Remote learning developments in undergraduate medical education in response to the COVID-19 pandemic: A BEME systematic review

Jennifer Stojan* (1), Mary Haas* (1), Satid Thammasitboon (2), Lina Lander (3), Sean Evans (3), Cameron Pawlik (1), Teresa Pawilkowska (4), Madelyn Lew (1), Deena Khamees (1), William Peterson (1), Ahmad Hider (1), Ciaran Grafton-Clarke (5), Hussein Uraiby (5), Morris Gordon (6,7), Michelle Daniel (3)

1. University of Michigan Medical School, USA
2. Texas Children's Hospital and Baylor College of Medicine, Houston, Texas, USA
3. University of California San Diego School of Medicine, USA
4. RCSI University of Medicine and Health Sciences, Ireland.
5. School of Medicine, University of Leicester, Leicester, United Kingdom
6. Blackpool Victoria Hospital, Blackpool, United Kingdom
7. School of Medicine, University of Central Lancashire, United Kingdom

* Joint first-author and equal contribution.

Corresponding author:

Jennifer N. Stojan, MD
Institution: University of Michigan
Phone: 586-770-6642, E-mail: jstojan@umich.edu
Address: 3119 Taubman Center, SPC5376
1500 East Medical Center Drive
Ann Arbor, MI 48109
Abstract:

Background

The COVID-19 pandemic resulted in an abrupt transition away from in-person educational activities. This systematic review investigated the pivot to online learning for nonclinical undergraduate medical education (UGME) activities and explored descriptions of educational offerings deployed, their impact, and lessons learned.

Methods

The authors systematically searched four online databases and hand searched MedEdPublish up to December 21, 2020. Two authors independently screened titles, abstracts and full texts, performed data extraction and assessed risk of bias from study methods and interventional reporting. A third author resolved discrepancies. Findings were reported in accordance with the STORIES (STructured apprOach to the Reporting in healthcare education of Evidence Synthesis) statement and BEME guidance.

Results

Fifty-six articles were included. The majority (n=41) described the rapid transition of existing offerings to online formats, whereas fewer (n=15) described novel developments. The majority (n=27) included a combination of synchronous and asynchronous components. Didactics (n=40) and small groups (n=26) were the most common instructional methods. Teachers largely integrated technology to replace and amplify rather than transform learning, though learner engagement was often interactive. Thematic analysis revealed unique challenges of online learning, as well as exemplary practices. The quality of study designs and reporting was modest,
with *underpinning theory* at highest risk of bias. Virtually all studies (n=54) assessed reaction/satisfaction, fewer than half (n=23) assessed changes in attitudes, knowledge or skills, and none assessed behavioral, organizational or patient outcomes.

**Conclusions**

UGME educators successfully transitioned face-to-face instructional methods online and implemented novel solutions during the COVID-19 pandemic. Although technology’s potential to transform teaching is not yet fully realized, the use of synchronous and asynchronous formats encouraged virtual engagement, while offering flexible, self-directed learning. As we transition from emergency remote learning to a post-pandemic world, educators must underpin new developments with theory, report additional outcomes and provide details that support replication.
Practice Points:

- A range of instructional methods were successfully transitioned online (e.g., didactics, small groups, PBL, TBL, clinical skills) and novel approaches implemented (e.g., mixed-reality).
- Synchronous, asynchronous and combined approaches provided opportunities for virtual engagement, active learning and connectivity, as well as flexibility and self-directed learning.
- Technology (e.g., video-conferencing software) was largely used to replace or amplify traditional teaching, but technology’s potential to transform teaching remains largely unrealized.
- Higher quality study designs and reporting are urgently needed, including studies that incorporate validity evidence into evaluation tools and those that underpin educational developments with theory.

Keywords

Best evidence medical education, remote learning, online learning, undergraduate medical education, COVID-19
Introduction:

Physical distancing requirements generated by the COVID-19 pandemic have resulted in a transition to emergency remote learning across the continuum of medical education (Gill et al. 2020; Hodges et al. 2020). Traditionally, learning in undergraduate medical education (UGME) has occurred in person in a variety of physical spaces, including classrooms for lectures and small group activities, simulated clinical environments for clinical and procedural skills, and laboratories for anatomic dissection and other lab-based skills. Students have also engaged in workplace-based learning in a variety of clinical settings.

