 2020;6(1):e14983
URL: http://mededu.jmir.org/2020/1/e14983/
doi:10.2196/14983
PMID:32297866

Original Paper

An Objective Structured Clinical Examination for Medical Student Radiology Clerkships: Reproducibility Study

Pedro Vinícius Staziaki[1], MD; Rutuparna Sarangi[1], MD; Ujas N Parikh[1], MD; Jeffrey G Brooks[1,2], MD; Christina Alexandra LeBedis[1], MD; Kitt Shaffer[1], MD, PhD

[1]Department of Radiology, Boston Medical Center, Boston University School of Medicine, Boston, MA, United States

[2]Department of Radiology, Milford Reginal Medical Center, Milford, MA, United States

Corresponding Author:
Pedro Vinícius Staziaki, MD
Department of Radiology
Boston Medical Center
Boston University School of Medicine
820 Harrison Ave, FGH Building, 4th Floor
Boston, MA
United States
Phone: 1 6174145135
Email: staziaki@gmail.com

Abstract

Background: Objective structured clinical examinations (OSCEs) are a useful method to evaluate medical students' performance in the clerkship years. OSCEs are designed to assess skills and knowledge in a standardized clinical setting and through use of a preset standard grading sheet, so that clinical knowledge can be evaluated at a high level and in a reproducible way.

Objective: This study aimed to present our OSCE assessment tool designed specifically for radiology clerkship medical students, which we called the objective structured radiology examination (OSRE), with the intent to advance the assessment of clerkship medical students by providing an objective, structured, reproducible, and low-cost method to evaluate medical students' radiology knowledge and the reproducibility of this assessment tool.

Methods: We designed 9 different OSRE cases for radiology clerkship classes with participating third- and fourth-year medical students. Each examination comprises 1 to 3 images, a clinical scenario, and structured questions, along with a standardized scoring sheet that allows for an objective and low-cost assessment. Each medical student completed 3 of 9 random examination cases during their rotation. To evaluate for reproducibility of our scoring sheet assessment tool, we used 5 examiners to grade the same students. Reproducibility for each case and consistency for each grader were assessed with a two-way mixed effects intraclass correlation coefficient (ICC). An ICC below 0.4 was deemed poor to fair, an ICC of 0.41 to 0.60 was moderate, an ICC of 0.6 to 0.8 was substantial, and an ICC greater than 0.8 was almost perfect. We also assessed the correlation of scores and the students' clinical experience with a linear regression model and compared mean grades between third- and fourth-year students.

Results: A total of 181 students (156 third- and 25 fourth-year students) were included in the study for a full academic year. Moreover, 6 of 9 cases demonstrated average ICCs more than 0.6 (substantial correlation), and the average ICCs ranged from 0.36 to 0.80 (P<.001 for all the cases). The average ICC for each grader was more than 0.60 (substantial correlation). The average grade among the third-year students was 11.9 (SD 4.9), compared with 12.8 (SD 5) among the fourth-year students (P=.005). There was no correlation between clinical experience and OSRE grade (−0.02; P=.48), adjusting for the medical school year.

Conclusions: Our OSRE is a reproducible assessment tool with most of our OSRE cases showing substantial correlation, except for 3 cases. No expertise in radiology is needed to grade these examinations using our scoring sheet. There was no correlation between scores and the clinical experience of the medical students tested.

(JMIR Med Educ 2020;6(1):e15444) doi:10.2196/15444

KEYWORDS

radiology; education; education methods; medical education; undergraduate

Introduction

Background

At our institution, there are approximately 160 to 180 medical students per graduating class, with 15 to 20 students in each 4-week radiology clerkship block, comprising predominantly third- and a few fourth-year medical students. Students receive 1 to 2 hours of daily didactic-style teaching directed toward a weekly rapid-fire quiz on topics including chest imaging, abdominal imaging, musculoskeletal imaging, pediatric radiology, neuroradiology, and nuclear medicine. Throughout the rotation, the medical students also observe residents and faculty in various reading rooms: general radiology, neuroradiology, body imaging, musculoskeletal imaging, pediatric radiology, breast imaging, and interventional radiology.

A variety of methods are used to assess medical students' performance during clerkships at different institutions. As a result, the final performance evaluation is often a combination of subjective and objective grading techniques. The subjective evaluation involves direct observation of the student performing duties and written assessments or presentations, whereas the objective evaluations include multiple-choice questions such as in Radiology ExamWeb examinations [1] and patient logs. Multiple-choice examinations are the most commonly used, albeit with an only limited assessment of a higher level of knowledge, which would require more complex questions [2,3], while also placing heavy emphasis on recognition and recall. Other limitations often found with multiple-choice examinations include the lack of feedback that test takers receive as well as poor validity [4]. In contradistinction, oral examinations may allow for assessment of a higher level of knowledge and reason but are limited by inconsistency in grading and potential bias [5].

The objective structured clinical examination (OSCE) has been proposed initially by Harden in 1975 as a standard for evaluating medical students' performance in the clerkship years [6]. The OSCE is intended to evaluate skills and knowledge in a standard clinical setting, and via a preset standard grading sheet, so that clinical knowledge can be evaluated at a high level and in a reproducible way. In a study by Morag et al [7], students' scores on an OSCE test were shown to increase with additional clinical knowledge. For that reason, many fields of medicine have since demonstrated the OSCE as a useful method to evaluate both medical students and residents [8-11], including radiology [7,12]. In the radiology setting, medical imaging requesting and ordering, imaging interpretation, and the next step in management can be tested for and graded in a single examination.

Objective

We proposed an OSCE assessment tool designed as an assessment tool for radiology clerkship students. Given the imaging-centered aspect of radiology clerkship, we called it objective structured radiology examination (OSRE). The goal of our proposed tool was to evaluate skills and knowledge in a structured manner, with reproducible results across different examples and different graders. This resource will advance the assessment of radiology clerkship medical students by providing an objective, structured, reproducible, and low-cost method to evaluate radiology clinical knowledge in an OSCE-like format.

Methods

Objective Structured Radiology Examination Design

We developed 9 radiology OSRE cases, each with a set of 5 questions for assessment. Initially, for 3 months, we gave these OSRE cases to medical students for preliminary testing. We then openly reviewed the student scorings and reformed the grading sheets to include as many correct and incorrect scorings as possible. For each OSRE case, we designed a scoring sheet with a set of checkboxes corresponding to correct and incorrect scorings. We assigned a point value to each correct or incorrect scoring.

Each OSRE case comprises 1 to 3 radiology images that covered basic radiology diagnoses, followed by a question sheet containing a detailed clinical history and 5 examination questions to be answered in the same sheet. We developed the 5 questions to simulate activities that nonradiology clinicians might perform in a structured fashion: selection of pertinent clinical history needed for filling out imaging requisitions, recognition of clinically important findings, formulation of an overall impression, as well as questions about recommendations and follow-up. We displayed images associated with each OSRE on a projector. The supervisor in the examination room, most commonly a radiology resident, ensured that the image was visible to all. All the case images consisted of radiographs except for a head computed tomography image.

Objective Structured Radiology Examination Cases

Case 1 included a posterior-anterior (PA) and lateral chest radiograph showing right upper lobe pneumonia. Case 2 included an upright and supine radiograph of the abdomen showing a small bowel obstruction. Case 3 included 3 axial noncontrast computed tomography images at different levels of the brain through a subdural hemorrhage. Case 4 included frontal and lateral radiographs of the wrist showing a distal radial fracture. Case 5 included a portable frontal chest radiograph showing a right pleural effusion. Case 6 included a supine radiograph of the abdomen showing a feeding tube in the right lower lobe bronchus. Case 7 included a single cross-table radiograph of the knee with a fat fluid level in a large suprapatellar effusion. Case 8 included a PA and lateral radiograph of the chest showing right lower lobe pneumonia. Finally, case 9 included an upright and supine radiograph of the abdomen, showing a small bowel obstruction. All these cases had been previously published at *MedEdPORTAL* as free downloadable resources [13]. Multimedia Appendix 1 is a template for an OSRE case.

Objective Structured Radiology Examination Grading

The OSRE scoring sheets comprised checklists with specific point values for correct and incorrect scorings. Each question's score was worth between 1 and 4 points. Many of the individual questions allowed for multiple scorings. For example, 1 question in an OSRE asked students to describe the pertinent positive and negative findings on the chest radiograph displayed on the projector. Students were given positive points for defined correct scorings and negative points for defined incorrect scorings. The highest possible score on the OSRE tests ranged from 24 to 26 points. However, we established that the lowest score possible for any individual OSRE case was 0 even when the number of points amassed was negative. Multimedia Appendix 2 is a template for a scoring sheet.

Study Design

We obtained institutional review board approval to conduct educational research using students enrolled in the radiology clerkship during an entire academic year, and the need to acquire consent from each medical student was waived. Our study was designed and performed following the Declaration of Helsinki. At the beginning of each block, we informed the students of the research project and told them that their scores from the OSRE cases would not count toward their final grade. The students in each block were taught the standard curriculum throughout their radiology rotation without any specific teaching toward the newly designed OSRE.

There were 11 four-week clerkship blocks (ie, classes) of students in total during the entire year, representing 11 months across the year. At the end of each of the initial 3 weeks of their 4-week block, all the students from the same class were given 1 OSRE. Therefore, each student completed a total of 3 OSRE cases during their rotation at the end of each of the first 3 weeks of the course. The exception was block 11, when these students had only 2 OSREs, instead of 3. All 9 OSREs were given in order. Blocks 1, 4, 7, and 10 took cases 1, 2, and 3; blocks 2, 5, and 8 took cases 4, 5, and 6; and blocks 3, 6, and 9 took cases 7, 8, and 9. Again, as an exception, block 11 was given only cases 4 and 6. We chose this design to spread out the 9 different OSCEs across the entire year in a uniform fashion.

Five different examiners graded each of the OSREs independently for every single medical student. Grader 1 and grader 3 had 1 to 2 years of experience in medical student education. Grader 2 had over 20 years of medical student education experience. Grader 4 was a second-year radiology resident, and grader 5 was a medical student. These graders were selected with the aim of sampling graders at various stages of medical education.

We graded a subset of the OSREs (3 random sets of OSRE tests) a second time, approximately 2 months after completion of the academic year, to assess internal consistency between the graders and reproducibility of our assessment tool.

There was no specific training or instruction for graders, as we designed the test and grading to be self-explanatory based on the scoring sheets. Each examination took approximately 30 seconds to 1 min to grade. We gave students their scores and individualized formative feedback on their OSRE performance at the midclerkship review and final course feedback session as part of the standard process at the radiology clerkship at our institution. Any questions the medical students had regarding the OSRE questions and scorings were answered.

Statistical Analysis

The reproducibility of our OSRE was assessed by performing interrater reliability with a two-way mixed effects intraclass correlation coefficient (ICC) to determine consistency between the 5 graders. Reproducibility for each grader was also evaluated with an ICC test 2 months later. Utilizing the classification system for ICCs by Landis and Koch [14], an ICC below 0.4 was classified as poor to fair, an ICC of 0.41 to 0.60 was considered moderate, an ICC of 0.6 to 0.8 was substantial, and an ICC greater than 0.8 was almost perfect. An ICC of 0.6 or more was considered a significant correlation.

We also sought to find if there was an association with a higher OSRE score and clinical experience with block number and with the medical school year. For this, we used a multivariate linear regression model in which the mean OSRE score was the outcome variable and the year block was the explanatory variable, with the medical student year as a controlling covariate. We then compared mean OSRE grades between the third- and fourth-year medical students using a two-sided Student *t* test.

Results

Summary of Scores

A total of 181 medical students were included in this study, 156 third-year medical students and 25 fourth-year medical students. OSRE score averages by blocks and cases are depicted in Tables 1 and 2, respectively.

Table 1. Mean objective structured radiology examination scores for blocks.

Block	Score, mean (SD)
1	13.0 (4.1)
2	8.3 (3.60)
3	12 9 (6.2)
4	15.1 (3.6)
5	11.5 (4.3)
6	12.6 (5.1)
7	13.3 (4.2)
8	9.2 (3.8)
9	12.2 (4.9)
10	14 (4.2)
11	9 (4.1)

Table 2. Mean objective structured radiology examination scores for cases.

Case	Score, mean (SD)
1	15.1 (4.1)
2	12.8 (4.5)
3	13.6 (3.5)
4	8.6 (4.0)
5	8.5 (3.2)
6	11.4 (4.3)
7	6.4 (3.2)
8	15.9 (4.0)
9	14.5 (3.0)

Reproducibility

Interrater reliability was shown to be ranging from poor to substantial average ICCs, with an average range of 0.36 to 0.80 (*P*<.001; Table 3). In most cases, 6 out of 9 showed correlation values of at least 0.6 (substantial correlation). However, case 3 had a poor correlation, and cases 6 and 7 showed moderate correlation.

Grader consistency on the 3 random OSRE cases (cases 1, 2, and 3) after 2 months showed that 4 out of the 5 graders had an ICC equal to or greater than 0.8 (substantial correlation), whereas grader 3 had an ICC of 0.68. Comparing these regraded exams, the range of mean raw score differences was −1 to 0.8 (Table 4). These data illustrate the reproducibility of the grading.

Table 3. Interrater reliability among graders with intraclass correlation coefficients.

OSRE[a]	Correlation	95% CI	*P* value
OSRE 1			
Block 1	0.80	0.64 to 0.92	<.001
Block 4	0.71	0.52 to 0.87	<.001
Block 7	0.72	0.53 to 0.88	<.001
Block 10	0.65	0.41 to 0.86	<.001
Average	0.72	N/A[b]	N/A
OSRE 2			
Block 1	0.79	0.64 to 0.9	<.001
Block 4	0.79	0.63 to 0.91	<.001
Block 7	0.90	0.81 to 0.96	<.001
Block 10	0.73	0.54 to 0.88	<.001
Average	0.80	N/A	N/A
OSRE 3			
Block 1	0.28	0.08 to 0.56	<.001
Block 4	0.31	0.1 to 0.6	<.001
Block 7	0.53	0.31 to 0.76	<.001
Block 10	0.32	0.11 to 0.61	<.001
Average	0.36	N/A	N/A
OSRE 4			
Block 2	0.76	0.58 to 0.89	<.001
Block 5	0.81	0.66 to 0.92	<.001
Block 8	0.80	0.62 to 0.92	<.001
Block 11	0.67	0.46 to 0.85	<.001
Average	0.76	N/A	N/A
OSRE 5			
Block 2	0.76	0.58 to 0.89	<.001
Block 5	0.54	0.32 to 0.76	<.001
Block 8	0.53	0.29 to 0.78	<.001
Average	0.61	N/A	N/A
OSRE 6			
Block 2	0.50	0.27 to 0.74	<.001
Block 5	0.37	0.16 to 0.64	<.001
Block 8	0.51	0.26 to 0.77	<.001
Block 11	0.50	0.27 to 0.74	<.001
Average	0.47	N/A	N/A
OSRE 7			
Block 3	0.32	0.12 to 0.59	<.001
Block 6	0.56	0.35 to 0.77	<.001
Block 9	0.64	0.45 to 0.81	<.001
Average	0.50	N/A	N/A
OSRE 8			

OSRE[a]	Correlation	95% CI	*P* value
Block 3	0.61	0.4 to 0.8	<.001
Block 6	0.54	0.34 to 0.75	<.001
Block 9	0.76	0.61 to 0.88	<.001
Average	0.64	N/A	N/A
OSRE 9			
Block 3	0.77	0.6 to 0.9	<.001
Block 6	0.61	0.41 to 0.8	<.001
Block 9	0.77	0.61 to 0.9	<.001
Average	0.72	N/A	N/A

[a]OSRE: objective structured radiology examination.

[b]N/A: not applicable.

Table 4. Reproducibility by grader consistency after 2 months.

Grader	Correlation	Difference	*P* value[a]
Grader 1			
OSRE[b] 1	0.88	0.50	.30
OSRE 2	0.92	0.39	.35
OSRE 3	0.85	0.50	.19
Average	0.88	0.46	N/A[c]
Grader 2			
OSRE 1	0.89	−1.34	*.02*
OSRE 2	0.85	−2.06	*<.001*
OSRE 3	0.73	0.25	.70
Average	0.82	−1.05	N/A
Grader 3			
OSRE 1	0.35	−0.84	.43
OSRE 2	0.89	1.06	*.03*
OSRE 3	0.79	−0.69	.11
Average	0.68	−0.16	N/A
Grader 4			
OSRE 1	0.95	0.06	.83
OSRE 2	0.82	0.50	.37
OSRE 3	0.78	−0.44	.39
Average	0.85	0.04	N/A
Grader 5			
OSRE 1	0.86	0.60	.24
OSRE 2	0.88	1.06	*.03*
OSRE 3	0.66	0.69	.25
Average	0.80	0.78	N/A

[a]Italicized values were statistically significant.

[b]OSRE: objective structured radiology examination.

[c]N/A: not applicable.

Scores and Clinical Experience

The average OSRE score among all students was 12 (SD 4.9). The average grade among third-year students was 11.9 (SD 4.9), compared with 12.8 (SD 5) among fourth-year students (*P*=.005). There was no correlation between the block number and OSRE score. On the multiple linear regression, the block had an effect of −0.02 (95% CI −0.08 to 0.04; *P*=.48), adjusting for the medical school year (Table 5).

Table 5. Multiple linear regression showing the association of block with the objective structured radiology examination score, adjusting for the medical school year.

Variables	Estimate (95% CI)	*P* value[a]
Intercept	9.3 (7.4 to 11.2)	*<.001*
Block	−0.02 (−0.08 to 0.04)	.48
Year	0.9 (0.3 to 1.5)	*.01*

[a]Italicized values were statistically significant.

Discussion

Presentation of the Material

In summary, our OSRE assessment resource comprises a set of 9 cases that include 1 to 3 images each, a clinical scenario, and structured questions, along with a standardized scoring sheet that allows for an objective, structured, and low-cost assessment of radiology clerkship medical students. The structured questions aim to assess medical students' ability to understand history and indication, to describe imaging findings, to give an imaging impression or diagnosis, and to come up with the next step in management.

We found that our OSREs achieve their goal of being objective, structured, reproducible, and low cost. Most cases demonstrated a substantial interrater correlation (6 out of 10 showing an ICC of 0.6 or more). However, the correlation varied from poor to substantial, ranging from 0.36 to 0.80. The graders provided reproducible scores, even after 2 months, with a substantial interrater correlation (above 0.6). Finally, we did not find a correlation between the OSRE scores and clerkship block, but we did see that fourth-year medical students scored better than third-year medical students.

What We Observed and Lessons Learned

In assessing the different OSRE cases, we found that OSRE cases 3, 6, and 7 had a poor-to-moderate correlation. As all the remaining OSRE cases had an ICC value of more than or equal to 0.6, we still feel that our OSRE is a reproducible testing resource. All graders were consistent, shown by the very small variability in average scores (−1 to 0.8) when the graders regraded the same subset of 3 cases 2 months apart. Our use of various graders with differing medical education backgrounds demonstrates that expertise in radiology is not necessary to grade these examinations. If a consistent and clear grading sheet is used, grading can be performed by anyone with knowledge of medical terminology.

Regarding the association of grade and clinical experience, scores are not supposed to improve with later blocks, as this would mean that they either depend on the overall clinical experience or that students could be sharing the cases or questions with future students, thereby giving them a leg up by providing examination information to their colleagues. We found that scores did not vary with the block, adjusting for the medical school year. However, the fourth-year students had a slightly better average grade than the third-year students, which makes intuitive sense.

Other studies have found the OSCE to provide valuable feedback as well [4,7,15]. In a study of 122 medical students by Morag et al [7], the authors concluded that the OSCE cases provided an opportunity for feedback, by uncovering deficits in individuals. Students were able to review their performance in different clinical topics (chest pain, abdominal pain, etc) as well as types of questions (selection of imaging modality and anatomy) with ease. An unforeseen benefit of our OSRE implementation was that having the OSRE results weekly allowed the clerkship director and assistant to carefully examine areas where students displayed deficiencies or gaps in knowledge as well as to give each student more information on areas of strength and weakness at both the midcourse feedback session and the final course feedback session.

Agarwal et al [15] point out that radiology should incorporate OSCEs as a part of its examination and explain how an ideal radiology OSCE could look like. Specifically, they describe an OSCE method with 10 to 20 stations, some manned and other unmanned, each evaluating activities related to specific radiology topics, for instance, a basic task such as loading a radiograph (radiography OSCE station) or demonstrating an examination technique, such as performing an ultrasound examination of the abdomen in a patient. Completion of a 5-min task within a single station would involve either demonstrating a task to an examiner, providing verbal answers, or writing specific objective answers in a response sheet [15]. Their approach is different than the one we propose here, as our OSRE is a much simpler proposal, albeit less expensive and difficult to implement.

Limitations

Although this assessment tool has several advantages, it is not devoid of limitations. For instance, our interrater correlation was not substantial for all cases. This could have been remediated with previous training of the graders. However, we opted not to train graders, as training would artificially increase the interrater correlation of the grading process. Although it would be ideal to train graders before they score examinations, graders in real-life settings (such as teachers) may not always get the appropriate training to score the OSCEs.

Another issue was that this evaluation occurred at a single institution. Despite being low cost, the successful implementation of this assessment model requires informatics facilities to hold OSRE documents, including images, cases, and scoring sheets, which need diligent organization. In our institution, we have a clerkship coordinator and 2 volunteering second-year radiology residents to help coordinate the evaluation of medical students. Finally, we should be aware that medical students can use a recall system to convey the OSRE case to the medical students of future blocks. For this reason, there is a need to constantly create new cases.

Our choices of cases and questions are also a limitation. Although multiple modalities were selected, there were no normal cases, and they were a very small selection of the medical students' radiology clerkship curriculum. In addition, each case had only a few questions, and several other questions could have been included. For example, we could ask students about normal anatomy or imaging pitfalls or even ask them to provide the appropriate history to order an imaging examination. Different OSREs can be created to assess different skills in the radiology specialty such as the use of clinical guideline algorithms (eg, Breast Imaging-Reporting and Data System or what to recommend for an incidental finding), dictation and descriptive skills, differential diagnosis, or next step in management, among others. However, these skills would be most appropriate to radiology residents, not to medical students.

OSCEs are an excellent method to evaluate medical students, but they work best when they aim to evaluate a clinical, especially manual, skill. In contradistinction, the output of a radiologist is usually a report, which can be written in subjective ways. Structured radiology reports and modern template standards are useful to make our reports more objective, but they do not reduce the inherent subjectivity of the radiologist evaluating an imaging examination. This means that the main activity of a radiologist cannot be evaluated with the OSRE described here. On the other hand, the goal of a radiology clerkship is not to train a radiologist. Rather, it aims to teach and evaluate concepts that underlie the foundations of radiology, which should be assessed more objectively whenever possible.

The Next Steps

Given these limitations, there are many areas of improvement and ways to refine this resource, for instance, by expanding our questions and our cases, as described above. Furthermore, this model can be enhanced by making it all computer based, with a cloud-based storage software on the web. If we create a large online database of hundreds of OSRE cases or more, a piece of software could download a random case for each student. This could lead to the expansion of this model to other institutions. If multiple institutions are interested in this endeavor, it could remain to be a low-cost model.

Finally, future studies are needed to assess the validity of this tool compared with the standard means of assessing knowledge, including multiple-choice questions. Additional research could also assess the validity of OSRE-style examinations in radiology clerkships with a larger number of institutions and medical students.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Template for a case.

[DOC File , 219 KB - mededu_v6i1e15444_app1.doc]

Multimedia Appendix 2

Template for a scoring sheet.

[DOC File , 245 KB - mededu_v6i1e15444_app2.doc]

References

1. The Association of University Radiologists (AUR). Radiology ExamWeb URL: https://www.aur.org/Secondary-Alliances.aspx?id=139 [accessed 2018-12-13]
2. Smart GA. The multiple choice examination paper. Br J Hosp Med 1976;1976:131-136 [FREE Full text]
3. Simkin MG, Kuechler WL. Multiple-choice tests and student understanding: what is the connection? Decis Sci J Innov Educ 2005 Jan;3(1):73-98. [doi: 10.1111/j.1540-4609.2005.00053.x]
4. Gunderman RB, Ladowski JM. Inherent limitations of multiple-choice testing. Acad Radiol 2013 Oct;20(10):1319-1321. [doi: 10.1016/j.acra.2013.04.009] [Medline: 24029066]
5. Cohen R, Reznick RK, Taylor BR, Provan J, Rothman A. Reliability and validity of the objective structured clinical examination in assessing surgical residents. Am J Surg 1990 Sep;160(3):302-305. [doi: 10.1016/s0002-9610(06)80029-2] [Medline: 2393060]
6. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. Br Med J 1975 Feb 22;1(5955):447-451 [FREE Full text] [doi: 10.1136/bmj.1.5955.447] [Medline: 1115966]

7. Morag E, Lieberman G, Volkan K, Shaffer K, Novelline R, Lang EV. Clinical competence assessment in radiology: introduction of an objective structured clinical examination in the medical school curriculum. Acad Radiol 2001 Jan;8(1):74-81. [doi: 10.1016/S1076-6332(03)80746-8] [Medline: 11201460]
8. Jefferies A, Simmons B, Tabak D, McIlroy JH, Lee K, Roukema H, et al. Using an objective structured clinical examination (OSCE) to assess multiple physician competencies in postgraduate training. Med Teach 2007 Mar;29(2-3):183-191. [doi: 10.1080/01421590701302290] [Medline: 17701631]
9. Dwyer T, Takahashi SG, Hynes MK, Herold J, Wasserstein D, Nousiainen M, et al. How to assess communication, professionalism, collaboration and the other intrinsic CanMEDS roles in orthopedic residents: use of an objective structured clinical examination (OSCE). Can J Surg 2014 Aug;57(4):230-236 [FREE Full text] [doi: 10.1503/cjs.018813] [Medline: 25078926]
10. Zayyan M. Objective structured clinical examination: the assessment of choice. Oman Med J 2011 Jul;26(4):219-222 [FREE Full text] [doi: 10.5001/omj.2011.55] [Medline: 22043423]
11. Turner JL, Dankoski M. Objective structured clinical exams: a critical review. Fam Med 2008 Sep;40(8):574-578. [Medline: 18988044]
12. van den Berk I, van de Ridder J, van Schaik J. Radiology as part of an objective structured clinical examination on clinical skills. Eur J Radiol 2011 Jun;78(3):363-367. [doi: 10.1016/j.ejrad.2008.10.042] [Medline: 19091506]
13. Brooks J, Lebedis C, Shaffer K, Coleman S. The 3rd year radiology medical student clerkship objective structured clinical examination (OSCE). MedEdPORTAL 2014;10:9692. [doi: 10.15766/mep_2374-8265.9692]
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
15. Agarwal A, Batra B, Sood A, Ramakantan R, Bhargava S, Chidambaranathan N, et al. Objective structured clinical examination in radiology. Indian J Radiol Imaging 2010 May;20(2):83-88 [FREE Full text] [doi: 10.4103/0971-3026.63040] [Medline: 20607015]

Abbreviations

ICC: intraclass correlation coefficient
OSCE: objective structured clinical examination
OSRE: objective structured radiology examination
PA: posterior-anterior

Edited by G Eysenbach; submitted 10.07.19; peer-reviewed by YH Kwan, IV George, S Purkayastha; comments to author 28.12.19; revised version received 05.02.20; accepted 10.02.20; published 06.05.20.

Please cite as:
Staziaki PV, Sarangi R, Parikh UN, Brooks JG, LeBedis CA, Shaffer K
An Objective Structured Clinical Examination for Medical Student Radiology Clerkships: Reproducibility Study
JMIR Med Educ 2020;6(1):e15444
URL: http://mededu.jmir.org/2020/1/e15444/
doi:10.2196/15444
PMID:32374267

Original Paper

Comparing Classroom Instruction to Individual Instruction as an Approach to Teach Avatar-Based Patient Monitoring With Visual Patient: Simulation Study

Julian Rössler[1], MD; Alexander Kaserer[1], MD; Benjamin Albiez[1], RN, BNSc; Julia Braun[2], PhD; Jan Breckwoldt[1], MD; Donat Rudolf Spahn[1], MD; Christoph Nöthiger[1], MD; David Werner Tscholl[1], MD

[1]University Hospital Zurich, Zurich, Switzerland

[2]Biostatistics and Prevention Institute, Departments of Epidemiology and Biostatistics, University of Zurich, Zurich, Switzerland

Corresponding Author:
David Werner Tscholl, MD
University Hospital Zurich
Raemistrasse 100
Zurich, 8091
Switzerland
Phone: 41 786636787
Email: davidtscholl@me.com

Related Article:

This is a corrected version. See correction statement: https://mededu.jmir.org/2020/2/e24459/

Abstract

Background: Visual Patient is an avatar-based alternative to standard patient monitor displays that significantly improves the perception of vital signs. Implementation of this technology in larger organizations would require it to be teachable by brief class instruction to large groups of professionals. Therefore, our study aimed to investigate the efficacy of such a large-scale introduction to Visual Patient.

Objective: In this study, we aimed to compare 2 different educational methods, one-on-one instruction and class instruction, for training anesthesia providers in avatar-based patient monitoring.

Methods: We presented 42 anesthesia providers with 30 minutes of class instruction on Visual Patient (class instruction group). We further selected a historical sample of 16 participants from a previous study who each received individual instruction (individual instruction group). After the instruction, the participants were shown monitors with either conventional displays or Visual Patient displays and were asked to interpret vital signs. In the class instruction group, the participants were shown scenarios for either 3 or 10 seconds, and the numbers of correct perceptions with each technology were compared. Then, the teaching efficacy of the class instruction was compared with that of the individual instruction in the historical sample by 2-way mixed analysis of variance and mixed regression.

Results: In the class instruction group, when participants were presented with the 3-second scenario, there was a statistically significant median increase in the number of perceived vital signs when the participants were shown the Visual Patient compared to when they were shown the conventional display (3 vital signs, *P*<.001; effect size –0.55). No significant difference was found for the 10-second scenarios. There was a statistically significant interaction between the teaching intervention and display technology in the number of perceived vital signs (*P*=.04; partial η^2=.076). The mixed logistic regression model for correct vital sign perception yielded an odds ratio (OR) of 1.88 (95% CI 1.41-2.52; *P*<.001) for individual instruction compared to class instruction as well as an OR of 3.03 (95% CI 2.50-3.70; *P*<.001) for the Visual Patient compared to conventional monitoring.

Conclusions: Although individual instruction on Visual Patient is slightly more effective, class instruction is a viable teaching method; thus, large-scale introduction of health care providers to this novel technology is feasible.

(JMIR Med Educ 2020;6(1):e17922) doi:10.2196/17922

KEYWORDS
avatar; computer-assisted; diagnosis

Introduction

Monitoring and continuous evaluation of vital signs by anesthesia providers is central to perioperative patient safety [1]. With 313 million surgeries performed worldwide every year, patient monitors are ubiquitous in perioperative health care [2]. However, there have been no recent substantial changes to the industry standard of displaying vital signs as numbers and curves, and some anesthesia providers report difficulties regarding this form of presentation [3]. Considering the design principles of situation awareness, Visual Patient was developed as an additional way to present vital signs [4-6]. The Visual Patient displays vital signs by modification of an animated avatar (Figure 1, Multimedia Appendix 1). The avatar, which corresponds to the patient, can display 11 vital signs; for example, it pulsates with different intensities and frequencies, breathes, and changes color on desaturation (Figure 2, Multimedia Appendix 1). Tscholl and colleagues were able to show that after briefly seeing a display of the Visual Patient, anesthesia providers were able to recall more vital signs than with conventional monitoring. They further reported improved confidence and reduced cognitive effort [4,7]. This may help healthcare providers gain situation awareness more efficiently and may increase patient safety [8-11]. However, the implementation of this technology may be difficult, as conventional monitoring is well known and established. Feasibility of Visual Patient training for widespread implementation would require the training to be deliverable to multiple participants at once, short in duration (eg, 30 minutes), and suitable for large auditoriums.

We designed a simulation study where participants who had no previous experience with Visual Patient underwent either individual or classroom-based instruction and were then asked to interpret conventional displays and avatar-based Visual Patient displays. We hypothesized that the 2 instruction methods would be comparable in efficacy as an introduction to avatar-based monitoring with Visual Patient.

Figure 1. Screenshots of the presented scenarios showing conventional monitoring (A) and avatar-based monitoring with the Visual Patient (B).

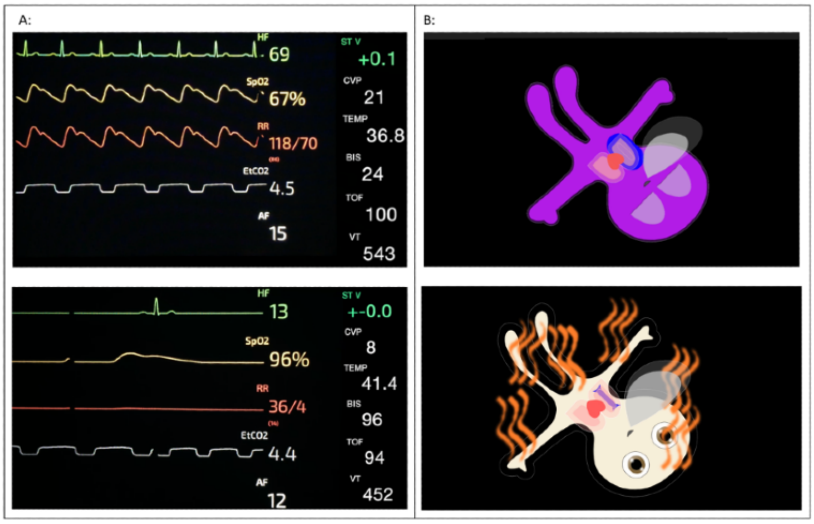
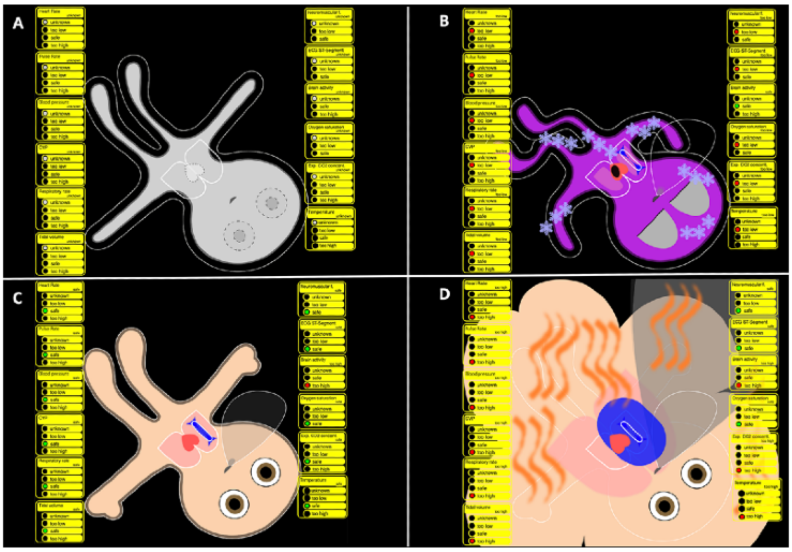


Figure 2. Vital sign parameters of the Visual Patient with a legend showing how each parameter is visualized. A: Visual Patient display when no vital sign data are received. B: Desaturated, hypothermic patient with ST-segment deviation. C: Visual Patient with all vital signs in a safe state and high brain activity (open eyes). D: Hypertensive, hyperthermic patient with high central line pressure.



Methods

Participants and Trial Design

On January 19, 2019, 42 nurse anesthetists were recruited to the classroom instruction group. Since the study did not include any real patient data or any human material, the research project did not fall into the scope of the Human Research Act and did not require ethics committee approval. However, we obtained written consent from all participants to use the collected data for scientific purposes.

We delivered a 30-minute plenary presentation to all participants in the classroom instruction group. The presentation included an introduction to the concept and technology of Visual Patient as well as an educational video on how the system is used (supplementary video 1 in Multimedia Appendix 1). Subsequently, the participants were shown 4 scenarios in a randomized order. In 2 scenarios, vital signs were presented with the Visual Patient, and in the other 2 scenarios, the vital signs were presented as in conventional monitoring. The display scenarios were projected on a screen for either 3 or 10 seconds, after which the screen was blacked out. After each scenario, the participants were asked to rate every presented vital sign as normal, abnormal, or not perceived. Data collection was simultaneous for all participants, as each individual's desk was equipped with an iPad (Apple, Inc) containing a questionnaire (iSURVEY, Harvest Your Data) for the participants to complete [12].

The individual instruction group consisted of a selected sample from a previously published study on Visual Patient [4]. We selected 16 participants who were shown the same scenarios as the class instruction group. The methodology of this study is described in the previous publication. In brief, the data collection was similar for these participants, except that each participant was individually introduced to the Visual Patient followed by presentation of the scenarios to the participant alone.

Outcomes

To assess the educational success of class instruction on the Visual Patient technology, each rating of a vital sign was graded as correct or incorrect. This enabled us to compare the correct and incorrect perceptions of the vital signs displayed with both technologies.

At the end of the study, participants rated their introduction to the Visual Patient on a 5-point Likert scale (1=insufficient, 2=inadequate, 3=O.K., 4=good, and 5=very good).

Statistical Analysis

Data are provided as medians and interquartile ranges (IQR) regardless of normality or estimated marginal means for linear models. Normality was assessed with the Shapiro-Wilks test and visual inspection of quantile-quantile plots of dependent variables. Binary variables are presented as frequencies with percentages. The Wilcoxon signed-rank test was conducted to determine the effects of the Visual Patient display on the ability to correctly perceive vital signs after seeing the display for either 3 or 10 seconds. The different scores were approximately symmetrically distributed, as assessed by box plots. For both scenarios, post hoc descriptive graphs were created detailing whether each vital sign was perceived correctly, incorrectly, or not at all.

To compare the effects of classroom instruction and individual instruction, 2-way mixed analysis of variance (ANOVA) was calculated with the factors of display technology (within-subject) and instruction method (between-subject). There was a single outlier, as assessed by inspection of a box plot for values greater than 1.5 box lengths from the edge of the box. As the studentized residual for this outlier was only 3.06, it was retained in the analysis. Homogeneity was observed for variances (P>.05) and covariances (P>.001), as assessed by the Levene test of homogeneity of variances and the Box M test, respectively.

We fitted a mixed logistic regression model for the correct perception of vital signs with a random intercept for each participant. The model included the instruction variable, which denoted whether the participant received individual instruction or classroom instruction. We additionally adjusted for the display mode (Visual Patient vs conventional monitoring), the duration of the task (3 seconds vs 10 seconds), and the previous experience of the participants.

Analyses were conducted in SPSS 25 (IBM Corporation) and R version 3.6.1 (R Foundation for Statistical Computing). Figures were created using GraphPad Prism 8.1.1 (GraphPad Software, Inc). As group differences were calculated separately for both scenarios, a Bonferroni adjusted *P* value <.025 was considered to indicate statistical significance.

Availability of Data and Material

The data sets used and analyzed during the current study are available from the corresponding author on reasonable request.

Results

Participants

The 42 nurse anesthetists participating in the study reported a median professional experience of 12 years (IQR 3-31). Of the participants, 28/42 (67%) were female. As they were presented in randomized order with 2 sets, each consisting of a Visual Patient scenario and a matched conventional display scenario, 84 direct within-subject comparisons were performed. After the study, most participants in the class instruction group rated the introductory presentation as very good (20/42, 48%) or as good (13/42, 31%), whereas 9 participants did not take part in the follow-up survey.

The selected sample of 16 participants from a previous study, who received individual instruction on the Visual Patient, consisted of 8 (50%) physician anesthetists and 8 (50%) nurse anesthetists, where 11 (69%) were female. Table 1 gives an extended overview of the characteristics of the participants.

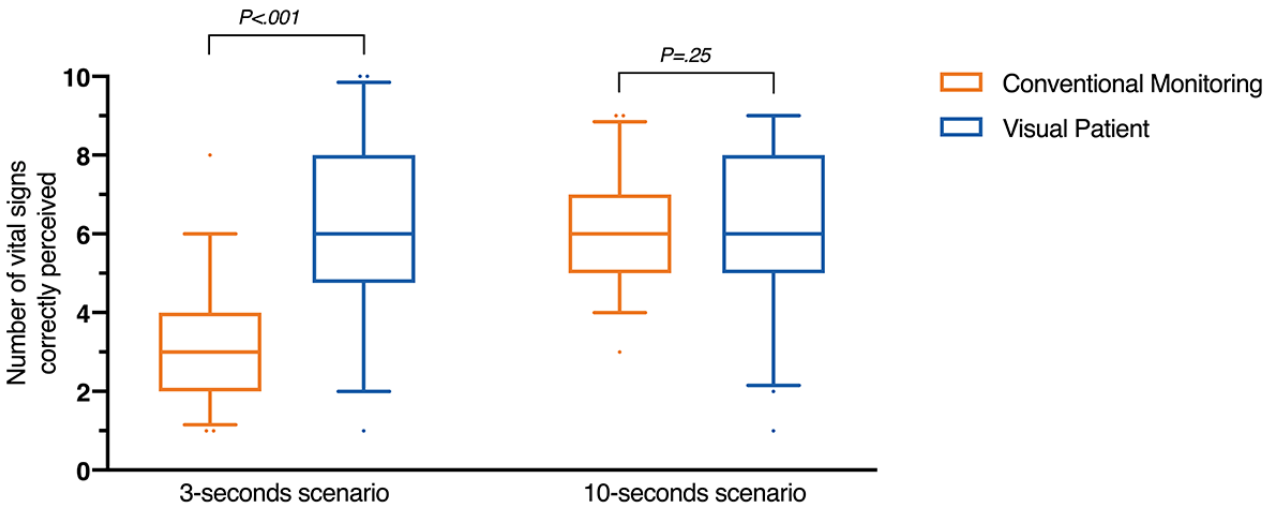
Table 1. Participant characteristics.

Characteristic	Class instruction group (n=42)	Individual instruction group (n=16)
Gender, n (%)		
Female	28 (67)	11 (69)
Male	14 (33)	5 (31)
Profession, n (%)		
Nurse anesthetist	42 (100)	8 (50)
Physician anesthetist	0 (0)	8 (50)
Experience, n (%)		
<1 year	2 (5)	1 (6)
1-5 years	9 (21)	5 (31)
>10 years	3 (7)	4 (25)
5-10 years	17 (40)	6 (38)
Unknown	11 (26)	0 (0)

Perception of Vital Signs

After the classroom instruction, when presented with the 3-second scenarios, participants were able to correctly perceive a median of 6 vital signs (IQR 4.8-8) with the Visual Patient and a median of 3 vital signs (IQR 2-4) with the conventional monitoring display. The Wilcoxon signed-rank test determined a significant median increase in the perception of vital signs (3) when participants were shown the Visual Patient compared to when they were shown the conventional display (z=–5.0; *P*<.001) with a large effect size of –0.55 (Figure 3).

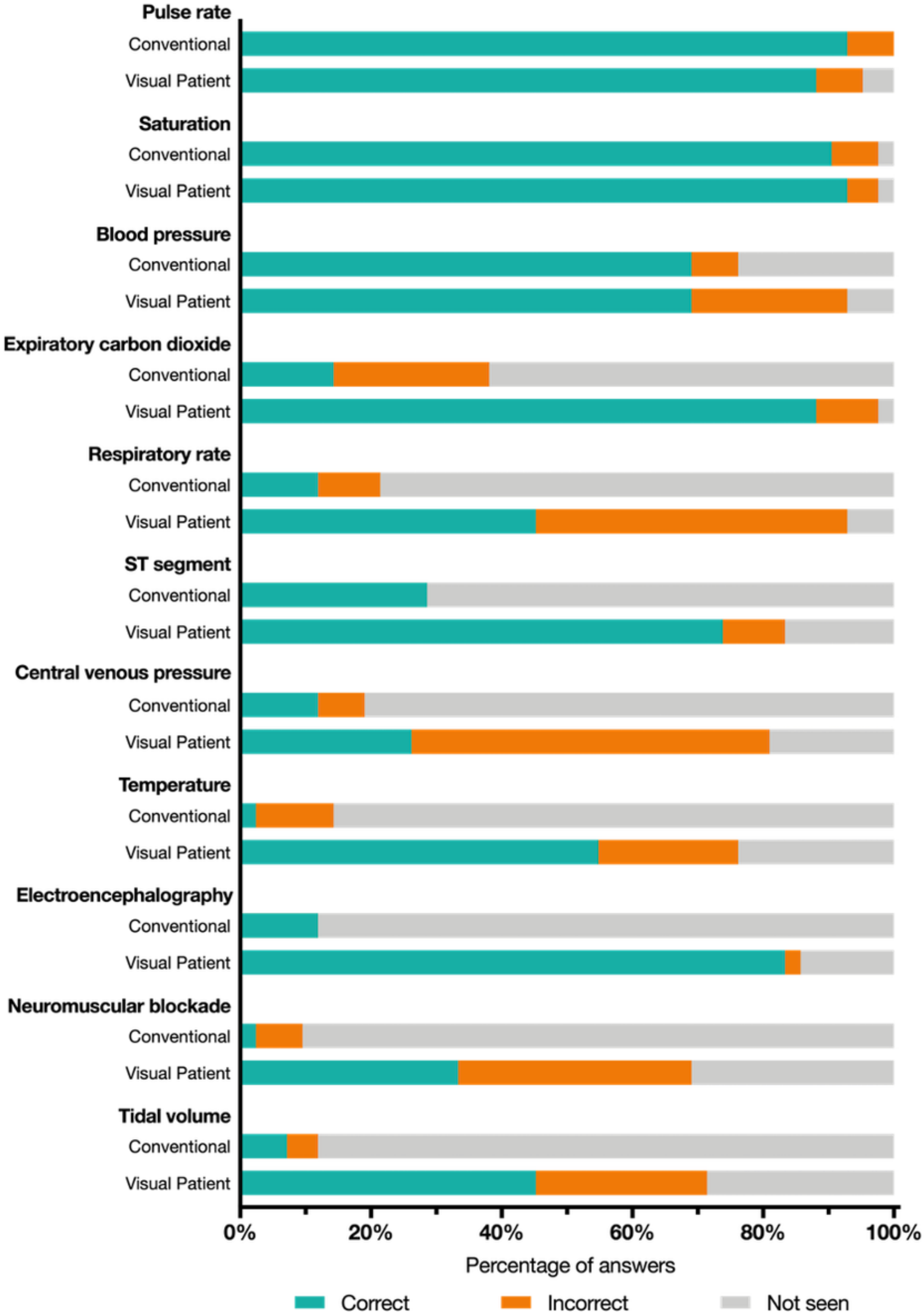
Figure 3. Box plots of the vital signs that were correctly perceived with both the Visual Patient and conventional monitoring. Participants were shown scenarios for either 3 or 10 seconds. Group differences were assessed by Wilcoxon signed-rank test. The whiskers indicate the 5th and 95th percentiles.