Although online learning has been gaining traction as an alternative to established methods or to enrich in-person educational activities for some time, the COVID-19 pandemic accelerated the transition away from physical locations (Daniel et al. 2021). A number of terms are used interchangeably to describe online learning, including e-learning, web-based learning, remote learning, computer-assisted instruction, and internet-based learning (Ruiz et al. 2006). Online learning may be synchronous, asynchronous or both. Synchronous online learning occurs “in real time”, whereas asynchronous learning does not require teachers and learners to be online simultaneously (Hrastinski 2008; Worthington 2013).

The benefits of online learning are well-established and include the ability to provide more flexible and personalized teaching, while making accessing, updating, standardizing and distributing content easier (Ruiz et al. 2006; Wentling et al. 2000; Rosenberg 2001). Additionally, online learning allows for collaborations across institutions that reduce duplicative efforts, enhance the quality of curricula and serve large numbers of learners (Chen et al. 2019).
When thoughtfully designed and implemented, online learning can help optimize principles of adult learning theory, by emphasizing adaptability to a variety of learning styles, autonomy, motivation, self-direction and reflection (Taylor and Hamdy 2013).

Evidence involving primary, secondary and higher education learners has shown that online learning is non-inferior and possibly superior to face-to-face instruction (Means et al. 2009). A scoping review of online lectures in UGME found that students reported high satisfaction and demonstrated improvement on knowledge tests (Tang et al. 2018). Some online learning proponents have argued that COVID-19 has provided medical educators with a golden opportunity to revamp or even eliminate traditional classroom didactics, to capitalize on new digital infrastructures and flexibility, to promote widespread use of flipped classroom formats and shorter lectures, and to encourage multi-institutional access to resources through technology (Chen and Mullen 2020; Emanuel 2020).

The transition to online learning emerged as a prominent theme in two recent systematic reviews investigating the impact of COVID-19 on medical education: a rapid review by Gordon et al. (2020) that included 49 articles, and a follow up scoping review by Daniel et al. (2021) that included an additional 127 articles published through mid-September 2020. These reviews identified a need to further investigate the transition to online learning and ultimately spurred three additional systematic reviews: one focused on nonclinical educational experiences in UGME (i.e., this review), one focused on nonclinical educational experiences in postgraduate medical education (PGME), and one focused on clinical experiences across the UGME to PGME continuum.
The aim of this review was to synthesize published reports of developments in UGME in response to the COVID-19 pandemic, focusing on the ‘pivot’ to online learning and de novo developments in remote learning for nonclinical educational activities. Our review included publications described in the past two reviews (Gordon et al. 2020; Daniel et al. 2021), as well as newly published works. We addressed the following:

- What novel solutions or developments were deployed as institutions pivoted from face-to-face to remote / online learning? (i.e., description, or ‘what was done’) (Cook et al. 2008)
- What was the impact of these changes? What educational (Kirkpatrick’s) outcomes have been reported for these medical education developments? (i.e., justification or ‘did it work?’)
- What lessons were learned by the teams who deployed these developments that can guide future practice? (i.e., implications or ‘what’s next?’)

**Methods:**

This review was conducted in a relatively rapid time frame with 15 weeks elapsing from inception to completion. The methodological rigor of the approach was not compromised by speed. Systematicity was emphasized from the search strategy to the synthesis methods (Gordon et al. 2019a), with our prior work (Gordon et al. 2020; Daniel et al. 2021) serving as guides. The study protocol was posted to the Best Evidence in Medical Education (BEME) website. Study
reporting was aligned with the STORIES (STructured apprOach to the Reporting In healthcare education of Evidence Synthesis) statement (Gordon and Gibbs 2014) and BEME guidance (Hammick et al. 2010).

**Search strategy**

Consistent with our prior reviews, an electronic search was performed in four databases (MEDLINE, EMBASE, CINAHL, and PsychINFO). Our search strategy was developed by a librarian using the Accelerator Polyglot search translation tool (Clark et al. 2020). The full search strategy can be found in Appendix 1. PubMed was searched from August 2020 to December 21, 2020. We overlapped the dates of our prior search (Daniel et al. 2021) by a few weeks to ensure that no articles were missed. The other databases were searched from January 1, 2020 to December 21, 2020, as there was no option to delineate by month in these databases. MedEdPublish was hand searched.