When the 10-second scenarios were shown after the class instruction, participants were able to correctly perceive a median of 6 vital signs (IQR 5-8) with the Visual Patient and a median of 6 vital signs (IQR 5-7) with the conventional monitoring display. Thus, there was no statistically significant median increase in the perception of vital signs (z=–1.2; *P*=.25) as determined by Wilcoxon signed-rank test (Figure 3).

Vital sign–specific descriptive analysis in the class instruction group showed that in the 3-second scenarios, nearly all participants were able to correctly perceive the pulse rate and oxygen saturation. Furthermore, most participants correctly recalled the blood pressure. The overall group difference was largest within the other parameters, as shown in Figure 4. Moreover, with the Visual Patient, the correct perceptions increased, but the incorrect perceptions also increased; therefore, the number of unperceived vital signs decreased (Figure 4, Multimedia Appendix 2). For the 10-second scenarios, vital sign–specific descriptive analysis is available in Multimedia Appendix 2.

Figure 4. Stacked bar graph indicating the perception of presented vital signs after the 3-second scenario. Percentages were calculated from the 4 possible answers to each vital sign: too high, normal, too low, and did not perceive. Depending on the presented scenario, the answers were rated as correct, incorrect, or not seen.

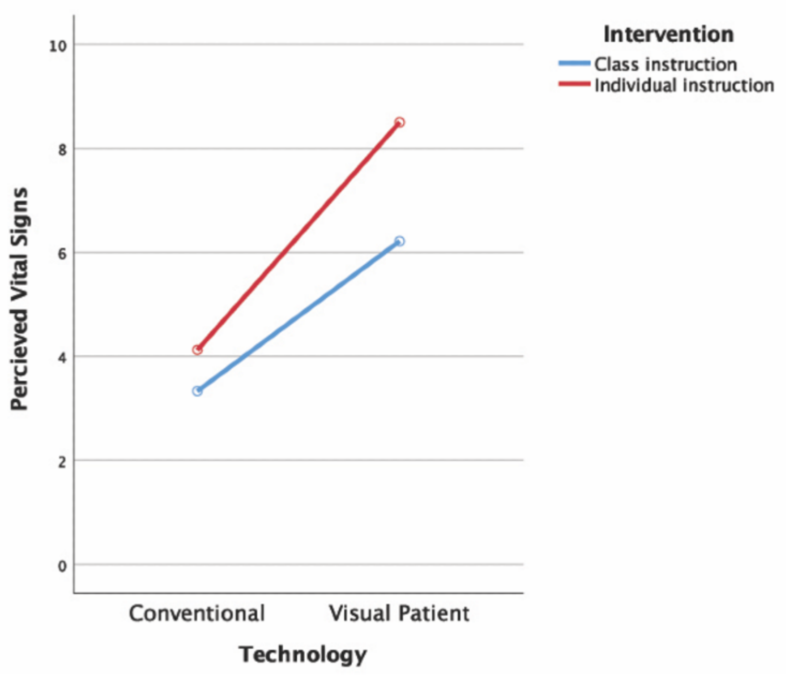


Effects of the Instruction Method

The 2-way mixed ANOVA indicated a statistically significant interaction between the teaching intervention and display technology for perceived vital signs ($F_{1,56}$ 4.61; *P*=.04; partial η^2=.076). Post-hoc univariate analysis yielded a statistically significant difference between the 2 teaching interventions for the Visual Patient ($F_{1,56}$ 14.42; *P*<.001; partial η^2=.205) but not for conventional monitoring ($F_{1,56}$ 3.06; *P*=.09; partial η^2=.052).

In the classroom instruction group, the estimated marginal means of the perceived vital signs increased from 3.3 (95% CI 2.9-3.8) with conventional monitoring to 6.2 (95% CI 5.6-6.8) with the Visual Patient. In the individual instruction group, the estimated marginal means of the perceived vital signs increased from 4.1 (95% CI 3.4-4.9) with conventional monitoring to 8.5 (95% CI 7.5-9.5) with the Visual Patient. As shown in Figure 5, this resulted in a mean difference of 2.3 between the number of vital signs perceived with the Visual Patient in the 2 instruction groups (95% CI 1.1-3.5; *P*<.001).

Figure 5. Marginal means of the perceived vital signs by the 2 instruction groups estimated by 2-way mixed ANOVA.



Mixed Logistic Regression

The mixed logistic regression model showed evidence of a difference between the teaching modes in favor of individual instruction, yielding OR 1.88 (95% CI 1.41-2.52; *P*<.001) for correct vital sign perception after individual instruction. Moreover, the model displayed very strong evidence for the superiority of Visual Patient, with OR 3.03 (95% CI 2.50-3.70; *P*<.001) for correct vital sign perception with the Visual Patient (Table 2).

Table 2. Mixed logistic regression for correct perception of vital signs with the random intercept for each participant.

Variable	OR[a] (95% CI)	*P* value
Teaching mode		
Class	Reference	N/A[b]
Individual	1.88 (1.41-2.52)	<.001
Display technology		
Conventional	Reference	N/A
Visual Patient	3.03 (2.50-3.70)	<.001
Scenario duration		
3 seconds	Reference	N/A
10 seconds	2.31 (1.91-2.80)	<.001
Experience		
<1 year	Reference	N/A
1-5 years	1.21 (0.67-2.16)	.53
5-10 years	1.01 (0.55-1.88)	.97
>10 years	0.86 (0.49-1.50)	.59

[a]OR: odds ratio.

[b]Not applicable.

Discussion

Principal Findings

Avatar-based patient monitoring is an alternative way to display vital signs. It can facilitate perception, reduce mental workload, and increase situation awareness [4,13]. This technology is generally well received by users and thought to be easy to learn; however, to implement it in larger health care systems, it must be trainable via class instruction [14].

In this study, we presented 42 anesthesia providers with 0.5 hours of class instruction on Visual Patient. Afterward, they were shown monitors with either conventional displays or Visual Patient displays and asked to interpret vital signs. If the participants saw the scenarios for 3 seconds, they were able to perceive significantly more vital signs with the Visual Patient. Further, the calculated effect size of the Visual Patient on correct perceptions was large (–0.55). No significant difference was found for the 10-second scenarios. These results are similar to those of a previous study by our research group on Visual Patient, where more vital signs were perceived with the Visual Patient after both 3 and 10 seconds [4]. However, in this study, the median difference was also less for the 10-second scenarios [4]. In the previous study, instruction was individual. To compare the efficacy of both instruction methods, we therefore compared the current sample with a selected historical sample from the previous study. While both instruction methods were successful, individual instruction yielded slightly better results.

In daily clinical practice, the superiority of the Visual Patient when seeing a monitor for 3 seconds may already be very relevant. It has been shown that anesthesia providers tend to look at patient monitors in short glances [9]. These glances become more frequent during critical situations, where vital signs can change rapidly and many can change at once [9]. In these cases, the median increase of 3 more vital signs perceived with the Visual Patient may make a crucial difference.

Participants were introduced to the Visual Patient according to our prespecified necessary criteria for general implementation. The teaching was conducted with 30 minutes of plenary classroom instruction (Figure 6), which was well received by the participants. The replication of results from previous studies, where each participant was introduced to the Visual Patient in a one-on-one setting, shows the feasibility of large-scale teaching. If avatar-based monitoring is implemented in health care systems or single hospitals, one-on-one teaching of each employee will not be practical. Employees will need to be trained to use the technology in a setting similar to that in our study [15]. Alternatively, e-learning may be considered or even no instruction at all, as the Visual Patient technology is generally perceived as intuitive to understand [14]. Animated avatars have been used to provide visual support in the education of patients with sensory impairment in the form of assistive computer vision [16,17] as well as in the education of children with autism, where an avatar can display emotions and support affective learning [18,19]. As an avatar is a manifestation of self and reflects already known images or movements in simplified ways, the avatar itself can be used as an educational tool [20]. Therefore, implementation of avatars without instruction may be a subject of future study.

Figure 6. The auditorium in which the introduction to the Visual Patient was conducted.

Post hoc analysis of perception of specific vital signs showed that with both technologies, the pulse rate and oxygen saturation were nearly always perceived correctly; also, the blood pressure was perceived correctly in approximately 69% of cases. In conventional monitoring displays, these figures are often the largest displayed and are thus easily perceived. The vital signs with more pronounced differences (expiratory carbon dioxide, respiratory rate, ST segment, central venous pressure, temperature, electroencephalography, neuromuscular blockade, and tidal volume) may be displayed in smaller sizes or in less prominent places. One advantage of the Visual Patient is that all vital signs are displayed in close proximity to each other and sometimes repeatedly (eg, the respiratory rate can be deduced from the lung movement or the expired “gas bubble”). This is supported by an eye-tracking study on Visual Patient, which showed that participants were able to visually fixate on more vital signs with Visual Patient than with conventional monitoring [21]. The close proximity further facilitates perception by peripheral vision [22]. Another advantage of Visual Patient is that due to the way the vital signs are displayed, parallel acquisition of information is possible. For example, users can recognize the pulsation frequency, color, and shape of an object in a single glance. To do the same in conventional number-based and waveform-based patient monitoring, users must read several numbers in several glances. [23]

Although the introduction to Visual Patient seemed to be sufficient and the monitoring capability improved, further progress may be possible with more detailed teaching or continued clinical use. The vital sign–based analysis showed that while correct perception of vital signs increased with the Visual Patient, incorrect perception increased as well. This may be due either to the design of these parameters or to inexperience with the Visual Patient. More detailed user perception studies are required to evaluate this result; however, it is more likely to be due to inexperience. The Visual Patient parameters and their display were calibrated using a Delphi process and generally show high interrater reliability, with a previously reported Fleiss kappa >.94 [4]. Participants seemed to be able to perceive these vital signs, as corroborated by the eye-tracking study [21]; however, their knowledge of Visual Patient may still have been insufficient to correctly interpret them. This implies that with further clinical use and practice, Visual Patient will yield even better situation awareness.

Strengths and Limitations

This study had some limitations. The study was simulation based; thus, translational evidence for clinical practice may be limited. Further studies in a high-fidelity simulation environment or in clinical practice are required. However, it is plausible that the effects would persist if used in a clinical setting, as the basic physiological specifications of information intake do not change. The results are in line with those of similar avatar-based monitoring systems, such as the Visual Clot, an animated blood clot that represents coagulation disorders [24]. This study also had particular strengths. The examined group was somewhat heterogenous, which increases the external validity. However, more physicians should be included in further studies. The study was not conducted in a sensory-sterile environment; both the instruction and data collection were performed with a large group, where the possible distractions are more similar to a real clinical atmosphere.

Conclusions

Although individual instruction on the Visual Patient is slightly more effective, class instruction is a viable teaching method; this increases the feasibility of large-scale introduction of health care providers to this novel technology. This study further contributes to the growing evidence of the superiority of avatar-based monitoring to conventional monitoring in certain situations.

Acknowledgments

This study was supported by funds from the Institute of Anesthesiology of the University Hospital of Zurich, Switzerland, and by a Proof of Concept Funding grant from the University of Zurich (UZ16/288PoC). DWT was additionally financed by a Filling the Gap career development grant from the University of Zurich. The external granting institutions had no further role in the design of the study; the collection, analysis, and interpretation of data; the writing of this report; or the decision to submit this article for publication.

Authors' Contributions

JR, AK, BA, DS, CN, and DT contributed to the study design. DT, BA, and CN collected the data. JR analyzed the data and created the figures. All authors contributed to the interpretation of the data. JR and AK wrote the first draft of the manuscript. All authors provided critical revisions to the manuscript before seeing and approving the final version.

Conflicts of Interest

The University of Zurich (Zurich, Switzerland) and Koninklijke Philips N.V. (Amsterdam, Netherlands) entered a joint development and licensing agreement to develop avatar-based monitoring software based on technology that is owned by the University and described in this manuscript. As part of their contract with the University, as designated inventors, the authors DWT and CBN may receive royalties.

Multimedia Appendix 1
Supplementary video 1: the Visual Patient educational video shown to all participants of this study.
[MOV File , 52399 KB - mededu_v6i1e17922_app1.mov]

Multimedia Appendix 2
Supplementary tables and figures.
[DOCX File , 999 KB - mededu_v6i1e17922_app2.docx]

References

1. WHO Guidelines For Safe Surgery 2009: Safe Surgery Saves Lives. Geneva: World Health Organization, Patient Safety; 2009.
2. Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T, et al. Size and distribution of the global volume of surgery in 2012. Bull. World Health Organ 2016 Mar 01;94(3):201-209F. [doi: 10.2471/blt.15.159293]
3. Tscholl DW, Handschin L, Rössler J, Weiss M, Spahn DR, Nöthiger CB. It's not you, it's the design - common problems with patient monitoring reported by anesthesiologists: a mixed qualitative and quantitative study. BMC Anesthesiol 2019;19:87. [doi: 10.21203/rs.2.238/v3]
4. Tscholl DW, Handschin L, Neubauer P, Weiss M, Seifert B, Spahn DR, et al. Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals. British Journal of Anaesthesia 2018 Sep;121(3):662-671. [doi: 10.1016/j.bja.2018.04.024]
5. Endsley MR. Toward a Theory of Situation Awareness in Dynamic Systems. Hum Factors 2016 Nov 23;37(1):32-64. [doi: 10.1518/001872095779049543]
6. Endsley M. Designing For Situation Awareness: An Approach To User-centered Design, Second Edition. Boca Raton, FL: CRC Press; 2011.
7. Garot O, Rössler J, Pfarr J, Ganter MT, Spahn DR, Nöthiger CB, et al. Avatar-based versus conventional vital sign display in a central monitor for monitoring multiple patients: a multicenter computer-based laboratory study. BMC Med Inform Decis Mak 2020 Feb 10;20(1). [doi: 10.1186/s12911-020-1032-4]
8. Schulz CM, Endsley MR, Kochs EF, Gelb AW, Wagner KJ. Situation Awareness in Anesthesia. Anesthesiology 2013;118(3):729-742. [doi: 10.1097/aln.0b013e318280a40f]
9. Schulz C, Schneider E, Fritz L, Vockeroth J, Hapfelmeier A, Brandt T, et al. Visual attention of anaesthetists during simulated critical incidents. Br J Anaesth 2011 Jun;106(6):807-813. [doi: 10.1093/bja/aer087]
10. Gaba DM, Howard SK, Small SD. Situation Awareness in Anesthesiology. Hum Factors 2016 Nov 23;37(1):20-31. [doi: 10.1518/001872095779049435]
11. Fioratou E, Flin R, Glavin R, Patey R. Beyond monitoring: distributed situation awareness in anaesthesia. British Journal of Anaesthesia 2010 Jul;105(1):83-90. [doi: 10.1093/bja/aeq137]
12. Tscholl DW, Weiss M, Spahn DR, Noethiger CB. How to Conduct Multimethod Field Studies in the Operating Room: The iPad Combined With a Survey App as a Valid and Reliable Data Collection Tool. JMIR Res Protoc 2016 Jan 05;5(1):e4. [doi: 10.2196/resprot.4713]
13. Drews FA, Westenskow DR. The Right Picture Is Worth a Thousand Numbers: Data Displays in Anesthesia. Hum Factors 2016 Sep 06;48(1):59-71. [doi: 10.1518/001872006776412270]
14. Tscholl DW, Weiss M, Handschin L, Spahn DR, Nöthiger CB. User perceptions of avatar-based patient monitoring: a mixed qualitative and quantitative study. BMC Anesthesiol 2018 Dec 11;18(1). [doi: 10.1186/s12871-018-0650-1]
15. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3(1). [doi: 10.1186/s40359-015-0089-9]
16. Leo M, Furnari A, Medioni G, Trivedi M, Farinella G. Deep Learning for Assistive Computer Vision. In: Champringer International Publishing. 2019 Presented at: European Conference on Computer Vision; 2018 Sept 8-14; Munich, Germany p. 3-14. [doi: 10.1007/978-3-030-11024-6_1]
17. Bouzid Y, Khenissi M, Jemni M. The Effect of Avatar Technology on Sign Writing Vocabularies Acquisition for Deaf Learners. 2016 Presented at: IEEE 16th International Conference on Advanced Learning Technologies; 2016 July 25-28; Austin, TX, USA p. 441-445. [doi: 10.1109/ICALT.2016.127]
18. Leo M, Coco M, Carcagni P. Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. In: Proceedings of the IEEE. Washington D.C., USA: IEEE Computer Society; 2015 Presented at: IEEE International Conference on Computer Vision Workshop; 2015 Dec 11-18; Santiago, Chile p. 537. [doi: 10.1109/iccvw.2015.76]
19. Johnson E, Hervás R, Gutiérrez López de la Franca C, Mondéjar T, Ochoa SF, Favela J. Assessing empathy and managing emotions through interactions with an affective avatar. Health Informatics J 2016 Sep 30;24(2):182-193. [doi: 10.1177/1460458216661864]
20. Falloon G. Using avatars and virtual environments in learning: What do they have to offer? Br J Educ Technol 2010;41:108-122. [doi: 10.1111/j.1467-8535.2009.00991.x]
21. Tscholl DW, Rössler J, Handschin L, Seifert B, Spahn DR, Nöthiger C. The Mechanisms Responsible for Improved Information Transfer in Avatar-Based Patient Monitoring: Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2020 Mar 16;22(3):e15070 [FREE Full text] [doi: 10.2196/15070] [Medline: 32175913]
22. Pfarr J, Ganter MT, Spahn DR, Nöthiger CB, Tscholl DW. Avatar-Based Patient Monitoring With Peripheral Vision: A Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2019 Jul 17;21(7):e13041. [doi: 10.2196/13041]
23. Tscholl DW, Rössler J, Handschin L, Seifert B, Spahn DR, Nöthiger CB. The Mechanisms Responsible for Improved Information Transfer in Avatar-Based Patient Monitoring: Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2020 Mar 16;22(3):e15070 [FREE Full text] [doi: 10.2196/15070] [Medline: 32175913]

24. Rössler J, Meybohm P, Spahn DR, Zacharowski K, Braun J, Nöthiger CB, et al. Improving decision making through presentation of viscoelastic tests as a 3D animated blood clot: the Visual Clot. Anaesthesia 2020 Feb 06. [doi: 10.1111/anae.14985]

Abbreviations

ANOVA: analysis of variance
IQR: interquartile range
OR: odds ratio

Edited by G Eysenbach; submitted 22.01.20; peer-reviewed by A Sribar, M Leo, IY Jang, JJ Mira, I Kedan; comments to author 12.02.20; revised version received 26.02.20; accepted 23.03.20; published 23.04.20.

Please cite as:
Rössler J, Kaserer A, Albiez B, Braun J, Breckwoldt J, Spahn DR, Nöthiger C, Tscholl DW
Comparing Classroom Instruction to Individual Instruction as an Approach to Teach Avatar-Based Patient Monitoring With Visual Patient: Simulation Study
JMIR Med Educ 2020;6(1):e17922
URL: http://mededu.jmir.org/2020/1/e17922/
doi:10.2196/17922
PMID:32205304

Original Paper

A Virtual 3D Dynamic Model of Caries Lesion Progression as a Learning Object for Caries Detection Training and Teaching: Video Development Study

Juan Sebastian Lara[1*], DDS, MSci, PhD; Mariana Minatel Braga[2*], DDS, PhD; Carlos Gustavo Zagatto[3*], BSE; Chao Lung Wen[3*], MD, PhD; Fausto Medeiros Mendes[2*], DDS, MSci, PhD; Pedroza Uribe Murisi[4*], DDS, MSci; Ana Estela Haddad[2*], DDS, MSci, PhD

[1]Department of Cariology, Operative Dentistry and Dental Public Health, Indiana University School of Dentistry, Indianapolis, IN, United States

[2]Department of Pediatric Dentistry, School of Dentistry, University of Sao Paulo, Sao Paulo, Brazil

[3]Discipline of Telemedicine, Department of Pathology, School of Medicine, University of Sao Paulo, Sao Paulo, Brazil

[4]Department of Pediatric Dentistry, Dental School, University of Guadalajara, Guadalajara, Mexico

[*]all authors contributed equally

Corresponding Author:
Juan Sebastian Lara, DDS, MSci, PhD
Department of Cariology, Operative Dentistry and Dental Public Health
Indiana University School of Dentistry
415 Lansing Street, 144C
Indianapolis, IN, 46202
United States
Phone: 1 3142745463
Email: laraseb@iu.edu

Abstract

Background: In the last decade, 3D virtual models have been used for educational purposes in the health sciences, specifically for teaching human anatomy and pathology. These models provide an opportunity to didactically visualize key spatial relations that can be poorly understood when taught by traditional educational approaches. Caries lesion detection is a crucial process in dentistry that has been reported to be difficult to learn. One especially difficult aspect is linking clinical characteristics of the different severity stages with their histological features, which is fundamental for treatment decision-making.

Objective: This project was designed to develop a virtual 3D digital model of caries lesion formation and progression to aid the detection of lesions at different severity stages as a potential complement to traditional lectures.

Methods: Pedagogical planning, including identification of objectives, exploration of the degree of difficulty of caries diagnosis–associated topics perceived by dental students and lecturers, review of the literature regarding key concepts, and consultation of experts, was performed prior to constructing the model. An educational script strategy was created based on the topics to be addressed (dental tissues, biofilm stagnation areas, the demineralization process, caries lesion progression on occlusal surfaces, clinical characteristics related to different stages of caries progression, and histological correlations). Virtual 3D models were developed using the Virtual Man Project and refined using multiple 3D software applications. In the next phase, computer graphic modelling and previsualization were executed. After that, the video was revised and edited based on suggestions. Finally, explanatory subtitles were generated, the models were textured and rendered, and voiceovers in 3 languages were implemented.

Results: We developed a 6-minute virtual 3D dynamic video in 3 languages (English, Spanish, and Brazilian Portuguese) intended for dentists and dental students to support teaching and learning of caries lesion detection. The videos were made available on YouTube; to date, they have received more than 100,000 views.

Conclusions: Complementary pedagogical tools are valuable to support cariology education. This tool will be further tested in terms of utility and usability as well as user satisfaction in achieving the proposed objectives in specific contexts.

(JMIR Med Educ 2020;6(1):e14140) doi:10.2196/14140

KEYWORDS

3d virtual models; dental education; e-learning; learning object; caries; cariology

Introduction

Dental caries is one of the most prevalent chronic diseases worldwide [1]. It is caused by the interaction of several factors that culminate in dissolution of the localized chemical tooth structure by metabolic events occurring in the oral biofilm [2]. This cumulative mineral loss is known as a caries lesion; these lesions can vary from simple changes in enamel translucency to extensive cavities involving the dentine and pulp [3].

Dental caries are detected by recognizing the signs and symptoms involved in the abovementioned process [4]. The importance of caries detection lies in the possibility of confirming the presence or absence of disease, assessing its prognosis, contributing to the decision-making process, informing the patient, and monitoring the clinical course of the disease [5]. In this sense, adequate caries detection is fundamental for planning and implementing health policies aimed to control the disease [6].

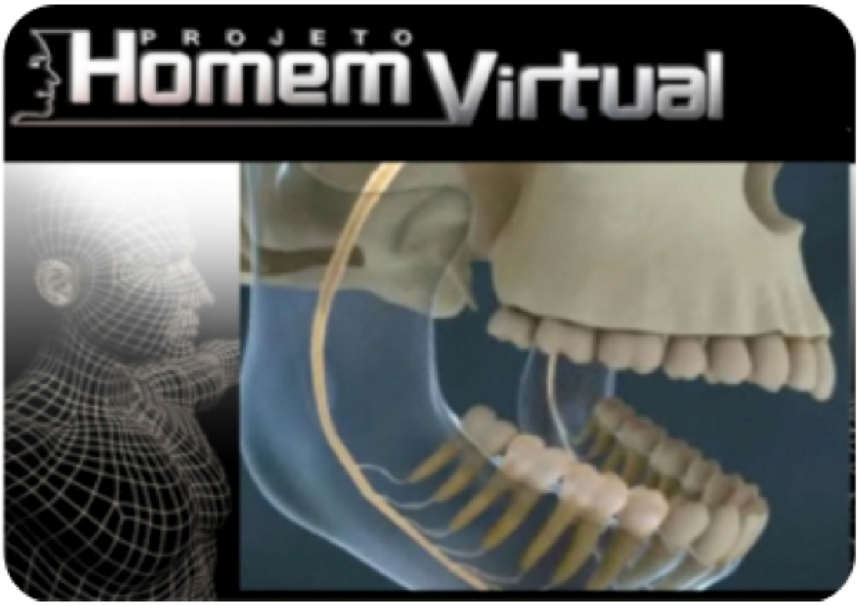
Several caries classification and detection methods have been developed to assess different stages of caries lesions [7]. However, the many differences and lack of standardization of these methods highlight the need to develop a defined, standardized, and validated caries detection system based upon current scientific evidence and the consensus of experts in the field of cariology [8]. In this regard, the International Caries Detection and Assessment System (ICDAS) [9] was designed to detect 6 stages of the caries process according to the disease severity, ranging from early visual changes in the enamel to extensive cavitation. Although this system has been widely used and has been shown to contribute to more accurate caries lesion detection [10], developing teaching tools is important and necessary to achieve and effectively disseminate new concepts and paradigms to facilitate their understanding and use [11]. These tools will reduce the difficulties of applying such concepts in a clinical scenario.

Based on this, the ICDAS Foundation designed an e-learning program to universalize and spread the use of their system [12]. This free 90-minute tool can be accessed online in 4 different languages to support training, provide dental examination protocols, and review the scoring system. Although the ICDAS e-learning program has been shown to improve the diagnostic skills of dental students for the detection of occlusal caries [13,14], specific clinical characteristics of different stages of caries lesion progression could not be linked to their respective histopathological features, which is important to understand the prognosis and influence of these stages on clinical decision-making.

3D animation models can show spatial and dynamical relationships from almost any angle; this can provide information that may be difficult to acquire using traditional static learning resources [15]. In this sense, the Virtual Man Project [16] developed at the Telemedicine Discipline of the University of Sao Paulo creates 3D images and animations of the human body that aid the comprehension of anatomy, physiology, pathologies, drug interactions, and surgical techniques in several areas (Figure 1).

Thus, the aim of the present project was to develop a digital, dynamic, and virtual 3D model of the formation, progression, and detection of caries lesions at different stages using the ICDAS at the Virtual Man Project Laboratory to complement traditional teaching resources.

Figure 1. The Virtual Man Project developed at the Telemedicine Discipline, School of Medicine, University of Sao Paulo, Brazil.



Methods

This descriptive study was developed in collaboration with the Discipline of Telemedicine (School of Medicine), the Teledentistry Centre, and the Department of Pediatric Dentistry (School of Dentistry) at the University of Sao Paulo, Brazil. The study was approved by the Ethics Committee of the Dental School (protocol 206.345/2013).

Pedagogic Planning

This phase comprised the initial steps to develop a learning object oriented toward the detection of caries lesions based on their developmental stages and the differences among their clinical characteristics. In this phase, we discussed the objectives, the topics to approach, and the best methodologies for knowledge transmission. Firstly, a team of experienced lecturers and researchers in the area of cariology was formed to discuss the learning object purpose, key topics representing the minimum skills a dental student should develop in this field as a future dental practitioner, and the possibility of using technology to achieve the proposed goals. Sources such as the First Consensus Workshop on the Development of a European Curriculum in Cariology [17] and a study on current cariology education in dental schools in Spanish-speaking Latin American countries [18] were considered at this point.

As a second step in the previous study, we assessed the degrees of difficulty of caries detection–related learning topics perceived by dental students and lecturers [19]. In this phase, we used a conjoint analysis survey to determine the most difficult topics to learn regarding the detection of caries lesions [20]. In this survey, respondents were asked to rate the perceived degree of difficulty not by individual subactions but in combinations known as profiles. Conjoint analysis allows the identification of subactions that are considered to be better or worse examples of each research engagement action by calculating numerical weights, which are called utilities. These utilities represent the score to be assigned to each subaction. The topic considered to be the most difficult by students and lecturers was the histology of caries lesions (the correlation between the clinical characteristics of a caries lesion and its histological depth). Therefore, we based the construction of the present learning object on this specific subject to address this difficulty didactically.

Graphic Design and Video Production

First, images of clinical and histological caries lesions were acquired to include them in the learning object as complementary supports. To that end, an examiner trained in the ICDAS criteria selected a sample of human teeth (N=12) with different caries lesion severity stages on occlusal surfaces (scores 0-6) from the Human Teeth Bank of the Dental School, University of Sao Paulo. Clinical images of these lesions were then obtained with a digital camera (DS126151 EOS Digital Rebel XTi, Canon) with a macro lens (EF 100mm 1:2.8, Canon). Then, the teeth were fixed with the crown exposed to Eppendorf tubes using utility wax and transparent acrylic resin. Longitudinal sections (100 micrometers) were made at the center of each lesion using a cutting machine (IsoMet 1000 precision sectioning saw, Buehler) with a diamond grinding disc (Extec 12205). The histological sections were analyzed under a stereomicroscope (M80, Leica), and the images were captured and processed with a digital camera (DFC 295, Leica) and QWin Plus software (Leica). This material was saved in digital files until its inclusion in the learning object.

A technological plan structure and interactive tele-education strategies were developed in partnership with a multidisciplinary team composed of cariology experts, digital designers, journalists, and tele-education strategists. This phase was carried out at the Virtual Man Project Laboratory using dual Pentium 4 graphic workstations (Xeon HT) with 4 gigabytes of RAM, a professional video board, tablets, the 3D Studio Max program (AutoDesk, Inc), and Photoshop and After Effects software (Adobe, Inc).

Before graphical production, a descriptive lecture regarding the caries process and the ICDAS criteria on occlusal surfaces was given to the team members who were not familiar with cariology. Concept descriptions as well as clinical and histological images of caries lesions were shown. Then, an educational script strategy was created to define the sequence in which the topics would appear in the video, emphasizing the most difficult topics (Textbox 1).

Textbox 1. Educational script showing the sequence of topics in the constructed video according to the pedagogical planning.

Sequence of topics

1. Presentation
2. Dental structures (enamel, dentin, pulp)
3. Dental enamel structure, microscopic view
4. Dentine structure, microscopic view
5. Plaque stagnation areas
6. Demineralization process
7. Caries lesion formation
8. ICDAS scores (clinical/histological)
9. ICDAS scores (histological correlation)
10. Credits

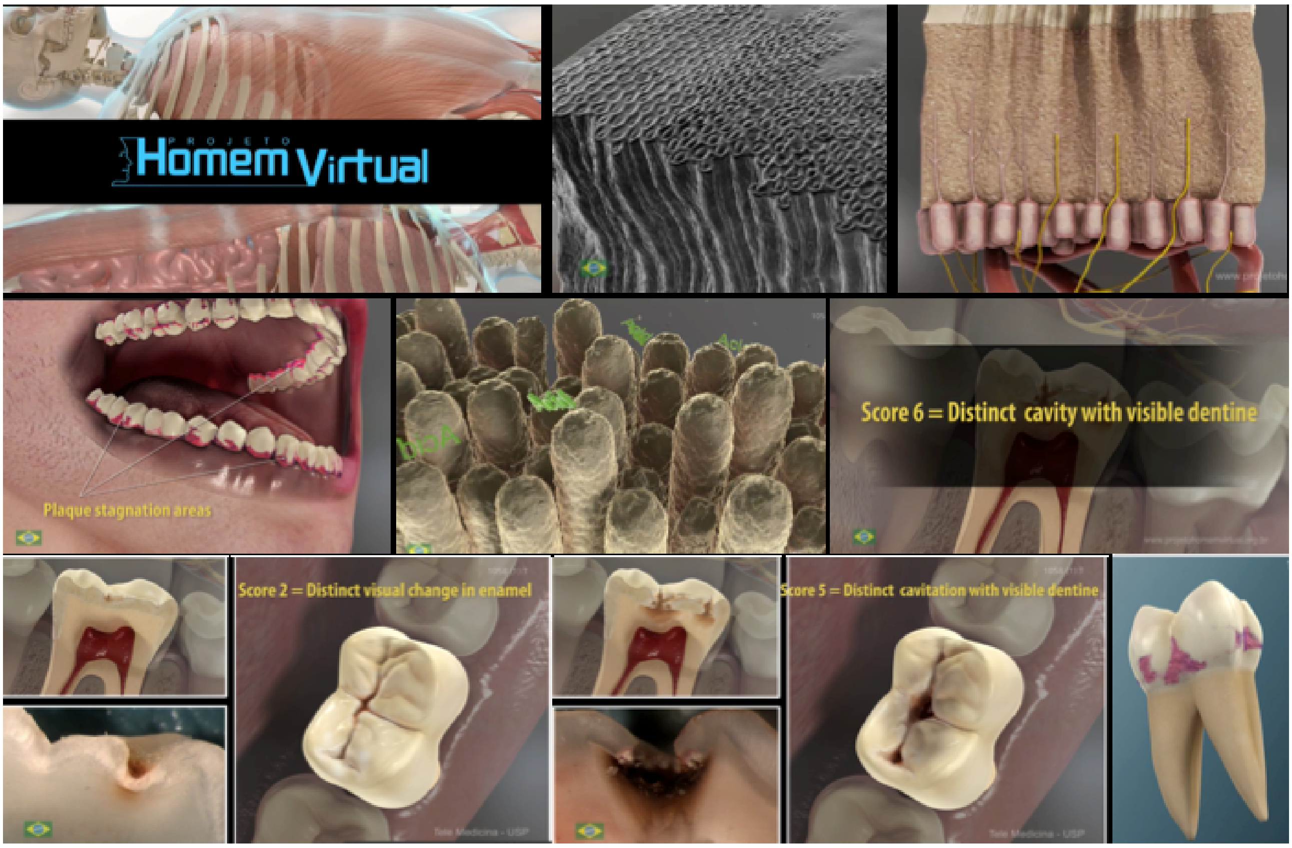
Graphical computer models (using the Virtual Man Project) of the clinical and histological images and a preview video were promptly generated. A group of experts in the field of cariology who were not involved in the video production reviewed this first version of the video. After that, the video was edited to correct some theoretical and technical inconsistencies. Then, a new version was produced that incorporated all the suggestions; after the team's approval, this video was rendered and textured. To achieve this, an image synthesis process was performed to generate photorealistic images (geometry, viewpoint, texture, lighting, shading, etc.) from the developed 3D models using computer programs. By this process, a silent 6-minute video was created using the Virtual Man Project (Figure 2).

Figure 2. Scenes of the video produced at the Virtual Man Laboratory. ICDAS: International Caries Detection and Assessment System.

Voiceover Recording

After the graphic design phase was complete, voiceovers were recorded to be played over the video to narrate the dynamic illustrations in an understandable way. Scripts in 3 languages (English, Spanish, and Brazilian Portuguese) were written and revised by native experts. After that, the scripts were adjusted and synchronized with the times and sequences in the video. We engaged a native-speaking narrator from the United Kingdom, Mexico, and Brazil. Each of the 3 narrators received a brief explanation about the video and its objectives together with the script. They were then taken into a recording studio and provided with instructions on how to record the voiceovers. Voice volume and speed tests were performed. Later, the narrator read the script while the voiceover was recorded. If the narrator made mistakes, they were required to start over from the previous paragraph. Repetitions were made when necessary. Finally, the recordings of the scripts were edited and incorporated into the body of the video. The final versions, a 6-minute video in each language, were sent to the scientific board for revision. Minimal corrections and additions were performed at this point (Figure 3).

Figure 3. Frames from the final English version of the 6-minute video about the caries process and the International Caries Detection and Assessment System.



Results

Using the process described in the Methods section, a 6-minute dynamic video was produced in 3 languages (English, Spanish, and Brazilian Portuguese) showing the dental structures, biofilm stagnation areas, caries lesion formation, demineralization process, caries lesion progression, and severity stages of caries lesions on occlusal surfaces according to the ICDAS (Figure 2 and Figure 3). This process, from conception to the final product, required approximately 2 years. The first 6 months were dedicated to idealization of the project, formation of the team, and design of the methodology to assess the topics to be included in the learning object. In the following months, the multidisciplinary team was formed and the audiovisual production proceeded. Finalization of the audiovisual production required approximately 1.5 years. This included product conception, design, production, editing, rendering, and voiceover incorporation. The project leader was exclusively dedicated to the project, and the Virtual Man Project Laboratory staff worked an average of 20 hours per week on the production of the material. All 3 videos [21-23] were uploaded to YouTube in 2016; since then, they have received more than 130,000 views (English version: 33,000, Spanish version: 28,000, and Brazilian Portuguese version: 72,000). The 3 versions of the learning object can be accessed on the YouTube platform using the keywords “ICDAS” and “caries”.

Discussion

Principal Findings

Interest is increasing in developing educational resources using information and communication technology to improve students’ understanding of human body processes [15]. The developed tool is presented in an audiovisual media format that is compatible with computers, tablets, and smartphones using the Virtual Man Project as an innovative, dynamic, and directed communication method. The tool implements 3D graphical modeling and uses a visual classification system to transmit knowledge associated with caries lesion formation and its clinical manifestations based on the developmental stages of caries, which may benefit the caries detection process [10].

This project represents an improvement in educational infographics, as it may facilitate and accelerate understanding related to a specific matter [20]. The Virtual Man Project Laboratory at the University of Sao Paulo had previously developed some dentistry-related content, such as tooth extraction and mandibular nerve anesthesia [24] and atraumatic restorative treatment [25]. This learning object represents the continuous production of learning tools in the area of dentistry, specifically cariology. When watching the video, viewers can observe 3D animated anatomical structures that simulate the demineralization process due to bacterial acid production, the caries process, and the severity stages of lesions as well as their histological correlations; these are difficult topics to assimilate by conventional methods [15,26].

Part of the pedagogical planning for the development of this learning object was based on the findings of a previous study [19]. Those findings were extremely important, as they guided the development of the learning object based on students' real expectations to stimulate ideal achievement of the knowledge and skills required to detect caries lesions in a clinical scenario. However, the impact that this learning object will have on students' learning and competence acquisition is a matter of future study.

We consider that this tool may have an impact on the theoretical understanding of caries lesion formation and progression and therefore may improve students´ knowledge and grades. This is supported by a study in which students who accessed virtual tools scored higher on assessments than students who did not [27]. This tool is important because the implementation of multimedia designs for anatomical teaching purposes reduces students' cognitive load [26] and permits dissemination of information to more students, with significant effects on improving their understanding of the relevant morphology [26].

As mentioned, the ICDAS e-learning program is an interactive resource that supports training in the use of the ICDAS criteria for dental education, examination protocols, and scoring systems [13] using static images, text, and voiceover recordings. The advantage of dynamically correlating the clinical process with the histopathological features of disease in the presented learning object may complement not only existing online resources but also traditional lectures, helping educators improve their teaching methodology [27].

One of our main goals is the dissemination of the produced tool. Open access is an ideal aspect of this process, and the first version of the video is already available on YouTube. However, additional steps should be performed to test the characteristics of the learning material as well as to address copyright issues; these steps are currently in progress. In the phase described in this paper, the authors created a learning object in 3 different languages: English, Spanish, and Brazilian Portuguese. This can be seen as an advantage in terms of dissemination, as these 3 languages have some of the largest populations of first-language speakers in the world [28].

As a next step to fully address the efficacy of the current learning object, the authors will test this tool in a student population in different contexts and countries. This testing will be conducted to validate and assess the potential benefits, obstacles, and user acceptability of the learning object as a novel pedagogical resource in the area of cariology. Therefore, we will perform a multicenter randomized study involving dental students from different countries with the aim of evaluating the impact of the 3D virtual model as a learning object on the training and teaching of undergraduate dental students to detect caries lesions using the ICDAS; the results of this study will be discussed in a future paper.

Conclusions

We produced a 6-minute virtual 3D video intended for dentists and dental students to support teaching and learning of the caries detection process. We suggest that complementary pedagogical tools, such as the one described here, are valuable to complement education in cariology. This learning object will be further tested in terms of utility and usability as well as user satisfaction in achieving the proposed objectives in specific contexts.

Acknowledgments

This study was funded by the Coordination for the Improvement of Higher Education Personnel of the Brazilian government, which granted scholarships to the MSc and PhD candidates during the conduction of the study. The authors are grateful to Diogo Miranda for his assistance with the voiceover recording and editing process and for lending his voice to the Portuguese version; we are also grateful to Harry Johnson for his participation in the English version. To Kim Ekstrand, Stefania Martignon, and Nigel Pitts, we offer special recognition of their support to make this project a reality.

Authors' Contributions

JSL, MMB, CLW, FMM, and AEH contributed substantially to the concept and design of this study. JSL, MMB, CGZ, CLW, FMM, and AEH designed the study and worked on it from its conception. JSL, MMB, PUM, and AEH were involved in the collection of the data (in the previous study) and of the materials for the 3D model development. JSL, CGZ, and CLW worked on the technical development and editing process of the 3D model. JSL and PUM recorded the voiceovers and drafted the manuscript. JSL, MMB, FMM, and AEH critically reviewed the intellectual content of the manuscript. All authors have provided approval of the final version to be published. JSL is responsible for the integrity of the work as a whole.

Conflicts of Interest

None declared.

References

1. Ismail AI, Tellez M, Pitts NB, Ekstrand KR, Ricketts D, Longbottom C, et al. Caries management pathways preserve dental tissues and promote oral health. Community Dent Oral Epidemiol 2013 Feb;41(1):e12-e40. [doi: 10.1111/cdoe.12024] [Medline: 24916676]
2. Fejerskov O. Pathology of dental caries. In: Fejerskov O, Nyvad B, Kidd EAM. (eds) Dental caries: the disease and its clinical management, 4th ed. Copenhagen, Denmark: Wiley-Blackwell, chapter. Copenhagen: Wiley Blackwell; 2015:5-1.

3. Fejerskov O, Nyvad B, Kidd E. Dental caries: the disease and its clinical management, 4th ed. In: Copenhagen, Denmark: Wiley-Blackwell, 480 p. Copenhagen: Wiley Blackwell; 2015:978-971.
4. Nyvad B, Machiulskiene V, Soviero V, Baelum V. Visual-tactile caries diagnosis. In: Fejerskov O, Nyvad B, Kidd EAM. (eds) Dental caries: the disease and its clinical management, 4th ed. Copenhagen, Denmark: Wiley-Blackwell, chapter. Copenhagen: Wiley Blackwell; 2015:11-11.
5. Knottnerus JA, van Weel C, Muris JWM. Evaluation of diagnostic procedures. BMJ 2002 Feb 23;324(7335):477-480 [FREE Full text] [Medline: 11859054]
6. Schulte AG, Buchalla W, Huysmans MDNJM, Amaechi BT, Sampaio F, Vougiouklakis G, et al. A survey on education in cariology for undergraduate dental students in Europe. Eur J Dent Educ 2011 Nov;15 Suppl 1:3-8. [doi: 10.1111/j.1600-0579.2011.00708.x] [Medline: 22023540]
7. Altarakemah Y, Al-Sane M, Lim S, Kingman A, Ismail AI. A new approach to reliability assessment of dental caries examinations. Community Dent Oral Epidemiol 2013 Aug;41(4):309-316. [doi: 10.1111/cdoe.12020] [Medline: 23278284]
8. Ismail AI. Visual and visuo-tactile detection of dental caries. J Dent Res 2004;83 Spec No C:C56-C66. [Medline: 15286124]
9. Ismail AI, Sohn W, Tellez M, Amaya A, Sen A, Hasson H, et al. The International Caries Detection and Assessment System (ICDAS): an integrated system for measuring dental caries. Community Dent Oral Epidemiol 2007 Jun;35(3):170-178. [doi: 10.1111/j.1600-0528.2007.00347.x] [Medline: 17518963]
10. Gimenez T, Piovesan C, Braga MM, Raggio DP, Deery C, Ricketts DN, et al. Visual Inspection for Caries Detection: A Systematic Review and Meta-analysis. J Dent Res 2015 Jul;94(7):895-904. [doi: 10.1177/0022034515586763] [Medline: 25994176]
11. Fejerskov O. Changing paradigms in concepts on dental caries: consequences for oral health care. Caries Res 2004 May;38(3):182-191 [FREE Full text] [doi: 10.1159/000077753] [Medline: 15153687]
12. ICDAS and ICCMS Foundation. ICCMS. 2019 Mar. International Caries Classification and Management System (ICCMS) URL: https://iccms-web.com/ [accessed 2019-03-25] [WebCite Cache ID 77931uDee]
13. El-Damanhoury HM, Fakhruddin KS, Awad MA. Effectiveness of teaching International Caries Detection and Assessment System II and its e-learning program to freshman dental students on occlusal caries detection. Eur J Dent 2014 Oct;8(4):493-497 [FREE Full text] [doi: 10.4103/1305-7456.143631] [Medline: 25512730]
14. Diniz MB, Lima LM, Santos-Pinto L, Eckert GJ, Zandoná AGF, de Cássia Loiola Cordeiro R. Influence of the ICDAS e-learning program for occlusal caries detection on dental students. J Dent Educ 2010 Aug;74(8):862-868 [FREE Full text] [Medline: 20679455]
15. Allen LK, Bhattacharyya S, Wilson TD. Development of an interactive anatomical three-dimensional eye model. Anat Sci Educ 2015 Sep;8(3):275-282. [doi: 10.1002/ase.1487] [Medline: 25228501]
16. Bhom GM, Chao LW. Projeto Homem Virtual URL: http://www.projetohomemvirtual.org.br [accessed 2019-03-25] [WebCite Cache ID 77950Jhqv]
17. Schulte AG, Pitts NB, Huysmans MCDNJM, Splieth C, Buchalla W. European Core Curriculum in Cariology for undergraduate dental students. Eur J Dent Educ 2011 Nov;15 Suppl 1:9-17. [doi: 10.1111/j.1600-0579.2011.00694.x] [Medline: 22023541]
18. Martignon S, Gomez J, Tellez M, Ruiz JA, Marin LM, Rangel MC. Current cariology education in dental schools in Spanish-speaking Latin American countries. J Dent Educ 2013 Oct;77(10):1330-1337 [FREE Full text] [Medline: 24098037]
19. Lara JS, Braga MM, Shitsuka C, Wen CL, Haddad AE. Dental students' and lecturers' perception of the degree of difficulty of caries detection associated learning topics in Brazil. J Educ Eval Health Prof 2015 Dec;12:56 [FREE Full text] [doi: 10.3352/jeehp.2015.12.56] [Medline: 26708573]
20. Makkar SR, Williamson A, Turner T, Redman S, Louviere J. Using conjoint analysis to develop a system to score research engagement actions by health decision makers. Health Res Policy Syst 2015 Apr 26;13:22 [FREE Full text] [doi: 10.1186/s12961-015-0013-z] [Medline: 25928693]
21. YouTube. Caries Lesion Formation and severity stages according to ICDAS URL: https://www.youtube.com/watch?v=hs-UCoSxvYA [accessed 2020-04-28]
22. YouTube. Formación de lesiones de caries en superfície oclusal y estadíos de severidade de acuerdo al ICDAS URL: https://www.youtube.com/watch?v=oz7GED4OnIY&t=8s [accessed 2020-04-28]
23. YouTube. Formação de lesões de cárie em superfície oclusal e estágios de severidade segundo o ICDAS URL: https://www.youtube.com/watch?v=kf9bYCab7RU [accessed 2020-04-28]
24. Alencar C, Sequeira E, Wen C, Haddad A. Development of modern learning objects in Teledentistry. Rev ABENO 2010;10(2):5-11. [doi: 10.30979/rev.abeno.v10i2.14]
25. Camargo LB, Raggio DP, Bonacina CF, Wen CL, Mendes FM, Bönecker MJS, et al. Proposal of e-learning strategy to teach Atraumatic Restorative Treatment (ART) to undergraduate and graduate students. BMC Res Notes 2014 Jul 17;7:456 [FREE Full text] [doi: 10.1186/1756-0500-7-456] [Medline: 25034167]
26. Roth JA, Wilson TD, Sandig M. The development of a virtual 3D model of the renal corpuscle from serial histological sections for E-learning environments. Anat Sci Educ 2015 Mar;8(6):574-583. [doi: 10.1002/ase.1529] [Medline: 25808044]
27. McNulty JA, Sonntag B, Sinacore JM. Evaluation of computer-aided instruction in a gross anatomy course: a six-year study. Anat Sci Educ 2009 Jan;2(1):2-8. [doi: 10.1002/ase.66] [Medline: 19217066]

28. Simons GF, Fennig CD. SIL International. Languages of the world URL: https://www.ethnologue.com/ [accessed 2019-03-25] [WebCite Cache ID 7796TYypp]

Abbreviations

ICDAS: International Caries Detection and Assessment System

Edited by G Eysenbach; submitted 25.04.19; peer-reviewed by M El Tantawi, B Worm, S Amantini, P Stourac; comments to author 12.12.19; revised version received 06.03.20; accepted 23.03.20; published 22.05.20.