Deduplication was conducted using the modified Bramer method (Bramer et al. 2016). Citations were uploaded into DistillerSR (Evidence Partners, Ottawa, Ontario, Canada), a data management system for conducting systematic reviews, wherein additional duplicates were removed.

**Study selection**

The following inclusion criteria were used:
• The study described a development in medical education explicitly deployed in response to COVID-19.

• The study involved a ‘pivot’ to online learning or a novel remote learning development intended to continue learning previously delivered face-to-face in a classroom, clinical skills suite, lab or other ‘non-clinical’ or ‘non-workplace’ environment.

• The study was in undergraduate medical education.

• The study included medical students.

• The study described Kirkpatrick’s outcomes (Level 1: satisfaction / reaction; Level 2a: changes in attitudes; Level 2b: changes in knowledge or skills; Level 3: behavioral change; Level 4a: change in organizational practice; Level 4b: change in clinical outcome) (Kirkpatrick JD and Kirkpatrick WL 2016).

• The study was in any language.

The following exclusion criteria were used:

• The study was an opinion piece, perspective, call for change, needs assessment or other study where no actual development was deployed.

• The study described other developments in response to COVID-19 that did not involve online / remote learning (e.g., in-person simulations, assessments, well-being, clinical service reconfigurations, interviews, service provision).

• The study was in postgraduate medical education only.

• The study described remote or distance learning explicitly deployed to replace workplace-based (clinical) learning.
Author pairs (JS, CP, AH, HU, CGC, DK, WP, MD) independently screened titles and abstracts. Cohen’s Kappa was used to calculate inter-rater reliability (McHugh 2012). Two authors independently reviewed full texts and documented reasons for exclusion. Discrepancies were mediated through discussion or involvement of a third author, until consensus was achieved.

Full text screening was conducted in three stages. In the first stage, we utilized the same full-text screening form used by Daniel et al. (2021). This allowed us to identify all new developments related to COVID-19 since our last review, helping build our database for future studies. In the second stage, we honed in on developments describing a ‘pivot’ to online learning or novel remote learning. At this juncture we added in studies from the two prior reviews by Gordon et al. (2020) and Daniel et al. (2021) that focused on the transition to online learning. In the third stage, we classified all of these developments into 3 categories, forming the basis of 3 parallel reviews: 1) remote learning intended to continue learning previously delivered face-to-face in non-clinical or non-workplace environments for undergraduate learners; 2) remote learning intended to continue learning previously delivered face-to-face in non-clinical or non-workplace environments for postgraduate learners; and 3) remote learning intended to continue learning previously occurring in clinical or workplace environments for learners across the UGME to PGME continuum.

**Data extraction**

A data extraction form was modified from our two prior reviews (Gordon et al. 2020; Daniel et al. 2021) and based on BEME Guidance (Hammick et al. 2010). This form was uploaded into
Google Sheets to allow for sharing of extracted data. Two studies were extracted by all authors to pilot the form, and a team meeting was held to ensure that all authors had a shared understanding of the data to be extracted. Pairs of authors were then assigned a group of primary studies. They independently extracted information for the table and then met to ensure consensus prior to placing their data in Google Sheets. Discrepancies were resolved with involvement of the lead or senior author (JS or MD).

Data extracted included:

- Article identifiers (authors, title, journal, type of article, length (# pages), month of publication)
- Context (type and number of learners, education focus or specialty, region of the world, organization responsible)
- Characteristics of the educational development (synchronous, asynchronous or both; approach to instruction; transition of established offerings online vs new educational developments)
- Stated purpose of development
- Brief summary (description) of development
- Techniques used to increase virtual engagement
- PICRAT code and intervention type (see below)
- Resources (cost, time, and material resources needed to implement)
- Explicit theories or frameworks underpinning the development
- Kirkpatrick’s outcomes
- Summary of results
PICRAT technology integration framework

According to Selwyn (2010), the promises of new technologies often go unrealized because it’s hard to imagine the possibilities created by new tools. Thus, we decided to use a technology integration framework known as PICRAT (Kimmons et al. 2020) to analyze how learners and teachers engaged with technology during the pandemic. We aimed to determine the extent to which educators integrated technology to enhance teaching during the rapid shift to online instruction. The tool is applied by completing two statements: 1) the student’s relationship to technology is Passive, Interactive or Creative (PIC); and 2) the teacher’s use of technology Replaces, Amplifies or Transforms (RAT) previous teaching practice. The answers create a 2 letter code on a matrix that describes a continuum of technologies’ potential to engage learners and transform instructional practices.

Quality assessment

To assess quality of the included studies, we addressed two distinct and mutually important elements: 1) risk of bias in study methodology and 2) risk of bias in study reporting.
The Medical Education Research Quality Instrument (MERSQI) was used to assess the quality of study methodology (Reed et al. 2007; Cook and Reed 2015). Points were assigned to various domains including study design, sampling, type of data, validity evidence for the evaluation instrument scores, data analysis and outcomes. Frequencies of scores by domain were tallied.

A visual RAG (red, amber, green) ranking system was used to assess risk of bias in study reporting. This tool was previously used by Gordon et al. (2018), Gordon et al. (2019b) and Gordon et al. (2020), and originally modified from Reed et al. (2005). The areas assessed included underpinning theories, resources, setting, educational methods, and content (Table 1). Items were judged to be of low risk of bias (green), moderate risk of bias (amber) or high risk of bias (red).

[Insert table 1 near here]

**Synthesis of evidence**

Data from the extraction form was summarized to provide a narrative summary (description). A visual summary of the data was developed, similar to that found in Daniel et al. (2021). Kirkpatrick’s outcomes were summarized (justification). A meta-analysis was considered, but the interventions were too heterogeneous to make comparison feasible. A thematic analysis was performed as outlined by Clarke and Braun (2013) concerning lessons learned (implications).
Results:

A total of 11,111 records were identified through database searching and an additional 23 records were found by hand searching MedEdPublish. After duplicates were removed, 7,164 records remained. These were screened by title and abstract and 6,742 records were excluded. Inter-rater reliability at this stage was $\kappa=0.91$. The full text articles were then assessed for eligibility and 283 were excluded with reasons. Of the 139 remaining articles, 58 were excluded because they were not focused on remote / online learning. The remaining 81 articles were combined with the 81 articles focused on remote / online learning previously identified by Gordon et al. (2020) and Daniel et al. (2021). This resulted in a total of 162 studies since December 2019, when COVID was first reported. Of these, 55 articles were excluded as they pertained to the pivot from clinical or workplace-based learning (Grafton-Clarke et al. 2021). An additional 51 articles were excluded because they focused on postgraduate medical education (Khamees et al. 2021). Fifty-six studies involving a pivot from the ‘classroom’ to remote / online learning in undergraduate medical education were included in our final analysis. The PRISMA flow diagram for article identification is shown in Figure 1.

[Insert Figure 1 near here]

Appendix 2 provides a written summary of all the primary studies included in this review. For the sake of brevity, in the results section, we have not listed specific articles if the associated data is easily identifiable in Appendix 2. Figure 2 provides a visual summary of this data.
**Geographic origin of studies**

The geographic distribution of the included studies is demonstrated in Figure 2, origin of publication and Appendix 2, column region. Twenty-two studies (39.3%) were performed in the United States, one (1.8%) in Canada, one (1.8%) in Central America, one (1.8%) in South America, five (8.9%) in Europe, 17 (30.3%) in Asia, seven (12.5%) in the Middle East and two (3.6%) in Oceania.

**Month of publication**

The distribution of publication month can be found in Figure 2, month of publication and Appendix 2, column month. The earliest articles were published online in March (n=2). August was the month with the most articles published (n=11).

**Type of publication and journals where studies were published**

The type of publication fell into four main categories: four (7.1%) were letters to the editor, eleven (19.7%) were brief reports/innovations, seven (12.5%) were articles/commentaries, and thirty-four (60.7%) were original research articles. Letters to the editor and brief reports/innovations were all 1-2 pages in length, whereas articles/commentaries and original research articles ranged from 3-15 pages, with the latter having a specific research focus. Thirty-one studies (55.6%) were published in medical education journals, yet the top ranked journals according to impact factors (e.g., Academic Medicine, Medical Education, Medical Teacher, Teaching and Learning in Medicine) were either disproportionately represented by brief reports
or not represented at all (Figure 2, type of publication / medical education journals and Appendix 2, organized by type of publication).

Participants, institutional setting and medical specialty

The number of learners involved in each study ranged from six to 875 (Figure 2, number of participants in each study and Appendix 2, column learners). Twenty-seven of the studies (48.2%) had less than 100 participants, 13 (23.2%) had between 100 and 299 participants, and ten (17.9%) had 300 or more participants. Six studies (10.7%) did not report the number of participants.