Please cite as:
Lara JS, Braga MM, Zagatto CG, Wen CL, Mendes FM, Murisi PU, Haddad AE
A Virtual 3D Dynamic Model of Caries Lesion Progression as a Learning Object for Caries Detection Training and Teaching: Video Development Study
JMIR Med Educ 2020;6(1):e14140
URL: http://mededu.jmir.org/2020/1/e14140/
doi:10.2196/14140
PMID:32441661

Review

Artificial Intelligence Education and Tools for Medical and Health Informatics Students: Systematic Review

A Hasan Sapci[1*], MD; H Aylin Sapci[*], MD

[1]Adelphi University, Garden City, NY, United States

[*]all authors contributed equally

Corresponding Author:
A Hasan Sapci, MD
Adelphi University
Nexus Building
1 South Avenue
Garden City, NY, 11530
United States
Phone: 1 5168338156
Email: sapci@adelphi.edu

Abstract

Background: The use of artificial intelligence (AI) in medicine will generate numerous application possibilities to improve patient care, provide real-time data analytics, and enable continuous patient monitoring. Clinicians and health informaticians should become familiar with machine learning and deep learning. Additionally, they should have a strong background in data analytics and data visualization to use, evaluate, and develop AI applications in clinical practice.

Objective: The main objective of this study was to evaluate the current state of AI training and the use of AI tools to enhance the learning experience.

Methods: A comprehensive systematic review was conducted to analyze the use of AI in medical and health informatics education, and to evaluate existing AI training practices. PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) guidelines were followed. The studies that focused on the use of AI tools to enhance medical education and the studies that investigated teaching AI as a new competency were categorized separately to evaluate recent developments.

Results: This systematic review revealed that recent publications recommend the integration of AI training into medical and health informatics curricula.

Conclusions: To the best of our knowledge, this is the first systematic review exploring the current state of AI education in both medicine and health informatics. Since AI curricula have not been standardized and competencies have not been determined, a framework for specialized AI training in medical and health informatics education is proposed.

(JMIR Med Educ 2020;6(1):e19285) doi:10.2196/19285

KEYWORDS

artificial intelligence; education; machine learning; deep learning; medical education; health informatics; systematic review

Introduction

Overview

Artificial intelligence (AI) is one of the most disruptive innovations in health care, and the topic has attracted the attention of physicians, clinicians, researchers, and medical device industry professionals. Recent advancements in machine learning (ML) and deep learning (DL) algorithms and cloud computing have increased the adoption of AI. Consequently, applications that can handle a large number of unstructured data sets and solve complex problems have become a part of daily clinical practice.

Most AI applications process data and run self-learning algorithms behind the scenes. Although some AI applications provide data-driven recommendations to clinicians, others may not offer an option to accept, reject, or modify the output. The recommendations AI applications provide through statistical correlations may not be the best option because human-made AI algorithms may be flawed. To use and screen AI-based decisions, clinicians and health informaticians who develop AI applications should have an excellent understanding of the underlying AI concepts. This paper will focus on the emerging need for formal AI education in medicine and health informatics.

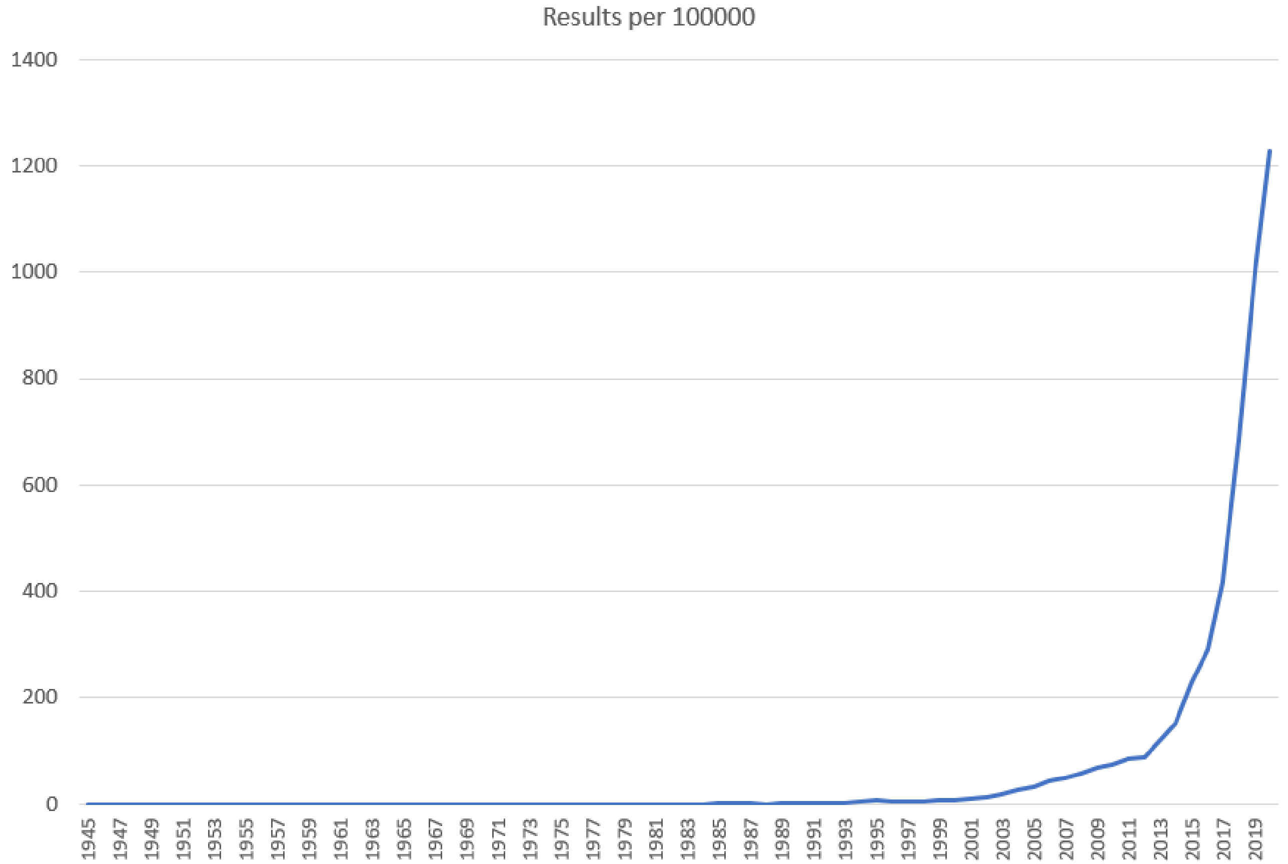
Background

Intelligence requires the capacity to perceive contexts, associate contexts to actions, and act. Even though the concept of machines that imitate intelligent human behavior is not new, AI has recently become a topic of interest [1]. As an academic discipline, the Dartmouth College Artificial Intelligence Conference that was organized by John McCarthy in 1956 was considered the birth of this field [2].

AI, ML, and DL are closely related, and the absence of universal definitions might be confusing; however, the difference between AI, ML, and DL is simple. AI is defined as "the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" [3]. AI-based devices can perceive the environment, simulate human intelligence, and solve problems. Their ability to adapt through progressive learning algorithms is what differentiates AI technologies from robotic and hardware-driven automation. In other words, computers can mimic human intelligence using AI techniques [4]. ML is the subset of AI that allows systems to learn from data and develop self-learning algorithms. ML applications can learn from data without being explicitly programmed; make predictions and recommendations using various tools; and enable computer applications to improve their performance [5]. DL is a subfield within ML that allows machines to use algorithms inspired by the structure of neural networks. A computer can learn how to classify images and how to assign labels to words in a sentence (semantic labeling) by using DL algorithms [5]. DL programming uses large quantities of unstructured data, calculates complex statistical models, and predicts outcomes without being explicitly programmed. Virtual assistants, chatbots, and facial recognition algorithms are some other practical examples of DL.

The main purpose of this study was to investigate peer-reviewed publications focused on AI education and to determine objective assessment methods for AI skills and competency training for medical and health informatics professionals. The impact of AI on the learning experience was evaluated to assess the need for AI education. As medical education, clinical informatics education, and health informatics education are closely related, medical and health informatics education trends were analyzed together. The American Medical Informatics Association has been working closely with the Commission on Accreditation for Health Informatics and Information Management Education to determine health informatics competencies; clinical informatics became a medical subspecialty in 2011 [6]. Although clinical and health informatics programs are designed for students who plan to pursue different career pathways, they use similar competencies. The implementation of ML in health care could result in unintended challenges and biased decisions, depending on the algorithms, data sources, and methodologies used [7]. Physicians who are not familiar with the evidence standards for AI might not be able to use the right approaches to integrate AI into clinical care. Even though there are several studies that explored how AI algorithms were helping enhance education [8-12], the number of peer-reviewed publications that focused on artificial intelligence education in medicine is limited (Figure 1).

Figure 1. The number of citations that appear on PubMed by year for the following search terms: ("Medical Education" OR "Medical Training") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning").



Methods

Using a replicable systematic search strategy, a full-text review was performed between November 2019 and February 2020. PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) systematic review methodology introduced by Moher et al [13] was used to identify and analyze reliable literature. The PRISMA-P method uses a structured procedure that consists of a 17-item checklist to facilitate systematic review protocols.

The combination of five groups of keywords was used to search PubMed, IEEE (Institute of Electrical and Electronics Engineers) Xplore Digital Library, CINAHL (Cumulative Index to Nursing and Allied Health Literature) Plus, and ScienceDirect databases: (1) medical education, (2) medical training, (3) artificial intelligence, (4) machine learning, and (5) deep learning (Figure 2 and Table 1). Overall, 2082 articles matched the search criteria. After removing duplicate studies and performing an abstract review, 76 full-text articles were selected for the review. All search results were entered into EPPI-Reviewer 4 text mining software (the EPPI-Centre, University of London), and the studies that met the inclusion criteria were identified. Two researchers performed the extraction independently and assessed quality.

Figure 2. Venn diagram for the PubMed search. A total of 363 results were found with the following search terms: ("Medical Education" OR "Medical Training") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning").

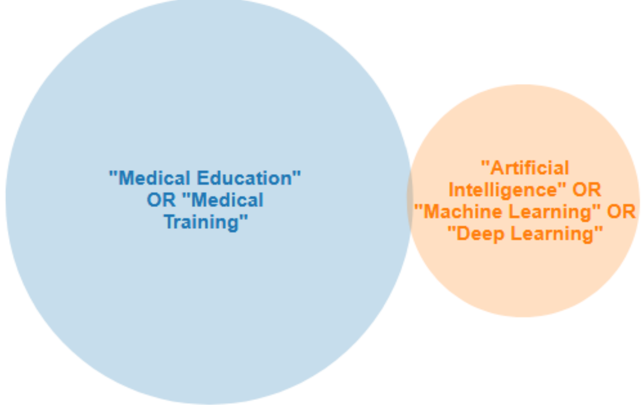
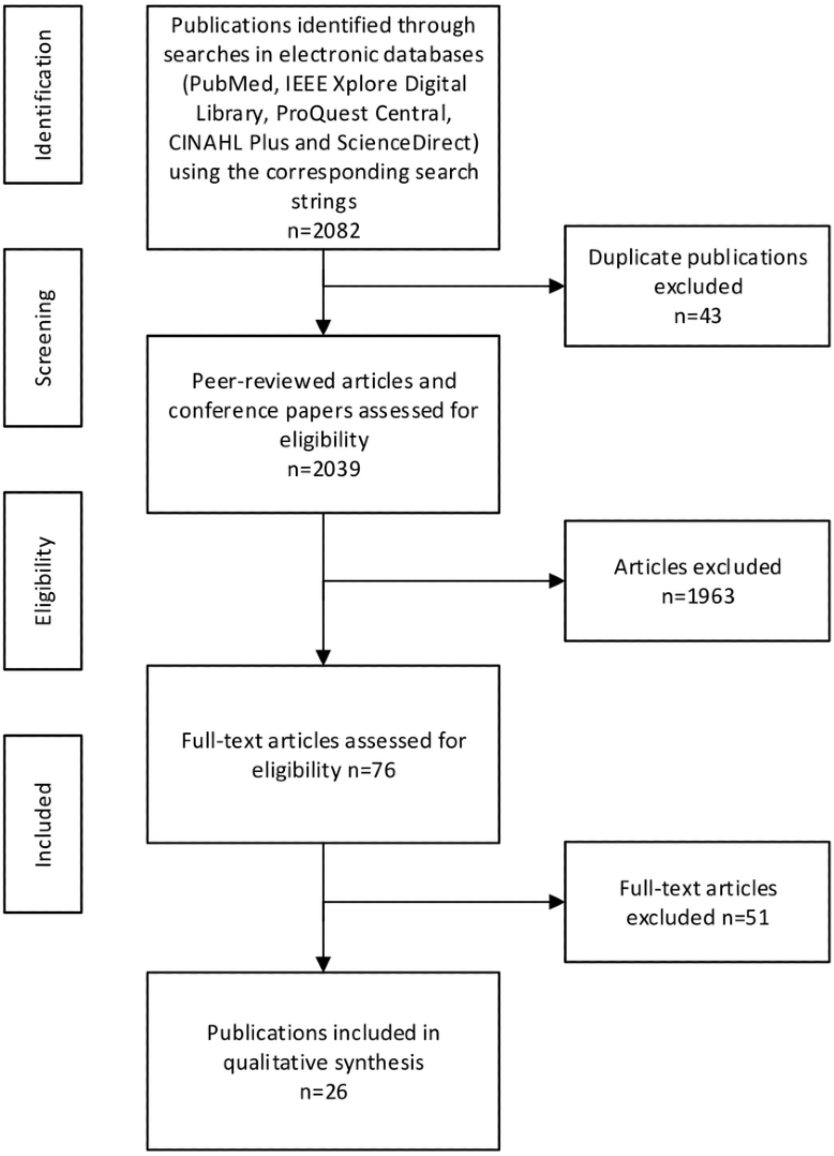


Table 1. Literature sources and keywords.

Search query and literature sources	Search in	Return value
("Medical Education" OR "Medical Training") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning")		
PubMed	All fields	363
IEEE (Institute of Electrical and Electronics Engineers) Xplore	Full text and metadata	60
ProQuest Central	Full text and peer reviewed	6271
CINAHL (Cumulative Index to Nursing and Allied Health Literature) Plus	All text (TX)	68
ScienceDirect	Title, abstract, author-specified keywords	1588

Based on the following inclusion and exclusion criteria, the selection process was applied. Peer-reviewed research articles, review papers, conference papers, case reports, correspondences, discussions, viewpoint papers, editorials, mini-reviews, and short communications papers that focused on AI tools to enhance the learning experience in medical and health informatics education or teach AI as a new competency published after 1990 were included (Table 1). Book chapters, news, and extended abstracts published before 1990 in languages other than English were excluded. To establish validity, disagreements were discussed until a consensus was reached. Overall, 26 articles matched the inclusion criteria of this research (Figure 3).

Figure 3. Search methodology. CINAHL: Cumulative Index to Nursing and Allied Health Literature; IEEE: Institute of Electrical and Electronics Engineers.



Results

One of the goals of this systematic literature review was to evaluate existing studies and determine the current state of AI education. The selected papers were used to identify the answers to three research questions.

The first research question is the following: what topics are discussed in peer-reviewed publications that focus on medical and health informatics education and AI? To answer this question, the selected publications were classified based on their education foci and the characteristics of included studies were summarized in Table 2. The publications that focused on the use of AI applications in medical and health informatics education were categorized as Category 1. These studies used various AI-based tools to enhance the learning experience. Several case studies and new initiatives to teach specific AI skills, such as ML programming languages and big data analytics software, were also identified. The publications that evaluated AI education were classified as Category 2. Category 1 studies discussed different AI applications to enhance education and summarized the impact of AI on medical and health informatics education, while Category 2 studies focused on the teaching of AI concepts.

Table 2. Characteristics of included studies.

Author(s), year, and reference	Country	Title or objective	Category[a]	Level of evidence[b]	Study objective	Comments and knowledge gap
Winkler-Schwartz et al, 2019 [14]	Canada	Artificial Intelligence in Medical Education: Best Practices Using Machine Learning to Assess Surgical Expertise in Virtual Reality Simulation	1	IV	The authors developed a checklist to assess surgical expertise in virtual reality simulation.	The study provided a general framework only. The authors emphasized the need to add further elements.
Chan and Zary, 2019 [15]	Singapore	Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review	1	IV	This review evaluated current applications of AI[c] in medical education and highlighted the main challenges.	The authors acknowledged that a low number of studies were reviewed and stated that conclusions might be inconsequential.
Lillehaug and Lajoie, 1998 [16]	Sweden	AI in medical education—another grand challenge for medical informatics	1	IV	This comprehensive review discussed the potential use of AI to enhance medical informatics education.	This article was published before the discovery of high-performance computing processors and recent advancements in data recording technology.
Frize and Frasson, 2000 [17]	Canada	Decision-support and intelligent tutoring systems in medical education	1	V	This study evaluated the use of intelligent tutoring systems in medical education.	This article discusses the potential use of decision support tools but emphasizes the need for further research to validate their usefulness.
Zhao et al, 2018 [18]	China	Research on Application of Artificial Intelligence in Medical Education	1	V	This article analyzed the application of AI in medical education.	This study evaluated the effect of AI technology on traditional medical education with a focus on personalized learning.
Chary et al, 2018 [19]	United States	A Review of Natural Language Processing in Medical Education	1	IV	This study reviewed the application of NLP[d] to medical education and identified concepts from NLP used in those applications.	The authors investigated the integration of NLP to medical education resources using published manuscripts and stated the potentially biased representation of the scope.
Caudell et al, 2003 [20]	United States	Virtual patient simulator for distributed collaborative medical education	1	IV	The study investigated the feasibility of using a real-time AI simulation engine in medical school curricula.	The study described an ongoing project and did not provide any data about the difference between problem-based learning using virtual patient simulators and standard paper case tutorials.
Guimarães et al, 2017 [21]	Portugal	Rethinking Anatomy: How to Overcome Challenges of Medical Education's Evolution	1	IV	This literature review evaluated the integration of complementary technology-based methodologies to medical instruction.	The authors discussed the potential of AI in learning analytics-oriented systems to predict behavior but did not make any recommendations about new research studies.
Bowyer et al, 2008 [22]	United States	Immersive Virtual Environments for Medical Training	1	IV	This study highlighted the role of advanced virtual environments and surgical simulators as a training platform for medical training.	The paper described various virtual reality environments where students can interact with AI-based simulators.
Sitterding et al, 2019 [23]	United States	Using Artificial Intelligence and Gaming to Improve New Nurse Transition	1	IV	This research discussed the preliminary pilot study data from a virtual reality simulation education intervention that compared virtual reality, augmented reality, serious gaming, and gamification.	The sample size and pending postintervention findings were stated as the limitations of the preliminary findings.

Author(s), year, and reference	Country	Title or objective	Category[a]	Level of evidence[b]	Study objective	Comments and knowledge gap
Boulet and Durning, 2019 [24]	United States	What we measure … and what we should measure in medical education	1	V	This paper focused on the validity of assessment scores and discusses the application of AI to automate the assessment process.	The authors recommended developing new competency assessment practices and highlighted the importance of the application of AI. They did not provide supporting evidence about AI's potential to eliminate the need for human ratings.
Conde et al, 2009 [25]	United States	Telehealth Innovations in Health Education and Training	1	V	This discussion paper indicated the potential of telehealth technologies for health education and training.	The authors recommended the development of AI applications for patient simulation and the integration of telehealth applications in health education, but the paper did not provide any evidence.
Kabassi et al, 2008 [26]	Greece	Specifying the personalization reasoning mechanism for an intelligent medical e-learning system on Atheromatosis: An empirical study	1	IV	The objective of this empirical study was to incorporate intelligent techniques in web-based medical education.	The authors described the specification of an intelligent medical learning system for atheromatosis that can interact with students. The design was based on the results of empirical data and the authors did not compare the e-learning system with traditional methods.
Klar and Bayer, 1990 [27]	Germany	Computer-assisted teaching and learning in medicine	1	IV	This article provided a comprehensive discussion of computer-assisted instruction systems and discussed expert systems' contribution to software for medical learning.	This paper was published before AI impacted multiple fields but the authors successfully envisioned how AI would transform decision making, simulation, and medical education.
Yang et al, 2019 [28]	Taiwan	An expert-led and artificial intelligence (AI) system-assisted tutoring course increase confidence of Chinese medical interns on suturing and ligature skills: prospective pilot study	1	IV	This paper examined the impact of an AI system tutoring course on clinical training.	This study compared regular, expert-led, and expert-led+AI groups and found an increased improvement in the expert-led+AI tutoring group. Authors recommended AI-assisted tutoring for novice medical interns.
Alonso-Silverio et al, 2018 [29]	Mexico	Development of a Laparoscopic Box Trainer Based on Open Source Hardware and Artificial Intelligence for Objective Assessment of Surgical Psychomotor Skills	1	IV	This study evaluated the effect of a laparoscopic trainer system that uses an AI algorithm.	The authors described the development of a low-cost intelligent simulator to improve laparoscopic skills and proposed the training as a validated training tool for surgical education programs.
Kolachalama and Garg, 2018 [30]	United States	Machine learning and medical education	2	V	This perspective article discussed the lack of student access to machine learning content and makes some suggestions to instructors.	This perspective paper only provided an outline and did not provide any evidence.
Park et al, 2019 [31]	Korea	What should medical students know about artificial intelligence in medicine?	2	IV	This short review emphasized the lack of direct access to machine learning education for clinicians and recommended the inclusion of focused content.	The review emphasized the need to identify correct information about AI.
Wartman and Combs, 2018 [32]	United States	Medical Education Must Move From the Information Age to the Age of Artificial Intelligence	2	V	This article discussed the need to develop new curricular components to teach the use of AI tools.	This commentary article summarized the authors' perspective and did not provide supporting evidence.

Author(s), year, and reference	Country	Title or objective	Category[a]	Level of evidence[b]	Study objective	Comments and knowledge gap
Wartman and Combs, 2019 [33]	United States	Reimagining Medical Education in the Age of AI	2	V	This paper indicated the need for a more sophisticated mathematical understanding of analytics.	The authors proposed a new curriculum that will include the skill sets required to use AI effectively.
Beregi, 2018 [34]	France	Artificial intelligence and medical imaging 2018: French Radiology Community white paper	2	IV	This review discussed current applications of AI in medical imaging and recommended AI education for radiology residents.	This position paper summarized AI principles, provided an update on research in the area of AI, and described radiologists' role in providing education about AI.
Tang et al, 2018 [35]	Canada	Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology	2	IV	This paper assessed the educational needs of radiologists and medical students, and provided recommendations.	The AI working group recommended the integration of health informatics and computer science courses to analyze the opportunities and challenges associated with new AI tools.
Masters, 2019 [36]	Oman	Artificial intelligence in medical education	2	V	This review highlighted the demand to learn how to work with AI systems and emphasized the need for AI training.	The authors identified new AI applications in medicine and recommended changes to medical curricula.
Chin-Yee and Upshur, 2017 [37]	Canada	Clinical judgement in the era of big data and predictive analytics	2	IV	This article explored different approaches to clinical judgment.	Authors indicated that data-driven and AI-based applications move medicine away from virtue-based approaches to clinical reasoning and recommended an integrative approach.
Santos et al, 2019 [38]	Germany	Medical students' attitude toward artificial intelligence: a multicenter survey	2	IV	This study investigated undergraduate medical students' attitudes toward AI.	The authors designed a survey to explore students' familiarity with AI concepts in radiology and concluded that they did not have an understanding of the basic technical principles underlying AI.
Paranjape et al, 2019 [39]	Netherlands	Introducing Artificial Intelligence Training in Medical Education	2	IV	This paper summarized the state of medical education and recommended a framework to include AI education.	This viewpoint paper suggested different AI-related content for different stages of medical education.

[a]Category 1: the use of AI applications in medical and health informatics education; Category 2: AI education.

[b]Evidence levels were as described by the Oxford Centre for Evidence-Based Medicine Levels of Evidence [40].

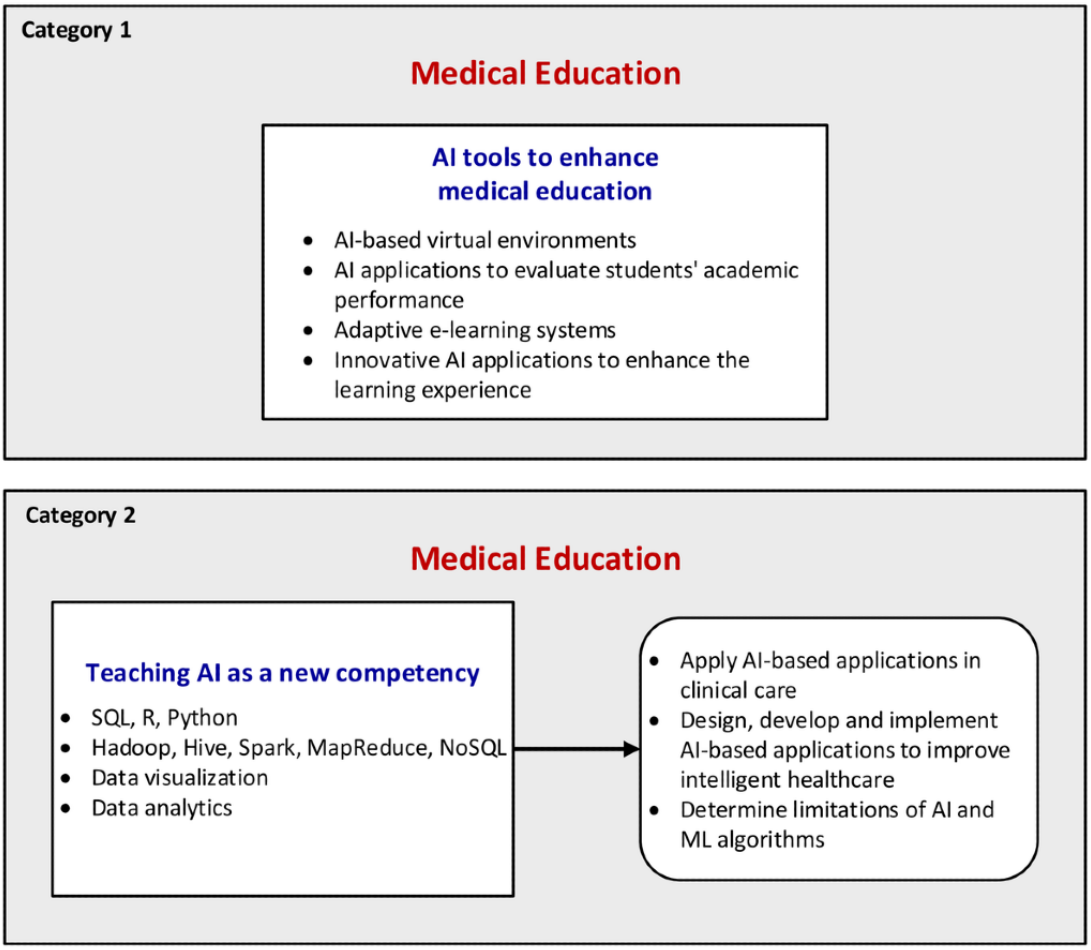
[c]AI: artificial intelligence.

[d]NLP: natural language processing.

Of 26 publications, 16 (61%) investigated the use of AI applications in medical education (Category 1) and 10 publications (39%) evaluated AI education in medicine (Category 2; Figure 4).

The first publications about the use of AI for medical applications were published in the early 1990s, and the capabilities of AI applications were restricted by technological limitations at that time. Klar and Bayer's paper [27], published in 1990, was one of the first publications about the application of AI, and they discussed the integration of expert knowledge into computer-assisted teaching in medicine. Lillehaug and Lajoie [16] proposed greater integration of AI in their 1998 paper, and they were among the first researchers who advocated for intelligent decision support systems and AI-based applications for medical education. Frize and Frasson [17] examined the role of decision support and intelligent tutoring systems in medical education, and recommended multidisciplinary studies in 2000.

Figure 4. Classification of selected AI publications. AI: artificial intelligence; ML: machine learning.



Most Category 1 studies explored innovative applications designed to improve the learning experience. For example, Chan and Zary [15] evaluated existing AI applications in medicine, determined that the primary reason to use AI in medical education was to provide feedback, and identified that significant challenges included the assessment of effectiveness and management of technical difficulties. Another systematic review conducted by Chary et al [19] identified 30 articles that assessed the application of natural language processing (NLP) to medical education. NLP is a subfield of AI and refers to intelligent communication methods using natural languages. The study revealed the benefits of NLP training in residency education and recommended strategies for its application.

Simulation-based learning has evolved over the last decade, and the virtual environment has become essential for education. Our review identified multiple case studies about advanced virtual environments. For instance, Winkler-Schwartz et al [14] analyzed virtual reality simulators that use AI, and developed a checklist to assess studies using ML algorithms to evaluate technical skills. This study concluded that the checklist had the potential to decrease the knowledge gaps in the use of AI in surgical education. Similarly, Zhao et al [18] concluded that virtual patient systems and other distance education systems that used AI increased the efficiency of medical education. Another case study that focused on virtual patient simulators identified educational and technical challenges to enhancing the learning process with AI virtual reality applications [20].

The case study published by Bowyer et al [22] described the role of advanced virtual environments in surgical training. Moreover, Conde et al [25] recommended the use of AI applications for training and education when simulated human patients were not an option.

Augmented reality is another form of virtual reality, and this technology superimposes images on top of the video viewer. Sitterding et al [23] described the differences between virtual reality, augmented reality, serious gaming, and gamification, and shared the preliminary findings of their pilot study, which determined that the simulation experience was similar to a real-life environment.

In our research, we noted different innovative AI applications for special learning activities. In their 2017 article, Guimarães et al [21] reviewed current education models for anatomy education, and recommended the use of AI analytic tools to personalize the learning process. Boulet and Durning [24] discussed the application of AI to replace human ratings and assess medical education competencies. Another study that used augmented reality and AI algorithms evaluated a case study about the development of a laparoscopic box trainer to assess surgical psychomotor skills, and concluded the proposed system had potential benefits [29].

Incorporating intelligent techniques into an adaptive e-learning system was another research group's focus. Kabassi et al [26] designed a web-based educational system for medical education, incorporated intelligent algorithms to individualize the learning experience, and shared the findings. The authors recommended further studies with more participants. Yang and Shulruf [28] designed a pilot study to demonstrate the value of AI-assisted tutoring, and determined that this additional tutoring significantly enhanced the performance of medical students.

Our second research question involved determining the highest level of evidence of current research. The articles were classified according to the Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence ranking scheme [40]. Of 26 articles, 18 (69%) were classified as Level IV evidence (case series), and 8 (31%) were classified as Level V evidence (expert opinions). The majority of the publications had Level IV and V evidence, which are considered poor reference standards. Hence, these findings emphasize the need to design new research studies.

Finally, our third research question was the following: what is the status of AI education in medicine and health informatics? In our study, 10 articles discussing AI education were identified. Since this topic is a relatively new area in medical and health informatics education, our findings were consistent with recent developments.

In their review, Paranjape et al [39] summarized multiple initiatives for AI in medical education, in which students worked with data experts and solved health care problems. The authors recommended familiarizing students with AI-based clinical applications, and introducing linear algebra, calculus, and probability during different stages of medical education. Similarly, Chin-Yee and Upshur [37] discussed the random error and biased data generated by AI and ML applications, and emphasized the effect of medical education for appraising clinical judgments.

A review conducted by Park et al [31] emphasized the importance of understanding AI to be able to validate the clinical accuracy of AI algorithms. Furthermore, multiple publications stated the need to move beyond traditional medical education, suggested a reform to align education with new practice requirements, and emphasized the role of academics and teachers in the development of appropriate AI application skills [32,33,36].

As discussed earlier, ML is an AI technique to process massive amounts of data and make predictions using computers. Kolachalama and Garg [30] proposed the integration of ML-related content in medical student, resident, and fellow education. The authors recommended the integration of real-world clinical examples into ML courses as well as practical guidelines for choosing the right tools.

The radiology community embraced AI and ML long before other medical specialties, and pioneered the usage of AI algorithms in advanced imaging applications. A multicenter survey of undergraduate medical students determined students' optimistic views about the implications of AI applications for radiology [38]. The French Radiology Community developed principles to regulate the use of AI tools, and recommended specific education to evaluate AI technologies [34]. Similarly, the Canadian Association of Radiologists recommended the integration of computer science, health informatics, and statistics training during residency education [35].

Discussion

Overview

The technological advancements in computer and software technologies; digitization of health care data; and methodological developments in information science, philosophy, mathematics, linguistics, and psychology disciplines accelerated medical research programs that focus on ML and commercialization. Maturation of AI technologies changed the roles of clinicians, and novel decision-making processes in medical settings and innovative AI-based protocols have the potential to provide diagnostic and treatment decisions by analyzing complex data sets [41].

Over the last decade, many AI researchers have concentrated on developing a proof of concept system for clinicians and patients. There is an increase in the number of studies that evaluate the effectiveness of intelligent reasoning. Intelligent monitoring technologies require new algorithms to detect anomalies, predict patterns, and make decisions.

A recent independent report prepared for the UK Secretary of State for Health and Social Care explored how the health care workforce could be prepared to use digital technology. The report emphasized the skills gap in the workforce, and made some recommendations about the integration of digital health care technologies, AI, and data analytics in undergraduate curricula [42,43].

Clinicians should have a realistic view of AI, and become familiar with the right tasks for AI in health care. Formal training for medical and health informatics students should enable them to develop AI algorithms, use AI technologies in a competent manner, and keep bias out of AI tools. Any new AI technology might encounter something new for which it has no experience, and therefore the physician should be able to assess problematic decisions and take necessary precautions when needed. Recent discussions about the need for medical education reform emphasize the shortcomings of the current model of education [44]. Overreliance on ML and AI technologies might have unintended severe adverse consequences, such as failure to recognize invalid test results [45].

While machine learning and AI algorithms are able to handle high-dimensional data classification problems and medical image interpretation, their success rates in risk prediction and diagnosis are lower. Consequently, there is a need to determine the most appropriate application areas for AI in health care [46].

There are different ways to implement AI in clinical practice, and clinicians and health informaticians need formal training to use the right approaches. Health informaticians and physician champions who design and develop AI-based protocols need to have a good understanding of complex algorithms, methodologies for data quality assessment, probabilistic forecasting, and comparative model assessment to work with engineers and develop reliable AI applications. Moreover, clinicians who use AI applications should become familiar with potential challenges.

AI-based relational time pattern analysis replaced simple threshold-based diagnostic rules. Current medical education and health informatics curricula still do not provide the ability to understand AI communications and necessary skill sets to develop AI systems that can detect and analyze relational time patterns [41].

The ability to interpret AI algorithms' mistakes and formulate the best strategies to correct these applications requires specialized training. Consequently, medical and health informatics education must emphasize algorithm-based platforms, and include relevant data analytics and AI topics in their curricula. Moreover, computer science and health informatics programs should consist of health care–focused digital skills training.

To the best of our knowledge, no previous research has investigated the use of AI tools to enhance the learning experience and AI education of medical and health informatics students. The main findings of this systematic review are as follows: (1) Although there are several recommendations on the integration of AI into medical and health informatics curricula and some academic institutions implemented experimental training programs, AI and ML education are not a part of traditional medical and health informatics curricula yet. (2) Current medical education and health informatics accreditation standards do not require AI training, and AI competencies have not been determined. (3) Using the OCEBM Levels of Evidence classification table, the majority of studies were classified as Level IV and V, which indicates poor reference standards.

Limitations

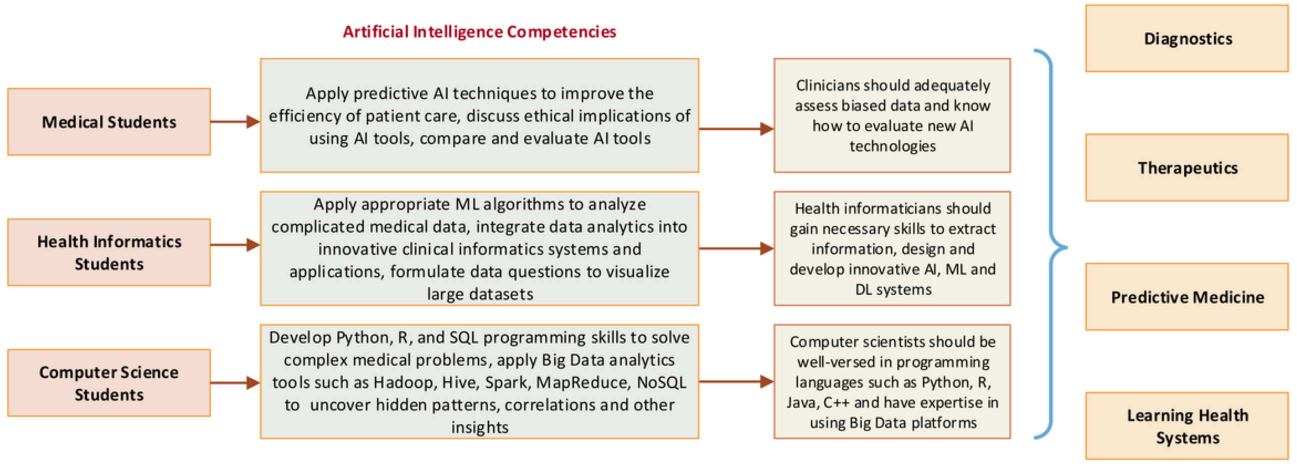
Several efforts were made to design an optimal systematic review process; however, there were still many limitations. It is probable that some studies might not be listed in the peer-reviewed academic literature databases or might be published in a non-English language. Although this is a systematic review of the field, AI is a new technical discipline, particularly in medicine, and therefore the number of articles that met the inclusion criteria was limited.

Future Directions

Overall, the selected publications did not provide specific details about different jobs' requirements and curriculum needs. The emergence of intelligent systems in health care requires new learning modalities. Even though several organizations, agencies, and work groups such as the International Medical Informatics Association [47], the Commission on Accreditation for Health Informatics and Information Management Education [48,49], the Health Informatics Society of Australia [50], the TIGER (Technology Informatics Guiding Education Reform) Initiative [51], and the Association of American Medical Colleges [52] published skill recommendations for health informatics curricula, they have not determined specific skill sets for AI education. Furthermore, a recent study evaluated health informatics students' skills in developing AI apps and emphasized the need to develop new competencies [53].

A specialized AI education framework for various professional fields would be useful; using the results of this systematic review, we propose a framework for specialized AI training for different domains (Figure 5). Medical students need to become familiar with clinical AI applications and predictive modeling techniques to assess biased data and evaluate innovative AI technologies. Health informatics students should become familiar with the application of appropriate ML algorithms and development of innovative clinical informatics systems. Furthermore, they should gain the hands-on skills required to extract data, manage large data sets to perform sophisticated data analytics, and develop innovative AI systems. Computer science students need specialized skill sets to work with data scientists and should become familiar with Python, R, and SQL programming languages, and data analytics tools (Figure 5).

Figure 5. Proposed framework for specialized AI training for different professional fields. AI: artificial intelligence; DL: deep learning; ML: machine learning.



Conflicts of Interest

None declared.

References

1. Mayor A. Gods and robots: myths, machines, and ancient dreams of technology. In: Princeton: Princeton University Press. Princeton, NJ, USA: Princeton University Press; 2018.
2. Moor J. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. AI Magazine 2006;27(4):87-91. [doi: 10.1609/aimag.v27i4.1911]
3. Artificial intelligence. Oxford Reference. URL: https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095426960 [accessed 2020-06-22]
4. Poole DL, Mackworth AK, Goebel R. Computational intelligence: a logical approach. New York, NY, USA: Oxford University Press; 1998:1-558.
5. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
6. Clinical Informatics Becomes a Board-certified Medical Subspecialty Following ABMS Vote. 2011. URL: https://www.amia.org/news-and-publications/press-release/ci-is-subspecialty [accessed 2020-03-01]
7. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. N Engl J Med 2018 Mar 15;378(11):981-983. [doi: 10.1056/nejmp1714229]
8. McArthur D, Lewis M, Bishary M. The Roles Of Artificial Intelligence In Education: Current Progress And Future Prospects. JET 2005 Mar 15;1(4):42-80. [doi: 10.26634/jet.1.4.972]
9. Popenici SAD, Kerr S. Exploring the impact of artificial intelligence on teaching and learning in higher education. Res Pract Technol Enhanc Learn 2017 Nov 23;12(1):22 [FREE Full text] [doi: 10.1186/s41039-017-0062-8] [Medline: 30595727]
10. Bayne S. Teacherbot: interventions in automated teaching. Teaching in Higher Education 2015 Apr 16;20(4):455-467. [doi: 10.1080/13562517.2015.1020783]
11. Petit dit Dariel OJ, Raby T, Ravaut F, Rothan-Tondeur M. Developing the Serious Games potential in nursing education. Nurse Educ Today 2013 Dec;33(12):1569-1575. [doi: 10.1016/j.nedt.2012.12.014] [Medline: 23332500]
12. Nye BD. Intelligent Tutoring Systems by and for the Developing World: A Review of Trends and Approaches for Educational Technology in a Global Context. Int J Artif Intell Educ 2014 Oct 9;25(2):177-203. [doi: 10.1007/s40593-014-0028-6]
13. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4(1):1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
14. Winkler-Schwartz A, Bissonnette V, Mirchi N, Ponnudurai N, Yilmaz R, Ledwos N, et al. Artificial Intelligence in Medical Education: Best Practices Using Machine Learning to Assess Surgical Expertise in Virtual Reality Simulation. J Surg Educ 2019 Nov;76(6):1681-1690. [doi: 10.1016/j.jsurg.2019.05.015] [Medline: 31202633]
15. Chan KS, Zary N. Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. JMIR Med Educ 2019 Jun 15;5(1):e13930 [FREE Full text] [doi: 10.2196/13930] [Medline: 31199295]
16. Lillehaug S, Lajoie SP. AI in medical education—another grand challenge for medical informatics. Artificial Intelligence in Medicine 1998 Mar;12(3):197-225. [doi: 10.1016/s0933-3657(97)00054-7]
17. Frize M, Frasson C. Decision-support and intelligent tutoring systems in medical education. Clin Invest Med 2000;23:266-269. [doi: 10.1007/3-540-45108-0]
18. Zhao H, Li G, Feng W. Research on Application of Artificial Intelligence in Medical Education. 2018 Aug 10 Presented at: International Conference on Engineering Simulation and Intelligent Control (ESAIC); 2018; Changsha, China p. 340-342. [doi: 10.1109/esaic.2018.00085]
19. Chary M, Parikh S, Manini A, Boyer E, Radeos M. A Review of Natural Language Processing in Medical Education. West J Emerg Med 2019 Jan 12;20(1):78-86 [FREE Full text] [doi: 10.5811/westjem.2018.11.39725] [Medline: 30643605]
20. Caudell TP, Summers KL, Holten J, Hakamata T, Mowafi M, Jacobs J, et al. Virtual patient simulator for distributed collaborative medical education. Anat Rec B New Anat 2003 Jan 13;270(1):23-29 [FREE Full text] [doi: 10.1002/ar.b.10007] [Medline: 12526063]
21. Guimarães B, Dourado L, Tsisar S, Diniz JM, Madeira MD, Ferreira MA. Rethinking Anatomy: How to Overcome Challenges of Medical Education's Evolution. Acta Med Port 2017 Feb 27;30(2):134-140 [FREE Full text] [doi: 10.20344/amp.8404] [Medline: 28527481]
22. Bowyer MW, Streete KA, Muniz GM, Liu AV. Immersive Virtual Environments for Medical Training. Seminars in Colon and Rectal Surgery 2008 Jun;19(2):90-97. [doi: 10.1053/j.scrs.2008.02.005]
23. Sitterding MC, Raab DL, Saupe JL, Israel KJ. Using Artificial Intelligence and Gaming to Improve New Nurse Transition. Nurse Leader 2019 Apr;17(2):125-130. [doi: 10.1016/j.mnl.2018.12.013]
24. Boulet JR, Durning SJ. What we measure … and what we should measure in medical education. Med Educ 2019 Jan 14;53(1):86-94. [doi: 10.1111/medu.13652] [Medline: 30216508]

25. Conde JG, De S, Hall RW, Johansen E, Meglan D, Peng GC. Telehealth innovations in health education and training. Telemed J E Health 2010 Feb;16(1):103-106 [FREE Full text] [doi: 10.1089/tmj.2009.0152] [Medline: 20155874]
26. Kabassi K, Virvou M, Tsihrintzis GA, Vlachos Y, Perrea D. Specifying the personalization reasoning mechanism for an intelligent medical e-learning system on Atheromatosis: An empirical study. IDT 2008 Oct 07;2(3):179-190. [doi: 10.3233/idt-2008-2304]
27. Klar R, Bayer U. Computer-assisted teaching and learning in medicine. International Journal of Bio-Medical Computing 1990 Jul;26(1-2):7-27. [doi: 10.1016/0020-7101(90)90016-n]
28. Yang YY, Shulruf B. Expert-led and artificial intelligence (AI) system-assisted tutoring course increase confidence of Chinese medical interns on suturing and ligature skills: prospective pilot study. J Educ Eval Health Prof 2019 Apr 10;16:7 [FREE Full text] [doi: 10.3352/jeehp.2019.16.7] [Medline: 30986892]
29. Alonso-Silverio GA, Pérez-Escamirosa F, Bruno-Sanchez R, Ortiz-Simon JL, Muñoz-Guerrero R, Minor-Martinez A, et al. Development of a Laparoscopic Box Trainer Based on Open Source Hardware and Artificial Intelligence for Objective Assessment of Surgical Psychomotor Skills. Surg Innov 2018 Aug 29;25(4):380-388. [doi: 10.1177/1553350618777045] [Medline: 29809097]
30. Kolachalama VB, Garg PS. Machine learning and medical education. NPJ Digit Med 2018 Sep 27;1(1):54 [FREE Full text] [doi: 10.1038/s41746-018-0061-1] [Medline: 31304333]
31. Park SH, Do KH, Kim S, Park JH, Lim Y. What should medical students know about artificial intelligence in medicine? J Educ Eval Health Prof 2019 Jul 03;16:18 [FREE Full text] [doi: 10.3352/jeehp.2019.16.18] [Medline: 31319450]
32. Wartman SA, Combs CD. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. Academic Medicine 2018;93(8):1107-1109. [doi: 10.1097/acm.0000000000002044]
33. Wartman SA, Combs CD. Reimagining Medical Education in the Age of AI. AMA J Ethics 2019 Feb 01;21(2):E146-E152 [FREE Full text] [doi: 10.1001/amajethics.2019.146] [Medline: 30794124]
34. SFR-IA Group, CERF, French Radiology Community. Artificial intelligence and medical imaging 2018: French Radiology Community white paper. Diagn Interv Imaging 2018 Nov;99(11):727-742 [FREE Full text] [doi: 10.1016/j.diii.2018.10.003] [Medline: 30470627]
35. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, Canadian Association of Radiologists (CAR) Artificial Intelligence Working Group. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. Can Assoc Radiol J 2018 May 01;69(2):120-135 [FREE Full text] [doi: 10.1016/j.carj.2018.02.002] [Medline: 29655580]
36. Masters K. Artificial intelligence in medical education. Medical Teacher 2019 Apr 21;41(9):976-980. [doi: 10.1080/0142159x.2019.1595557]
37. Chin-Yee B, Upshur R. Clinical judgement in the era of big data and predictive analytics. J Eval Clin Pract 2018 Jun 13;24(3):638-645. [doi: 10.1111/jep.12852] [Medline: 29237237]
38. Pinto Dos Santos D, Giese D, Brodehl S, Chon SH, Staab W, Kleinert R, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. Eur Radiol 2019 Apr 6;29(4):1640-1646. [doi: 10.1007/s00330-018-5601-1] [Medline: 29980928]
39. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing Artificial Intelligence Training in Medical Education. JMIR Med Educ 2019 Dec 03;5(2):e16048 [FREE Full text] [doi: 10.2196/16048] [Medline: 31793895]
40. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. Oxford Centre of Evidence-Based Medicine. 2014. URL: https://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf [accessed 2020-03-15]
41. Lynn LA. Artificial intelligence systems for complex decision-making in acute care medicine: a review. Patient Saf Surg 2019 Feb 1;13(1):6 [FREE Full text] [doi: 10.1186/s13037-019-0188-2] [Medline: 30733829]
42. Rimmer A. Technology will improve doctors' relationships with patients, says Topol review. BMJ 2019 Feb 11;364:l661. [doi: 10.1136/bmj.l661] [Medline: 30745296]
43. National Health Service. The Topol Review: Preparing the healthcare workforce to deliver the digital future. 2019. URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf [accessed 2020-03-01]
44. Mehta NB, Hull AL, Young JB, Stoller JK. Just Imagine. Academic Medicine 2013;88(10):1418-1423. [doi: 10.1097/acm.0b013e3182a36a07]
45. Cabitza F, Rasoini R, Gensini GF. Unintended Consequences of Machine Learning in Medicine. JAMA 2017 Aug 08;318(6):517-518. [doi: 10.1001/jama.2017.7797] [Medline: 28727867]
46. Maddox TM, Rumsfeld JS, Payne PRO. Questions for Artificial Intelligence in Health Care. JAMA 2019 Jan 01;321(1):31-32. [doi: 10.1001/jama.2018.18932] [Medline: 30535130]
47. Mantas J, Ammenwerth E, Demiris G, Hasman A, Haux R, Hersh W, et al. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics - First Revision. ejbi 2011. [doi: 10.24105/ejbi.2011.07.2.2]
48. European Commission. The EU Framework Programme for Research and Innovation. Horizon 2020. 2013. URL: https://www.uv.es/operuv/docs_h2020/InfoKit_UK_240214_Final.pdf [accessed 2020-03-15]
49. EU US eHealth Work Consortium. TIGER Initiative. HIMSS Foundation. 2018. URL: http://www.himss.org/professionaldevelopment/tigers-euus-ehealth-work-project [accessed 2020-03-15]

50. Martin-Sanchez F. Health Informatics Competencies Framework. Health Informatics Society of Australia. 2018. URL: http://www.healthinformaticscertification.com/wp-content/uploads/2016/02/CHIA-competencies-Framework_FINAL.pdf [accessed 2020-03-15]
51. Gugerty B, Delaney C. TIGER Informatics Competencies Collaborative (TICC) Final Report. TIGER Summit. 2009. URL: https://tigercompetencies.pbworks.com/f/TICC_Final.pdf [accessed 2020-03-15]
52. Souze A, Fyfe-Mills K. Bridging the Skills Gap - Workforce Development and the Future Work. ATD Public Policy Advisory Group 2018:1-56 [FREE Full text]
53. Sapci AH, Sapci HA. Teaching Hands-On Informatics Skills to Future Health Informaticians: A Competency Framework Proposal and Analysis of Health Care Informatics Curricula. JMIR Med Inform 2020 Jan 21;8(1):e15748 [FREE Full text] [doi: 10.2196/15748] [Medline: 31961328]

Abbreviations

AI: artificial intelligence
CINAHL: Cumulative Index to Nursing and Allied Health Literature
DL: deep learning
IEEE: Institute of Electrical and Electronics Engineers
ML: machine learning
NLP: natural language processing
OCEBM: Oxford Centre for Evidence-Based Medicine
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols
TIGER: Technology Informatics Guiding Education Reform

Edited by G Eysenbach; submitted 12.04.20; peer-reviewed by T Virgona, W Jacobowitz, E Frontoni, C Liao; comments to author 01.05.20; revised version received 06.05.20; accepted 14.06.20; published 30.06.20.

Please cite as:
Sapci AH, Sapci HA
Artificial Intelligence Education and Tools for Medical and Health Informatics Students: Systematic Review
JMIR Med Educ 2020;6(1):e19285
URL: http://mededu.jmir.org/2020/1/e19285/
doi:10.2196/19285
PMID:32602844

Original Paper

Self-Perceptions of Readiness to Use Electronic Health Records Among Medical Students: Survey Study

Lina Lander[1], SD; Sally L Baxter[2,3], MD, MSc; Gary L Cochran[4], PharmD, SM; Helena E Gali[5], MD; Kristen Cook[4], PharmD; Thomas Hatch[6], PharmD; Regan Taylor[7], MD; Linda Awdishu[8], PharmD, MAS

[1]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, United States
[2]Department of Biomedical Informatics, University of California San Diego Health, La Jolla, CA, United States
[3]Shiley Eye Institute and Viterbi Family Department of Ophthalmology, University of California San Diego, La Jolla, CA, United States
[4]College of Pharmacy, University of Nebraska Medical Center, Omaha, NE, United States
[5]School of Medicine, University of California San Diego, La Jolla, CA, United States
[6]Information Services Department, University of California San Diego Health, La Jolla, CA, United States
[7]Department of Internal Medicine, University of Nebraska Medical Center, Omaha, NE, United States
[8]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Lina Lander, SD
Department of Family Medicine and Public Health
University of California San Diego
6500 Gillman Drive
La Jolla, CA, 92093-0606
United States
Phone: 1 8582601917
Email: linalander@health.ucsd.edu

Abstract

Background: Although several national organizations have declared the ability to work with electronic health records (EHRs) as a core competency of medical education, EHR education and use among medical students vary widely. Previous studies have reported EHR tasks performed by medical students, but students' self-perceived readiness and comfort with EHRs are relatively unknown.

Objective: This study aimed to better understand medical students' self-perceived readiness to use EHRs to identify potential curricular gaps and inform future training efforts based on students' perspectives.

Methods: The authors deployed a survey investigating self-perceived comfort with EHRs at 2 institutions in the United States in May 2019. Descriptive statistics were generated regarding demographics, comfort level with various EHR-related tasks, and cross-institutional comparisons. We also assessed the impact of extracurricular EHR experience on comfort level.

Results: In total, 147 medical students responded, of which 80 (54.4%) were female, with equal distribution across all 4 years of training. Overall confidence was generally higher for students with longer extracurricular EHR experience, even when adjusted for age, gender, year of training, and institution. Students were most comfortable with tasks related to looking up information in the EHR and felt less comfortable with tasks related to entering new information and managing medications. Fourth-year students at both schools reported similar levels of comfort with EHR use, despite differences in preclinical EHR training. Open-ended comments emphasized the value of experiential training over didactic formats.

Conclusions: Information entry and medication management in the EHR represent areas for future curricular development. Experiential training via extracurricular activities and early clinical exposure may be high-yield approaches to help medical students achieve critical EHR competencies.

(JMIR Med Educ 2020;6(1):e17585) doi:10.2196/17585

KEYWORDS

electronic health record; medical student; education; training; residency

Introduction

Electronic health records (EHRs) have become widely adopted across the United States [1-4]. EHR use has become an increasingly prominent component of physician work time and effort across multiple specialties, in some cases equaling or surpassing time spent in face-to-face interactions with patients [5-12]. This widespread integration of EHRs has generated a national discussion regarding the role of EHR training in medical school curricula. In response, several national organizations, such as the Association of American Medical College (AAMC) [13], the Alliance for Clinical Education [14], and the Liaison Committee on Medical Education [15], have issued guidance stating that the ability to work in an EHR is a core competency for medical students before beginning residency. In fact, 2 of the AAMC's core entrustable professional activities for entering residency relate to EHR use: (1) *enter and discuss orders and prescriptions* and (2) *document a clinical encounter in the medical record* [16]. The need for medical student preparation and readiness in EHRs has also been echoed by professional medical societies, such as the American Medical Association [17], the American College of Surgeons [18], and the American Academy of Family Physicians [19].

Despite the recognized need for directed EHR training during medical school, EHR education and use among medical students vary widely among different institutions and clerkships [20-24], and a lack of focused initiatives to engage students in EHR education may leave gaps in the competencies expected of residents [25]. Previous studies have focused on the types of tasks performed within the EHR by medical students during their clinical rotations, such as accessing information, entering information, entering notes, and entering orders. However, knowledge about medical students' self-perceptions of their comfort and readiness in using EHRs as well as how this sense of readiness may relate to varying curricular approaches is lacking.

To fill this gap, we deployed a survey to medical students at 2 different institutions to gain an understanding about self-perceptions of EHR readiness. We hypothesized that perceived readiness would be higher among students with prior extracurricular EHR experience as well as among students who participated in an integrated EHR curriculum. The purpose of this study was to better understand medical students' perceived readiness for EHR use and identify gaps in curriculum and training that could be addressed to improve medical school curricula and address this critical competency of EHR education.

Methods

Study Population

Eligible participants included all medical students from the University of Nebraska Medical Center College of Medicine (UNMC) and the University of California San Diego School of Medicine (UCSD). Currently enrolled medical students at all levels of training were eligible. Both medical schools are affiliated with academic medical centers that use the same EHR vendor. The institutional review boards of both UNMC and UCSD approved this study.

Curricula and Electronic Health Record Training

UNMC is a 4-year Doctor of Medicine degree–granting program, affiliated with Nebraska Medicine, which enrolls approximately 135 students per year. Students acquire EHR skills throughout the 4-year curriculum. Medical students begin learning to use the EHR during the first week of school, and formal training continues over the first 18 months before entering clerkships. The UNMC preclinical phase consists of 10 organ systems–based blocks, each containing an EHR exercise. These range from a scavenger hunt, in which students learn where to locate information in the health record, to specific cases designed to help them learn order entry or how to type notes in the EHR. All preclinical exercises are performed in the EHR training environment and are supervised by a faculty member. Early sessions also have information technology staff support to help troubleshoot issues with access or functionality. The initial sessions are used to learn navigation skills—finding specific patient information, laboratory and imaging results, or searching encounter notes and discharge summaries. Once basic skills are established, EHR-based cases are used to improve students' skills in the medical record to include documenting clinical encounters; entering orders; documenting medical history, allergies, and medications; reviewing pertinent medical information from prior notes; and using alerts and reminders in the EHR to complete health care maintenance tasks. Formal preclinical EHR training was initiated in 2017. During the preclinical phase, students also have the opportunity to practice their skills at the student-run free clinic (SHARING Clinic). Approximately 50% of the preclinical students volunteer for this clinic, which uses the same EHR as the main medical campus. Immediately before clerkships, students have a session to learn shortcuts offered within the EHR—specifically, using templates for note writing. These training exercises are designed to increase competence and confidence with the use of EHR and to promote active participation in the delivery of care during clerkships.

UCSD is a 4-year Doctor of Medicine degree–granting program, affiliated with the UC San Diego Health, which enrolls approximately 134 students per year. EHR exposure is limited in the preclinical years. There is no formal EHR training in the preclinical curriculum, unless a student elects to participate in the student-run free clinic. Approximately 80% of first- and second-year students volunteer at the student-run free clinic, where students use a clinic-specific EHR from the same vendor as the academic medical center. Students who volunteer in the free clinic undergo a 2-hour orientation session on EHR functionality and clinic-specific workflows. All medical students, irrespective of free clinic participation, receive formal EHR education during a clinical transition week at the beginning of the third year. Students review several web-based modules provided by the vendor and participate in one 2-hour didactic session that covers both inpatient and ambulatory clinic tools and workflows. Examples of skills taught include finding patients on a clinic schedule and navigating summary reports, demographics, patients' problem lists, notes, and labs.

At both institutions, third- and fourth-year medical students are actively engaged in EHR use in their clinical clerkships and elective rotations. Both institutions also require subinternships where students function as interns under the close supervision of faculty physicians and senior residents. EHR tasks include placing orders throughout the hospitalization; performing medication reconciliation; and writing admission notes, progress notes, and discharge summaries. However, the main contrast in curricula is that UCSD has little formal EHR training during the preclinical years, whereas UNMC has integrated EHR training throughout the preclinical curriculum.

Survey

We modified EHR competency assessment tools provided by the EHR vendor for formal training sessions to develop the self-perceived readiness survey (full survey instrument available in Multimedia Appendix 1). Instead of asking students to perform specific tasks such as finding allergies, immunizations, and others, we rephrased the questions to ask about self-perceived comfort levels while performing each task. Students rated the comfort level with using various EHR components on a 5-point Likert scale, ranging from 1 (very uncomfortable) to 5 (very comfortable). We also asked students to provide their gender, age group, year of training, and content and length of preclinical extracurricular EHR experience. The survey concluded with an open-ended item asking students for general comments about their EHR training and preparedness to work in the EHR. In total, 3 School of Medicine faculty members and 2 medical students from both institutions assessed the survey for face validity, readability, and understanding.

The electronic survey was administered anonymously via email to all current medical students in May 2019 at UCSD and in July 2019 at UNMC, with 2 reminder emails at 7 and 14 days after the initial invitation. Owing to the timing of the survey administration at UNMC, preclinical students were students between years 1 and 2 of the curriculum; no incoming first-year students were surveyed, as they had not yet started the curriculum. The survey remained open for a total of 30 days. Survey completion required approximately 10 minutes and did not affect students' grades or evaluations. Survey data were collected using Qualtrics (Qualtrics, Provo, UT, USA).

Statistical Analyses

Statistical analysis consisted of descriptive statistics using the mean and SD or counts/frequencies where appropriate. To compare categorical data between institutions, we used the Pearson chi-square test for independence. We used the Student *t* test to compare mean Likert scores for survey items. Although Likert scale data are classically analyzed with nonparametric testing such as the Mann-Whitney-Wilcoxon test, we chose to compare mean Likert values using the Student *t* test [26] to facilitate data interpretation. For any *t* test that generated a *P* value of less than .10, we conducted the Mann-Whitney-Wilcoxon test as a sensitivity analysis. For clarity, only *P* values from *t* tests are reported. In all cases, we reached the same conclusion regarding statistical significance, regardless of whether we used a *t* test or a Mann-Whitney-Wilcoxon test. For all hypothesis tests and models, statistical significance was defined as a *P* value of less than .05. Statistical analyses were conducted using Stata version 11 (StataCorp LLC, College Station, TX, USA) and R (RStudio Inc, Boston, MA) [27].

Results

General Demographics

In total, 147 medical students responded to the survey on EHR readiness at the 2 institutions. Of the 506 medical students who received survey invitations at UCSD, 95 (19%) responded. The response rate at UNMC was 13.4% (52/386). About half of the respondents were female (80/147, 54.4%; Table 1). The majority (27/52, 52%) of respondents at UNMC were aged <25 years, whereas the most well-represented group among UCSD respondents were those in the 25 to 27 years age range (*P*=.02). The gender distribution of the survey respondents was generally consistent with the overall enrollment at the 2 institutions—the proportion of females in the overall UCSD student population was 52.7% (369/700), whereas at UNMC, it was 45.0% (175/389). The age distribution of the survey respondents at UCSD corresponded with that of the overall student population (188/700, 26.9% aged <25 years; 291/700, 41.6% aged 25-27 years; and 221/700, 31.6% aged ≥28 years). The survey respondents at UNMC had a greater proportion of individuals younger than 25 years (27/52, 52% of the survey respondents compared with 104/389, 26.7% in the overall student population).

Table 1. Characteristics of medical students from the University of California San Diego School of Medicine and the University of Nebraska Medical Center College of Medicine responding to a survey on self-perceived electronic health record readiness.

Characteristics	University of California San Diego (n=95), n (%)	University of Nebraska Medical Center (n=52), n (%)	Total (N=147), n (%)	*P* value[a]
Gender				
Female	55 (58)	25 (48)	80 (54)	.25
Year of training				**<.001**
1	26 (27)	0 (0)	47 (32)	
2	16 (17)	30 (58)	25 (14)	
3	29 (31)	11 (21)	40 (27)	
4	24 (25)	11 (21)	35 (24)	
Age (years)				**.02**
<25	25 (26)	27 (52)	52 (35)	
25-27	46 (49)	15 (29)	61 (41)	
≥28	24 (25)	10 (19)	34 (23)	
Extracurricular EHR[b] experience (≥1 month)				
Student-run free clinic	64 (67)	14 (27)	78 (53)	<.001
Inpatient setting	45 (47)	21 (40)	66 (45)	.42
Ambulatory clinic	62 (65)	17 (33)	79 (54)	<.001

[a]Pearson chi-square test was used to evaluate differences between the UCSD School of Medicine and the University of Nebraska Medical Center College of Medicine.

[b]EHR: electronic health record.

Impact of Extracurricular Electronic Health Record Experience

Medical students at both UCSD and UNMC reported engagement with EHRs via extracurricular activities (Table 1). These activities were undertaken by medical students outside of their formal medical school curricula (ie, not formal clinical rotations). Examples included volunteering in student-run free clinics for underserved populations as well as volunteering in inpatient settings or ambulatory clinics. Significantly higher proportions of UCSD medical students reported having 1 month or more of EHR experience in the student-run free clinic and in ambulatory settings compared with respondents at UNMC (67% vs 53% and 65% vs 54%, respectively; *P*<.001 for both comparisons). The proportion of respondents with 1 month or more of extracurricular EHR experience in inpatient settings was similar between the 2 institutions (47% vs 40%; *P*=.42).

We specifically investigated the impact of extracurricular EHR experience on overall confidence using EHRs. Overall confidence was a single Likert score to gauge the students' overall self-perceived confidence in using EHRs. This overall confidence score was compared between students who had less than 1 month of extracurricular EHR experience and those who had 1 month or more of extracurricular EHR experience (Table 2).

All settings (ie, free clinic, inpatient, and ambulatory) were evaluated, for the cohort overall and individually at each institution. As expected, the mean Likert scores for overall confidence were generally higher for students with longer extracurricular EHR experience at both institutions. Specifically, medical students who had longer exposures to EHR interactions in inpatient settings and ambulatory clinics reported significantly higher overall confidence in using EHRs, with average Likert scores of 3.5 or greater at both institutions. These differences did not reach statistical significance for longer exposure to the student-run free clinic. Those with less than a month of extracurricular EHR experience in any of the settings had mean Likert scores of less than 3 for overall confidence.

The effects of extracurricular activities were also evaluated in a multivariable model. The extracurricular ambulatory clinic EHR experience of 1 month or more was associated with significantly higher overall confidence using the EHR after adjusting for institution, year of training, age, and gender (average increase in the Likert score of 0.57 compared with those with <1 month of experience; *P*=.004). The effect of inpatient experience, however, was borderline significant (average increase in the Likert score of 0.38; *P*=.06). Similar to the unadjusted analysis, students with ≥1 month EHR experience in the student-run free clinic were not significantly more confident in the multivariable model (*P*=.14).

Table 2. Impact of extracurricular electronic health record experience on mean Likert scale score for overall confidence.

Extracurricular electronic health record experience	Overall			University of California San Diego			University of Nebraska Medical Center		
	Average rating of overall confidence[a]		*P* value	Average rating of overall confidence[a]		*P* value	Average rating of overall confidence[a]		*P* value
	<1 month	≥1 month		<1 month	≥1 month		<1 month	≥1 month	
Student-run free clinic	2.7	3.2	.06	3.1	3.3	.55	2.9	3.2	.10
Inpatient setting	2.4	3.8	<.001	2.8	3.6	.003	2.6	3.7	<.001
Ambulatory clinic	2.3	3.5	<.001	2.9	3.6	.008	2.6	3.5	<.001

[a]Average rating of overall confidence in using electronic health record among students with <1 month versus ≥1 months of experience.

Perceptions of Electronic Health Record Readiness Across Various Domains

The survey included 15 items asking medical students to rate their comfort level with various tasks in the EHR as well as an item to rate their overall confidence in working with EHRs. We grouped task-related items into 3 domains: (1) looking up information, (2) entering new information, and (3) medication management. We compared the mean Likert scores for self-perceived comfort or readiness for each item between the 2 institutions (Table 3).

In the domain of *looking up information*, there were no significant differences between UCSD and UNMC. Medical students at both institutions reported high levels of comfort (Likert scores >4) for looking up laboratory values and finding progress notes. Students at both institutions felt less comfortable with identifying clinical documentation errors in the EHR (mean score of 2.3 at UCSD and 2.4 at UNMC).

Compared with looking up information, medical students at both institutions felt less confident while entering new information, as no mean Likert scores exceeded 4 in this domain at either institution. Of the 8 EHR tasks included in this domain, medical students at UNMC were significantly more comfortable than medical students at UCSD with 4 of these tasks: entering a new diagnosis, updating a patient's problem list to include a new problem, documenting immunizations in the EHR, and documenting allergies in the EHR (Table 3). UCSD medical students were significantly more comfortable with messaging other providers within the EHR (2.7 vs 1.8 at UNMC; *P*<.001). Medical students from both institutions had similar comfort levels with documenting past medical history and past social history; documenting clinical encounters using templates within the EHR; and completing documentation of notes such as progress notes, admission notes, and discharge summaries.

Medical students at both institutions reported lower comfort levels with medication management in the EHR compared with looking up information and entering new information, as no mean scores exceeded 3.5 in this domain. Although UNMC medical students had a significantly greater comfort level with entering new medication orders (3.2 vs 2.5 at UCSD; *P*=.005), there were no statistically significant differences in the remaining items, such as verifying medication orders, reviewing history and scheduled medications, and performing medication reconciliation.

In response to the item "Overall, I feel prepared to use the EHR," medical students from both institutions endorsed a midlevel comfort score, with UCSD students having a mean score of 3.1 and UNMC students having a mean score of 3.2 (*P*=.65).

Table 3. Average self-reported feeling of comfort with using various components of the electronic health record by institution.

Electronic health record task[a], mean (SD)	University of California San Diego (n=95)	University of Nebraska Medical Center (n=52)	P value[b]
Looking up information			
Laboratory	4.0 (1.3)	4.2 (0.8)	.19
Progress note	4.1 (1.2)	4.2 (0.9)	.44
Clinical documentation errors	2.3 (1.3)	2.4 (1.2)	.54
Entering new information			
Diagnosis	2.7 (1.3)	3.5 (1.0)	<.001
Problem reported by patient	3.1 (1.4)	3.7 (1.1)	.007
Immunizations	2.5 (1.3)	2.9 (1.2)	.045
Allergies	2.9 (1.4)	3.5 (1.1)	.008
Past medical/social history	3.4 (1.3)	3.8 (1.0)	.06
Clinical encounter documentation using template[c]	3.8 (1.3)	3.9 (1.1)	.47
Notes[d]	3.7 (1.4)	3.8 (1.3)	.81
Message other providers	2.7 (1.4)	1.8 (1.2)	<.001
Medication management			
Entering new medication orders	2.5 (1.4)	3.2 (1.2)	.005
Verifying medication orders	2.4 (1.3)	2.7 (1.2)	.12
Reviewing history and scheduled medications	3.2 (1.4)	3.4 (1.1)	.35
Reconciliation	2.4 (1.2)	2.5 (1.1)	.89
Overall, feeling prepared to work in EHR[e]	3.1 (1.3)	3.2 (1.0)	.65

[a]Average rating of comfort level using various EHR components on a scale from 1 (very uncomfortable) to 5 (very comfortable). The full survey instrument is available in Multimedia Appendix 1.

[b]Student *t* test was used to evaluate differences between UCSD and UNMC.

[c]Documenting the clinical encounter using prespecified note templates in the EHR.

[d]Documenting notes, including history and physical examination on admission, progress notes, and discharge summaries.

[e]EHR: electronic health record.

Comparisons of Electronic Health Record Readiness Among Preclinical Students

The primary curricular difference between the 2 institutions was that UCSD did not offer formal preclinical EHR training, whereas UNMC did. Therefore, we aggregated data from first- and second-year students to analyze perceptions of EHR readiness among preclinical students (n=131, with 77 from UCSD and 44 from UNMC). The mean Likert scores for comfort and readiness to perform various EHR tasks were compared by institution (Multimedia Appendix 2). There were no significant differences in tasks related to looking up information. Tasks where UNMC preclinical students reported significantly higher levels of readiness included entering new information, such as diagnoses (*P*<.001), problems (*P*=.002), immunizations (*P*=.007), allergies (*P*<.001), prior medical/social history (*P*=.003), and medication orders (*P*<.001). The only task where UCSD preclinical students reported greater levels of comfort than UNMC preclinical students was messaging other providers (*P*=.002). Despite differences in comfort level with individual tasks, preclinical students from the 2 institutions did not have significantly different overall levels of comfort with the EHR (*P*=.14).

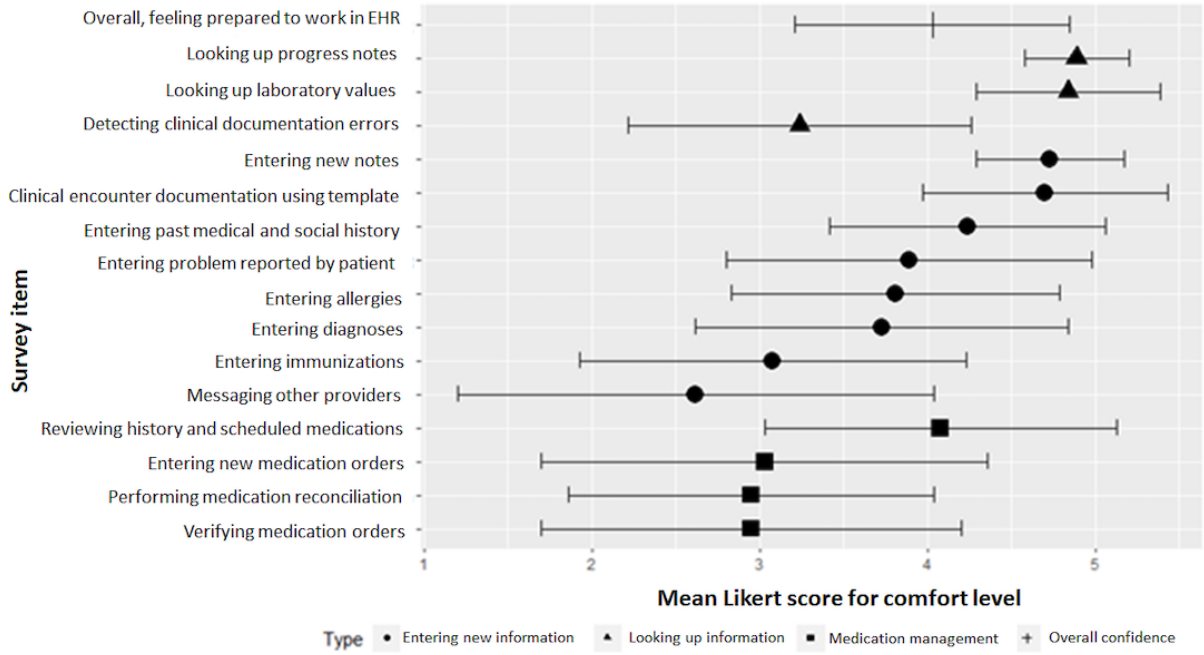
Perceptions of Electronic Health Record Readiness Among Fourth-Year Medical Students

To measure self-perceived readiness in performing EHR-related tasks at the end of undergraduate medical training, we specifically analyzed data from fourth-year medical students. For nearly all survey items, there were no significant differences between fourth-year medical students at the 2 institutions (Multimedia Appendix 3). There were significant differences for only 2 survey items: UCSD fourth-year students were significantly more comfortable with messaging other providers within the EHR (3.0 vs 1.5 at UNMC; *P*=.001) and exhibited greater overall confidence (4.2 vs 3.5 at UNMC; *P*=.04). As there were no differences in all other items and because of the relatively small sample size, the data for fourth-year medical students were combined from the 2 institutions to examine general trends among the overall cohort.

Fourth-year medical students generally felt comfortable with EHR-related tasks, reporting a mean Likert score of 3 or higher for about three-fourths of the EHR-related tasks (11/15, 73%;

Figure 1). Tasks for which the mean Likert scores for comfort level were less than 3 for fourth-year medical students were as follows: entering immunizations, messaging other providers, verifying medication orders, and medication reconciliation (Figure 1).

Figure 1. Mean Likert scores for comfort level with electronic health record–related tasks among senior fourth-year medical students from the University of California San Diego School of Medicine and the University of Nebraska Medical Center College of Medicine, 2019. EHR: electronic health record. Error bars indicate standard deviations.



Students' Comments

Of the 147 respondents, 37 (25.1%) provided open-ended comments (UCSD, n=23 and UNMC, n=14). Comments from the preclinical UCSD students (15/23, 65% of all UCSD commenters) reflected the lack of formal EHR training, for example:

> *I have received no training.*
>
> *We don't get any training? Needs to change.*
>
> *I only feel very comfortable because I have worked at Epic...Otherwise I have received very little/no training.*

Overall, 2 UCSD students (9% of all UCSD commenters) mentioned exposure to EHR in the student-run free clinic, although some felt it was insufficient. For example, one third-year student wrote:

> *Free Clinic was good exposure but I still feel like I could have used more training on note writing before MS3.*

Similarly, a first-year student wrote:

> *I feel very unready to work in EHR. I have to struggle through it every time I am at free clinic. Even the 4th years at free clinic struggle to help me sometimes because they are not as well-versed in EHR as they could be.*

Furthermore, 5 third- and fourth-year medical students (22% of all UCSD commenters) stated that they had learned how to use the EHR through prior experience as a scribe before medical school, resident coaching while on rotations, or just "doing the work." Several students outlined specific areas that could be addressed by training, where they felt relatively less well prepared. These areas included placing orders, more emphasis on inpatient training, and "training focused on common pitfalls or more efficient use of the EHR."

At UNMC, 14 students provided free-text comments, evenly split between preclinical students in year 2 and clinical students in years 3 and 4. There were no comments on the lack of formal training, reflecting the structured preclinical EHR curriculum at UNMC. Moreover, 2 students (14% of all UNMC commenters) stated that this training was helpful, such as:

> *The allergy small group was really good.*
>
> *There are so many tips and tricks that you don't learn unless someone shows you.*

However, the remaining students emphasized the importance of experiential learning over didactic training. For example, the comments included:

> *The EHR is a learn by doing process. I'm in the first three months of clinical rotations, and I have learned more than any of the training sessions.*
>
> *I think plain old practice has made the biggest difference for me.*
>
> *We can get all the "trainings" you want but if we don't actually practice what we "learn," it's gone by the following week.*
>
> *I feel like a lot of what I know how to do is through trial and error.*

Similar to UCSD medical students, several UNMC students (4/14, 29%) cited prior work experience and the critical role of residents in helping them feel comfortable with the EHR.

Furthermore, 11 students (30% of all commenters between both institutions) provided suggestions for improving training, such as the need for formal training at UCSD, interactive small group sessions, request for shorter but more frequent sessions, desire for “training focused on common pitfalls or more efficient use of the EHR,” and “training to personalize EHR.”

Discussion

Principal Findings

The complexity of EHRs was reflected in students’ self-perceived comfort, which varied by the specific EHR components. Across 4 years of medical school, students felt more comfortable with lower-complexity tasks such as looking up existing information and less comfortable with more complex tasks such as entering new information and medication management. Students at both institutions reported lower comfort with looking up clinical documentation errors, entering information on immunizations, reconciling medications, and messaging other providers. This discrepancy persisted even among graduating seniors, with an average comfort with looking up progress notes of 4.9 (out of 5) compared with comfort with reconciling medications of 2.8. Overall, students’ self-reported comfort with working in the EHR was 3.1 at UCSD and 3.2 at UNMC.

EHR training in medical school curricula could benefit from a combination of lower- and higher-complexity tasks. If the current training focuses primarily on navigation and data acquisition from the chart, more emphasis on information entry and medical *decision making* (ie, medication management) for medical students may improve their comfort level with these tasks. Case studies or simulated exercises, where students both look up and enter information into a training EHR environment, could be one strategy to increase comfort with higher-complexity tasks.

For medical trainees, practical experience with EHRs continues to be an important factor in becoming proficient, which has also been highlighted by national guidelines [13-16]. Mean overall confidence was generally higher among students with 1 month or more experience of extracurricular EHR compared with those with less than 1 month of experience. We asked about working or volunteering in different settings because EHR tasks and experience vary by location of usage. With the exception of a student-run free clinic, a longer experience of working in both inpatient and ambulatory settings was associated with higher average comfort of working in the EHR (3.8 vs 2.4 and 3.5 vs 2.3, respectively). The results were similar for both institutions. These effects persisted even after adjusting for other factors such as age, gender, year of training, and institution.

Although both UCSD and UNMC are similar in class size and general curriculum, there were several notable differences among students. Across all stages of training, scores of UNMC students were consistently higher than those of UCSD students on individual items, although not all items were significant. This could be because of the more structured EHR curriculum at UNMC in the preclinical years. This was reinforced by a subanalysis of preclinical students, where UNMC preclinical students expressed significantly higher levels of comfort across multiple tasks in the EHR than UCSD students. Comfort levels, however, were similar among senior fourth-year students at both UNMC and UCSD, suggesting that firsthand experience during the clinical rotation years using the EHR in the context of patient care closed this initial gap. In fact, overall, graduating UCSD students felt more prepared to work with EHR compared with UNMC students (4.2 vs 3.5; P=.04), despite not having any formal preclinical EHR training.

For fourth-year medical students, it is possible that tasks with mean comfort Likert scores of <3, such as for entering immunizations, messaging other providers, verifying medication orders, and medication reconciliation, were reflective of difficulty performing the task itself rather than difficulty with performing the task using the EHR. This could be because of hesitation to make permanent changes in the EHR that impact a patient’s medical record outside of the encounter under which the medical student documents.

Students’ comments emphasized the importance of practical experience, and UCSD respondents reported the lack of formal training. Despite having had a formal preclinical EHR curriculum, UNMC students still emphasized the importance of experiential training. Although the value of didactic training may not be as high as experiential training, it can still provide the foundation and basic familiarity to help students feel more comfortable and confident as they approach their clinical rotations. Students can feel anxious about transitioning to the clinical environment, as they continue to develop their medical knowledge and skills, adjust to working in a clinical environment, and learn to interact with new team members [28-30]. Thus, greater familiarity with EHR could help mitigate some anxiety inherent with this transition to clinical rotations.

As with clinical skills, mentors and, specifically, residents play a large role in students’ learning of the EHR, which was also highlighted in students’ comments. Although much of the resident-led training happens organically, there could be high variability in experience among students depending on the residents’ own familiarity with the EHR. Graduate medical education programs may consider providing formal evaluations of residents’ EHR competencies and, if needed, training.

Limitations

Our study has several limitations, including relatively low survey response rates at both institutions and a cross-sectional design. As such, the results show associations only, and we could not evaluate causality for factors such as the institution and, thus, preclinical EHR curriculum or extracurricular experience. The survey respondents at UNMC tended to be younger than the overall student population, but the demographics of survey respondents at both institutions were consistent with the overall population. We believe that 1 month is generally sufficient to learn basic workflows in 1 setting and decided to use that time frame as a cutoff point for EHR experience. Given the importance of experience in comfort using EHRs, future studies may consider collecting detailed information on prior EHR experience. A longitudinal follow-up may help evaluate how students' responses change as they progress through the curriculum and help determine the most impactful opportunities for EHR training. In addition, in the survey instrument, we did not collect data regarding absence from medical training, such as that for extended research projects, additional degrees, health issues, or parental leave.

Strengths

Our study strengths included multiple institutions and participants in all 4 years of training. Both institutions had vastly different preclinical EHR training curricula. Previous studies that examined EHR use among medical students focused on the types of EHR tasks performed by medical students during their clinical clerkships [20-22,24,31]. We included medical students across the training spectrum, examining the variations in preclinical curricula. In addition, we measured medical students' self-reported comfort rather than specific tasks they could complete in the EHR. Similar to the movement toward patient-reported outcomes rather than objective clinical outcomes in the realm of clinical research, medical education research should consider the subjective experience of medical students.

Conclusions

Medical schools worldwide strive to continue improving the education and well-being of their students. Considering the impact of extracurricular experience on EHR readiness, we should provide more practical opportunities embedded within the preclinical curriculum rather than putting the onus on students to seek out appropriate experiences. Even student-run free clinics may be insufficient to allow all students ample clinical exposure that involves EHR practice. Some medical schools are introducing clinical exposure as formal longitudinal clerkships and introducing clinical rotations earlier in their curricula [32-36]. A combination of didactic and practical experiences combined with structured mentorship and personalization will help provide better EHR training for medical students and address this critical competency in medical education.

Acknowledgments

The authors would like to thank all the medical students who participated in this study. This study was supported by the National Institutes of Health/National Library of Medicine (SB; grant T15LM011271).

Authors' Contributions

LL and GC conceived and designed the study. LL, GC, KC, RT, LA, and HG acquired the data. LL, GC, and SB analyzed the data. LL and SB drafted the manuscript. All authors contributed to the interpretation of the data and critically revised the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey instrument for assessing self-perceived readiness for using electronic health records among medical students.
[PDF File (Adobe PDF File), 63 KB - mededu_v6i1e17585_app1.pdf]

Multimedia Appendix 2
Mean Likert scores for self-perceived readiness to perform various electronic health record–related tasks for preclinical (first- and second-year) medical students by institution.
[DOCX File , 22 KB - mededu_v6i1e17585_app2.docx]

Multimedia Appendix 3
Mean Likert scores for self-perceived readiness to perform various electronic health record–related tasks for senior fourth-year medical students by institution.
[DOCX File , 22 KB - mededu_v6i1e17585_app3.docx]

References

1. Adler-Milstein J, des Roches CM, Kralovec P, Foster G, Worzala C, Charles D, et al. Electronic health record adoption in US hospitals: progress continues, but challenges persist. Health Aff (Millwood) 2015 Dec;34(12):2174-2180. [doi: 10.1377/hlthaff.2015.0992] [Medline: 26561387]
2. Jha AK, Burke MF, des Roches C, Joshi MS, Kralovec PD, Campbell EG, et al. Progress toward meaningful use: hospitals' adoption of electronic health records. Am J Manag Care 2011 Dec;17(12 Spec No):SP117-SP124 [FREE Full text] [Medline: 22216770]
3. Jha AK, DesRoches CM, Kralovec PD, Joshi MS. A progress report on electronic health records in US hospitals. Health Aff (Millwood) 2010 Oct;29(10):1951-1957. [doi: 10.1377/hlthaff.2010.0502] [Medline: 20798168]

4. Henry J, Pylypchuk Y, Searcy T, Patel V. Health IT Dashboard. 2016. Adoption of Electronic Health Record Systems among US Non-Federal Acute Care Hospitals: 2008-2015 URL: http://dashboard.healthit.gov/evaluations/data-briefs/non-federal-acute-care-hospital-ehr-adoption-2008-2015.php [accessed 2016-09-20]
5. Wachter R. The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age. New York, USA: McGraw-Hill Education; 2017.
6. Tipping MD, Forth VE, O'Leary KJ, Malkenson DM, Magill DB, Englert K, et al. Where did the day go?-a time-motion study of hospitalists. J Hosp Med 2010;5(6):323-328. [doi: 10.1002/jhm.790] [Medline: 20803669]
7. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Dec 6;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
8. Robertson SL, Robinson MD, Reid A. Electronic health record effects on work-life balance and burnout within the I population collaborative. J Grad Med Educ 2017 Aug;9(4):479-484 [FREE Full text] [doi: 10.4300/JGME-D-16-00123.1] [Medline: 28824762]
9. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc 2005;12(5):505-516 [FREE Full text] [doi: 10.1197/jamia.M1700] [Medline: 15905487]
10. Joukes E, Abu-Hanna A, Cornet R, de Keizer NF. Time spent on dedicated patient care and documentation tasks before and after the introduction of a structured and standardized electronic health record. Appl Clin Inform 2018 Jan;9(1):46-53 [FREE Full text] [doi: 10.1055/s-0037-1615747] [Medline: 29342479]
11. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426 [FREE Full text] [doi: 10.1370/afm.2121] [Medline: 28893811]
12. Tai-Seale M, Olson CW, Li J, Chan AS, Morikawa C, Durbin M, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. Health Aff (Millwood) 2017 Apr 1;36(4):655-662 [FREE Full text] [doi: 10.1377/hlthaff.2016.0811] [Medline: 28373331]
13. Kirch D. Compliance Advisory: Electronic Health Records (EHRs) in Academic Health Centers. Allowing Medical Student Documentation in the Electronic Health Record URL: https://www.aamc.org/em/aamc/compliance_advisory.pdf [accessed 2019-09-17]
14. Hammoud MM, Dalymple JL, Christner JG, Stewart RA, Fisher J, Margo K, et al. Medical student documentation in electronic health records: a collaborative statement from the alliance for clinical education. Teach Learn Med 2012;24(3):257-266. [doi: 10.1080/10401334.2012.692284] [Medline: 22775791]
15. The Liaison Committee on Medical Education. 2019. Functions and Structure of a Medical School: Standards for Accreditation of Medical Education Programs Leading to the MD Degree URL: https://lcme.org/wp-content/uploads/filebase/standards/2020-21_Functions-and-Structure_2019-10-04.docx [accessed 2019-11-30]
16. AAMC Store. Core Entrustable Professional Activities for Entering Residency URL: https://store.aamc.org/core-entrustable-professional-activities-for-entering-residency.html [accessed 2019-09-13]
17. American Medical Association. 2015. AMA Supports Medical Student Access to Electronic Health Records URL: https://www.ama-assn.org/press-center/press-releases/ama-supports-medical-student-access-electronic-health-records [accessed 2019-09-18]
18. The Bulletin. 2018. Statement on Medical Student Use of the Electronic Health Record URL: http://bulletin.facs.org/2018/04/statement-on-medical-student-use-of-the-electronic-health-record/ [accessed 2019-09-18]
19. American Academy of Family Physicians. Medical Student Access to Electronic Medical Record (EMR) URL: https://www.aafp.org/about/policies/all/medical-student.html [accessed 2019-09-18]
20. Wittels K, Wallenstein J, Patwari R, Patel S. Medical student documentation in the electronic medical record: patterns of use and barriers. West J Emerg Med 2017 Jan;18(1):133-136 [FREE Full text] [doi: 10.5811/westjem.2016.10.31294] [Medline: 28116025]
21. Virden RA, Sonnett FM, Khan AN. Medical student documentation in the emergency department in the electronic health record era-a national survey. Pediatr Emerg Care 2019 Mar;35(3):220-225. [doi: 10.1097/PEC.0000000000001095] [Medline: 28291152]
22. Chi J, Bentley J, Kugler J, Chen JH. How are medical students using the electronic health record (EHR)?: an analysis of EHR use on an inpatient medicine rotation. PLoS One 2019;14(8):e0221300 [FREE Full text] [doi: 10.1371/journal.pone.0221300] [Medline: 31419265]
23. Biagioli FE, Elliot DL, Palmer RT, Graichen CC, Rdesinski RE, Kumar KA, et al. The electronic health record objective structured clinical examination: assessing student competency in patient interactions while using the electronic health record. Acad Med 2017 Jan;92(1):87-91 [FREE Full text] [doi: 10.1097/ACM.0000000000001276] [Medline: 27332870]
24. Foster LM, Cuddy MM, Swanson DB, Holtzman KZ, Hammoud MM, Wallach PM. Medical student use of electronic and paper health records during inpatient clinical clerkships: results of a national longitudinal study. Acad Med 2018 Nov;93(11S Association of American Medical Colleges Learn Serve Lead: Proceedings of the 57th Annual Research in Medical Education Sessions):S14-S20. [doi: 10.1097/ACM.0000000000002376] [Medline: 30365425]

25. Rajaram A, Hickey Z, Patel N, Newbigging J, Wolfrom B. Training medical students and residents in the use of electronic health records: a systematic review of the literature. J Am Med Inform Assoc 2020 Jan 1;27(1):175-180. [doi: 10.1093/jamia/ocz178] [Medline: 31592531]
26. Sullivan GM, Artino AR. Analyzing and interpreting data from likert-type scales. J Grad Med Educ 2013 Dec;5(4):541-542 [FREE Full text] [doi: 10.4300/JGME-5-4-18] [Medline: 24454995]
27. R-Project. The R Project for Statistical Computing URL: https://www.r-project.org/ [accessed 2018-11-12]
28. Wenrich M, Jackson MB, Scherpbier AJ, Wolfhagen IH, Ramsey PG, Goldstein EA. Ready or not? Expectations of faculty and medical students for clinical skills preparation for clerkships. Med Educ Online 2010 Aug 6;15(1):5295 [FREE Full text] [doi: 10.3402/meo.v15i0.5295] [Medline: 20711483]
29. Small RM, Soriano RP, Chietero M, Quintana J, Parkas V, Koestler J. Easing the transition: medical students' perceptions of critical skills required for the clerkships. Educ Health (Abingdon) 2008 Dec;21(3):192. [Medline: 19967639]
30. Haglund ME, aan het Rot M, Cooper NS, Nestadt PS, Muller D, Southwick SM, et al. Resilience in the third year of medical school: a prospective study of the associations between stressful events occurring during clinical rotations and student well-being. Acad Med 2009 Feb;84(2):258-268. [doi: 10.1097/ACM.0b013e31819381b1] [Medline: 19174682]
31. White J, Anthony D, WinklerPrins V, Roskos S. Electronic medical records, medical students, and ambulatory family physicians: a multi-institution study. Acad Med 2017 Oct;92(10):1485-1490. [doi: 10.1097/ACM.0000000000001673] [Medline: 28379934]
32. Gheihman G, Jun T, Young GJ, Liebman D, Sharma K, Brandes E, et al. A review of longitudinal clinical programs in US medical schools. Med Educ Online 2018 Dec;23(1):1444900 [FREE Full text] [doi: 10.1080/10872981.2018.1444900] [Medline: 29542394]
33. Latessa R, Beaty N, Royal K, Colvin G, Pathman DE, Heck J. Academic outcomes of a community-based longitudinal integrated clerkships program. Med Teach 2015;37(9):862-867. [doi: 10.3109/0142159X.2015.1009020] [Medline: 25693796]
34. Royan R, Wu C, Theyyunni N, Montas S, Cranford J, House J, et al. Anything but shadowing! Early clinical reasoning in emergency department improves clinical skills. West J Emerg Med 2018 Jan;19(1):177-184 [FREE Full text] [doi: 10.5811/westjem.2017.10.36691] [Medline: 29383078]
35. Whipple ME, Barlow CB, Smith S, Goldstein EA. Early introduction of clinical skills improves medical student comfort at the start of third-year clerkships. Acad Med 2006 Oct;81(10 Suppl):S40-S43. [doi: 10.1097/00001888-200610001-00011] [Medline: 17001132]
36. Kossoff EH, Hubbard TW, Gowen CW. Early clinical experience enhances third-year pediatrics clerkship performance. Acad Med 1999 Nov;74(11):1238-1241. [doi: 10.1097/00001888-199911000-00019] [Medline: 10587688]

Abbreviations

AAMC: Association of American Medical College
EHR: electronic health record
UCSD: University of California San Diego School of Medicine
UNMC: University of Nebraska Medical Center College of Medicine

Edited by G Eysenbach; submitted 22.12.19; peer-reviewed by J Youm, KH Tan; comments to author 24.03.20; revised version received 05.04.20; accepted 09.04.20; published 12.06.20.

Please cite as:
Lander L, Baxter SL, Cochran GL, Gali HE, Cook K, Hatch T, Taylor R, Awdishu L
Self-Perceptions of Readiness to Use Electronic Health Records Among Medical Students: Survey Study
JMIR Med Educ 2020;6(1):e17585
URL: http://mededu.jmir.org/2020/1/e17585/
doi:10.2196/17585
PMID:32442135

Original Paper

Wikipedia in Vascular Surgery Medical Education: Comparative Study

Michael Yacob[1], MD, FRCSC; Shamim Lotfi[2], MD, FRCSC; Shannon Tang[1]; Prasad Jetty[2], MD, FRCSC

[1]Queen's University, Kingston, ON, Canada
[2]University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
Michael Yacob, MD, FRCSC
Queen's University
76 Stuart St, Victory 3
Kingston, ON, K7L 2V7
Canada
Phone: 1 6135482573
Email: michaelyacob@gmail.com

Abstract

Background: Medical students commonly refer to Wikipedia as their preferred online resource for medical information. The quality and readability of articles about common vascular disorders on Wikipedia has not been evaluated or compared against a standard textbook of surgery.

Objective: The aims of this study were to (1) compare the quality of Wikipedia articles to that of equivalent chapters in a standard undergraduate medical textbook of surgery, (2) identify any errors of omission in either resource, and (3) compare the readability of both resources using validated ease-of-reading and grade-level tools.

Methods: Using the Medical Council of Canada Objectives for the Qualifying Examination, 8 fundamental topics of vascular surgery were chosen. The articles were found on Wikipedia using Wikipedia's native search engine. The equivalent chapters were identified in Schwartz Principles of Surgery (ninth edition). Medical learners (n=2) assessed each of the texts on their original platforms to independently evaluate readability, quality, and errors of omission. Readability was evaluated with Flesch Reading Ease scores and 5 grade-level scores (Flesch-Kincaid Grade Level, Gunning Fog Index, Coleman-Liau Index, Simple Measure of Gobbledygook Index, and Automated Readability Index), quality was evaluated using the DISCERN instrument, and errors of omission were evaluated using a standardized scoring system that was designed by the authors.

Results: Flesch Reading Ease scores suggested that Wikipedia (mean 30.5; SD 8.4) was significantly easier to read (*P*=.03) than Schwartz (mean 20.2; SD 9.0). The mean grade level (calculated using all grade-level indices) of the Wikipedia articles (mean 14.2; SD 1.3) was significantly different (*P*=.02) than the mean grade level of Schwartz (mean 15.9; SD 1.4). The quality of the text was also assessed using the DISCERN instrument and suggested that Schwartz (mean 71.4; SD 3.1) had a significantly higher quality (*P*=.002) compared to that of Wikipedia (mean 52.9; SD 11.4). Finally, the Wikipedia error of omission rate (mean 12.5; SD 6.8) was higher than that of Schwartz (mean 21.3; SD 1.9) indicating that there were significantly fewer errors of omission in the surgical textbook (*P*=.008).

Conclusions: Online resources are increasingly easier to access but can vary in quality. Based on this comparison, the authors of this study recommend the use of vascular surgery textbooks as a primary source of learning material because the information within is more consistent in quality and has fewer errors of omission. Wikipedia can be a useful resource for quick reference, particularly because of its ease of reading, but its vascular surgery articles require further development.

(JMIR Med Educ 2020;6(1):e18076) doi:10.2196/18076

KEYWORDS

medical education; Wikipedia; vascular surgery; medical student

Introduction

Medical education has changed drastically with the increasing use of technology. In particular, internet resources are used by doctors, students, and patients alike to answer clinical questions. Web 2.0 resources such as Wikipedia are rapidly evolving because of their open-source editing community. Currently, there are more than 6 million English articles that are actively monitored and updated by a community of Wikipedia editors [1]. This vast community attracts readers from all backgrounds, from patients seeking medical information to medical professionals needing a quick reference.

In 2009, a survey showed that 80% of physicians use Google, 70% of physicians routinely used Wikipedia, and 53% of physician internet visits involved user-generated web 2.0 resources [2]. Medical students have similarly been observed searching for information online and have identified Wikipedia as a preferred learning resource because of Wikipedia's ease of access (98% of respondents) and ease of understanding (95% of respondents) [3]. Another study [4] showed that, in addition to university resources, first year medical students used Google and Wikipedia most frequently and rarely accessed recommended journal articles and online textbook chapters.

Generally, Wikipedia has been regarded for its readability, although research [5] has identified a lack of consistency in some subjects. Previous research [6] has found that neurosurgery articles on Wikipedia have worse readability when compared to that of national information articles, and that they do not meet the Centers for Disease Control and Prevention clear communication guidelines for patients. A study [7] observed reading engagement in medical students by tracking their eye movements while reading various online resources and found better engagement with Wikipedia. Another study [8] reported that preclerkship medical students showed improved short-term knowledge acquisition after reading general medical topics on Wikipedia compared to that shown when reading textbooks and UpToDate (an online clinical decision support resource).

However, previous research [5] has also suggested that Wikipedia articles across medical and scientific topics vary in quality. While one study's [9] findings supported the use of Wikipedia in answering specific questions about pharmacotherapy, another study [10] found gross inaccuracies in articles on topics in otolaryngology when compared with the same topic in a standard textbook of surgery. In recent years, there has been increasing focus on whether it is appropriate for medical students to use web 2.0 resources for learning in specific subspecializations; assessments of cardiology, gastroenterology, and respirology articles have determined that quality and errors of omission are of significant concern when considering Wikipedia as a medical education resource [11-13]. Similar findings were reported when Wikipedia articles were compared to Grant's Atlas of Anatomy for musculoskeletal anatomy [14].

In this paper, we provide an analysis of the readability and quality of Wikipedia vascular surgery articles by comparing them to a standard surgery textbook. Through this analysis, we hope to better understand whether Wikipedia is suitable as an academic resource for medical students and for junior trainees in the field of vascular surgery.

Methods

Identification and Assessment of Content

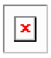
Common diagnoses in vascular surgery (8 different topics) were identified from the Medical Council of Canada Objectives for the Qualifying Examination [15]. These diagnoses were used as default search terms in Wikipedia's [16] native search engine; corresponding chapters of the same title were identified from the table of contents of Schwartz Principles of Surgery [17]. Any discrepancies in article identification were resolved by discussion among authors.

Assessment of Readability

To evaluate the readability of each resource, validated ease-of-reading tools were used. The Flesch Reading Ease score was used to measure reading ease, and the grade levels of each article were determined using 5 different scoring systems—Flesch-Kincaid Grade Level, Gunning Fog Index, Coleman-Liau Index, Simple Measure of Gobbledygook (SMOG) Index, and Automated Readability Index.

The Flesch Reading Ease score [17], represented as a number from 0 to 100, determines the degree of textual difficulty and is calculated as *Flesch Reading Ease score* = 206.835 – (1.015 *ASL*) – (84.5 *ASW*), where *ASL* is the average sentence length and *ASW* is the average number of syllables per word. Scores from 90 to 100 suggest the content is easily understood at a fifth grade level, scores from 80 to 90 suggest the content is easily understood at a sixth grade level, scores from 70 to 80 suggest the content is easily understood at a seventh grade level, scores from 60 to70 suggest the content is easily understood at eighth to ninth grade levels, scores from 50 to 60 suggest the content is easily understood at 10th to 12th grade levels, scores from 30 to 50 suggest the content is easily understood at a college (or university) level, and scores from 0 to 30 suggest the content is at the level of university graduates [6].

Using the Flesch Reading Ease score as a baseline, the other scoring systems were used to determine the education level that the average reader should have in order to understand the text presented. Similarly based on average sentence length and average number of syllables per word, the Flesch-Kincaid Grade Level system [18] calculates a resultant age level that corresponds with respective United States academic grade levels using the equation *Flesch-Kincaid Grade Level* = (0.39 *ASL*) + (11.8 *ASW*) – 15.59. The Gunning-Fog Index calculates grade level using the equation *Gunning-Fog Index* = 0.4 (*ASL* + *PHW*) that considers the average number of words per sentence, *ASL*, and the percentage of hard words, *PHW*. Hard words were considered to be those with 3 or more syllables that were not proper nouns, combinations of easy words, hyphenated words, or two-syllable verbs with -es or -ed endings [6,18]. The Coleman-Liau Index [19] uses the equation *Coleman-Liau Index* = 0.0588 (*average number of letters per 100 words*) – 0.296 (*average number of sentences per 100 words*) – 15.8 that considers the average number of letters and sentences per 100 words. The SMOG Index grade level equation [20] uses the number of polysyllabic words and the total number of sentences:



Gunning-Fog Index, Coleman-Liau Index, and SMOG Index scores correspond with American grade levels with scores greater than 13 indicating college (or university) level and above. The Automated Readability Index [18] is calculated using the total number of characters, words, and sentences in the equation *Automated Readability Index* = 4.71 (*characters*/*words*) + 0.5 (*words*/*sentence*) – 21.43 and provides a score from 1 to 14 that corresponds to grade level, where 13 is equivalent to college (or university) level and 14 is at the level of a college (or university) graduate.

These calculations were performed using Office Word software (version 2010; Microsoft Inc) by copying the full text of each article into the word processor and removing all formatting, images, and tables.

Assessment of Quality

Overall quality and completeness of the resources were assessed with the DISCERN instrument [21] and by identifying errors of omission. To determine the quality of information within, articles were rated on each of the DISCERN instrument's 16 questions with 5-point responses. DISCERN questions address two key domains and the instrument is typically used to help consumers of health information assess the quality of published information regarding treatment choices. DISCERN questions focus on the reliability of the publication source and whether the information is complete [22]. The final question of the DISCERN instrument considers the reader's overall impression of quality. For each question, a score of 1 indicated that the resource had serious or extensive shortcomings, a score of 3 indicated some potentially important but not serious shortcomings, and a score of 5 indicated minimal shortcomings. The total score could range from 16 to 80, where 63 to 80 suggested excellent quality, 51 to 62 was good quality, 39 to 50 was fair quality, and 16 to 38 was poor quality [22].

Assessment of Errors of Omission

Errors of omission were assessed by considering the completeness of the information under categories of epidemiology, pathogenesis, natural history, presenting symptoms, signs on physical examination, noninvasive investigations, invasive investigations, conservative management, medical management, surgical management, endovascular management, and references. Each category, with the exception of references, was assigned a score of 0 if not present, 1 if present but incomplete, or 2 if complete. References were assigned 0 if they were absent or incomplete, or 1 if they were complete. A maximum score of 23 could be assigned to each passage (indicating no errors of omission in any of the aforementioned categories). Errors in each category were counted and totaled per article. To compare between articles, an error of omission rate was calculated by dividing the number of errors by the total word count per article. There was an inverse relationship between the score and the number of errors of omission (or error of omission rate).

Statistical Analysis

Readability scores were analyzed by taking the mean of the grade level or grade-level equivalent obtained from the Flesch Reading Ease score, Flesch-Kincaid Grade Level, Gunning-Fog Index, Coleman-Liau Index, SMOG Index, and Automated Readability Index. Overall readability of each source (Wikipedia or Schwartz) was calculated by taking the mean of all article scores from that source.

Each passage from Wikipedia and Schwartz Principles of Surgery was independently read and assessed for quality and errors of omission by two junior medical trainees with similar levels of education at the time of the study. The articles were read in their native formats to simulate typical reading circumstances. Mean DISCERN score and mean error of omission rate were calculated for each article by averaging between the two readers prior to further analysis. To ensure the quality of these scores, interobserver concordances (kappa value, κ) were calculated, and any concordance values less than 0.8 were discussed by the research team to clarify the discrepancy. The quality and error of omission rate of each source was determined by taking the mean of all articles from that source.

Readability scores, DISCERN scores, and error of omission rates were compared between corresponding Wikipedia and Schwartz articles using two-tailed independent *t* tests with unequal variances. Statistical significance was defined as $P<.05$. All statistical analyses were performed in Excel (version 2020; Microsoft Inc) using the statistical package add-on.

Results

Content Characteristics

Articles on 8 vascular surgery topics were analyzed—carotid artery disease, critical limb ischemia, claudication, acute limb ischemia, aortic dissection, abdominal aortic aneurysm, venous insufficiency, and mesenteric ischemia. At the time of the initial Wikipedia search in July 2013, a search for "carotid artery disease" redirected to an article titled Carotid Artery Stenosis, a search for "critical limb ischemia" redirected to an article titled Peripheral Vascular Disease, and a search for "claudication" redirected to an article titled Intermittent Claudication. The other search terms produced eponymous articles.

An updated search in March 2020, showed that the article titled Peripheral Vascular Disease redirected to an article titled Peripheral Artery Disease, while a search for "critical limb ischemia" redirected to an article titled Chronic Limb Threatening Ischemia. Interestingly, a search for "claudication" resulted in 2 articles—1 entitled Claudication and the other entitled Intermittent Claudication. All other searches resulted in the same pages.

The aforementioned search terms were used to identify subchapters in three vascular surgery–specific chapters of Schwartz Principles of Surgery (ninth edition). The subchapter titled Lower Extremity Arterial Occlusive Disease included Critical Limb Ischemia, Claudication, and Acute Limb Ischemia subheadings. Chronic Venous Insufficiency and Mesenteric Artery Disease subheadings contained the information pertaining to venous insufficiency and mesenteric ischemia, respectively. A comparison with the 11th edition of Schwartz Principles of Surgery yielded no changes.

Readability Scores

Wikipedia had a mean Flesch Reading Ease score of 30.5 (SD 8.4) across all the articles, while Schwartz had a mean score of 20.2 (SD 9.0) which suggested that Wikipedia articles can be understood by readers at a college (or university) level, while Schwartz content was at the level of a college (or university) graduate.

Using 5 different indices to determine approximate grade level, Wikipedia had a mean grade level of 14.2 (SD 1.3), while the Schwartz had a mean grade level of 15.9 (SD 1.4). These were in agreement with the Flesch Reading Ease scores and placed the Wikipedia articles at a lower grade level than the Schwartz Principles of Surgery text. Both Wikipedia and Schwartz content was for readers at the postsecondary level.

The differences between Wikipedia and Schwartz readability scores (Table 1) were statistically significant for Flesch Reading Ease score (*P*=.03), Gunning Fog Index (*P*=.02), Coleman-Liau Index (*P*=.02), SMOG Index (*P*=.04), and Automated Readability Index (*P*=.04), but not for the Flesch-Kincaid Grade Level (*P*=.06). This suggests that the Wikipedia articles were consistently easier to read, and that Schwartz Principles of Surgery was written for a more advanced audience. These relationships are further illustrated in Figures 1 and 2.

Table 1. Comparison between Wikipedia and Schwartz Principles of Surgery readability.

Readability assessment	Wikipedia, mean (SD)	Schwartz, mean (SD)	*t* test (*df*)	*P* value
Flesch Reading Ease	30.5 (8.4)	20.2 (9.0)	2.14 (14)	.03
Flesch-Kincaid Grade Level	13.8 (1.6)	15.5 (1.7)	2.14 (14)	.06
Gunning Fog Index	16.6 (1.6)	19.1 (2.0)	2.14 (14)	.02
Coleman-Liau Index	15.0 (1.4)	16.5 (1.0)	2.16 (13)	.02
SMOG Index	12.5 (1.2)	13.9 (1.3)	2.14 (14)	.04
Automated Readability Index	12.9 (1.6)	14.7 (1.4)	2.14 (14)	.04
Mean grade level	14.2 (1.3)	15.9 (1.4)	2.14 (14)	.02

Figure 1. Comparison of Wikipedia and Schwartz average Flesch Reading Ease scores.

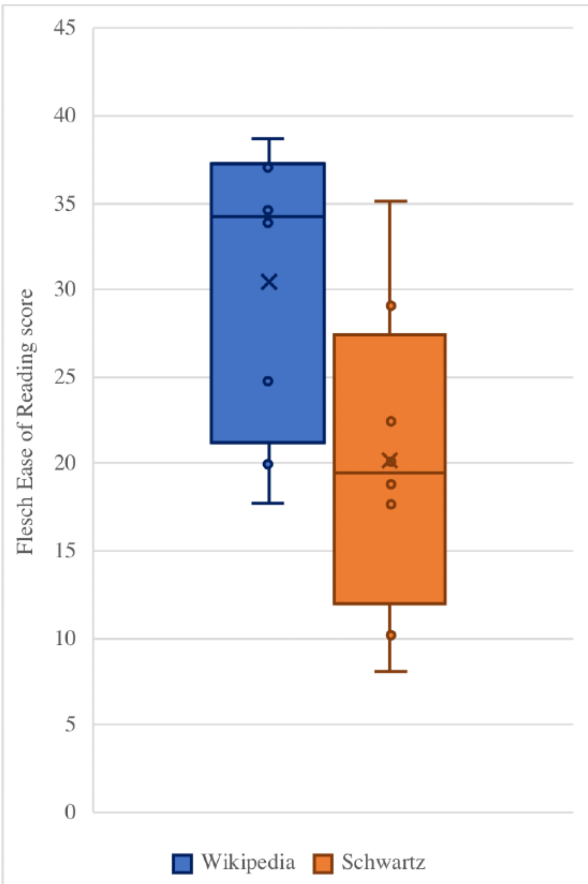
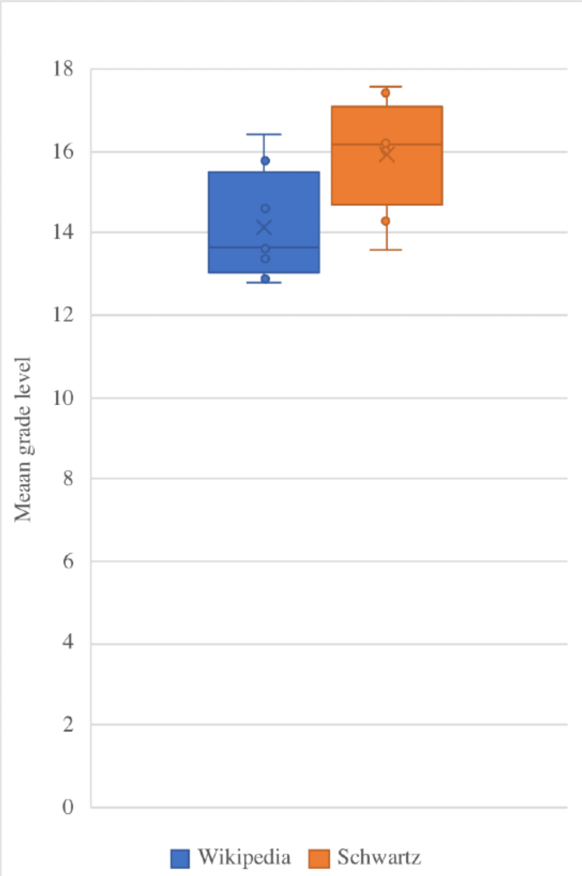


Figure 2. Comparison of mean grade level using Flesch-Kincaid Grade Level, Gunning Fog Index, Coleman-Liau Index, SMOG Index, and Automated Readability Index scores.



Quality Assessment

DISCERN scores are shown in Table 2. All of the Schwartz subchapters had scores between 65.0 and 74.5 (ie, all were classified as excellent by DISCERN criteria). In contrast, only the Wikipedia article entitled Peripheral Vascular Disease received a score that was classified as excellent. Of the 8 topics, 4 Wikipedia articles were classified as good, 2 were classified as fair, and 1 was classified as poor according to their DISCERN score. On the whole, Schwartz Principles of Surgery subchapters (mean 71.4; SD 3.1) performed significantly better than Wikipedia (mean 52.9; SD 11.4) on the DISCERN scoring criteria (P=.002). This suggests that the content of the text from Schwartz Principles of Surgery is superior in quality to that of the text from Wikipedia.

Interobserver concordance ranged from κ=0.68 to κ=0.96 for Wikipedia and from κ=0.64 to κ=1.00 for Schwartz. Interobserver concordance values less than 0.8 were observed for the Wikipedia article on aortic dissection and the Schwartz subchapters on acute limb ischemia and aortic dissection; however, the overall final scores were similar and were thus deemed acceptable.

Table 2. Comparison of Wikipedia and Schwartz Principles of Surgery DISCERN scores indicating content quality.

Article topics	DISCERN, mean of both readers	
	Wikipedia	Schwartz
Carotid artery disease	62.0	74.0
Critical limb ischemia	64.5	73.0
Claudication	45.0	73.5
Acute limb ischemia	56.5	71.0
Aortic dissection	57.5	70.0
Abdominal aortic aneurysm	61.0	74.5
Venous insufficiency	31.0	70.5
Mesenteric ischemia	45.5	65.0

Errors of Omission

The total number of errors of omission for each article are demonstrated in Table 3. Overall, Wikipedia articles contained a significantly greater number of errors of omission (P=.008) compared to Schwartz Principles of Surgery. Notably, the Wikipedia article on critical limb ischemia (Peripheral Vascular Disease) was extremely incomplete with a score of 1 out of a maximum of 23. The mean errors of omission scores for Wikipedia was 12.5 (SD 6.8)).The highest scoring article on Wikipedia was on the topic of abdominal aortic aneurysm with a score of 21.5 points. In contrast, most articles for Schwartz Principles of Surgery scored high with a mean of 21.3 (SD 1.9) points, and the subchapter on chronic venous insufficiency scored the lowest at 17 points. From these results, we can reasonably infer that Schwartz Principles of Surgery is a more complete resource compared to Wikipedia for vascular surgery topics.

Interobserver concordance for the Wikipedia assessments ranged from κ=0.77 to κ=1.00, while interobserver concordance for Schwartz Principles of Surgery ranged from κ=0.08 to κ=1.00 (Table 4). The subchapter of Venous Insufficiency from Schwartz had an interobserver concordance of κ=0.08 which reflected a number of incomplete sections in the text; however, the final errors of omission scores were ultimately similar between raters, so the mean score was nevertheless used in the statistical analysis. Further scoring breakdowns are available in Multimedia Appendix 1.

Table 3. Comparison of Wikipedia and Schwartz Principles of Surgery errors of omission.

Article topics	Errors of omission, mean of both readers	
	Wikipedia	Schwartz
Carotid artery disease	12.0	22.0
Critical limb ischemia	1.0	23.0
Claudication	9.0	22.0
Acute limb ischemia	14.5	23.0
Aortic dissection	19.5	21.0
Abdominal aortic aneurysm	21.5	20.5
Venous insufficiency	6.5	17.0
Mesenteric ischemia	16.0	21.5

Table 4. Interobserver concordance values for DISCERN and error of omission assessments.

Article topics	DISCERN interobserver concordance		Errors of omission interobserver concordance	
	Wikipedia	Schwartz	Wikipedia	Schwartz
Carotid artery disease	0.90	1.00	0.92	1.00
Critical limb ischemia	0.94	1.00	1.00	1.00
Claudication	0.84	0.86	0.82	1.00
Acute limb ischemia	0.84	0.64	0.90	1.00
Aortic dissection	0.68	0.66	0.81	0.36
Abdominal aortic aneurysm	0.82	0.82	0.77	0.67
Venous insufficiency	0.96	1.00	0.97	0.08
Mesenteric ischemia	0.96	1.00	0.90	0.77

Discussion

Common vascular surgery topics (8 topics) were selected from the Medical Council of Canada Objectives for the Qualifying Examination, and their equivalent articles were identified on Wikipedia and in subchapters in Schwartz Principles of Surgery. Through analysis of the readability and quality of the content of these sources, we investigated the suitability of Wikipedia as a vascular surgery resource for medical students.

We found that the quality of Wikipedia articles was mostly classified as good or fair by DISCERN criteria (Table 2), and that Wikipedia articles were written at the college (or university) level (Figure 1). There were numerous errors of omission in many of the Wikipedia articles, and some articles did not contain fundamental subsections (Table 3). Schwartz Principles of Surgery was found to contain consistently higher quality content (classified as excellent by DISCERN criteria) and contained fewer errors of omission (Tables 2 and 3) but was also found to have lower readability (Figure 1). This was attributed to the length explanations of concepts in the text which contained additional supporting figures and tables and likely contributes to the higher grade level that was associated with the text.

While the Wikipedia articles were found to have a lower grade level and higher reading ease compared to that of Schwartz

(Table 1 and Figure 2), both sources still required that the reader be comfortable with reading at the college (or university) level. The difference in reading level required for Wikipedia articles compared to that required of textbook passages, as well as Wikipedia's ease of access likely is the reason it and other online resources are appealing to users [2-4].

Existing research [7] suggested that preclerkship medical students performed better on short-term knowledge acquisition tests after studying from Wikipedia compared to when they studied from a textbook. This could be as a result of the presence of superfluous background information in the textbook, and the fact that Wikipedia is written for the general population. Because of their general target audience, Wikipedia can be written in such a way that key points are summarized simply without the need for extensive explanation which allows for faster knowledge acquisition, especially for newer medical students.

It is important to consider that Scaffidi et al's study [7] on Wikipedia content and quality focused on general medical topics, whereas other studies [11-13] have focused on specialized medical fields such as cardiology, respirology, and gastroenterology and have found findings contradictory to those of Scaffidi et al. Another study [14] focusing on musculoskeletal anatomy content on Wikipedia discovered inaccuracies and errors of omission as well as reporting that many references were not appropriate. In fact, these studies [11-14] do not recommend Wikipedia as a learning resource for medical students because of the significant errors of omission and low-quality content.

In vascular surgery topics, many Wikipedia articles were incomplete, and the many inconsistencies that were noted on the website that continue to be unresolved at the time this paper was written. An example of an inconsistency observed on Wikipedia is the presence of two separate articles titled Claudication and Intermittent Claudication, both describing the condition associated with vascular claudication in peripheral artery disease [23,24]. In addition, there are a number of statements that lack appropriate references (for example, in the article titled Claudication, numerous statements are note-cited). Citations on Wikipedia are recommended to be from reliable secondary or tertiary resources, and this is not true in a number of the vascular surgery articles [25].

The open-source nature of Wikipedia and other web 2.0 resources is one of its greatest strengths, but it can also be one of its greatest weaknesses. Wikipedia is a vast resource with over 100,000 actively contributing editors from around the world [26]. It is moderated by these editors who also strive to maintain the quality of the resource; however, in a specialized field such as vascular surgery there may be difficulty in validating content accuracy due a dearth of available or knowledgeable editors. In this era of rapid information turnover, it could be argued that textbooks cannot be updated as quickly as an open-source website such as Wikipedia. This ability to be kept up-to-date in combination with its ease of access makes Wikipedia a valuable resource for medical information and is the reason it is widely used by the general population and medical professionals alike [2-4].

In this study, we critically examined the quality, completeness, and readability of Wikipedia articles on common vascular surgery topics and compared them to corresponding excerpts from Schwartz Principles of Surgery. This study used validated tools, such as the Flesch Reading Ease scoring system and the DISCERN instrument. Five readability indices were used to minimize bias when calculating grade level and similar results were obtained which supports their validity. The DISCERN instrument and error of omission assessment introduce some subjectivity into the study as a result of individual interpretation. In addition, the items on the errors of omission rating scale were developed based on standard presentation in medical education and requires further assessment to demonstrate its validity. Presentation of information in textbooks and Wikipedia may vary between the sources for better readability in that particular source. For example, information regarding epidemiology may be combined with natural history to aid the reader in connecting these concepts. These variations were often the reason for the low interobserver concordance values in Table 4.

Ultimately, the use of Wikipedia in medical education should not be disregarded. It has the potential to serve as a powerful reference for all users, but medical professionals and students should be aware that articles on Wikipedia are written with the general population in mind. This study demonstrates that further development is required for vascular surgery topics on Wikipedia before it can be reliably recommended as a resource for medical trainees. Currently, surgical textbooks are more likely to reflect the depth and breadth of knowledge required for medical learners in the field of vascular surgery. Further research could examine how web 2.0 resources are utilized depending on the level of the trainee, and the motivations for choosing a particular resource at a particular stage of clinical knowledge acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Raw data set and statistical analysis details.

[XLSX File (Microsoft Excel File), 65 KB - mededu_v6i1e18076_app1.xlsx]

References

1. Statistics. Wikipedia. En URL: https://en.wikipedia.org/wiki/Wikipedia:Statistics [accessed 2020-03-02]

2. Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. Int J Med Inform 2009 Oct;78(10):645-655. [doi: 10.1016/j.ijmedinf.2009.04.008] [Medline: 19501017]
3. Allahwala UK, Nadkarni A, Sebaratnam DF. Wikipedia use amongst medical students - new insights into the digital revolution. Med Teach 2013 Apr;35(4):337 [FREE Full text] [doi: 10.3109/0142159X.2012.737064] [Medline: 23137251]
4. Judd T, Elliott K. Selection and use of online learning resources by first-year medical students: cross-sectional study. JMIR Med Educ 2017 Oct 02;3(2):e17. [doi: 10.2196/mededu.7382]
5. Mesgari M, Okoli C, Mehdi M, Nielsen F, Lanamäki A. "The sum of all human knowledge": a systematic review of scholarly research on the content of Wikipedia. J Assn Inf Sci Tec 2014 Dec 02;66(2):219-245. [doi: 10.1002/asi.23172]
6. Modiri O, Guha D, Alotaibi NM, Ibrahim GM, Lipsman N, Fallah A. Readability and quality of wikipedia pages on neurosurgical topics. Clinical Neurology and Neurosurgery 2018 Mar;166:66-70. [doi: 10.1016/j.clineuro.2018.01.021]
7. Saparova D, Nolan NS. Evaluating the appropriateness of electronic information resources for learning. J Med Libr Assoc 2016 Jan;104(1):24-32. [doi: 10.3163/1536-5050.104.1.004]
8. Scaffidi MA, Khan R, Wang C, Keren D, Tsui C, Garg A, et al. Comparison of the impact of Wikipedia, UpToDate, and a digital textbook on short-term knowledge acquisition among medical students: randomized controlled trial of three web-based resources. JMIR Med Educ 2017 Oct 31;3(2):e20. [doi: 10.2196/mededu.8188]
9. Kupferberg N, Protus BM. Accuracy and completeness of drug information in Wikipedia: an assessment. J Med Libr Assoc 2011 Oct;99(4):310-313. [doi: 10.3163/1536-5050.99.4.010]
10. Volsky PG, Baldassari CM, Mushti S, Derkay CS. Quality of Internet information in pediatric otolaryngology: a comparison of three most referenced websites. Int J Pediatr Otorhinolaryngol 2012 Sep;76(9):1312-1316. [doi: 10.1016/j.ijporl.2012.05.026] [Medline: 22770592]
11. Azer SA, AlSwaidan NM, Alshwairikh LA, AlShammari JM. Accuracy and readability of cardiovascular entries on Wikipedia: are they reliable learning resources for medical students? BMJ Open 2015 Oct 06;5(10):e008187 [FREE Full text] [doi: 10.1136/bmjopen-2015-008187] [Medline: 26443650]
12. Azer SA. Evaluation of gastroenterology and hepatology articles on Wikipedia: are they suitable as learning resources for medical students? Eur J Gastroenterol Hepatol 2014 Feb;26(2):155-163. [doi: 10.1097/MEG.0000000000000003] [Medline: 24276492]
13. Azer SA. Is Wikipedia a reliable learning resource for medical students? evaluating respiratory topics. Advances in Physiology Education 2015 Mar;39(1):5-14. [doi: 10.1152/advan.00110.2014]
14. London DA, Andelman SM, Christiano AV, Kim JH, Hausman MR, Kim JM. Is Wikipedia a complete and accurate source for musculoskeletal anatomy? Surg Radiol Anat 2019 Jul 1;41(10):1187-1192. [doi: 10.1007/s00276-019-02280-1]
15. Medical Council of Canada. Examination Objectives Overview URL: https://mcc.ca/objectives/ [accessed 2020-05-27]
16. Wikipedia. Welcome to Wikipedia URL: https://en.wikipedia.org/wiki/Main_Page [accessed 202-05-27]
17. Brandt M, Brunicardi F, Schwartz S. Schwartzs Principles of Surgery, Ninth Edition, ABSITE and Board Review. New York: McGraw-Hill Medical; 2011.
18. Kinkaid P, Fishburne R, Rogers R, Chissom B. Institute for Simulation and Training. 1975. Derivation Of New Readability Formulas For Navy Enlisted Personnel URL: https://stars.library.ucf.edu/istlibrary/56 [accessed 2020-03-02]
19. Coleman M, Liau TL. A computer readability formula designed for machine scoring. Journal of Applied Psychology 1975;60(2):283-284. [doi: 10.1037/h0076540] [Medline: 28425725]
20. McLaughlin G. SMOG grading: a new readability formula. Journal of Reading 1969;12(8):639-646.
21. DISCERN online. The DISCERN instrument URL: http://www.discern.org.uk/discern_instrument.php [accessed 2020-05-27]
22. Rees CE, Ford JE, Sheard CE. Evaluating the reliability of DISCERN: a tool for assessing the quality of written patient information on treatment choices. Patient Education and Counseling 2002 Jul;47(3):273-275. [doi: 10.1016/s0738-3991(01)00225-7]
23. Wikipedia. Claudication URL: https://en.wikipedia.org/wiki/Claudication [accessed 2020-04-13]
24. Wikipedia. Intermittent Claudication URL: https://en.wikipedia.org/wiki/Intermittent_claudication [accessed 2020-04-13]
25. Wikipedia. Reliable Sources URL: https://en.wikipedia.org/wiki/Wikipedia:Reliable_sources#Primary,_secondary,_and_tertiary_sources [accessed 2020-04-13]
26. Wikipedia. Wikipedians URL: https://en.wikipedia.org/wiki/Wikipedia:Wikipedians [accessed 2020-04-13]

Abbreviations

SMOG: Simple Measure of Gobbledygook index

Edited by G Eysenbach; submitted 01.02.20; peer-reviewed by LA Lee, M Laurent, J Heilman; comments to author 24.02.20; revised version received 13.04.20; accepted 14.05.20; published 19.06.20.

Please cite as:
Yacob M, Lotfi S, Tang S, Jetty P
Wikipedia in Vascular Surgery Medical Education: Comparative Study
JMIR Med Educ 2020;6(1):e18076
URL: http://mededu.jmir.org/2020/1/e18076/
doi:10.2196/18076
PMID:32417754

Original Paper

A Mobile Medical Knowledge Dissemination Platform (HeadToToe): Mixed Methods Study

Ido Zamberg[1,2*], MD; Olivier Windisch[3*], MD; Thomas Agoritsas[1,2,4], MD, PhD; Mathieu Nendaz[1,2], MD, MPH; Georges Savoldelli[2,5], MD, MEd; Eduardo Schiffer[2,5], MD, PhD

[1]Division of General Internal Medicine, Department of Medicine, University Hospitals of Geneva, Geneva, Switzerland

[2]Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3]Division of Urology, Department of Surgery, University Hospitals of Geneva, Geneva, Switzerland

[4]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[5]Division of Anesthesiology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care, and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Ido Zamberg, MD
Division of General Internal Medicine
Department of Medicine
University Hospitals of Geneva
Rue Gabrielle-Perret-Gentil 4
Geneva, 1205
Switzerland
Phone: 41 768035683
Email: idozamberg@gmail.com

Abstract

Background: Finding readily accessible, high-quality medical references can be a challenging task. HeadToToe is a mobile platform designed to allow easy and quick access to sound, up-to-date, and validated medical knowledge and guidance. It provides easy access to essential clinical medical content in the form of documents, videos, clinical scores, and other formats for the day-to-day access and use by medical students and physicians during their pre- and postgraduate education.

Objective: The aim of this paper is to describe the architecture, user interface, and potential strengths and limitations of an innovative knowledge dissemination platform developed at the University of Geneva, Switzerland. We also report preliminary results from a user-experience survey and usage statistics over a selected period.

Methods: The dissemination platform consists of a smartphone app. Through an administration interface, content is managed by senior university and hospital staff. The app includes the following sections: (1) main section of medical guidance, organized by clinical field; (2) checklists for history-taking and clinical examination, organized by body systems; (3) laboratory section with frequently used lab values; and (4) favorites section. Each content item is programmed to be available for a given duration as defined by the content's author. Automatic notifications signal the author when the content is about to expire, hence, promoting its timely updating and reducing the risk of using obsolete content. In the background, a third-party statistical collecting tool records anonymous utilization statistics.

Results: We launched the final version of the platform in March 2019, both at the Faculty of Medicine at the University of Geneva and at the University Hospital of Geneva in Switzerland. A total of 622 students at the university and 613 health professionals at the hospital downloaded the app. Two-thirds of users at both institutions had an iOS device. During the practical examination period (ie, May 2019) there was a significant increase in the number of active users (P=.003), user activity (P<.001), and daily usage time (P<.001) among medical students. In addition, there were 1086 clinical skills video views during this period compared to a total of 484 in the preceding months (ie, a 108% increase). On a 10-point Likert scale, students and physicians rated the app with mean scores of 8.2 (SD 1.9) for user experience, 8.1 (SD 2.0) for usefulness, and 8.5 (SD 1.8) for relevance of content. In parallel, postgraduate trainees viewed more than 6000 documents during the first 3 months after the implementation in the Division of Neurology at our institution.

Conclusions: HeadToToe is an educator-driven, mobile dissemination platform, which provides rapid and user-friendly access to up-to-date medical content and guidance. The platform was given high ratings for user experience, usefulness, and content quality and was used more often during the exam period. This suggests that the platform could be used as tool for exam preparation.

(JMIR Med Educ 2020;6(1):e17729) doi:10.2196/17729

KEYWORDS

clinical skills; clinical competence; clinical practice guidelines; medical education; smartphone; innovation; medical guidance; mobile phone

Introduction

The teaching of clinical skills and choice of management plans are important cornerstones in medical schools' curricula [1]. Skills like medical and personal history taking, physical examination, and procedural skills are taught in parallel to more theoretical aspects and prepare the medical student for his or her clinical years. A medical student is expected to be able to examine patients displaying a large variety of complaints, as well as to perform many different medical procedures. Throughout medical school, students' clinical skills are evaluated and in the Switzerland medical federal licensing examination, passing an Objective Structured Clinical Examination [2] test is mandatory in order to become a fully certified physician. Like any skill, clinical skills improve with experience, repetition, and constructive feedback provided by supervisors. Nevertheless, students rely on references (ie, documents and videos) to prepare themselves for the clinical environment [3,4]. Even though a multitude of references do exist for clinical skills, the information is vast and scattered between a large number of sources, and finding high-quality and validated references can be a time-consuming and frustrating process [5,6].

In the clinical environment, students and residents rely on clinical practice guidelines and recommendations to choose a treatment plan for a patient, which are, as mentioned, not always readily available [6]. Finding optimal guidance regarding clinical skills, procedural skills, or patient management is challenging for all clinicians. It is even more so for students and postgraduate trainees, who seek not only sound and trustworthy guidance [7,8] but also perhaps the most didactic and one that fits the expectations of their educators and senior clinical staff. Thus, learners can benefit from a dissemination platform whose content is authored or endorsed by their local faculty, educators, and senior hospital staff.

Moreover, the use of smartphones and other mobile devices in the medical environment, commonly termed mobile health (mHealth) [9], has increased rapidly throughout the last several years [10-12]. mHealth can offer not only patient-centered solutions for chronic illness management, behavioral change, or self-monitoring [13-22] but also learning opportunities for pre- and postgraduate medical education [23]. The online environment plays an important role as well in continuous education, and health professionals rely on easy-to-access and high-quality medical content to improve patient care [24-27]. Furthermore, it was shown that health professionals favor significantly well-known targeted medical resources to more general web browsing in their search for what they perceive as validated medical content [28].

This increasing use of online resources and smartphones and the need for easy-to-access, validated, and high-quality medical information urged the development of HeadToToe, a mobile platform intended for the dissemination of medical knowledge and tailored guidance for pre- and postgraduate health professionals.

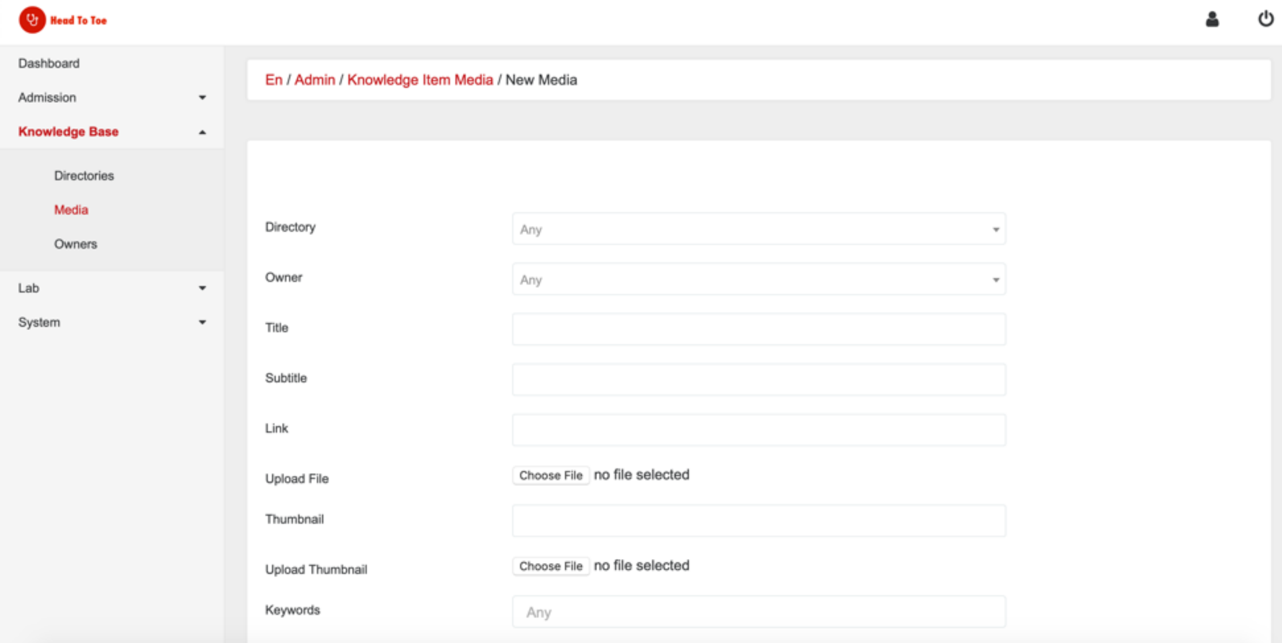
The purpose of this article is to describe the architecture of the tool and its potential strengths and limitations compared to existing tools, as well as to present preliminary evaluation of usage statistics and user experience. Our hypothesis at this stage of development is that users would perceive our platform as a rapid and user-friendly way to access up-to-date medical content and guidance, and that its content, being validated by local senior educators, would be therefore perceived as trustworthy and useful, both for education and practice.

Methods

Platform Description

The HeadToToe [29] platform consists of an iOS and Android front end that is compatible with iPhones, iPads, Android phones, and tablets. Both front ends connect to the same back-end server, which includes server-side code and a database allowing the centralization of all content and easy updating of the platform. The platform is managed through an administration interface (see Figure 1), which is hosted on the same server as the database. A blank version (ie, content-free version) of the platform can be made available to other institutions for the purpose of academic collaborations. The current operational platform is in French, but the platform does support a multilingual environment.

Figure 1. The administration interface of the HeadToToe platform. Educational content managers can use the interface to insert, update, and delete new content. Through the interface, administrators can also create the admission checklist for the admission checklist section as well as the laboratory values section.



Back End: Database and Administration Interface

Basic Architecture

The server is hosted on Heroku and contains the database, the code, and the application programming interface (API) for its administration. The API and the administration interface were written with Ruby on Rails, version 2.5.3 (MIT). The administration interface provides a user-friendly interface for adding, deleting, and editing content. Access to the administration interface is made through a weblink, restricted by a username and a password for educational content managers (ECMs) (ie, faculty educators and senior hospital staff managing content in HeadToToe).

Content Management

The cornerstone of the platform is that the ECMs are responsible for identification, endorsement, and updating of the content for the medical field under their responsibility. This provides the validation that learners need to be able to trust that the content provided by the platform meets their educational needs. The users themselves do not have access to the management interface and cannot upload their own content. Each time the app is launched or refreshed by the users, new or updated content is highlighted with a specific icon.

Content Metadata

Each content item has several attributes, including title, subtitle, search keywords, link, thumbnail, update date, and expiration date. An item can be a link to a file, a video, or a website. In addition, each item's metadata includes the full name and contact details of the ECM, providing the users the possibility to communicate with content managers.

Content Item Responsibility

The content item responsibility can be defined by clinical field, subfield, or by individual content items. For example, the head of neurology could be the ECM for the whole neurology section or could delegate the responsibility for certain subfields or items to other senior medical staff or educators.

Duration of Validated Content and Updating Mechanism

Medical knowledge and guidance, whether international, regional, or local, is always evolving. For each item, the ECM selects an expiration date, thus defining the duration over which the content will be accessible to users. Automatic notifications signal the author when the content is about to expire, thus promoting its timely updating and reducing the risk of using obsolete content. Namely, 2 weeks and 24 hours before a content item reaches its expiration date an automatic email will be sent to the ECM to remind him or her about the expiration of the specific content item. The ECM will then be able to decide if the item is still valid, thus prolonging its availability on the platform. If the item is considered not valid, the ECM may update it with a newer version or delete it. Eventually, expired items will be deleted automatically after a predefined period if no action confirms the validity of the resource.

Front End: User Interface

The Front End of the iOS and Android App Includes Several Sections

The Knowledge Base

The knowledge base is the heart of the app (see Figure 2) and contains the medical content and guidance organized by clinical specialty with folder hierarchy defined through the administration interface by the ECM of each section. Folder hierarchy is flexible and can differ between clinical fields. This section allows the user to navigate through each medical field's sections and find the needed content. After navigating to the desired section, the user can access files or media (ie, videos or website links). An information button is available for each document and media item. A click on the button will open a pop-up view containing the item's metadata: contact details for the ECM, last update date, and expiration date. The expiration date will show in green within the duration of its availability and in red when it is expired; it will stay red until it is either updated, manually deleted by the ECM, or automatically deleted after a predefined period. For documents, a download button is present and allows download and offline use.

Figure 2. The knowledge base of the HeadToToe platform (in French). Users can navigate through different medical specialties organized as folders.

Admission

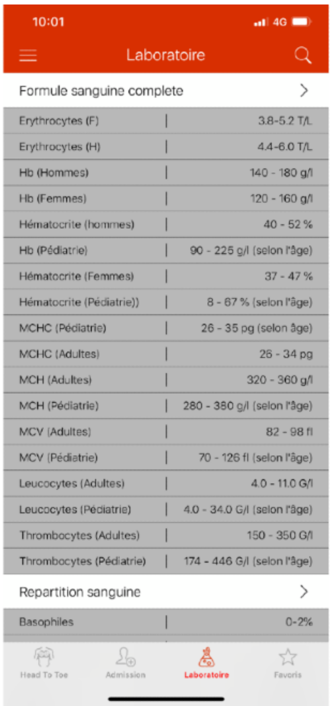
In this section (see Figure 3), users can find an extensive checklist with questions to be asked during history-taking of a patient and actions to perform during clinical examination. Items can be checked, comments can be added to each item, and the user can generate a PDF containing checked items and comments.

Figure 3. History-taking and clinical examination checklist section in the HeadToToe plaform (in French). Users can consult the list, add comments, and create a PDF with all checked items and added information.

Laboratory Values

This section (see Figure 4) contains local reference values for frequently used laboratory values, such as complete blood count, chemistry, basic metabolic panel, arterial blood gas, and more.

Figure 4. Laboratory values in the HeadToToe platform (in French). Users can consult basic laboratory values validated for local reference values.

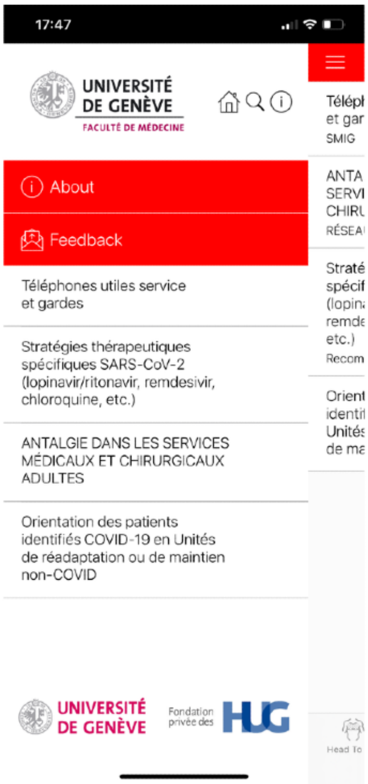


Favorites

Documents in the knowledge base can be marked as favorites. This section shows the list of favorites and allows users to unmark them (see Figure 5). Favorite documents can also be found on the side menu of the app (see Figure 6).

Figure 5. Favorites screen in the HeadToToe platform (in French). Through this screen, users can quickly access marked content.

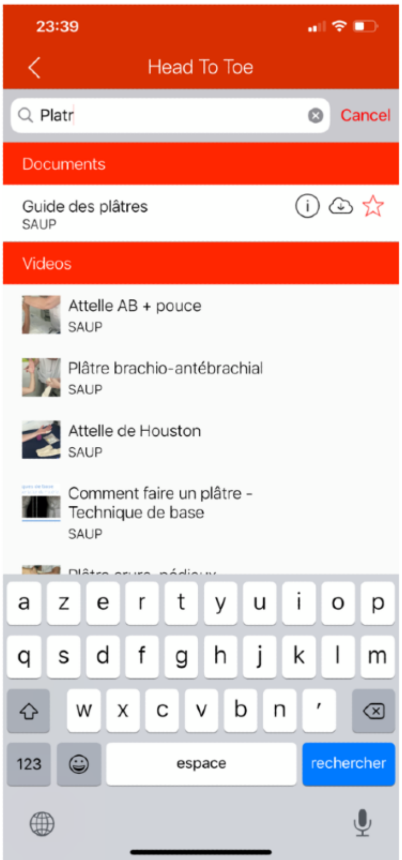
Figure 6. Side menu in the HeadToToe platform (in French). Through this menu, users can quickly access their favorite content and the *About* menu. Users may also access the Feedback screen and send feedback and questions concerning app use and content.



Search Screen

The search screen (see Figure 7) is accessible throughout the app and allows a transversal search through all content (ie, documents, media, and lab values) for quick access to needed content. A search is made by the item's title, subtitle, and keywords defined in the administration interface to improve search accuracy.

Figure 7. The search screen of the HeadToToe platform (in French). Users can search through all types of content from this screen and quickly access needed information.



Implementation and Dissemination of the App in the Institution

Target users of the platform are medical students in the Faculty of Medicine at the University of Geneva and postgraduate residents from the University Hospital of Geneva, Switzerland. The app is currently available for download exclusively through a private website in order to preserve the content's copy and the authors' rights of local protocols as well as content imported from external sources. The iOS app is distributed through Apple's Enterprise distribution methods. The Android app is distributed through the same private website with an APK (Android Package Kit) file. Students and physicians have separate versions for distinct statistical analysis. A student does not have the credentials to download the physicians' version, and vice versa, to ensure statistical separation.

In the pregraduate level, the app was implemented at the University of Geneva in 2015 as a beta version and the final version was released in March 2019. The medical curriculum consists of a fully integrated bachelor's, master's, and medical degree 6-year program. The teaching of clinical skills starts from as early as the second year onward; clinical skills are taught in parallel and with relation to basic sciences studies until the end of the sixth year [30]. Therefore, the target population was medical students from the second to the sixth year of medical school, which includes roughly 750 students at a given time. Since the launch of the final version of the app in March 2019, to date, the app was downloaded 622 times, counting multiple user devices.

At the postgraduate level, we first performed a pilot test from November 2018 to February 2019 and disseminated the app primarily in the Division of Neurology at the University Hospital of Geneva, Switzerland. In March 2019, we launched the final version of the app and made it progressively available to the Divisions of General Pediatrics, Pediatric Emergency, Neonatology, Neurosurgery, Pre-Hospital Emergency Care, Anesthesiology, Urology, Nephrology, Hematology, Diabetology, and Primary Care. Our roadmap includes the implementation of HeadToToe in more divisions, and eventually the whole institution, in coordination with all relevant medical leadership. So far, the app was downloaded 613 times by a total of approximately 1900 doctors from the above-mentioned divisions in our institution (see Table 1).

Table 1. User distribution between university students and hospital medical personnel since the launch of the final version of the HeadToToe platform in March 2019.

Users and devices	App downloads[a], n (%)	Potential users[b], n	Potential yearly renewals[c], n
University of Geneva students			
Total	622 (100)	750	150
Users with Apple devices	430 (69.1)	N/A[d]	N/A
Users with Android devices	192 (30.9)	N/A	N/A
University Hospital of Geneva health professionals			
Total	613 (100)	1900	300
Users with Apple devices	409 (66.7)	N/A	N/A
Users with Android devices	204 (33.3)	N/A	N/A

[a]Includes multiple-device downloads by the same user.

[b]Represents an estimated number of members of the target audience in the mentioned institutions.

[c]An estimate of new institutional students or employees, based on mean number of new medical students per year as well as mean number of new medical staff in our institution.

[d]N/A: Not applicable.

Educational Content Managers: Authoring or Endorsing Content

To ensure validated and high-quality content to the best of our abilities, we contacted senior physicians in different medical fields in our institution. Pregraduate content was chosen by head faculty members responsible for clinical skills teaching, and postgraduate content was selected by either heads of divisions and units or by other physicians from the respective divisions or units to whom the mission was delegated. Each of these specialists was responsible and contributed to his or her field's section in the platform with international and local content that he or she deemed updated, validated, and useful for his or her fellow colleagues. Each department, division, and unit in the institution can have its own section in the app without any content-type or size restriction. The goal of combining pre- and postgraduate content was to allow continuous usage of the platform from the student level, through residency, and the rest of medical training.

Platform Development Process

The platform's concept emerged from an urgent need and difficulties mentioned above by some of the authors to find high-quality and locally validated information concerning clinical skills during their medical university years. Several members of our team have dual training in computer science and medicine, which puts them in a position to be both code and architecture developers, as well as end users of the platform. This position made the design of the app and its architecture user-centered from its basic core, as we were not only developing a product but a tool that we ourselves use in our daily practice.

To ensure an even more user-centered development process, before launching the final version, we held several trial periods for pre- and postgraduate users in order to try and understand each group's specific needs and to receive feedback. In addition to the postgraduate trial period described earlier, a beta version was launched in 2015 and used by 93 fifth-year medical students in the University of Geneva before distributing the app to the rest of the faculty. In addition, both users and ECMs have the possibility to send us feedback and ideas directly from the app and from the administration interface. Automatic utilization statistics allow us to identify the most-used type of content and, thus, to add similar content as well as to eliminate unused content. Lastly, as discussed further in the Platform Assessment section, satisfaction surveys were sent to platform users with the possibility to write free-form text feedback.

These measures allowed us to improve user experience, fix bugs, and add requested features and content as suggested by medical students and health professionals during each step of the development.

Platform Assessment

Evaluation of the platform utilization and usage was made in two different ways. First, a factual analysis was conducted with data obtained from statistical collection by Yahoo's Flurry Analytics to allow quantitative analysis of the utilization of the tool. Four different Flurry API keys were created to distinguish between iOS and Android apps and between medical students and physicians. Collection was made continuously, anonymously, and automatically.

Events sent to Flurry by the iOS and Android apps were the title of the item, the medical field it belongs to, and whether it was accessed by search or directly through the main knowledge base. Statistics are separate for medical students and doctors. This allows us to record summary statistics about how many times each item was used. Flurry automatically collects information about the usage of the app, including the number of active devices per day, number of sessions per day, number of sessions per device per day, time of usage per device per day, median session length, and more. The user's journeys are also recorded, which means we are able to follow each user's session and actions made in order the find needed information as well as the amount of time spent to retrieve it.

A second assessment focused on user satisfaction, and qualitative utility was conducted through a survey. One survey was addressed to all medical students, from the second to the sixth year of medical school, and another one was addressed to hospital physicians and nurses from the Division of Neurology, as they were the first to use the app during a trial run. The surveys focused on user experience, general usefulness of the app, and the relevance of the content (see Multimedia Appendix 1 for survey questions). Surveys sent to both students and physicians were identical.

Statistical Analysis

P values were calculated using unpaired *t* tests for continuous quantitative variables. Calculations were made using Stata statistical software, version 16 (StataCorp). *P*<.05 was considered statistically significant.

Results

User Demographics

As automatic statistics collection was anonymous, we do not have exact demographic knowledge regarding age and sex of users. We can estimate that the mean age for university medical students is between 20 and 30 years and that for medical doctors is between 25 and 65 years. All users live in Geneva, Switzerland, or its surroundings, and are either medical students at the University of Geneva or health professionals (ie, doctors or nurses) at the University Hospital of Geneva. Distribution of downloads between the university and the hospital, as well as between iOS and Android devices, is summarized in Table 1.

Pregraduate Analysis

Assessment of Utilization

In the pregraduate level, during the period from March to June 2019, a total of 251 students downloaded the app (iOS and Android combined). There was a significant rise in daily users and usage time with an average of 24.5 (SE 1.8) students per day during the exam period, compared to 16.5 (SE 1.9) from March to April 2019 (*P*=.003) (see Table 2). This resulted in an average of 8.2 (SE 0.8) minutes per day during the exam period compared to 5.1 (SE 0.5) during the control period (*P*<.001) (see Table 2), and almost double the total number of sessions during the exam period as compared to the months before (see Figure 8). Number of sessions per day increased significantly as well, with a mean of 89.5 (SE 8.3) sessions per day during the exam period, compared to 56.7 (SE 5.4) sessions per day during the control period (*P*<.001) (see Table 2). The median session length, which may reflect the time each user needs to find the requested content, was 35 seconds. This fact can be cross-referenced with *user journeys* gathered by Yahoo's Flurry Analytics for each user's session, which shows the exact navigation path each user made through the app and time spent. During this period, 3756 documents and 1570 videos were consulted (see Table 2).

Table 2. Pregraduate assessment of utilization during the control period compared to the exam period within 251 HeadToToe app downloads among medical students at the University of Geneva, Switzerland.

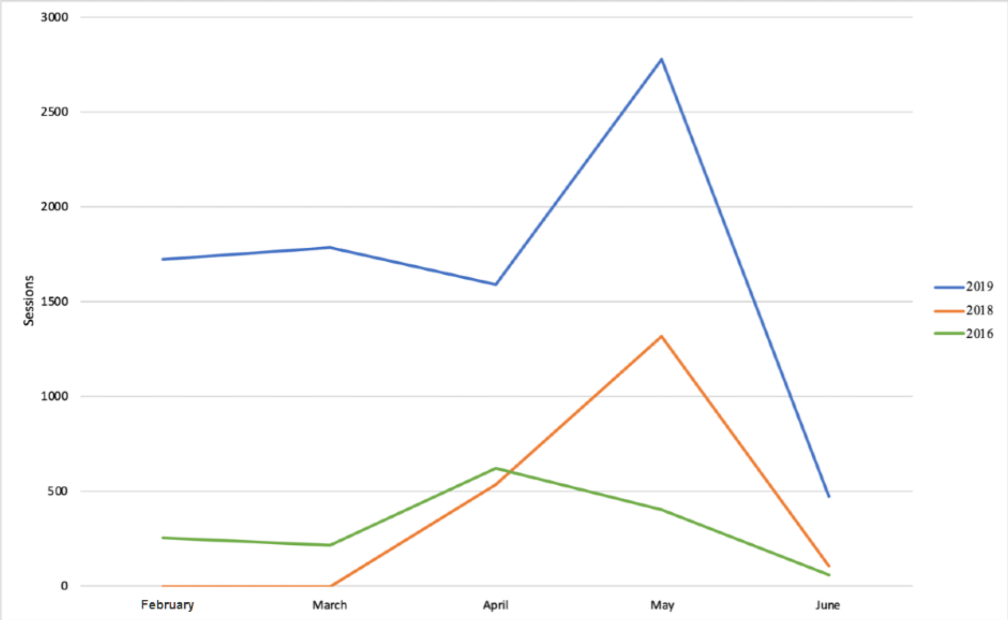
Assessment measure[a]	Exam period (May 2019)	Control period (March and April 2019)	Difference	*P* value[b]
Daily activity, mean (SE)				
Active devices	24.5 (1.8)	16.5 (1.9)	8.0 (2.8)	.003
Minutes per user per day	8.2 (0.8)	5.1 (0.5)	3.1 (0.95)	<.001
Total sessions per day	89.5 (8.3)	56.7 (5.4)	32.8 (9.5)	<.001
Content usage, n				
Documents consulted	1451	2305	N/A	N/A[c]
Videos consulted	1086	484	N/A	N/A

[a]Data were from automatic statistics collection with Yahoo's Flurry Analytics.

[b]Calculated with an unpaired *t* test for continuous quantitative variables.

[c]N/A: Not applicable.

Figure 8. Exam period usage patterns from 2016 to 2019. The exam period is usually at the end of the spring semester, either at the beginning or end of May each year.



Assessment of Utility

In the pregraduate level, 138 students answered the survey. Out of a top score of 10, students rated the app with a mean score of 8.2 (SD 1.9) for user experience, 8.1 (SD 2.0) for usefulness, and 8.5 (SD 1.8) for the relevance of content (see Table 3). A total of 39.1% (54/138) of students considered the video section as the most useful and 42.0% (58/138) considered the document section as the most useful. A total of 48.6% (67/138) of students said they would like to see more procedural skills videos and 30.4% (42/138) said they would like to see more clinical scores.

Table 3. Pre- and postgraduate assessment of utility.

Institution and users	N	User experience, mean (SD)[a]	Usefulness, mean (SD)[a]	Relevance of content, mean (SD)[a]
University of Geneva medical students	138	8.2 (1.9)	8.1 (2.0)	8.5 (1.8)
Division of Neurology, University Hospital of Geneva physicians	28	7.8 (2.2)	8.6 (1.7)	8.5 (1.5)

[a]Values represent weighted mean scores on a 10-point Likert scale survey answered by users at both institutions.

Postgraduate Analysis

Assessment of Utilization

At the postgraduate level, the app was downloaded 98 times, counting multiple devices per health professional, during the trial period of 3 months (ie, between November 2018 and February 2019). Regarding usage, there was an average of 10.1 (SD 5.1) users per day for an average of 8.1 (SD 8.6) minutes per user per day with a mean of 55 (SD 35) sessions per day (see Table 3). Median session length was 29 seconds. During this period, 6494 files and 144 videos were consulted (see Table 4). Documents viewed were mainly local guidance for acute stroke management.

Table 4. Postgraduate assessment of utilization during a trial period among health professionals in the Division of Neurology at the University Hospital of Geneva, Switzerland.

Assessment measure[a]	Mean (SD)	Range[b]	Total, n	Mean per month (SD)
Daily activity				
Active devices per day	10.1 (5.1)	2-26	N/A[c]	N/A
Minutes per user per day	8.1 (8.6)	0.8-66	N/A	N/A
Total sessions per day	55.0 (35.5)	6-162	N/A	N/A
Content usage				
Documents viewed	N/A	N/A	6494	2134 (480.5)
Videos viewed	N/A	N/A	144	48 (12.2)

[a]Data were from automatic statistics collection with Yahoo's Flurry Analytics between November 2018 and February 2019.

[b]Represents minimum and maximum values for reported parameters.

[c]N/A: Not applicable.

Assessment of Utility

A total of 28 health professionals answered the survey (see Table 3). Out of a top score of 10, user-friendliness was rated with a mean score of 7.8 (SD 2.2), usefulness had a mean score of 8.6 (SD 1.7), and content relevance had a mean score of 8.5 (SD 1.5). When asked for the most useful section of the app, documents were elected at 100%, compared to admission, videos, and laboratory values. When asked for further content addition, 71% (20/28) wanted to have additional protocols, 21% (6/28) asked for clinical scores, and 7% (2/28) elected additional clinical skills videos as the most useful.

Discussion

Principal Findings

Finding easy-to-access, high-quality, and validated evidence-based medical content is a challenging task. The use of the online and mobile environment is increasingly growing and becoming a relevant tool for pre- and postgraduate medical education and clinical practice. Health professionals tend to favor the use of targeted medical online and mobile resources as compared to nonspecific web browsing. In order to address these difficulties and to adapt to modern medical content–consuming habits, we have developed the HeadToToe platform.

Medical students and doctors using the app found our platform to be user-friendly and generally useful, and they perceived the content as practical and relevant for daily clinical practice. For example, during a pilot implementation phase in the Division of Neurology, 6500 documents were consulted in 3 months (ie, about 72 per day).

Continuous and automatic user-based statistics, as gathered in our platform, can provide understanding of the students' learning process and helps identify the most frequently used content items. During an evaluation period of 3 months since the distribution of the final version through March and June 2019, as well as statistics collection from the beta version since 2015, we noted daily constant use of the app with increased activity during exam periods with the same activity patterns for the last 4 years (see Figure 8). Usage in 2017 is not shown as no statistics were gathered due to further development of the platform. We noticed a progressive increase in use during the month preceding the exam period and were able to identify frequently used content during this period. These findings can allow future focus on students' needs and studying habits, as frequently used content can be further developed. These findings might also suggest that the platform may be used as a tool for exam preparation, specifically in the clinical skills field.

Numerous online and mobile platforms exist and provide highly validated and high-quality medical information. Geeky Medics [31] and Bates' Guide to Clinical Examination and History Taking [32] are examples of highly used and validated online and mobile platforms that provide access to clinical skills material. The MDCalc website and app is an example of a useful tool for clinical scores calculation [33]. UpToDate and PubMed are other tools used for retrieving evidence-based medical knowledge and treatment plans [34]. These sources are just a few among others, and their use and trustworthiness is widely recognized. Each of these tools provides quality information about a specific aspect of medical knowledge and they are complementary with regard to a specific clinical question [35].

However, local institutions often lack dissemination for clinically relevant content endorsed by local medical educators. Different file managers, such as Dropbox and Google Drive, can be used [36,37] but have obvious limitations. For example, they lack specific medical-relevant structure and user experience and do not provide specific solutions for obsolete content management.

In contrast, our platform has several potential strengths. HeadToToe provides an easy solution for the dissemination of selected medical references within an institution. Content selection is educator driven, which means that ECMs from each medical field in a given institution can be responsible for content management and endorsement, thus ensuring content validation and quality. Information can be retrieved from several validated knowledge bases and is not restricted to a single source. Users can then easily access information from a variety of medical fields through a mobile app, and frequently used content can be tagged and downloaded for offline use. Through the administration interface, adding and updating content is simple, and users will only have access to the most recent version of an item and are automatically notified of the arrival of new content. Moreover, HeadToToe contains a programmatically determined expiration date for each item, which ensures up-to-date content by automatically notifying the ECM when content is about to expire and requires new validation or updating of the item. When no action is taken, expired items will be automatically deleted. This is particularly important in the medical environment where evidence is evolving constantly, and this feature provides a continuously updated platform avoiding the dissemination of obsolete medical content. This feature would be harder to achieve with a simple file manager, as obsolete content is difficult to identify, manage, and suppress and would require a manual updating system. Furthermore, the platform provides metadata for each item, displaying information and contact details of the ECM as well as visibility of the updated date and expiration date.

The centralization of institutional information within a single tool may help with the integration of new medical staff at any stage of their medical career by providing quick and easy access to local practice guidance and practical information, such as important phone numbers, call schedule, and more. For example, we observed that including both pre- and postgraduate material seems to facilitate implementation of the platform, as medical students are exposed to the platform early in their curriculum and continue to use it during clinical rounds and after graduation. This can help them transfer and apply knowledge and skills in the clinical environment and promote continuous education.

From an educational and academic point of view, the tool provides automatic monitoring of content usage and user activity and can provide information to ECMs about content relevance and learners' needs. Usage data gathered may provide interesting insights for research in the field of medical education. Another advantage of the platform's unified architecture is that it also offers a simple way to disseminate local medical knowledge to other academic or private partners. Indeed, we are developing targeted versions of the app, which can be customized for specific medical content intended for external use.

Data collected using the platform could help in presenting more evidence for the utility of mHealth solutions in clinical practice. In fact, while many mobile solutions targeted to health professionals are being used on a daily basis, evidence is still lacking concerning their impact on clinicians' adherence to validated guidance and on the clinical impact on reducing unwarranted variation in practice. By monitoring user activity and content usage patterns, and linking it with clinical indicators, we might be able to present more concrete, real-world data supporting the use of mobile platforms in clinical settings.

The platform might present a cost-effective and ecological solution [38,39] for knowledge dissemination, as it may make obsolete the need for printing medical and practical information for new and current staff members, even more so for information that is updated frequently. In addition, the platforms' use might present as time-efficient and, therefore, cost-saving for health professionals, as knowledge is centralized and rapidly accessible within a mobile app, thus eliminating the need to find an available computer and to browse the web. The economic and ecological impact of this type of intervention was not yet studied and would be of better value when the platform is fully deployed within the institution.

The platform presents several limitations. The first and main limitation is that it requires the validation and triage of a large amount of medical content as well as the coordination between several educators. The platform also requires content-quality check, which could become time-consuming, as more content accumulates on the platform. The designation of an ECM is a crucial part of the process and does require active participation in the creation of the platform. Nevertheless, institutions are often required to locally adapt or endorse the use of international guidance and to translate it into local practice due to differences in populations and local resources. Thus, triage and centralization of medical content can be useful for identifying existing content and promoting it, as well as identifying fields where content and evidence is lacking and need to be addressed in a given institution. Our experience and the uptake in our institution has so far been very positive but remains to be further assessed. Implementation and evaluation would also need to be assessed in other clinical environments and institutions.

To date, we did not include patients in our platform development and design process, as they were not the target users of the mobile app. Nevertheless, as there is growing evidence for the importance of patients' input in mHealth interventions, especially for patient-centered resources [19,40-42], inclusion of patients in future developments could be of interest, specifically concerning their perception of the utilization of mobile devices by health professionals in daily practice.

The quality of the content can also be criticized as subjective and be a matter of debate. Indeed, we do not offer a technological solution for the measurement of the content's quality. However, ECMs, as mentioned, are senior medical staff and specialists who are, by definition, responsible for local strategies, medical education, and treatment plans. Thus, content quality and relevance are guaranteed by the ECMs' institutional roles. Moreover, the app has the merit of making validated content transparent to all partners and, thus, helps identify either information needs or conflicting guidance on similar topics. Institutions and local ECMs would still need to use sound methods for critical appraisal of content to include. For content that amounts to recommendations for clinical practice, ECMs should use trustworthiness criteria, such as the ones published by the Institute of Medicine [8,43] or by leading experts, such as the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) Working Group [44]. Finally, all online content needs to meet copyright regulations, both for published content as well as for locally created content.

Conclusions

The HeadToToe platform allows medical educators to create a validated, high-quality, and up-to-date reference platform for simplified pre- and postgraduate medical education knowledge dissemination, to the benefit of their students and medical staff. It is built for easy and quick access. Users found the app to be user-friendly, relevant, and useful for clinical practice. Implementation in different universities and clinical settings would be the next natural step for assessing its relevance in broader settings and scalability. Utilization patterns should be further examined in light of students' and residents' information needs and learning habits, both as a tool for exam preparation and for daily clinical activity. The potential impact on the reduction of unwarranted variations in practice, quality of care, and economic outcomes should be further studied, and randomized trials could compare the use of such integrated dissemination platform to current available tools.

Acknowledgments

This project received financial support from the Informatics Committee of the University of Geneva, Switzerland, and the private foundation of the University Hospital of Geneva, Switzerland. Both are nonprofit organizations.

Authors' Contributions

IZ was the main investigator, wrote the manuscript, was cofounder of the project and idea, programmed the platform code, took part in the analyses of the results and literature review, and took part in the implementation at the pre- and postgraduate levels. OW was the cofounder of the project and idea, took part in the analyses of the results and literature review, wrote the manuscript, and took part in the implementation at the pre- and postgraduate levels. ES was the team leader and cofounder of the project and idea, wrote the first draft and critically revised the manuscript, and supervised the project. TA critically revised the manuscript, took part in the analyses of the results and literature review, and took part in the writing of the final draft. MN critically revised the manuscript and helped in the implementation of the platform at the university. GS critically revised the manuscript and helped and supervised the implementation of the platform at the university and hospital.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questions.
[DOCX File , 14 KB - mededu_v6i1e17729_app1.docx]

References

1. Elder A. Clinical skills assessment in the twenty-first century. Med Clin North Am 2018 May;102(3):545-558. [doi: 10.1016/j.mcna.2017.12.014] [Medline: 29650075]
2. Zayyan M. Objective structured clinical examination: The assessment of choice. Oman Med J 2011 Jul;26(4):219-222 [FREE Full text] [doi: 10.5001/omj.2011.55] [Medline: 22043423]
3. Kim JY, Myung SJ. Could clinical experience during clerkship enhance students' clinical performance? BMC Med Educ 2014 Oct 02;14:209 [FREE Full text] [doi: 10.1186/1472-6920-14-209] [Medline: 25273978]
4. Vogel D, Harendza S. Basic practical skills teaching and learning in undergraduate medical education: A review on methodological evidence. GMS J Med Educ 2016;33(4):Doc64 [FREE Full text] [doi: 10.3205/zma001063] [Medline: 27579364]
5. Thiele RH, Poiro NC, Scalzo DC, Nemergut EC. Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: Results of a randomised trial. Postgrad Med J 2010 Aug;86(1018):459-465. [doi: 10.1136/pgmj.2010.098053] [Medline: 20709767]
6. Kritz M, Gschwandtner M, Stefanov V, Hanbury A, Samwald M. Utilization and perceived problems of online medical resources and search tools among different groups of European physicians. J Med Internet Res 2013 Jun 26;15(6):e122 [FREE Full text] [doi: 10.2196/jmir.2436] [Medline: 23803299]
7. Institute of Medicine, Board on Health Care Services, Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. In: Steinberg E, Greenfield S, Miller Wolman D, editors. Clinical Practice Guidelines We Can Trust. Washington, DC: National Academies Press; 2011.
8. Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: A new era is imminent. Chest 2013 Aug;144(2):381-389. [doi: 10.1378/chest.13-0746] [Medline: 23918106]
9. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: Conceptualization of a global framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
10. Gavali MY, Khismatrao DS, Gavali YV, Patil KB. Smartphone, the new learning aid amongst medical students. J Clin Diagn Res 2017 May;11(5):JC05-JC08 [FREE Full text] [doi: 10.7860/JCDR/2017/20948.9826] [Medline: 28658804]
11. Morris CG, Church L, Vincent C, Rao A. PDA usage and training: Targeting curriculum for residents and faculty. Fam Med 2007 Jun;39(6):419-424. [Medline: 17549651]

12. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: A systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
13. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
14. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: A systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
15. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: A systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
16. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: A meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
17. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: A systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
18. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
19. Badawy SM, Thompson AA, Liem RI. Technology access and smartphone app preferences for medication adherence in adolescents and young adults with sickle cell disease. Pediatr Blood Cancer 2016 May;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]
20. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: Systematic review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
21. Osborn J, Ajakaiye A, Cooksley T, Subbe CP. Do mHealth applications improve clinical outcomes of patients with cancer? A critical appraisal of the peer-reviewed literature. Support Care Cancer 2020 Mar;28(3):1469-1479 [FREE Full text] [doi: 10.1007/s00520-019-04945-4] [Medline: 31273501]
22. Prada P, Zamberg I, Bouillault G, Jimenez N, Zimmermann J, Hasler R, et al. EMOTEO: A smartphone application for monitoring and reducing aversive tension in borderline personality disorder patients, a pilot study. Perspect Psychiatr Care 2017 Oct;53(4):289-298. [doi: 10.1111/ppc.12178] [Medline: 27439663]
23. Davies BS, Rafique J, Vincent TR, Fairclough J, Packer MH, Vincent R, et al. Mobile Medical Education (MoMEd). How mobile information resources contribute to learning for undergraduate clinical students: A mixed methods study. BMC Med Educ 2012 Jan 12;12:1 [FREE Full text] [doi: 10.1186/1472-6920-12-1] [Medline: 22240206]
24. McCord G, Smucker WD, Selius BA, Hannan S, Davidson E, Schrop SL, et al. Answering questions at the point of care: Do residents practice EBM or manage information sources? Acad Med 2007 Mar;82(3):298-303. [doi: 10.1097/ACM.0b013e3180307fed] [Medline: 17327723]
25. Casebeer L, Bennett N, Kristofco R, Carillo A, Centor R. Physician internet medical information seeking and on-line continuing education use patterns. J Contin Educ Health Prof 2002;22(1):33-42. [doi: 10.1002/chp.1340220105] [Medline: 12004639]
26. Foster CC, Macy ML, Simon N, Stephen R, Lehnig K, Bohling K, et al. Emergency care connect: Extending pediatric emergency care expertise to general emergency departments through telemedicine. Acad Pediatr 2020 Feb 26:1. [doi: 10.1016/j.acap.2020.02.028] [Medline: 32112864]
27. Dumestre DO, Fraulin F. Avoiding breach of patient confidentiality: Trial of a smartphone application that enables secure clinical photography and communication. Plast Surg (Oakv) 2020 Feb;28(1):12-18. [doi: 10.1177/2292550319880910] [Medline: 32110641]
28. De Leo G, LeRouge C, Ceriani C, Niederman F. Websites most frequently used by physician for gathering medical information. AMIA Annu Symp Proc 2006:902 [FREE Full text] [Medline: 17238521]
29. Windisch O, Zamberg I, Iselin C, Schiffer E. Head To Toe, a medical knowledge distribution platform: A practical example in urology [Article in French]. Rev Med Suisse 2019 Nov 27;15(673):2205-2208. [Medline: 31778051]
30. Vu NV, Bader CR, Vassalli JD. The current curriculum in medicine at the University of Geneva [Article in French]. Rev Med Suisse Romande 1997 Jan;117(1):65-68. [Medline: 9082540]
31. Frere M, Tepper J, Fischer M, Kennedy K, Kropmans T. Measuring situation awareness in medical education objective structured clinical examination guides. Educ Health (Abingdon) 2017;30(3):193-197 [FREE Full text] [doi: 10.4103/efh.EfH_306_16] [Medline: 29786019]
32. Bickley LS. Bates' Guide to Physical Examination and History Taking. 7th edition. Philadelphia, PA: Lippincott Williams & Wilkins; 1999.
33. Elovic A, Pourmand A. MDCalc medical calculator app review. J Digit Imaging 2019 Oct;32(5):682-684. [doi: 10.1007/s10278-019-00218-y] [Medline: 31025219]

34. Hoogendam A, Stalenhoef AFH, de Vries Robbé PF, Overbeke AJPM. Answers to questions posed during daily patient care are more likely to be answered by UpToDate than PubMed. J Med Internet Res 2008 Oct 03;10(4):e29 [FREE Full text] [doi: 10.2196/jmir.1012] [Medline: 18926978]
35. Agoritsas T, Vandvik PO, Neumann I, Rochwerg B, Jaeschke R, Hayward R, et al. Finding current best evidence. In: Guyatt G, Rennie D, Meade MO, Cook DJ, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd edition. New York, NY: McGraw-Hill Education; 2015.
36. Yoldi-Negrete M, Morales-Cedillo IP, Navarro-Castellanos I, Fresán-Orellana A, Panduro-Flores R, Becerra-Palars C. Combining Adobe Forms and Dropbox to obtain a low-cost electronic data collection system. Telemed J E Health 2019 Mar;25(3):250-256. [doi: 10.1089/tmj.2018.0029] [Medline: 29932811]
37. Luan A, Momeni A, Lee GK, Galvez MG. Cloud-based applications for organizing and reviewing plastic surgery content. Eplasty 2015;15:e48 [FREE Full text] [Medline: 26576208]
38. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
39. Badawy SM, Kuhns LM. Economic evaluation of text-messaging and smartphone-based interventions to improve medication adherence in adolescents with chronic health conditions: A systematic review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
40. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: A think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
41. Badawy SM, Thompson AA, Kuhns LM. Medication adherence and technology-based interventions for adolescents with chronic health conditions: A few key considerations. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
42. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
43. Institute of Medicine, Board on Health Care Services, Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. In: Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical Practice Guidelines We Can Trust. Washington, DC: National Academies Press (US); 2011.
44. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]

Abbreviations

API: application programming interface
APK: Android Package Kit
ECM: educational content manager
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
mHealth: mobile health

Edited by G Eysenbach; submitted 09.01.20; peer-reviewed by T Ungar, S Badawy; comments to author 26.02.20; revised version received 19.03.20; accepted 20.03.20; published 27.05.20.

Please cite as:
Zamberg I, Windisch O, Agoritsas T, Nendaz M, Savoldelli G, Schiffer E
A Mobile Medical Knowledge Dissemination Platform (HeadToToe): Mixed Methods Study
JMIR Med Educ 2020;6(1):e17729
URL: http://mededu.jmir.org/2020/1/e17729/
doi:10.2196/17729
PMID:32249758

Original Paper

A Peer-Led Social Media Intervention to Improve Interest in Research Careers Among Urban Youth: Mixed Methods Study

Christianah Ogunleye[1], BSA; Jeanne M Farnan[1], MD, MHPE; Shannon K Martin[1], MD, MS; Audrey Tanksley[1], MD; Samantha Ngooi[1], MPP; Laura Ruth Venable[1], MPH, MS; Samantha Anderson[1], BA; Jhonatan Marte[1], BS; David O Meltzer[1], MD, PhD; Vineet M Arora[1], MD, MAPP

Pritzker School of Medicine, University of Chicago, Chicago, IL, United States

Corresponding Author:
Vineet M Arora, MD, MAPP
Pritzker School of Medicine
University of Chicago
5841 South Maryland Ave
W216
Chicago, IL, 60637
United States
Phone: 1 773 702 8157
Email: varora@medicine.bsd.uchicago.edu

Abstract

Background: Novel methods to boost interest in scientific research careers among minority youth are largely unexplored. Social media offers a unique avenue toward influencing teen behavior and attitudes, and can therefore be utilized to stimulate interest in clinical research.

Objective: The aim of this study was to engage high-achieving minority youth enrolled in a science pipeline program to develop a targeted social media marketing campaign for boosting interest in clinical research careers among their peers.

Methods: Students enrolled in the Training Early Achievers for Careers in Health program conducted focus groups in their communities to inform themes that best promote clinical research. They then scripted, storyboarded, and filmed a short video to share on social media with a campaign hashtag. Additionally, each student enrolled peers from their social circle to be subjects of the study. Subjects were sent a Career Orientation Survey at baseline to assess preliminary interest in clinical research careers and again after the campaign to assess how they saw the video, their perceptions of the video, and interest in clinical research careers after watching the video. Subjects who did not see the video through the online campaign were invited to watch the video via a link on the postsurvey. Interest change scores were calculated using differences in Likert-scale responses to the question "how interested are you in a career in clinical research?" An ordinal logistic regression model was used to test the association between watching a peer-shared video, perception of entertainment, and interest change score controlling for underrepresented minorities in medicine status (Black, American Indian/Alaska Native, Native Hawaiian, or Pacific Islander), gender, and baseline interest in medical or clinical research careers.

Results: From 2014 to 2017, 325 subjects were enrolled as part of 4 distinct campaigns: #WhereScienceMeetsReality, #RedefiningResearch, #DoYourResearch, and #LifeWithoutResearch. Over half (n=180) of the subjects watched the video via the campaign, 227/295 (76.9%) found the video entertaining, and 92/325 (28.3%) demonstrated baseline interest in clinical research. The ordinal logistic regression model showed that subjects who viewed the video from a peer (odds ratio [OR] 1.56, 95% CI 1.00-2.44, *P*=.05) or found the video entertaining (OR 1.36, 95% CI 1.01-1.82, *P*=.04) had greater odds of increasing interest in a clinical research career. Subjects with a higher baseline interest in medicine (OR 1.55, 95% CI 1.28-1.87, *P*<.001) also had greater odds of increasing their interest in clinical research.

Conclusions: The spread of authentic and relevant peer-created messages via social media can increase interest in clinical research careers among diverse teens. Peer-driven social media campaigns should be explored as a way to effectively recruit minority youth into scientific research careers.

(JMIR Med Educ 2020;6(1):e16392) doi:10.2196/16392

KEYWORDS

pipeline programs; minorities; underrepresented minorities; adolescents; teens; social media; peer led

Introduction

Despite significant shifts in the demographic landscape of the United States over the past 50 years, minorities remain disproportionally underrepresented in medical practice and research [1]. As researchers often rely on personal experience and background to formulate questions, inadequate representation of underrepresented minority leaders in clinical research is an impediment to the adequate study of health conditions relevant to minority groups [2]. Attempts to alleviate this inequity typically focus on pipeline programs that directly target youth and adolescents. However, few studies have explored how to initially reach and recruit teens with low baseline interest or knowledge about science, technology, engineering, and mathematics (STEM) careers.

Enhanced interest in research careers has been previously described as a strong predictor of the "attitude" component of the theory of aligned ambition [3]. Conceptualized by Barbara Schneider, aligned ambition is a framework used to predict an adolescent's future success in a desired career based upon their knowledge, behavior, and aforementioned attitude about said career path [4]. Because pipeline programs target recruitment at such an early stage, these programs influence all three aspects of aligned ambition, giving teens a more holistic picture of STEM fields. This type of pointed exposure aims to improve both the retention and persistence of minorities through scientific career paths [5].

In considering how to encourage interest in a scientific research career, the growing use of social media among adolescents offers new avenues toward influencing minority teens [6]. Both descriptive and experimental studies have supported targeted social marketing disseminated on the internet as an effective way to influence the health behaviors of teens. A 2014 systematic review found that 9 out of 10 online interventions reported significant improvements in some aspect of health behavior change [7]. From smoking cessation to health and sex education, peer-created social media campaigns are a powerful tool in engaging adolescents not only in the United States but also in international spheres [8-11]. However, no study has explored social media as a means to impact attitudes and behaviors regarding career interest.

Accordingly, the aim of this study was to engage high-achieving minority youth enrolled in a pipeline program with the Spreading Teen Research Inspired Videos to Engage Schoolmates (STRIVES) intervention. The purpose of STRIVES was to facilitate the creation of a video and peer-led social media campaign aimed at promoting interest in clinical research careers among their friends. We hypothesized that (1) a peer-shared video, compared to the same video shared by the investigative team, would be more effective in increasing interest in pursuing a career in research among peers of high-achieving minority youth enrolled in a pipeline program; and (2) videos that were perceived as entertaining would also be more effective in increasing interest in a clinical research career.

Methods

Study Design

Ethics

This study (IRB13-0848) was approved by the University of Chicago Biological Sciences Division Institutional Review Board (Chicago, IL, USA).

Setting

Training Early Achievers for Careers in Health (TEACH) is a pipeline initiative for rising high school juniors under the Collegiate Scholars Program, a partnership between the University of Chicago and Chicago Public Schools (CPS). The Collegiate Scholars Program is an intensive 3-year enrichment program designed to prepare talented high school students for academic success at the best colleges and universities. Starting in the summer after ninth grade, collegiate scholars select classes in literature, math, science, social sciences, and writing taught by University of Chicago faculty. Selection into this program is highly competitive, targeting students from ethnically and demographically underrepresented groups. Over 50% of collegiate scholars are underrepresented minorities (41% African American and 24% Latino/Hispanic), and 47.4% qualify to receive free or reduced lunch. Moreover, 41.56% of collegiate scholars will be first-generation college students [3].

Approximately 60 Collegiate Scholars Program students who expressed interest were selected as TEACH participants each year. These students were randomly assigned to one of two research groups: the clinical research group involving an immersive clinical research experience, or the field research group involving a more traditional basic science research program [3]. The 5-week clinical research summer program provided participants with several major experiences, including an introduction to the basics of clinical research through a classroom experience with faculty, an opportunity to work in a multitiered research team, and a hands-on research experience in a clinical environment performing patient interviews and physician surveys. The field research intervention consisted of science and nonscience lectures, a hands-on lab component led by PhD students, and field trips to local museums with high-quality science immersion.

Participant Recruitment

Each student enrolled in either the field or clinical research group recruited and provided contact information for approximately 10 "friends," defined as peers from their general community. Peers were recruited though emails and postcards. The student that ultimately recruited the most peers was given 50 CPS learning hours and an Amazon gift card. We obtained written informed consent from each friend and a parent/guardian before dissemination of baseline assessments.

Spreading Teen Research Inspired Video to Engage Schoolmates

The STRIVES intervention within the clinical research program engaged students with the creation of a short video and the subsequent social media campaign aimed at promoting clinical research as a viable career among peers. To accomplish this, students had weekly STRIVES meetings with project managers, medical faculty and students, technology experts, and research assistants to guide them through the campaign process.

Focus Groups

We coached each cohort of students in the clinical research group to conduct one or two focus groups with 5-6 peers from the Collegiate Scholars Program to identify themes related to careers in clinical research as highlighted by the Appreciative Inquiry 4D Model. This model encompasses (1) Discovery, or identifying the best way to achieve the goal; (2) Dream, or imagining new means of achieving the goal; (3) Design, or how to operationalize a change to reach the goal; and (4) Destiny, or anticipating the best practice [12]. Focus groups were conducted at University of Chicago, which were moderated by the teens and transcribed by research assistants. Deductive analysis was utilized to identify themes as they related to the 4D model.

Video Creation

Using themes that emerged in focus groups about the perception of research careers and how to best engage peers, students discussed, storyboarded, and scripted preliminary video ideas for the STRIVES campaign. The investigative team reviewed the scripts for relevance and accuracy. Technology experts presented lectures to the students on video recording and editing tactics. Students recorded the videos on the University of Chicago Pritzker School of Medicine campus. Video editing, with assistance from project managers, was completed using Apple Final Cut Pro X (Cupertino, CA, USA).

Social Media Campaign

The investigative team trained the students in viral marketing techniques based on the usage of already established social networks to spread information. We taught the Activation Theory of Information Exposure [13], a social media marketing concept that highlights that the most effective messages are not only informative but also captivating [14]. Using the 5M Model [15] consisting of mission, market, money, message, and methodology, the students organized a strategy aimed at maximizing the influence and effectiveness of their social media campaign. We had the students complete a worksheet that related to the 5M model to craft a social media campaign. For example, each individual campaign had a message that correlated to the content of the video.

Data Collection

Study subjects or "friends" completed an abbreviated online Career Orientation Survey following study enrollment to assess baseline knowledge, interest, and intent to pursue a variety of careers, including clinical research (see Multimedia Appendix 1). The Career Orientation Survey was originally developed and used in the Sloan Study of Youth and Social Development, a nationally representative longitudinal sample of American youth [16]. Adapted for the TEACH project, the screening instrument used in this study included standard and modified occupation survey items [3]. After the campaign, subjects once again completed the survey. Subjects who had not yet viewed the video through a peer-shared source by the time they received the postcampaign survey were given the opportunity to watch the video via the postsurvey. Of note, subjects who completed the postcampaign Career Orientation Survey in 2014 were not given the option to view the video via postsurvey; thus, all recorded responses were from subjects who viewed the video from a peer-disseminated source. We electronically sent friends who successfully submitted surveys an Amazon gift card and 5 CPS service learning hours. Study data were collected and managed using a Research Electronic Data Capture (REDCap) tool hosted at the University of Chicago [17].

The primary outcomes of interest were (1) perceptions of the video and (2) change in interest in pursuing a career in clinical research. Perceptions of the video were assessed by response on a 5-item Likert-type scale from "strongly disagree" to "strongly agree" in response to the statements "I found the video entertaining" and "I think this video is a good way to educate peers." Career interest was assessed using the question "how interested are you in pursuing a career in clinical research?" with 5 Likert-type responses ranging from "definitely not interested" to "definitely interested."

Data Analysis

Descriptive and comparative statistics were used to analyze subjects' demographic information and social media usage by year. An interest change score variable was calculated for each subject by taking the difference of Likert scores for researcher interest responses between the postcampaign and baseline Career Orientation Surveys.

An ordinal logistic regression was used to test the association between interest change score, a peer-shared vs survey-shared video, and perception of entertainment [18]. Model 1 controlled for (1) whether the subject was an underrepresented minority in medicine (ie, Black, Pacific Islander, Native American), (2) gender, (3) baseline interest in clinical research, (4) baseline interest in a medical career, and (5) whether the recruiting friend was in the clinical or field research group. Model 2 controlled for the same covariates listed above in addition to campaign year.

Results

Participant Characteristics

A total of 325 student-recruited subjects completed the pre and postcampaign Career Orientation Survey between 2014 and 2017. Of these subjects, 91.7% (n=298) watched the STRIVES video created by their respective year's cohort from any source before completing the postcampaign Career Orientation Survey. Demographic characteristics (Table 1) varied by year. Among all subjects, the majority were female and 23.1% (75/325) were underrepresented minorities in the medical field (Black, American Indian/Alaska Native, Native Hawaiian or Pacific Islander).

Perceptions About the Videos

A summary of STRIVES video topics by year, inspired by themes that arose in focus groups, are provided in Figure 1. Across all four cohorts, 55.4% (180/325) of the subjects watched a peer-shared video and 44.6% (145/325) watched a study-shared video. The majority of subjects from all cohorts had positive perceptions of their respective videos. For example, 76.9% (227/295) found the video entertaining and 81.1% (236/291) thought the video was a good way to educate peers. Among the subjects who watched a peer-shared video and completed the survey item, 80.9% (131/162) reported the video as entertaining, which was significantly higher (*P*=.04) than those that watched the video through the study (72.2%, 96/133). There was no difference in perception of whether videos were a good way to educate peers about careers based on the source of the video (study-shared 82% vs peer-shared 80%, *P*=.60).

Figure 1. Spreading Teen Research Inspired Videos to Engage Schoolmates (STRIVES) video topics and summaries by year [19-22], including the percentage of subjects who watched the video through peer-disseminated sources during the social media campaign rather than through a link on the post-campaign Career Orientation Survey and the percentage of subjects who responded “agree” or “strongly agree” to the Likert-style question “I found the video entertaining” in the postcampaign Career Orientation Survey. *P* values were generated through Chi square analysis in STATA software.

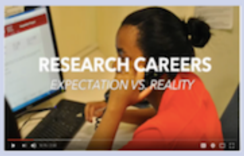
Year	Campaign hashtag and video screenshot	Summary of Video	Peer-Shared Video	Perceived as Entertaining
2014 (N=46)	#WhereScienceMeetsReality [19] 	• Inspired by Jimmy Kimmel’s “Lie Witness News” segment. • Interviewed teens on the street for thoughts on research relating to daily lives.	100%	84.2%
2015 (N=38)	#RedefiningResearch [20] 	• Inspired by Proctor & Gamble’s “Like a Girl” campaign. • Stressed the importance of “redefining” research to be more inclusive/diverse. • Accompanied by infographic and petition	52.6%	76.3%
2016 (N=147)	#DoYourResearch [21] 	• Inspired by Buzzfeed expectation vs reality videos • Aimed to myth bust teens’ perceptions of research careers and highlight the “social good” of a career.	67.4%	82.8%
2017 (N=94)	#LifeWithoutResearch [22] 	• Inspired by live animation video concepts. • Challenges teens to imagine a world without advanced scientific discoveries	28.7%	66.7%
Total (N=325)			55.4%	77.0%
P value			<.001	.03

Table 1. Demographic information of recruited friends by year.

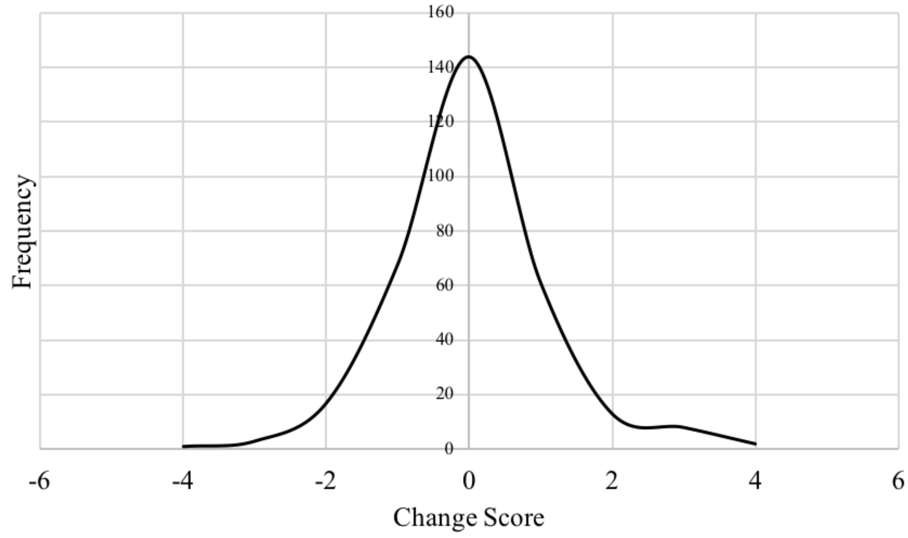
Variable	Total (N=325)	2014 (N=46)	2015 (N=38)	2016 (N=147)	2017 (N=94)	*P* value[a]
Gender, n (%)						.03
Male	100 (30.8)	12 (26.1)	10 (26.3)	52 (35.4)	26 (27.7)	
Female	214 (65.8)	33 (71.7)	28 (73.7)	85 (57.8)	68 (72.3)	
No Answer	11 (3.4)	1 (2.2)	0 (0)	10 (6.8)	0 (0)	
Race, n (%)						<.001
American Indian/Alaska Native	3 (1.0)	0 (0)	0 (0)	3 (2.1)	0 (0)	
Asian	78 (24.0)	15 (32.6)	10 (26.3)	24 (16.3)	29 (30.8)	
Native Hawaiian or Other Pacific Islander	3 (1.0)	3 (6.5)	0 (0)	0 (0)	0 (0)	
Black or African American	68 (20.9)	14 (30.4)	2 (5.3)	29 (19.7)	23 (24.5)	
White	77 (23.6)	10 (21.8)	14 (36.8)	36 (24.5)	17 (18.1)	
More than 1 race	29 (8.9)	0 (0)	2 (5.3)	16 (10.9)	11 (11.7)	
Unknown/Prefer Not to Answer	67 (20.6)	4 (8.7)	10 (26.3)	39 (26.5)	14 (14.9)	
Ethnicity, n (%)						.03
Hispanic	101 (31.1)	7 (15.2)	17 (44.7)	52 (35.4)	25 (26.6)	
Non-Hispanic	219 (67.4)	39 (84.8)	21 (55.3)	91 (61.9)	68 (72.3)	
No Answer	5 (1.5)	0 (0)	0 (0)	4 (2.7)	1 (1.1)	

[a]*P* values were generated through Chi square analysis in STATA (College Station, TX, USA) software.

Interest in Clinical Research

Before being exposed to the STRIVES campaign, 28.3% (92/325) of the subjects showed interest in a clinical research career compared to 25.8% (84/325) after the STRIVES campaign. Change scores were approximately normally distributed (Figure 2). Of the study subjects, 45.4% (144/317) did not change their interest in a clinical research career after watching the STRIVES video, 28.1% (89/317) expressed less interest, and 26.5% (84/317) showed greater interest. A significantly higher percentage of subjects who watched a peer-shared video showed postcampaign interest than those who watched a study-shared video (58/180, 32.2% and 30/145, 20.7%, respectively; *P*=.04).

Figure 2. Change in career interest over time (2014-2017). Career interest was assessed using the Likert-type question, "how interested are you in pursuing a career in clinical research?" with responses ranging from "definitely not interested" to "definitely interested." Positive interest was indicated by a response of "definitely interested" or "very interested." The change score variable was calculated by taking the difference of Likert scores for researcher interest responses between the postcampaign and baseline Career Orientation Surveys.



In the multivariate ordinal logistic regression (Table 2), watching a peer-shared video vs a study-shared video was significantly associated with increased odds of higher interest in a clinical research career. Those who found the video to be entertaining also had greater odds of increased interest in clinical research. In addition, subjects who had higher baseline interest in medicine were associated with greater odds of increased interest in clinical research, whereas higher baseline interest in clinical research was associated with lower odds of showing increased interest. Gender, underrepresented minority in medicine status, and whether the recruiting friend was in the clinical research program were not significantly associated with an increasing interest in a clinical research career.

In model 2 of the multivariate ordinal logistic regression, controlling for campaign year, those who found the video to be entertaining had greater odds of increased interest in clinical research, whereas there was no association based on whether the video was peer-shared or study-shared. Similar to model 1, subjects who had higher baseline interest in medicine were associated with greater odds of increased interest in clinical research, whereas a higher baseline interest in clinical research was associated with lower odds of showing increased interest after watching the video. Watching a peer-shared video, gender, campaign year, underrepresented minority in medicine status, and whether the recruiting friend was in the clinical research program were not significantly associated with a change in clinical research interest.

Table 2. Ordinal logistic regression testing the association between predictors and change in clinical research interest.

Variable (N=287)	Model 1			Model 2[a]		
	Odds ratio (SE)	95% CI	*P* value	Odds ratio (SE)	95% CI	*P* value
Peer-shared video[b]	1.56 (0.36)	1.00-2.44	.05	1.50 (0.36)	0.94-2.41	.09
Perceived video as entertaining	1.36 (0.20)	1.01-1.82	.04	1.35 (0.20)	1.00-1.82	.04
Baseline interest in clinical research	0.32 (0.04)	0.25-0.41	<.001	0.32 (0.04)	0.25-0.41	<.001
Baseline interest in medicine	1.55 (0.15)	1.28-1.87	<.001	1.55 (0.15)	1.29-1.87	<.001
Recruited by clinical research student[c]	1.08 (0.25)	0.69-1.69	.73	1.09 (0.25)	0.70-1.71	.70
Underrepresented minority in medicine[d]	0.88 (0.25)	0.50-1.53	.65	0.89 (0.25)	0.51-1.55	.67
Male gender	0.95 (0.23)	0.59-1.52	.82	0.95 (0.23)	0.59-1.52	.82

[a]Model 2 controlled for campaign year in addition the listed variables.

[b]Subjects watched the video from a peer-shared source vs a survey-shared source (ie, through a link provided on the postcampaign Career Orientation Survey).

[c]Subjects were recruited by friends participating in the clinical research program vs friends participating in the field research program.

[d]Underrepresented minority in medicine characterized as Hispanic, Black, Pacific Islander, or Native American.

Discussion

The results of this study suggest that peer-led social media campaigns paired with short, entertaining videos are an effective way to increase interest in clinical research careers among peers. Specifically, we found that videos shared via peers on social media were more effective in improving interest in clinical research compared to the same video watched through a link provided by the investigative team. Videos that subjects perceived as entertaining were also more likely to be associated with an increase in interest in clinical research careers among subjects.

Approximately half (55.4%, 180/325) of our study subjects viewed a STRIVES video through an online social media source as shared by a peer in the TEACH program. It is likely that these subjects who came across the videos on social media were already well-embedded within the social network of the students who not only created the videos but played an acting role within them. This results in a sense of perceived familiarity with the content, which in turn allowed the subjects to be more positively influenced by the videos [6]. These findings are consistent with other works showing how peer-to-peer interactions among adolescents are key to influencing teen cognition and behavior [23]. In fact, this strategy is often evoked by corporate marketing campaigns in the age of digital media, particularly when utilizing celebrity sponsorships or social media influencers with whom the targeted audiences feel more connected to [24]. Information received from peers is deemed to be not only credible but also relatable, particularly when paired with the interactivity of social media [25]. This influential effect of peer communication on adolescents has been shown to be even stronger among both female and underrepresented minority teens, thereby strengthening the efficacy of social media campaigns in engaging minority teens [26].

Subjects who watched a peer-shared video were significantly more likely to find the video entertaining and were also more likely to report increased interest in clinical research. This correlation can be explained within the context of entertainment value being crucial to how messages are perceived and processed. Content that is deemed to be amusing becomes more compelling to the viewer, resulting in increased internalization of the message and more engagement with the post [27]. Social media algorithms function in such a way that posts with the most interactions become more visible to those outside of the intended network [28]. This increase in visibility then gives way to a larger spread of the post and eventually its influence. In essence, entertaining videos are more likely to be shared, subsequently increasing the visibility and engagement, leading to a larger effect size.

Future considerations for interventions such as STRIVES require exploration of how these efforts translate into pipeline program enrollment. Campaigns and other social marketing endeavors act to change attitudes (ie, interest in scientific research careers), whereas more immersive counterparts such as pipeline programs are needed to influence the action of pursuing clinical research as a career. Essentially, social media campaigns are an important step toward "priming the pump." Forthcoming efforts to recruit teens to clinical research pipeline programs should leverage these findings in an attempt to bolster the number of interested applicants. Similar efforts are worth testing among college students and medical students given concerns of the physician-scientist pipeline.

The primary limitation of this study was the inability to randomize subjects into peer-shared vs study-shared video groups. This led to significant variability in the number of subjects exposed to peer-shared videos across the campaign years. For example, in 2014, 100% of subjects who completed the postsurvey viewed the STRIVES campaign from a peer-shared source. This inherently conflated associations between the two groups as there was no comparison available for peer-shared viewers that year. This likely attributed to the video source no longer being a significant predictor of changing interest in research when controlling for campaign year as seen in model 2 of the ordinal logistic regression.

Overall, this study demonstrated that peer-shared videos that are perceived as entertaining are significantly associated with increasing interest in a clinical research career among peers of high-achieving minority youth in a scientific research pipeline program. Authentic and relevant peer-driven messages have the potential to engage and activate minority youth, thus "priming the pump" into clinical research pipeline programs. The ultimate hope is that early exposure will translate into fruitful careers that will help diversify the scientific research workforce.

Acknowledgments

This project was funded by the National Institute of General Medical Sciences (grant R01GM107721).

Conflicts of Interest

None declared.

Multimedia Appendix 1

The Career Orientation Survey (COS) was originally developed and used in the Sloan Study of Youth and Social Development, a nationally representative longitudinal sample of American youth. This survey was sent to study subjects at baseline and again after the social media campaign to gain insight on social media usage and career alignment.

[PDF File (Adobe PDF File), 56 KB - mededu_v6i1e16392_app1.pdf]

References

1. Laurencin CT, Murray M. An American Crisis: the Lack of Black Men in Medicine. J Racial Ethn Health Disparities 2017 Jun 22;4(3):317-321 [FREE Full text] [doi: 10.1007/s40615-017-0380-y] [Medline: 28534304]
2. Benabentos R, Ray P, Kumar D. Addressing health disparities in the undergraduate curriculum: an approach to develop a knowledgeable biomedical workforce. CBE Life Sci Educ 2014 Dec;13(4):636-640 [FREE Full text] [doi: 10.1187/cbe.14-06-0101] [Medline: 25452486]
3. Arora V, Schneider B, Thal R, Meltzer D. Design of an Intervention to Promote Entry of Minority Youth into Clinical Research Careers by Aligning Ambition: The TEACH (Training Early Achievers for Careers in Health) Research Program. J Social Issues 2011;67(3):580-598. [doi: 10.1111/j.1540-4560.2011.01715.x]
4. Schneider B, Stevenson D. Why Aligned Ambitions Matter. In: The Ambitious Generation: America's Teenagers, Motivated But Directionless. New Haven and London: Yale University Press; 1999:263-264.
5. Flowers SK, Beyer KM, Pérez M, Jeffe DB. Early Environmental Field Research Career Exploration: An Analysis of Impacts on Precollege Apprentices. CBE Life Sci Educ 2016 Dec;15(4):ar67 [FREE Full text] [doi: 10.1187/cbe.15-11-0230] [Medline: 27909017]
6. Chan A, Kow R, Cheng JK. Adolescents' Perceptions on Smartphone Applications (Apps) for Health Management. JournalMTM 2017 Aug;6(2):47-55. [doi: 10.7309/jmtm.6.2.6]
7. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb;16(2) [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
8. Bottorff JL, Struik LL, Bissell LJ, Graham R, Stevens J, Richardson CG. A social media approach to inform youth about breast cancer and smoking: an exploratory descriptive study. Collegian 2014 Jun;21(2):159-168. [doi: 10.1016/j.colegn.2014.04.002] [Medline: 25109215]
9. Rogers EA, Fine S, Handley MA, Davis H, Kass J, Schillinger D. Development and early implementation of the bigger picture, a youth-targeted public health literacy campaign to prevent type 2 diabetes. J Health Commun 2014 Oct 14;19(Suppl 2):144-160 [FREE Full text] [doi: 10.1080/10810730.2014.940476] [Medline: 25315590]
10. Sun WH, Wong CKH, Wong WCW. A Peer-Led, Social Media-Delivered, Safer Sex Intervention for Chinese College Students: Randomized Controlled Trial. J Med Internet Res 2017 Dec 09;19(8):e284 [FREE Full text] [doi: 10.2196/jmir.7403] [Medline: 28793980]

11. Medich M, Swendeman DT, Comulada WS, Kao UH, Myers JJ, Brooks RA, Special Projects of National Significance Social Media Initiative Study Group. Authorship Correction: Promising Approaches for Engaging Youth and Young Adults Living with HIV in HIV Primary Care Using Social Media and Mobile Technology Interventions: Protocol for the SPNS Social Media Initiative. JMIR Res Protoc 2019 Sep 09;8(8):e15660 [FREE Full text] [doi: 10.2196/15660] [Medline: 31502543]
12. Michael S. The promise of appreciative inquiry as an interview tool for field research. Develop Pract 2005 Apr;15(2):222-230. [doi: 10.1080/09614520500042094]
13. Donohew L, Lorch EP, Palmgreen P. Applications of a theoretic model of information exposure to health interventions. Hum Commun Res 1998 Mar;24(3):454-468. [doi: 10.1111/j.1468-2958.1998.tb00425.x] [Medline: 12293438]
14. Helme DW, Donohew RL, Baier M, Zittleman L. A classroom-administered simulation of a television campaign on adolescent smoking: testing an activation model of information exposure. J Health Commun 2007 Jun 14;12(4):399-415. [doi: 10.1080/10810730701325947] [Medline: 17558790]
15. Reeves C. gaebler.com Resources for Entrepreneurs. 2012 Oct 07. Marketing Mix and the Five Ms of Marketing - Small Business Marketing URL: http://www.gaebler.com/Marketing-Mix-and-the-Five-Ms-of-Marketing.htm [accessed 2018-07-03]
16. Schneider B. Inspiring Youth to Careers in Science and Medicine. J Public Health Manag Pract 2009;15(Supplement):S102-S106. [doi: 10.1097/phh.0b013e3181be0031]
17. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
18. Wolfson RK, Alberson K, McGinty M, Schwanz K, Dickins K, Arora VM. The Impact of a Scholarly Concentration Program on Student Interest in Career-Long Research. Acad Med 2017;92(8):1196-1203. [doi: 10.1097/acm.0000000000001486]
19. YouTube. Teens on Research URL: https://youtu.be/BS_ywSGa9c4 [accessed 2020-04-09]
20. YouTube. Teens Redefine Research URL: https://youtu.be/jE0Iu6uwxsA [accessed 2020-04-09]
21. YouTube. Research Careers: Expectation vs. Reality URL: https://youtu.be/sgdqjMIaiyw [accessed 2020-04-09]
22. YouTube. Life Without Research: STRIVES 2017 URL: https://youtu.be/uxCNic-ZJpw [accessed 2020-04-09]
23. Yonker LM, Zan S, Scirica CV, Jethwani K, Kinane TB. "Friending" teens: systematic review of social media in adolescent and young adult health care. J Med Internet Res 2015 Jan 05;17(1):e4 [FREE Full text] [doi: 10.2196/jmir.3692] [Medline: 25560751]
24. Pilgrim K, Bohnet-Joschko S. Selling health and happiness how influencers communicate on Instagram about dieting and exercise: mixed methods research. BMC Public Health 2019 Aug 06;19(1):1054 [FREE Full text] [doi: 10.1186/s12889-019-7387-8] [Medline: 31387563]
25. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
26. Preszler RW. Replacing lecture with peer-led workshops improves student learning. CBE Life Sci Educ 2009 Sep;8(3):182-192 [FREE Full text] [doi: 10.1187/cbe.09-01-0002] [Medline: 19723813]
27. Evans WD. Social marketing campaigns and children's media use. Future Child 2008;18(1):181-203. [doi: 10.1353/foc.0.0009] [Medline: 21338010]
28. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]

Abbreviations

CPS: Chicago Public Schools
REDCap: Research Electronic Data Capture
STEM: science, technology, engineering, and mathematics
STRIVES: Spreading Teen Research Inspired Videos to Engage Schoolmates
TEACH: Training Early Achievers for Careers in Health

Edited by M Focsa; submitted 27.09.19; peer-reviewed by K Aguirre, E Da Silva; comments to author 27.11.19; revised version received 10.01.20; accepted 14.02.20; published 14.05.20.

Please cite as:
Ogunleye C, Farnan JM, Martin SK, Tanksley A, Ngooi S, Venable LR, Anderson S, Marte J, Meltzer DO, Arora VM
A Peer-Led Social Media Intervention to Improve Interest in Research Careers Among Urban Youth: Mixed Methods Study
JMIR Med Educ 2020;6(1):e16392
URL: http://mededu.jmir.org/2020/1/e16392/
doi:10.2196/16392
PMID:32406859

Original Paper

Translating Clinical Questions by Physicians Into Searchable Queries: Analytical Survey Study

Aurélie Seguin[1], MD; Robert Brian Haynes[2], MD, PhD; Sebastian Carballo[1], MD, PhD; Alfonso Iorio[2], MD, PhD; Arnaud Perrier[1], MD; Thomas Agoritsas[1,2], MD, PhD

[1]Division of General Internal Medicine, Department Medicine, University Hospitals of Geneva, Geneva, Switzerland

[2]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Thomas Agoritsas, MD, PhD
Division of General Internal Medicine
Department Medicine
University Hospitals of Geneva
Rue Gabrielle-Perret-Gentil 4
Geneva, 1205
Switzerland
Phone: 41 79 55 34 543
Fax: 41 22 372 90 35
Email: thomas.agoritsas@unige.ch

Abstract

Background: Staying up to date and answering clinical questions with current best evidence from health research is challenging. Evidence-based clinical texts, databases, and tools can help, but clinicians first need to translate their clinical questions into searchable queries. MacPLUS FS (McMaster Premium LiteratUre Service Federated Search) is an online search engine that allows clinicians to explore multiple resources simultaneously and retrieves one single output that includes the following: (1) evidence from summaries (eg, UpToDate and DynaMed), (2) preappraised research (eg, EvidenceAlerts), and (3) non-preappraised research (eg, PubMed), with and without validated bibliographic search filters. MacPLUS FS can also be used as a laboratory to explore clinical questions and evidence retrieval.

Objective: Our primary objective was to examine how clinicians formulate their queries on a federated search engine, according to the population, intervention, comparison, and outcome (PICO) framework. Our secondary objective was to assess which resources were accessed by clinicians to answer their questions.

Methods: We performed an analytical survey among 908 clinicians who used MacPLUS FS in the context of a randomized controlled trial on search retrieval. Recording account log-ins and usage, we extracted all 1085 queries performed during a 6-month period and classified each search term according to the PICO framework. We further categorized queries into background (eg, "What is porphyria?") and foreground questions (eg, "Does treatment A work better than B?"). We then analyzed the type of resources that clinicians accessed.

Results: There were 695 structured queries, after exclusion of meaningless queries and iterations of similar searches. We classified 56.5% (393/695) of these queries as background questions and 43.5% (302/695) as foreground questions, the majority of which were related to questions about therapy (213/695, 30.6%), followed by diagnosis (48/695, 6.9%), etiology (24/695, 3.5%), and prognosis (17/695, 2.5%). This distribution did not significantly differ between postgraduate residents and medical faculty physicians (P=.51). Queries included a median of 3 search terms (IQR 2-4), most often related to the population and intervention or test, rarely related to the outcome, and never related to the comparator. About half of the resources accessed (314/610, 51.5%) were summaries, 24.4% (149/610) were preappraised research, and 24.1% were (147/610) non-preappraised research.

Conclusions: Our results, from a large sample of real-life queries, could guide the development of educational interventions to improve clinicians' retrieval skills, as well as inform the design of more useful evidence-based resources for clinical practice.

Trial Registration: ClinicalTrials.gov NCT02038439; https://www.clinicaltrials.gov/ct2/show/NCT02038439

(JMIR Med Educ 2020;6(1):e16777) doi:10.2196/16777

KEYWORDS

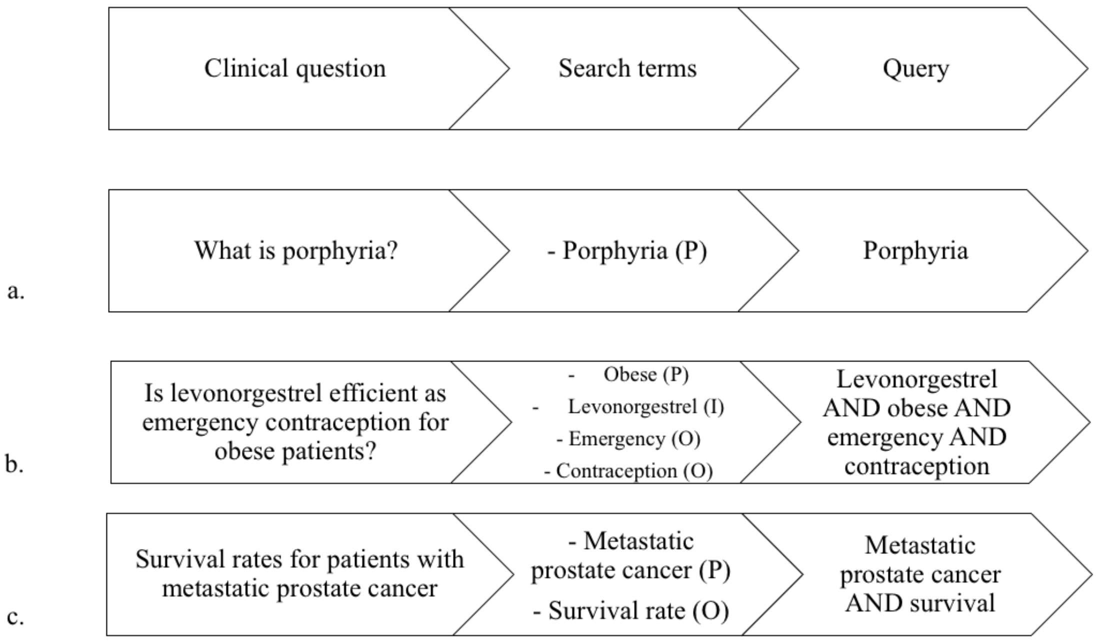
evidence-based medicine; evidence retrieval; Web-based resources; search engines; search taxonomy; clinical information science

Introduction

Web-based searches have become the norm when looking for information and answers to most of our questions in daily life. This has also become true in the practice of medicine; online medical resources to access evidence are increasingly considered "as essential as the stethoscope" [1]. While famous search engines, such as Google, or information sources, such as Wikipedia, are used in both medical and nonmedical worlds, answering clinical questions to inform point-of-care decisions has additional challenges and implications [2]. Triggered by more than 20 years of evidence-based medicine (EBM) [3,4], the unit of information in medicine comes mostly in the form of research evidence, published across thousands of medical journals and indexed in numerous databases (eg, MEDLINE, Embase, and the Cumulative Index to Nursing and Allied Health Literature [CINAHL]). The volume of this new evidence through all these channels is rapidly increasing at the pace of 3000-4000 new publications per day, compiled or processed in hundreds of EBM summaries and resources [5-7].

Physicians are typically familiar with only a few of these resources, likely those to which they have been exposed in training or by peers, and often ignore most of the ecosystem and architecture of published evidence. Yet, their daily practice triggers, on average, five to eight questions every 10 patients [8-10]. Clinical questions can be classified as background and foreground questions (see Figure 1). Background questions (eg, "What is porphyria?") are typically about the nature of a disorder, a measure, a treatment, or a test. They are easily answered through online textbooks. Foreground questions are more directly related to the diagnosis, prognosis, and treatment of a given patient population (eg, "How effective would levonorgestrel be as emergency contraception for an obese patient?") [11]. The teaching of EBM recommends that foreground questions be formulated according to the population, intervention, comparison, and outcome (PICO) framework, or the population, exposure, comparison, and outcome (PECO) framework, and answered by research evidence [12].

Figure 1. The path from a clinical question to a query using the population (P), intervention (I), comparison (C), and outcome (O) (PICO) framework. Examples are shown for (a) a background question, (b) a foreground therapy question, and (c) a foreground prognosis question.

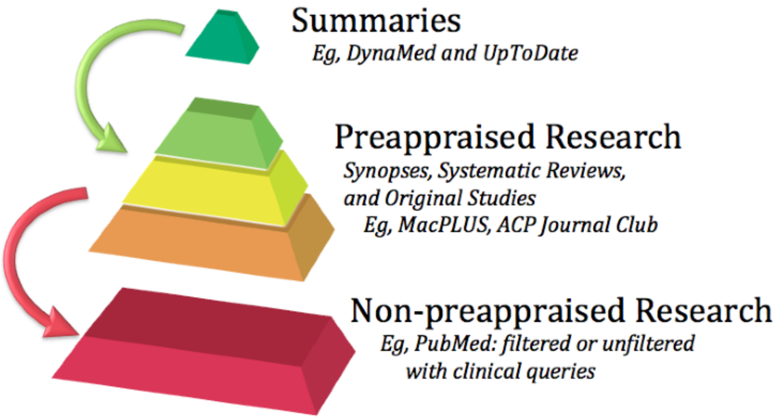


How physicians translate their clinical questions into searchable queries remains poorly known. How many search terms do they use? How often do their queries fit the PICO framework [12,13]? Do experienced and fully trained clinicians differ from residents in training? Do queries differ according to the medical specialty? We aimed to examine these questions in a large sample of practicing clinicians of various levels of training and specialty type.

The type of search engine or evidence resource may also influence the way we conduct queries. Google and Wikipedia tend to retrieve relevant answers, albeit selective, with intuitive, less-structured search strategies [14-16]. Some EBM online textbooks and evidence summaries may provide a similar user experience to clinicians. By contrast, searching PubMed or other databases requires more training and structure, is less intuitive, and tends to produce large and diluted outputs for similar clinical questions [12].

We, therefore, explored how clinicians formulate their queries in a federated online search engine, namely MacPLUS FS (McMaster Premium LiteratUre Service Federated Search). MacPLUS FS allows clinicians to explore multiple resources simultaneously, retrieving one single output that includes the following: (1) evidence from evidence-based summaries (eg, UpToDate and DynaMed), (2) preappraised research (eg, EvidenceAlerts), and (3) non-preappraised research (eg, PubMed), with and without validated search filters (see Figure 2). In this study, we will outline how we used MacPLUS FS, which functions as a laboratory, to explore clinical questions, the taxonomy of queries, and evidence retrieval (ie, what resources clinicians access to answer their questions when provided with a wide array of EBM resources) (see Multimedia Appendix 1) [5]. While MacPLUS FS functions as a laboratory for evidence retrieval research, its exact twin—ACCESSSS search engine—is freely available online [17].

Figure 2. Synopses, systematic reviews, and select studies of evidence-based medicine resources provided in the federated search engine MacPLUS FS (McMaster Premium LiteratUre Service Federated Search); adapted from Agoritsas et al, 2014. ACP: American College of Physicians.



Methods

Study Design and Clinician Sample

We conducted an analytical survey of clinical search queries among 431 postgraduate medical trainees and 477 medical faculty members registered to a federated search engine, MacPLUS FS. The service was freely available to registered users from any computer with an internet browser throughout the clinical setting or elsewhere.

Participating clinicians consented to be enrolled in 6-month, MacPLUS FS, randomized controlled trials [5], which tested three interventions to enhance the quantity and quality of searching for current best evidence in order to answer clinical questions in a factorial design. As described with more detail in the published protocol of the trials [5], we tested the following three interventions embedded in MacPLUS FS: (1) a Web-based clinical question recorder, (2) an evidence retrieval coach composed of eight short educational videos, and (3) an audit, feedback, and gamification approach to evidence retrieval, based on the allocation of *badges* and *reputation scores*. Participating clinicians were randomized to each of the three interventions in a factorial design (A × B × C).

For each clinician, utilization of MacPLUS FS was recorded through accounts tracking log-ins and usage, including their detailed search queries. Registration to the service was free, and access to each evidence resource was through clinicians' academic institutions, mostly McMaster University, Hamilton, Canada. Clinicians were categorized according to their baseline search levels and specialty types [5].

Sample of Search Queries

We extracted all 1085 search queries performed by clinicians during the conduct of the MacPLUS FS trials. Two authors (AS and TA) assessed each query individually, counting the number of search terms—counting all words (eg, the query "porphyria" contains 1 term)—and documenting all abbreviations and Boolean terms (ie, logical operators such as "AND," "OR," or "NOT"). Search queries were then classified into (1) structured searches, (2) searches for specific articles (eg, when clinicians typed in the title of a given study), (3) iteration of structured searches, namely a group of related structured queries with a similar PICO question within the same log-in session, and (4) undetermined searches (eg, "Scimitar").

Assessment of Search Queries and Evidence Resources Access

The same two authors (AS and TA) classified structured queries into background or foreground questions (see Figure 1), according to the PICO framework, and blinded the participants' characteristics, except the log-in session. Queries that included only terms related to population or intervention were classified as background questions. Those including several terms related to population and intervention and/or outcome and/or comparator were categorized as foreground, and further categorized into therapy, diagnosis, etiology, and prognosis. For each query, we examined the distribution of access to each evidence resource from the federated search: *summaries*, *preappraised research*, and *non-preappraised research* (see Figure 2).

Statistical Analysis

We examined types of questions (ie, background, foreground, and type of foreground) according to the level of training as well as clinicians' specialties and baseline frequencies of search (ie, in the prior months since their registration to MacPLUS FS). We then examined the number and type of search terms across each type of question. We compared distributions using chi-square parametric tests when relevant and Kruskall-Wallis tests for nonnormal distributions. Data abstraction was done using Microsoft Excel 2016, version 15.29, and data analysis was performed using SPSS Statistics for Windows, version 23.0 (IBM Corp).

Results

Clinicians

Participants were postgraduate residents and medical faculty members who had registered in MacPLUS FS prior to the trial. Of the 678 postgraduate residents and 753 medical faculty members, 431 (63.6%) and 477 (63.3%), respectively, were deemed eligible after the exclusion of 247 postgraduate residents and 266 medical faculty members, who either never logged in to MacPLUS FS during the year prior to the study or quit the institutions served by MacPLUS FS [5]. Searchers were further classified, depending on their baseline average search frequencies during the 6 months prior to the trial [5], as *regular searchers* (≥1 search per month), *occasional searchers* (<1 search per month), or *alert-only users* (no searches).

From Clinicians to Queries

The 908 clinicians made 1085 search queries, of which 235 (21.66%) were subsequent iterations of the same search, 124 (11.43%) were a search for a specific article, and 31 (2.86%) could not be classified and remained undetermined. A total of 695 out of 1085 queries (64.06%) were structured queries following the PICO format, with 480 out of 695 (69.1%) single queries, whereas 215 (30.9%) included a group of related queries. This corresponds to an average of 2.1 attempts per group query.

Table 1 summarizes the distributions of the 695 structured queries. We classified 56.5% (393/695) as background and 43.5% (302/695) as foreground questions, the majority of which were related to therapy (213/695, 30.6%), followed by diagnosis (48/695, 6.9%), etiology (24/695, 3.5%), and prognosis (17/695, 2.4%). Distributions did not differ according to level of training (*P*=.51) (see Table 1).

Table 1. Type of structured queries according to level of training.

Query type	Level of training, n (%)		
	Postgraduate residents (n=409)	Medical faculty members (n=286)	Total (n=695)[a]
Background	239 (58.4)	154 (53.8)	393 (56.5)
Foreground	170 (41.6)	132 (46.2)	302 (43.5)
Therapy	112 (27.4)	101 (35.3)	213 (30.6)
Diagnosis	34 (8.3)	14 (4.9)	48 (6.9)
Etiology	15 (3.7)	9 (3.1)	24 (3.5)
Prognosis	9 (2.2)	8 (2.8)	17 (2.4)
Total	409/695 (58.8)	286/695 (41.2)	695 (100)

[a]There were 695 structured queries among 1085 queries, the remaining being 235 iterations of the same search, 124 specific article searches, and 31 undetermined searches.

Table 2 shows the distributions of queries related to background and foreground clinical questions, with respect to the clinicians' levels of training, specialty types (ie, family medicine, internal medicine, internal medicine specialties, pediatrics, psychiatry, surgery, anesthesiology, and others detailed in Multimedia Appendix 2), and categories of search frequency. Internal and family medicine physicians made 48.5% (337/695) of structured queries, 55.2% (186/337) of which were related to background content (see Table 2). However, there were differences regarding the frequencies of searches with regular searchers looking for significantly more background questions (*P*=.009). There were no differences between specialty types (*P*=.67).

Table 2. Background versus foreground queries with respect to characteristics of clinicians.

Characteristic	Question type, n (%)		
	Background	Foreground	Total
Training			
Postgraduate residents	239 (58.4)	170 (41.6)	409 (100)
Medical faculty members	154 (53.8)	132 (46.2)	286 (100)
Specialty type			
Family medicine	114 (54.0)	97 (46.0)	211 (100)
Internal medicine	72 (57.1)	54 (42.9)	126 (100)
Other specialties[a]	207 (57.8)	151 (42.2)	358 (100)
Categories of search frequency			
≥1 (regular searchers)	164 (62.8)	97 (37.2)	261 (100)
<1 (occasional searchers)	88 (51.8)	82 (48.2)	170 (100)
0 (alert-only users)	141 (53.4)	123 (46.6)	264 (100)
Total	393 (56.5)	302 (43.5)	695 (100)

[a]Other specialties includes internal medicine specialties, pediatrics, psychiatry, surgery, anesthesiology, and others detailed in Multimedia Appendix 2.

Table 3 details the components of the queries. Queries included a median number of 3 search terms (IQR 2-4). There were significantly more terms with foreground questions compared to background questions (*P*<.001). Indeed, there were 70.2% (276/393) of background questions with 2 or fewer terms versus 18.2% (55/302) of the foreground questions, and 81.8% (247/302) of foreground questions with 3 or more terms versus 29.8% (117/393) of the background questions.

Overall, 72.5% (504/695) of structured queries (see Table 3) contained at least 1 term related to population, and 45.9% (319/695) contained at least 1 term related to an intervention. Few queries contained terms about etiology, diagnostic tests, or outcome. No query included the comparator. Background queries included a median of 2 search terms (IQR 1-3). Of these queries, 71.2% (280/393) included a population term, 24.7% (97/393) included an intervention term, 1.0% (4/393) included an etiology term, 6.1% (24/393) included a diagnostic term, and 2.5% (10/393) included an outcome term. Foreground queries included a median of 4 search terms (IQR 3-5). Of these queries, 74.2% (224/302) included a population term, 73.5% (222/302) included an intervention term, 21.5% (65/302) included an outcome term, 16.2% (49/302) included a diagnostic term, and 7.6% (23/302) included an etiology term. Clinicians made no use of explicit Boolean search terms to link various PICO elements.

Table 3. Number of terms with respect to type of structured queries.

Type of query	Number of terms, median (IQR)	Distribution of terms[a] within each type of query, n (%)				
		≥1 population term	≥1 intervention term	≥1 etiology term	≥1 diagnostic term	≥1 outcome term
Background (n=393)	2 (1-3)	280 (71.2)	97 (24.7)	4 (1.0)	24 (6.1)	10 (2.5)
Foreground						
All foreground (n=302)	4 (3-5)	224 (74.2)	222 (73.5)	23 (7.6)	49 (16.2)	65 (21.5)
Intervention (n=213)	3 (3-5)	173 (81.2)	210 (98.6)	1 (0.5)	2 (0.9)	43 (20.2)
Diagnostic (n=48)	4 (2-5)	24 (50)	6 (13)	0 (0)	46 (96)	0 (0)
Etiology (n=24)	4 (3-5)	12 (50)	2 (8)	22 (92)	0 (0)	17 (71)
Prognostic (n=17)	5 (4-5)	15 (88)	4 (24)	0 (0)	1 (6)	5 (29)
Total structured queries (n=695)	3 (2-4)	504 (72.5)	319 (45.9)	27 (3.9)	73 (10.5)	75 (10.8)

[a]The distribution of terms is significantly different between background and foreground queries (*P*<.001). No query included the comparator search term in either type of query, so we did not include a column for the comparator term.

The number of evidence-based resources that clinicians accessed for each type of query (ie, by clicking the available links in the search output) are displayed in Table 4. The distribution of accessed resources is significantly different across categories (*P*<.001). Although 35.7% (248/695) of structured queries did not result in any resource access, 39.9% (277/695) led to one resource accessed, 11.8% (82/695) led to two, and 12.7% (88/695) led to three or more resources. Across all 1085 queries, the average number of resources accessed was 0.88 (SD 1.42). When users attempted a second search on the same clinical question (ie, similar PICO concepts but revised search terms), 7.2% (17/235) resulted in one or more resources accessed, while 92.8% (218/235) led to an end of their search query with no additional resource accessed. When searching for a specific article, 37.9% (47/124) led to one resource accessed and 12.9% (16/124) led to two or more resources accessed.

Table 4. Number of accessed sites across 1085 queries.

Accessed site	Query type, n (%)				
	Structured search (n=695)	Iteration of a structured search[a] (n=235)	Specific article search (n=124)	Undetermined search (n=31)	Total (N=1085)
0	248 (35.7)	218 (92.8)	61 (49.2)	22 (71)	549 (50.60)
1	277 (39.9)	12 (5.1)	47 (37.9)	5 (16)	341 (31.43)
2	82 (11.8)	3 (1.3)	13 (10.5)	1 (3)	99 (9.12)
3	45 (6.5)	1 (0.4)	1 (0.8)	2 (6)	49 (4.52)
4	15 (2.2)	1 (0.4)	1 (0.8)	1 (3)	18 (1.66)
≥5	28 (4.0)	0 (0)	1 (0.8)	0 (0)	29 (2.67)
Total	695 (100)	235 (100)	124 (100)	31 (100)	1085 (100)

[a]A group of related structured queries with similar population, intervention, comparison, and outcome (PICO) concepts, but revised search terms within the same log-in session.

Table 5 shows types of accessed resources with respect to level of training, type of query, and specialty. Across the 695 structured queries, there were 610 accessed resources, with half of them (314/610, 51.5%) being summaries, 24.4% (149/610) being preappraised research, and 24.1% (147/610) being non-preappraised research. When comparing the distribution of resources that were accessed across the federated search output, medical faculty members looked at significantly more summaries than did postgraduate trainees (P<.001), and family physicians looked at significantly more resources than did internists and specialized physicians (P<.001).

Table 5. Resources accessed across structured queries that led to at least one evidence resource.

Training and specialty	Resources accessed[a], n (%)				*P* value
	Summaries	Preappraised research	Non-preappraised research	Total	
Training					**<.001**
Postgraduate residents	150 (43.5)	85 (24.6)	110 (31.9)	345 (100)	
Medical faculty members	164 (61.9)	64 (24.2)	37 (14.0)	265 (100)	
Specialty					**.001**
Family medicine	103 (64.0)	31 (19.3)	27 (16.8)	161 (100)	
Internal medicine	60 (46.2)	41 (31.5)	29 (22.3)	130 (100)	
Other specialty	151 (47.3)	77 (24.1)	91 (28.5)	319 (100)	
Total	314 (51.5)	149 (24.4)	147 (24.1)	610 (100)	

[a]A total of 610 resources were accessed by 695 structured queries. Access was recorded through the number of links accessed from MacPLUS (McMaster Premium LiteratUre Service) search output.

Discussion

Principal Findings

Among 1085 queries made by 908 clinicians, 695 were structured queries. A small majority were related to background questions, and most foreground questions were questions about therapy, rather than diagnostic or prognostic questions. Structured queries included a median of 3 terms, most often related to the population and intervention or test, rarely related to the outcome, and never related to the comparator. Explicit Boolean terms were rarely used; of note, the search engine assumed by default a Boolean "AND" between search terms. About half of the resources accessed were summaries, while the rest were equally divided between preappraised and non-preappraised resources.

We found no difference between searches made by postgraduate resident trainees and medical faculty members. As they are in training, one could have expected postgraduate residents to have more background questions, whereas faculty members were expected to have more foreground questions, for example, in comparing the effectiveness or risks of management strategies.

Our results did not confirm this assumption, as faculty members had more than half of their searches on background questions as well. This may be due to the complexity of patient care. A given faculty member may be an expert in a given field but adopt a learning strategy to rapidly get the big picture, to understand uncommon situations. Their high level of access for summary resources, such as UpToDate or DynaMed, likely supports this explanation. Similarly, family doctors also accessed more summary resources, not only because of their need for quick clear answers to questions arising within short appointments with patients, but also, perhaps, because they provide care for patients across an entire age spectrum.

Another issue relates to the frequency of searches clinicians are able to perform in daily life. In our study, 908 clinicians performed only 1085 queries in 6 months. Other studies have shown that clinicians tend not to search in order to answer the questions that arise in their daily clinical practice [10,18-20]. In our study, a third of the structured searches led to no resource access through the platform, for which we have no explanation. More than 20 years ago, Ely et al [19,20] already showed that clinicians spend less than 2 minutes looking to answer a question—a finding probably even more accurate nowadays with increased access to information online—and suggested that searching for evidence may not fit with clinicians' multiple tasks and training [21]. It is also possible that clinicians have looked for answers in other resources (eg, PubMed or UpToDate), or even in Google, Google Scholar, or Wikipedia. Alternatively, clinicians may often not conduct searches online but, rather, directly ask their peers or use local guidelines [22-26]. Reasons include convenience and time constraints to access ready-to-care information that conforms with local knowledge rather than challenging it. Although looking for answers on a general search engine or via colleagues or guidelines is easier, it does not guarantee or promote a fully EBM approach to health care [27,28]. Clinicians could, therefore, benefit from information specialists available to help at the point of care [29] and from the design of more intuitive tools to navigate the complexity of the evidence ecosystem.

Another observation from our study is that clinicians' queries tend to remain relatively simple: few search terms, often covering few PICO concepts, mostly population and intervention. While simple strategies work well for high-level summaries, they are much less efficient with large databases like PubMed. Our daily habits for searching on the Web may explain clinicians' tendencies for simple queries. Strictly from a user's perspective, we have all become very efficient in searching for information mostly through Google and Wikipedia, just by typing a few intuitive keywords in the free-text bar at the top of a webpage. Medical search engines may misguide the user in having them assume the engine will work similarly to Google [30].

One area for improvement of search engines could be to invite users to structure their queries according to the PICO framework. Schardt et al [31] have found that searchers using the PICO format had more precise results than users searching with the standard interface on PubMed; in that study, precision scores were defined as the number of relevant or gold-standard articles retrieved in a result set compared to the total number of articles retrieved in that set. Unfortunately, and possibly due to the small sample, the difference between the search groups was not statistically significant [31]. An alternative could be to improve search engine functionalities, with the remaining challenge, however, of avoiding any cherry picking of the evidence and, thus, potentially biased conclusions for clinical practice. A potential solution lies with federated search engines like MacPLUS FS, which complement summary-level evidence with other preappraised and non-preappraised research. Indeed, we have shown that physicians access all types of resources translating an interest into different layers of the EBM when these layers are displayed together on one page (see Figure 2). The use of a federated search engine may thus help clinicians navigate across EBM resources, allowing them to look at and compare different resources simultaneously and to identify the current best evidence that is most adapted to their information needs.

Limitations and Strengths of the Study

The main limitation of our study was that clinicians likely used other means than MacPLUS FS to answer some of their daily questions. Our design also did not assess the clinical impact of the answers retrieved. This would have required mixed methods approaches to estimate the number of patients needed to benefit from information (ie, number needed to benefit from information [NNBI]), as described by Pluye et al [32].

Finally, our sample of searches was recorded in the context MacPLUS FS randomized controlled trials [5], and it remains unclear how search queries may have differed without the possible influence of the interventions tested. The second intervention— the evidence retrieval coach—included eight short educational videos, of which only one was providing advice on the PICO formulation of clinical queries. However, only a small group of participants would have been exposed to that short video, and none of the other interventions were specifically aimed at improving the formulation of queries.

Strengths include the direct record of queries in one of the largest samples of physicians from different specialties and levels of practice. It is also the first study on a federated search engine, which allowed us to show that clinicians access all resources and not only summary-level evidence.

Conclusions

A constant flow of new articles overwhelms clinicians who are continuously exposed to them. To keep up and to answer our clinical questions, it is essential to clarify and translate clinical questions into searchable queries. Our results could lead to the development of educational and clinical interventions on how to increase searching skills [2]. These could include workshops and tools to translate clinical questions into queries and to better structure and adapt them to each type of resource.

Our findings also highlight the potential role of federated search engines over the use of single resources to meet clinicians' needs [23]. A federated search engine retrieves evidence and may help clinicians get answers to their questions with current best evidence, even with short time frames and limited experience and skills for searching.

Other avenues of research include the improvement of search functionalities and clinical interventions to meet users' expectations in navigating through the evidence, in order to rapidly find the most relevant and least-biased answers for better clinical practice and patient care.

Conflicts of Interest

TA, AI, and RBH are editors of American College of Physicians (ACP) Journal Club. The McMaster PLUS service was developed and is maintained by the Health Information Research Unit at McMaster University, Hamilton, Canada, which owns the intellectual property.

Multimedia Appendix 1

Evidence-based medicine (EBM) resources accessible through MacPLUS FS (McMaster Premium LiteratUre Service Federated Search).

[DOC File , 63 KB - mededu_v6i1e16777_app1.doc]

Multimedia Appendix 2

Characteristics of clinicians by specialty types.

[DOC File , 84 KB - mededu_v6i1e16777_app2.doc]

References

1. Glasziou P, Burls A, Gilbert R. Evidence based medicine and the medical curriculum. BMJ 2008 Sep 24;337:a1253. [doi: 10.1136/bmj.a1253] [Medline: 18815165]
2. Agoritsas T, Vandvik PO, Neumann I, Rochwerg B, Jaeschke R, Hayward R, et al. Finding current best evidence. In: Guyatt G, Renni D, Meade M, Cook D, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd edition. New York, NY: McGraw-Hill Education; 2014:29-49.
3. Agoritsas T, Guyatt GH. Evidence-based medicine 20 years on: A view from the inside. Can J Neurol Sci 2013 Jul;40(4):448-449. [doi: 10.1017/s0317167100014499] [Medline: 23786723]
4. Montori VM, Guyatt GH. Progress in evidence-based medicine. JAMA 2008 Oct 15;300(15):1814-1816. [doi: 10.1001/jama.300.15.1814] [Medline: 18854545]
5. Agoritsas T, Iserman E, Hobson N, Cohen N, Cohen A, Roshanov PS, et al. Increasing the quantity and quality of searching for current best evidence to answer clinical questions: Protocol and intervention design of the MacPLUS FS Factorial Randomized Controlled Trials. Implement Sci 2014 Sep 20;9:125 [FREE Full text] [doi: 10.1186/s13012-014-0125-9] [Medline: 25239537]
6. Moja L, Kwag KH. Point of care information services: A platform for self-directed continuing medical education for front line decision makers. Postgrad Med J 2015 Feb;91(1072):83-91 [FREE Full text] [doi: 10.1136/postgradmedj-2014-132965] [Medline: 25655251]
7. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? PLoS Med 2010 Sep 21;7(9):e1000326 [FREE Full text] [doi: 10.1371/journal.pmed.1000326] [Medline: 20877712]
8. González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, et al. Information needs and information-seeking behavior of primary care physicians. Ann Fam Med 2007;5(4):345-352 [FREE Full text] [doi: 10.1370/afm.681] [Medline: 17664501]
9. Graber MA, Randles BD, Ely JW, Monnahan J. Answering clinical questions in the ED. Am J Emerg Med 2008 Feb;26(2):144-147. [doi: 10.1016/j.ajem.2007.03.031] [Medline: 18272092]
10. Green ML, Ciampi MA, Ellis PJ. Residents' medical information needs in clinic: Are they being met? Am J Med 2000 Aug 15;109(3):218-223. [doi: 10.1016/s0002-9343(00)00458-7] [Medline: 10974185]
11. Guyatt G, Meade M, Agoritsas T, Richardson WS, Jeaschke R. What is the question? In: Guyatt G, Rennie D, Meade M, Cook D, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd edition. New York, NY: McGraw-Hill Education; 2014:19-28.
12. Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, et al. Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. J Med Internet Res 2012 Jun 12;14(3):e85 [FREE Full text] [doi: 10.2196/jmir.2021] [Medline: 22693047]
13. Hoogendam A, de Vries Robbé PF, Overbeke AJ. Comparing patient characteristics, type of intervention, control, and outcome (PICO) queries with unguided searching: A randomized controlled crossover trial. J Med Libr Assoc 2012 Apr;100(2):121-126 [FREE Full text] [doi: 10.3163/1536-5050.100.2.010] [Medline: 22514508]
14. Thiele RH, Poiro NC, Scalzo DC, Nemergut EC. Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: Results of a randomised trial. Postgrad Med J 2010 Aug;86(1018):459-465. [doi: 10.1136/pgmj.2010.098053] [Medline: 20709767]

15. Haase A, Follmann M, Skipka G, Kirchner H. Developing search strategies for clinical practice guidelines in SUMSearch and Google Scholar and assessing their retrieval performance. BMC Med Res Methodol 2007 Jun 30;7:28 [FREE Full text] [doi: 10.1186/1471-2288-7-28] [Medline: 17603909]
16. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, et al. Retrieving clinical evidence: A comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res 2013 Aug 15;15(8):e164 [FREE Full text] [doi: 10.2196/jmir.2624] [Medline: 23948488]
17. ACCESSSS. Hamilton, ON: Health Information Research Unit, McMaster University URL: http://www.accessss.org [accessed 2020-03-10]
18. Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: A systematic review. JAMA Intern Med 2014 May;174(5):710-718. [doi: 10.1001/jamainternmed.2014.368] [Medline: 24663331]
19. Ely JW, Osheroff JA, Ebell MH, Bergus GR, Levy BT, Chambliss ML, et al. Analysis of questions asked by family doctors regarding patient care. BMJ 1999 Aug 07;319(7206):358-361 [FREE Full text] [doi: 10.1136/bmj.319.7206.358] [Medline: 10435959]
20. Ely JW, Osheroff JA, Ebell MH, Chambliss ML, Vinson DC, Stevermer JJ, et al. Obstacles to answering doctors' questions about patient care with evidence: Qualitative study. BMJ 2002 Mar 23;324(7339):710 [FREE Full text] [doi: 10.1136/bmj.324.7339.710] [Medline: 11909789]
21. Haynes B, Glasziou P, Straus S. Advances in evidence-based information resources for clinical practice. ACP J Club 2000;132(1):A11-A14. [Medline: 10689385]
22. Slawson DC, Shaughnessy AF. Obtaining useful information from expert based sources. BMJ 1997 Mar 29;314(7085):947-949 [FREE Full text] [doi: 10.1136/bmj.314.7085.947] [Medline: 9099121]
23. Trevena LJ, Irwig L, Isaacs A, Barratt A. GPs want tailored, user friendly evidence summaries--A cross sectional study in New South Wales. Aust Fam Physician 2007 Dec;36(12):1065-1069 [FREE Full text] [Medline: 18075638]
24. Allan GM, Ma V, Aaron S, Vandermeer B, Manca D, Korownyk C. Residents' clinical questions: How are they answered and are the answers helpful? Can Fam Physician 2012 Jun;58(6):e344-e351 [FREE Full text] [Medline: 22700742]
25. Kritz M, Gschwandtner M, Stefanov V, Hanbury A, Samwald M. Utilization and perceived problems of online medical resources and search tools among different groups of European physicians. J Med Internet Res 2013 Jun 26;15(6):e122 [FREE Full text] [doi: 10.2196/jmir.2436] [Medline: 23803299]
26. Fiander M, McGowan J, Grad R, Pluye P, Hannes K, Labrecque M, et al. Interventions to increase the use of electronic health information by healthcare practitioners to improve clinical practice and patient outcomes. Cochrane Database Syst Rev 2015 Mar 14(3):CD004749. [doi: 10.1002/14651858.CD004749.pub3] [Medline: 25770311]
27. Ramos K, Linscheid R, Schafer S. Real-time information-seeking behavior of residency physicians. Fam Med 2003 Apr;35(4):257-260. [Medline: 12729309]
28. Clark E, Donovan EF, Schoettker P. From outdated to updated, keeping clinical guidelines valid. Int J Qual Health Care 2006 Jun;18(3):165-166. [doi: 10.1093/intqhc/mzl007] [Medline: 16613986]
29. Davidoff F, Florance V. The informationist: A new health profession? Ann Intern Med 2000 Jun 20;132(12):996-998. [doi: 10.7326/0003-4819-132-12-200006200-00012] [Medline: 10858185]
30. Montori VM, Wilczynski NL, Morgan D, Haynes RB, Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: Analytical survey. BMJ 2005 Jan 08;330(7482):68 [FREE Full text] [doi: 10.1136/bmj.38336.804167.47] [Medline: 15619601]
31. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 2007 Jun 15;7:16 [FREE Full text] [doi: 10.1186/1472-6947-7-16] [Medline: 17573961]
32. Pluye P, Grad RM, Johnson-Lafleur J, Granikov V, Shulha M, Marlow B, et al. Number needed to benefit from information (NNBI): Proposal from a mixed methods research study with practicing family physicians. Ann Fam Med 2013;11(6):559-567 [FREE Full text] [doi: 10.1370/afm.1565] [Medline: 24218380]

Abbreviations

ACP: American College of Physicians
CINAHL: Cumulative Index to Nursing and Allied Health Literature
EBM: evidence-based medicine
MacPLUS FS: McMaster Premium LiteratUre Service Federated Search
NNBI: number needed to benefit from information
PECO: population, exposure, comparison, and outcome
PICO: population, intervention, comparison, outcome

Edited by C Eichenberg; submitted 23.10.19; peer-reviewed by M Rethlefsen, R Grad; comments to author 03.12.19; revised version received 27.01.20; accepted 19.02.20; published 20.04.20.

Please cite as:
Seguin A, Haynes RB, Carballo S, Iorio A, Perrier A, Agoritsas T
Translating Clinical Questions by Physicians Into Searchable Queries: Analytical Survey Study
JMIR Med Educ 2020;6(1):e16777
URL: http://mededu.jmir.org/2020/1/e16777/
doi:10.2196/16777
PMID:32310137

Original Paper

Finding the Best Way to Deliver Online Educational Content in Low-Resource Settings: Qualitative Survey Study

Lucy Kynge[1,2], MA

[1]Interburns, Swansea University, Swansea, United Kingdom

[2]The Centre for Global Burn Injury, Policy and Research, Swansea University, Swansea, United Kingdom

Corresponding Author:
Lucy Kynge, MA
Interburns
Swansea University
Haldane Building
Swansea, SA2 8PP
United Kingdom
Phone: 44 07742437420
Email: lucy.kynge@interburns.org

Abstract

Background: The reach of internet and mobile phone coverage has grown rapidly in low- and middle-income countries (LMICs). The potential for sharing knowledge with health care workers in low-resource settings to improve working practice is real, but barriers exist that limit access to online information. Burns affect more than 11 million people each year, but health care workers in low-resource settings receive little or no training in treating burn patients. Interburns' training programs are tailor-made to improve the quality of burn care in Asia, Africa, and the Middle East; the challenge is to understand the best way of delivering these resources digitally toward improved treatment and care of burn patients.

Objective: The aim of the study, funded by the National Institute for Health Research (NIHR), was to understand issues and barriers that affect health care worker access to online learning in low-resource settings in order to broaden access to Interburns' training materials and improve burn-patient care.

Methods: A total of 546 participants of Interburns' Essential Burn Care (EBC) course held in Bangladesh, Nepal, Ethiopia, and the West Bank, the occupied Palestinian Territories, between January 2016 and June 2018 were sent an online survey. EBC participants represent the wide range of health care professionals involved with the burn-injured patient. A literature review was carried out as well as research into online platforms.

Results: A total of 207 of 546 (37.9%) participants of the EBC course did not provide an email address. Of the 339 email addresses provided, 81 (23.9%) "bounced" back. Surgeons and doctors were more likely to provide an email address than nurses, intern doctors, or auxiliary health care workers. A total of 258 participants received the survey and 70 responded, giving a response rate of 27.1%. Poor internet connection, lack of time, and limited access to computers were the main reasons for not engaging with online learning, along with lack of relevant materials. Computers were seen as more useful for *holding information*, while mobile phones were better for communicating and *sharing knowledge*. Health care workers in LMICs use mobile phones professionally on a daily basis. A total of 80% (56/70) felt that educational content on burns should be available through mobile apps.

Conclusions: Health care workers in low-resource settings face a variety of barriers to accessing educational content online. The reliance on email for sign-up to learning management systems is a significant barrier. Materials need to be relevant, localized, and easy to consume offline if necessary, to avoid costs of mobile phone data. Smartphones are increasingly used professionally every day for communication and searching for information, pointing toward the need for tailored educational content to be more available through mobile- and web-based apps.

(JMIR Med Educ 2020;6(1):e16946) doi:10.2196/16946

KEYWORDS

online education; digital content; health care; burns; low-resource settings

Introduction

The reach and proliferation of the internet and mobile phones has expanded rapidly across low- and middle-income countries (LMICs), enabling broader access to educational opportunities and an attitude shift in how content can be successfully delivered. However, reaching specific audiences such as health care workers in low-resource settings to influence working practice is a multilayered challenge. Materials need to be contextually appropriate, interesting, and formatted to engage the user on their chosen device, but they must also be easy to access with minimal barriers.

Since 2012 when Massive Open Online Courses (MOOCs) were launched, there has been huge growth in online educational content and providers, but course completion rates for *industrial* or *mass media*-type e-learning models have been low, and dropout rates high, particularly in developing countries. Reasons range from socioeconomic, cultural, and administrative factors to contextual barriers, such as content that is inappropriate, or different pedagogical approaches. A study in 2018 states that while "the primary factors of adoption are perceived usefulness and perceived ease of use...there are several other factors which work along with these two factors to explain technology adoption" [1].

Health care workers in low-resource settings face particular challenges and often feel professionally isolated. There is little time, incentive, or, in many cases, connectivity, to go online to study, which means that materials need to be especially meaningful and accessible if they are to impact working practice. In the case of LMICs, the *small-scale* model is encouraged. Professor Deidre Carabine of the Virtual University of Uganda has coined the term "global knowledge for local action," which encourages the contextualization and local creation of content in which quality and adaptability of educational programs are the key to success: "A major challenge...is to make online learning first-class, that's: relevant, interesting, engaging, and more exciting than the traditional classroom..." [2].

Understanding how health care workers in resource-poor settings can access online content has been the focus of a study by Interburns, a UK charity working closely with the Centre for Global Burn Injury and Policy Research (CGBIPR) at Swansea University and funded by the National Institute for Health Research (NIHR). Interburns works with partners in Asia, Africa, and the Middle East to develop and deliver contextualized face-to-face training programs to tackle the global burden of burn injuries through capacity development and quality improvement. Burns affect more than 11 million people annually, with 95% occurring in poor countries, and the majority of those affected are children. Interburns' tailored training programs focus on the reality of burn care services on the ground in low-resource settings and emphasize the need for *knowledge turned into action*.

The challenge is how to expand the reach of Interburns' face-to-face training materials in an easily acceptable online format. This research focused on understanding barriers that health care workers involved with burn-injured patients faced when accessing online content in Bangladesh, Ethiopia, Nepal, and the West Bank, the occupied Palestinian Territories. The information gathered has been used to reformat the Essential Burn Care (EBC) Training Manual as an online resource and to inform a digital strategy for all of Interburns' training resources.

Methods

An online survey was carried out between July and September 2018 to understand the way in which health care workers in low-resource settings engage with digital content and to understand common barriers to access. Questions focused on previous experiences with online education, reasons why online education had been seen to be successful or not, attitudes and access issues around computers and mobile phones, and suggestions of ways in which information on burns could best be delivered.

A total of 546 participants of Interburns' face-to-face course in EBC from Bangladesh, Nepal, Ethiopia, and the Palestinian West Bank between January 2016 and June 2018 were selected to participate in the survey. These health care workers represent the wide range of professions involved with the burn-injured patient; they are also the intended target of the digital version of the EBC Training Manual. Consent was given at the time of the training course and in the survey.

Following a review of survey results in December 2018 and a literature review, the EBC Training Manual was reformatted for an identified online platform and trialed with health care workers in 2019.

Results

A total of 207 of the 546 participants (37.9%) of Interburns' EBC training courses did not provide an email address when asked to leave contact details at the end of those courses. Out of 546 participants, 339 (62.1%) gave an email address, but 81 of those 339 emails (23.9%) with the survey link "bounced" back. This suggests that the email address given was incorrectly remembered or no longer in use. This is likely to show that a substantial number of health care workers in these settings do not regularly use email. Surgeons and doctors were more likely to give an email address than nurses, intern doctors, auxiliary health care workers, or medical officers. Out of 220 participants from Bangladesh, 118 (53.6%) did not give an email address; of these, 88 (74.6%) were nurses. Out of 281 health care workers from Nepal, 74 (26.3%) did not give an email address; 31 of these (41.8%) were nurses.

Learning management systems (LMSs) require an email address for sign-up and verification of individual users. If online educational content is held on an LMS, a large proportion of health care workers in these settings will be unable to sign up for a course unless they have an email address. This is an immediate, though hidden, barrier to access.

Of 258 participants who received the survey 70 responded, giving a response rate of 27.1%. Responses were balanced between male and female health care workers from Bangladesh (13/70, 19%), Nepal (42/70, 60%), Ethiopia (6/70, 9%), and the Palestinian West Bank (3/70, 4%). A total of 6 respondents (9%) selected *other*. See Table 1 for details of the professions represented by the survey.

A total of 74% (52/70) of respondents said they had used computers for online learning, but poor internet connection, lack of time, and limited access to computers were cited as reasons for not engaging well or regularly with online learning. With the rise in mobile phone ownership, better connectivity, and the increasing use of smartphones, participants expressed the view that computers were useful for *holding information*, while mobile phones were better for communicating and *sharing knowledge*, as expressed in the following quotes:

> *Because of the development of mobile phones, people rarely need access to use computers. In contrast, mobile phones nowadays are an essential accessory.* [Medical student, Chittagong, Bangladesh]

> *Information is safely stored in computer, and information of mobile phone is portable.* [Nurse, Central Development Region, Nepal]

> *I prefer mobile phone because we always carry mobile phone everywhere, so we can get information easily anywhere.* [Nurse, Kirtipur, Nepal]

This was backed up by the data, which showed that mobile phones are increasingly used professionally by health care workers in low-resource settings.

A total of 83% (58/70) of participants use mobile phones to find information for personal use at least two or three times each week, and 30% (21/70) also use their phones to look for professional information. A total of 80% (56/70) felt that mobile apps could be used to convey information and education on burns. This suggests that content created for mobile phones and apps would have a strong chance of uptake. However, 33% (23/70) of respondents worried about the high costs of data that can be incurred when accessing content. Representative quotes are given below:

> *Messenger systems are used by almost everyone that has a smartphone; short messages and information can effectively be transferred and accessed.* [Auxiliary health care worker, Oromia Region, Ethiopia]

> *It will be more helpful to use messenger systems for burn care. It will be quick and immediate and also recent.* [Health care manager, Bangladesh]

> *Cost-benefit ratio is not satisfactory.* [Surgeon, Comilla, Bangladesh]

> *It may cause financial problem.* [Physiotherapist, Tigray Region, Ethiopia]

> *[There is a cost] when using mobile data for seeing informative videos.* [Nurse, Kathmandu, Nepal]

Table 1. Participant professions.

Profession	Value (N=70), n (%)
Nurse	25 (36)
Doctor	19 (27)
Intern doctor	1 (1)
Therapist	6 (9)
Surgeon	14 (20)
Registrar	1 (1)
Auxiliary health care	1 (1)
Other	3 (4)
Total	70 (100)

Discussion

Principal Findings

The online survey showed that a significant number of health care workers in low-resource settings do not use email. This is a major barrier to accessing educational materials, which are hosted on LMSs that rely on email for sign-up and verification. The survey also found that mobile phones are increasingly used for professional purposes, although the high cost of data is a potential barrier for users.

Creating meaningful and contextualized educational content is the vital first step, and content needs to be formatted to be easily accessible in the online environment; Depover and Orivel stated, “The quality, relevance, and adaptability of educational programmes are becoming more important than ever” [3]. Lack of time, poor connectivity, irrelevant content, and limited access to computers are common reasons for low uptake of online courses, but each microsetting can throw up its own barriers in LMICs.

A study in Bangladesh applied a critical realist approach, in which *entity*, *agency*, and *causality* were used to explain the layers of context influencing online learning in the workplace. The study found that factors such as no desks, no internet connection, background noise, and no technical support hindered success. Other demotivating factors included low salary, too much work, lack of monitoring, and lack of support from family members [4]. Veletsianos stated, “...technology and certain practices associated with it are often expected to revolutionize the way individuals learn and teach. Yet scholars and practitioners alike are wise to maintain some skepticism about promises of transformation that ignore the environmental factors that surround innovations" [5].

In LMICs, there are a variety of barriers to accessing online content and reaching a specific population within a particular setting; in this case, the health care worker treating burn patients is a special challenge. Material must be formatted for the screen of choice and delivered with as few barriers as possible. Internet access and cost of data are high on the list of barriers. Mobile phone coverage is still far from uniform even within localities, though this is improving. In Bangladesh, the number of internet subscribers has risen year on year by approximately 10 million, reaching 83 million in 2018. Penetration by mobile broadband has seen similarly rapid growth. However, internet use in Nepal is low by international standards at 55% penetration in 2017, although the mobile phone market has grown quickly.

With greater connectivity, mobile devices can be used to communicate educational health messages, including at the community or rural level. Mobile phone penetration has risen considerably among community health care workers (CHWs) in low-resource settings, where multiple studies use smartphone apps to monitor and record patient details. A study into HIV data-gathering across five districts in Malawi found that mobile phone ownership was 100%, with smartphone ownership at 80% among decision makers and 50% among CHWs [6].

There have been experiments linking traditional LMSs, designed for computer-first online learning, with mobile messenger apps. A study in China piloted an online LMS linked to the WeChat messenger system as the training platform. Course completion rates rose from over 60% to 100%, and 63% of participating nurses wanted to receive *push* notifications through their phones for upcoming training courses [7].

Malhotra states, "Mobile phones are among the most coveted items among youth" [8], which increases the potential for greater acceptance of educational content delivered on a handset. Potential drawbacks, however, can be out-of-date devices, the potential for distraction, and the blurring of barriers between personal and professional use. However, benefits of using a mobile-first approach include the potential for greater access, more situated and contextualized learning, convenience, communication, and interaction. While learners will increasingly expect all online learning to work seamlessly on a mobile phone, there are challenges to scaling up approaches, not least of which is sustainable financing for the large-scale use of mobile phone technology in resource-limited settings [9].

This online survey showed a common but potentially hidden barrier, which is that sign-up processes for LMSs require email links and passwords to verify user information. Email is less used in many LMICs, meaning that registering for an online course is instantly challenging. Furthermore, where training is focused on the acquisition of knowledge-based technical skills, such as those required in burn treatment and care, online achievements may not reflect actual ability in the workplace [10]. Blended programs of regular face-to-face support are necessary to translate online learning into daily practice and will also encourage sustained use of digital materials.

Conclusions

Technological improvements have expanded the ways in which health care workers in low-resource settings can access knowledge to improve working practice, but each local setting can throw up barriers. It is vital to understand the context within which the individual accesses digital information so that materials can be tailored to the setting and screen of choice. LMSs use email for sign-up and verification, and passwords for course access. This is a significant barrier for health care workers in low-resource settings who are not used to using email in their daily lives.

Health care workers in LMICs increasingly use mobile messaging professionally, and the majority of those surveyed felt that educational information on burns should be made available on both computers and mobile devices. Digital content needs to be contextually appropriate, formatted to be interesting, and accessible by computer, but with a *mobile-first* approach to developing content. Blended programs of face-to-face and online training will encourage greater use of online resources.

Acknowledgments

This research was funded by the NIHR (project number: 16/137/110) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author and not necessarily those of the NIHR or the UK Department of Health and Social Care.

Conflicts of Interest

None declared.

References

1. Panigrahi R, Srivastava PR, Sharma D. Online learning: Adoption, continuance, and learning outcome—A review of literature. Int J Inf Manage 2018 Dec 06;43:1-14 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2018.05.005]
2. Carabine D. The Herald Zimbabwe. 2017 May 03. Traditional varsities on the way out URL: https://www.herald.co.zw/traditional-varsities-on-the-way-out/ [accessed 2020-04-20]
3. Depover C, Orivel F, UNESCO: International Institute for Educational Planning. Developing Countries in the E-Learning Era: Fundamentals of Educational Planning Series No. 96. Paris, France: United Nations Education, Scientific and Cultural Organization (UNESCO); 2013. URL: https://tinyurl.com/yb8wugla [accessed 2020-04-24]

4. Jamil MG. Technology-enhanced teacher development in rural Bangladesh: A critical realist evaluation of the context. Eval Program Plann 2018 Aug;69:1-9. [doi: 10.1016/j.evalprogplan.2018.04.002] [Medline: 29631146]
5. Veletsianos G. Emergence and Innovation in Digital Learning: Foundations and Applications. Edmonton, AB: AU Press, Athabasca University; 2016. The defining characteristics of emerging technologies and emerging practices in digital education URL: http://www.giseldacosta.com/wordpress/wp-content/uploads/2016/08/Veletsianos.pdf [accessed 2020-04-24]
6. Cooper HJ, Smith TR. The KUUNIKA Project: Data for Action. 2016 Aug. Data users study: Initial findings URL: https://static1.squarespace.com/static/548487dce4b08bf981fe60d5/t/580136a5e6f2e1ccc7c72d4b/1476474537679/ [accessed 2017-05-02]
7. Liu X, Cheng J, Huang S. Mobile phone training platform for the nursing staff in the emergency department. Telemed J E Health 2019 Jan;25(1):66-70. [doi: 10.1089/tmj.2017.0317] [Medline: 29742034]
8. Malhotra A, Sharma R, Srinivasan R, Mathew N. Widening the arc of indigenous communication: Examining potential for use of ICT in strengthening social and behavior change communication efforts with marginalized communities in India. Electron J Inf Syst Dev Ctries 2018 May 30;84(4):e12032. [doi: 10.1002/isd2.12032]
9. Leon N, Schneider H, Daviaud E. Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa. BMC Med Inform Decis Mak 2012 Nov 05;12:123 [FREE Full text] [doi: 10.1186/1472-6947-12-123] [Medline: 23126370]
10. Green J, Huntington A. Online professional development for digitally differentiated nurses: An action research perspective. Nurse Educ Pract 2017 Jan;22:55-62. [doi: 10.1016/j.nepr.2016.11.009] [Medline: 27940391]

Abbreviations

CGBIPR: Centre for Global Burn Injury and Policy Research
CHW: community health care worker
EBC: Essential Burn Care
LMIC: low- and middle-income country
LMS: learning management system
MOOC: Massive Open Online Course
NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 06.11.19; peer-reviewed by S Sarbadhikari, C Lederer; comments to author 09.12.19; revised version received 29.02.20; accepted 22.03.20; published 26.05.20.

Please cite as:
Kynge L
Finding the Best Way to Deliver Online Educational Content in Low-Resource Settings: Qualitative Survey Study
JMIR Med Educ 2020;6(1):e16946
URL: http://mededu.jmir.org/2020/1/e16946/
doi:10.2196/16946
PMID:32452810

Viewpoint

Using Consumer Perceptions of a Voice-Activated Speaker Device as an Educational Tool

Siubak Chung[1], BA; Benjamin KP Woo[1], MD

University of California, Los Angeles, Sylmar, CA, United States

Corresponding Author:
Siubak Chung, BA
University of California, Los Angeles
14445 Olive View Drive
Sylmar, CA,
United States
Phone: 1 7472103830
Email: emitam90@gmail.com

Abstract

Voice-activated smart speakers, with their ease of setup, low cost, and versatility, could be an affordable and accessible way to improve health and mental health outcomes. In 2018, there were a total of 320 comments generated from verified purchases of a voice-activated smart speaker. These comments revealed there could be potential benefits of reducing loneliness and social isolation for adult users, especially for the older population. However, further research is warranted to determine whether using such devices would be harmful to children's physical or mental development.

(JMIR Med Educ 2020;6(1):e17336) doi:10.2196/17336

KEYWORDS

consumer perceptions; voice-activated speaker device

Introduction

Voice-activated speaker devices have recently gained popularity with the release of commercial products such as Amazon Echo and Google Home. In the United States, Amazon Echo and Google Home were released in 2014 and 2016, respectively. In China, Xiaomi launched a Chinese-language, voice-activated smart speaker, XiaoAI, in July 2017. XiaoAI has functionalities such as checking the weather, controlling smart home devices, playing music, and translating foreign languages [1]. In the second quarter of 2018, 2 million units of XiaoAI Speaker Mini were sold [2]. As voice-activated smart speakers are easy to set up, low-cost, and versatile, they could be an affordable and accessible way of improving health and mental health outcomes. At the same time, it should be noted that recent studies have documented the deficits of various artificial intelligence-powered voice assistants in responding to questions about interpersonal violence, mental health, and physical health [3,4].

Customer reviews from verified buyers on electronic commerce websites such as Amazon.com (United States) and Taobao.com (China) could help potential buyers learn more about a product. These customer reviews could also assist medical students and health professionals in understanding how technology could impact the mental health of device users. Therefore, we evaluated Taobao customer reviews, written in simplified Chinese, of a Xiaomi XiaoAI voice-activated smart speaker to gain a better understanding of the users of this type of technology.

Methods

Leveraging user-generated textual data in the form of Taobao.com verified purchase reviews of a Xiaomi XiaoAI Speaker Mini in 2018, a retrospective review was performed. The recorded parameters included the number of verified purchase reviewer comments and the number of likes and dislikes. Among a total of 320 comments from verified buyers in 2018, there were 299 likes and 19 dislikes. The content of the positive and critical comments was also analyzed qualitatively. Negative comments regarding the weaknesses and defects of the device were excluded from this analysis.

Results

Four comments indicated how the voice-activated speaker device could potentially impact mental health for adult users. For example, one reviewer wrote:

> *Having her [Xiaomi XiaoAI], I am no longer lonely. But sometimes she ignores me, and I don't know why...But having her, I am no longer lonely.*

Five comments suggested that the smart speaker may decrease loneliness and social isolation for older users. For example, an elderly user wrote:

> *My son came back home to install [Xiaomi XiaoAI] for me...Now my grandson wants to come to my house all the time, and he will sing with me and XiaoAI.*

Moreover, there were an overwhelming 34 comments that highlighted the benefits of the voice-activated speaker device for children. For example, one parent wrote:

> *Xiaomi XiaoAI can speak to my child all day!*

Another parent remarked:

> *My child loves it, and they need to create one for outdoor use.*

Nevertheless, there was one parent who warned against its use:

> *Please do not buy for your child. Why do you want your child to talk to a dead object every minute, every second?*

Discussion

Principal Findings

This preliminary qualitative research into consumer comments provides insight into consumers' perception of voice-activated smart speakers for Chinese-language users. Our analysis revealed that there could be potential benefits including the reduction of loneliness and social isolation for Chinese-speaking adult users, especially for the older population. Although Chinese parents seem to acknowledge the benefits of having voice-activated speaker devices at home, further research is warranted to determine whether using such devices is harmful to children's physical or mental development. Nevertheless, no matter the age group, users of smart speakers should practice moderation. Previous research studies have highlighted the advantages and disadvantages of social media and technology for Chinese-speaking participants [5-8]. Future studies could explicitly focus on the advantages and disadvantages of using voice-activated smart speakers among Chinese-speaking individuals of different age groups.

Limitations and Future Directions

A few limitations should be considered when interpreting the results of this viewpoint. As noted, this viewpoint was preliminary and exploratory in nature. It only indicated the feasibility of using customer reviews to further understand a voice-activated smart speaker. In addition, some of the consumer comments were not from the direct users. As such, the comments from parents could be biased. Another limitation of this viewpoint was the exclusion of negative comments for the analysis. This preliminary viewpoint excluded negative comments because they mainly focused on the weaknesses and mechanical defects of the device. Further research is needed to determine the potential downsides of voice-activated smart speakers on mental health and mental health care. The use of focus groups in future studies could provide insights into the negative health and mental effects of such devices. Finally, future research could focus on how the presence of voice-activated speaker devices may enhance student learning. Medical students and trainees need to have a deeper understanding of how technology will change the future of health care, and understanding consumer perceptions of such products could be a way to supplement medical education for future physicians and health care professionals.

Conflicts of Interest

None declared.

References

1. Fingas, J. Engadget. Xiaomi's take on a voice assistant is built for China URL: https://www.engadget.com/2018/04/03/xiaomi-unveils-xiao-ai-voice-assistant/ [accessed 2018-04-03]
2. Sheng, W. Technode. China's tech giants battle for smart speaker supremacy as price war rages on URL: https://technode.com/2019/07/09/chinas-tech-giants-battle-for-smart-speaker-supremacy-as-price-war-rages-on/ [accessed 2019-07-09]
3. Miner AS, Milstein A, Schueller S, Hegde R, Mangurian C, Linos E. Smartphone-Based Conversational Agents and Responses to Questions About Mental Health, Interpersonal Violence, and Physical Health. JAMA Intern Med 2016 May 01;176(5):619-625 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0400] [Medline: 26974260]
4. Tsiang JT, Woo BK. Comparison of Online Dementia Information in Chinese and in English Languages. Cureus 2017 Oct 28;9(10):e1808 [FREE Full text] [doi: 10.7759/cureus.1808] [Medline: 29308336]
5. Woo BKP. Facebook Advertising for Disseminating Early Psychosis Education to the Chinese-Speaking General Public. Health Equity 2019;3(1):19-21 [FREE Full text] [doi: 10.1089/heq.2018.0087] [Medline: 30783633]
6. Lam NHT, Tsiang JT, Woo BKP. Exploring the Role of YouTube in Disseminating Psychoeducation. Acad Psychiatry 2017 Dec;41(6):819-822. [doi: 10.1007/s40596-017-0835-9] [Medline: 29022242]
7. Lam NHT, Woo BKP. YouTube as a New Medium for Dementia Education Among Chinese Americans. Community Ment Health J 2020 Apr;56(3):435-439. [doi: 10.1007/s10597-019-00493-7] [Medline: 31641910]
8. Lam NH, Woo BK. Digital Media Recruitment for Fall Prevention Among Older Chinese-American Individuals: Observational, Cross-Sectional Study. JMIR Aging 2018 Nov 01;1(2):e11772 [FREE Full text] [doi: 10.2196/11772] [Medline: 31518249]

Edited by G Eysenbach; submitted 07.12.19; peer-reviewed by N Lam, J Chung, X Zheng, N Nazari; comments to author 02.01.20; revised version received 11.02.20; accepted 26.02.20; published 24.04.20.

Please cite as:
Chung S, Woo BKP
Using Consumer Perceptions of a Voice-Activated Speaker Device as an Educational Tool
JMIR Med Educ 2020;6(1):e17336
URL: http://mededu.jmir.org/2020/1/e17336/
doi:10.2196/17336
PMID:32329740

Viewpoint

COVID-19 Can Catalyze the Modernization of Medical Education

Cathy Hsi Chen[1*], BSE, MD; Alexander Joseph Mullen[1*], BS, MD

Department of Medicine, University of Mississippi Medical Center, Jackson, MS, United States

[*]all authors contributed equally

Corresponding Author:
Alexander Joseph Mullen, BS, MD
Department of Medicine
University of Mississippi Medical Center
2500 N State St
Jackson, MS, 39216
United States
Phone: 1 6019841000
Email: ajmullen7@gmail.com

Abstract

Amid the coronavirus disease (COVID-19) crisis, we have witnessed true physicianship as our frontline doctors apply clinical problem-solving to an illness without a textbook algorithm. Yet, for over a century, medical education in the United States has plowed ahead with a system that prioritizes content delivery over problem-solving. As resident trainees, we are acutely aware that memorizing content is not enough. We need a preclinical system designed to steer early learners from "know" to "know how." Education leaders have long advocated for such changes to the medical school structure. For what may be the first time, we have a real chance to effect change. In response to the COVID-19 pandemic, medical educators have scrambled to conform curricula to social distancing mandates. The resulting online infrastructures are a rare chance for risk-averse medical institutions to modernize how we train our future physicians—starting by eliminating the traditional classroom lecture. Institutions should capitalize on new digital infrastructures and curricular flexibility to facilitate the eventual rollout of flipped classrooms—a system designed to cultivate not only knowledge acquisition but problem-solving skills and creativity. These skills are more vital than ever for modern physicians.

(JMIR Med Educ 2020;6(1):e19725) doi:10.2196/19725

KEYWORDS

medical education; health professions education; medical school; curriculum design; flipped classroom; preclinical education; COVID-19; coronavirus; medical student

Amid the coronavirus disease (COVID-19) crisis, we have witnessed true physicianship as our frontline doctors apply clinical problem-solving to an illness without a textbook algorithm. Yet, for over a century, medical education in the United States has plowed ahead with a system that prioritizes content delivery over problem-solving and passive learning over active learning. Trainees develop problem-solving skills despite our preclinical education system, not because of it. A smattering of institutions has begun to reinvent, but for what may be the first time, we have a chance to push through necessary change on a broader scale. In response to the pandemic, medical educators have scrambled to conform curricula to social distancing mandates. The resulting online infrastructures could enable us to achieve what education leaders have long advocated [1]—eliminate the traditional classroom lecture in favor of active learning.

According to the Association of American Medical Colleges, lectures comprise half of medical school teaching, with 86 percent occurring in the first 2 years [2]. Despite their ubiquity, in-person lectures are increasingly rejected by students. In 2019, 49% of preclinical students reported "never" or "only occasionally" attending lectures, up from 41% 2 years prior [3]. Instead, they are turning to online material.

And why not? Online lecture videos allow students to peruse content at their own pace, as well as pause, review, and adjust playback speed. Research supports what students implicitly understand: online lectures are noninferior for learning, and often actually better [4].

For disheartened faculty teaching to half-filled auditoriums, the instinct may be to incentivize attendance. But it is time to teach the way modern physicians learn rather than how traditional educators teach. Class time should be used for active learning and learner-educator interaction—not content delivery. Many students feel their lecture-based, preclinical education poorly prepares them for clinical rotations [5]. As resident trainees, we are acutely aware that knowing content is not enough. The path from "know" to "know how" can be treacherous.

This is where educators can provide value beyond lecturing—by engaging learners, guiding their clinical problem-solving, integrating preclinical material into the clinical context, and providing corrective feedback. Traditional lectures are inherently unable to catch and address these individual reasoning deficits in real time. When we commit class time to lectures, we waste faculty expertise and eschew evidence-based learning.

In "flipped classrooms," students consume content online before working with educators to apply knowledge in group sessions (eg, through problem-based learning). Compared to traditional medical lectures, flipped classrooms produce better learning outcomes [6], especially in higher-order thinking [7]. Perks include greater class attendance [7] and teacher-student satisfaction [8]. Meta-analyses are limited because implementations vary, but research suggests that as the methodology matures, outcomes will further improve [7].

As problem- and case-based learning sessions become more prevalent, we will undoubtedly witness them evolve in their effectiveness, with space to explore other educational approaches (eg, patient simulators, augmented reality) as well. In addition, with the right analytics, flipped classrooms can exploit big data in ways traditional curricula cannot. Educators can track progress and target active learning sessions to actual knowledge gaps identified by frequent no-stakes tests.

Instructor time and cost-effectiveness are perhaps the leading critiques of flipped classrooms. The pandemic, however, is an inescapable impetus to transition content online. With infrastructure in place and costs already sunk, the barriers to enacting flipped classrooms (once social distancing guidelines relax) will be the lowest they have ever been. A key remaining concern—that more faculty are needed to coordinate group sessions—may be addressed in several ways. With content online, faculty can afford to meet with students less frequently. They might also enlist more teaching assistants: upper-level medical students, residents, fellows, and clinical faculty, all of whom spend disappointingly little time interacting with early learners.

Flipped classrooms are not new, and medical schools have been moving in their direction, but slowly. Change is difficult in storied institutions. In the setting of a crisis, however, change is the new normal. The Liaison Committee on Medical Education, the accrediting body for US and Canadian Medical Schools, has acknowledged that broad changes to the mechanisms of learning need to occur [9]. As such, they are granting institutions significant curricular flexibility, which can be harnessed to implement novel pedagogy.

In 1913, Dr William Osler said, "The lecture has its value, but its day has gone, and it should give place to other methods better adapted to modern conditions" [10]. We should not emerge from the pandemic only to revert to a preclinical education system even Dr Osler found outdated. We must make the most of our new digital infrastructures and curricular flexibility to facilitate the eventual rollout of flipped classrooms—a system deliberately designed to cultivate not only knowledge acquisition but problem-solving skills and creativity. These skills are more vital than ever for modern physicians.

Conflicts of Interest

None declared.

References

1. Prober CG, Heath C. Lecture halls without lectures--a proposal for medical education. N Engl J Med 2012 May 03;366(18):1657-1659. [doi: 10.1056/NEJMp1202451] [Medline: 22551125]
2. American Association of Medical Colleges. Instructional Methods Used by US and Canadian Medical Schools: Curriculum Inventory, 2012-2019 URL: https://www.aamc.org/data-reports/curriculum-reports/interactive-data/instructional-methods-used-us-and-canadian-medical-schools [accessed 2020-04-15]
3. American Association of Medical Colleges. 2019. Medical School Year Two QuestionnaireAll Schools Summary Report Internet URL: https://www.aamc.org/system/files/2019-08/2019-gq-all-schools-summary-report.pdf [accessed 2020-04-08]
4. Means B, Toyama Y, Murphy R, Bakia M, Jones K. Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. Washington, DC: US Department of Education, Office of Planning, Evaluation, and Policy Development; 2009.
5. O'Brien B, Cooke M, Irby DM. Perceptions and attributions of third-year student struggles in clerkships: do students and clerkship directors agree? Acad Med 2007 Oct;82(10):970-978. [doi: 10.1097/ACM.0b013e31814a4fd5] [Medline: 17895662]
6. Hew KF, Lo CK. Flipped classroom improves student learning in health professions education: a meta-analysis. BMC Med Educ 2018 Mar 15;18(1):38 [FREE Full text] [doi: 10.1186/s12909-018-1144-z] [Medline: 29544495]
7. Chen K, Monrouxe L, Lu Y, Jenq C, Chang Y, Chang Y, et al. Academic outcomes of flipped classroom learning: a meta-analysis. Med Educ 2018 Jun 25 [FREE Full text] [doi: 10.1111/medu.13616] [Medline: 29943399]
8. Ding C, Li S, Chen B. Effectiveness of flipped classroom combined with team-, case-, lecture- and evidence-based learning on ophthalmology teaching for eight-year program students. BMC Med Educ 2019 Nov 14;19(1):419 [FREE Full text] [doi: 10.1186/s12909-019-1861-y] [Medline: 31727043]

9. Barzansky B, Catanese V. Liaison Committee on Medical Education. 2020. Medical Students, Patients, and COVID‐19: A Community Conversation about Education and Safety URL: https://lcme.org/wp-content/uploads/filebase/March-12-2020-LCME-principles-from-joint-webinar_FINAL.pdf [accessed 2020-04-15]
10. Osler W. An Introductory Address on Examinations, Examiners, and Examinees. The Lancet 1913 Oct;182(4702):1047-1050. [doi: 10.1016/s0140-6736(01)77509-5]

Abbreviations

COVID-19: coronavirus disease

Edited by G Eysenbach; submitted 30.04.20; peer-reviewed by M Triola, T Chan, H Schröder; comments to author 27.05.20; revised version received 28.05.20; accepted 04.06.20; published 12.06.20.

Please cite as:
Chen CH, Mullen AJ
COVID-19 Can Catalyze the Modernization of Medical Education
JMIR Med Educ 2020;6(1):e19725
URL: http://mededu.jmir.org/2020/1/e19725/
doi:10.2196/19725
PMID:32501809

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org