Sixteen studies (28.6%) reported on remote interventions as a substitute for face-to-face ‘classroom’ activities for preclinical learners, 16 (28.6%) for clinical learners, and two (3.5%) for both pre-clinical and clinical learners. Twenty-two studies (39.3%) did not specify the level of learner (Figure 2, medical student level and Appendix 2, column learners). One study included both medical students and residents (Chandrasinghe et al. 2020), and one study (Newcomb et al. 2021) included medical students and faculty. Seven studies included medical students working with learners from various allied health professions (e.g., dentistry, nursing, pharmacy, social work, physical therapy, counselling, speech therapy, nutrition, midwifery, athletic training).

Fifty of the studies (89.3%) were performed in a university setting, whereas five (8.9%) occurred in an academic hospital and one (1.8%) was in a multi-institutional setting (Figure 2, who is responsible for educational delivery and Appendix 2, column organization responsible).
Several content areas and medical specialties were represented. Thirteen of the studies (23.2%) focused on basic science (e.g., anatomy, physiology, pathology, pharmacology), three (5.4%) on health systems science (e.g., health equity, quality improvement, health policy), five (8.9%) on interprofessional education, and six (10.7%) on clinical skills (e.g., communication skills, physical exam skills, oral presentations). Five studies (8.9%) occurred in surgery and two (3.6%) in surgical subspecialties (neurosurgery and otolaryngology). One study (1.8%) occurred in internal medicine, with one additional study (1.8%) in a medical subspecialty (cardiology). Two studies (3.6%) occurred in pediatrics and one (1.8%) in ophthalmology. Seventeen studies (30.3%) did not report an area of focus or medical specialty (Figure 2, education focus or specialty and Appendix 2, column education focus or specialty).

**Transition of established offerings online versus new educational developments**

Forty-one studies (73.2%) described the rapid transition of in-person educational offerings to online formats utilizing similar instructional materials and/or approaches (Figure 2, established vs new and Appendix 2, column transition of established offerings online vs new educational developments). In these studies, face-to-face ‘classroom’ experiences (e.g., lectures and small groups) were replicated using video-conferencing software to achieve the same learning objectives among similarly sized groups of students. Simulated experiences (e.g., communication and physical exam skill building) were replicated by leveraging breakout rooms to facilitate remote standardized patient (SP) interactions, role-plays and telesimulations. Laboratory experiences (e.g., anatomy dissections and histology / pathology slide reviews) were largely replaced with online dissection videos, lectures and small group discussion. One unique
study utilized mixed-reality technology to simulate three-dimensional space, allowing for physical interaction with the content studied (Wish-Baratz et al. 2020).

Fifteen studies (26.8%) described new educational developments created in response to the pandemic. A few of these developments provided alternative educational experiences in lieu of activities that had ceased: Clemmons et al. (2020) built a pandemic course, covering basic, clinical and health systems science content related to COVID-19; Walton et al. (2020) developed a health policy course that involved writing policy briefs on current topics; and Prasad et al. (2020) developed an interprofessional telesimulation focused on newly pertinent content, including the use of personal protective equipment (PPE) and crisis resource management. A handful of developments specifically embraced novel technologies or features of the online environment: Wintraub et al. (2020) investigated device-accessory pairings (e.g., a chest mount and a GoPro) to facilitate remote visualization of patient encounters, Moro and Stromberga (2020) used serious games (e.g., Kahoot and the King’s Request) to teach anatomy, Iqbal et al. (2020) used a cloud-based messaging and file sharing application called Telegram, and Mendez-Reguera and Lopez Cabrera (2020) designed a contest wherein students created and voted on their favorite immunology memes.

One article (1.8%) described both a transition of established content online as well as a new innovation: Garg et al. (2020) moved an existing health systems science and social justice course online, and designed new content specifically related to COVID-19.

*Synchronous vs. asynchronous vs. a combination of synchronous and asynchronous formats*
Twenty-five (44.7%) of the educational offerings described synchronous learning, four (7.1%) asynchronous learning and 27 (48.2%) some combination of the two (Figure 2, synchronous / asynchronous learning and Appendix 2, column synchronous and/or asynchronous).

Synchronous educational offerings utilized video-conferencing platforms to provide teaching in real-time using a variety of instructional methods (see below). While some synchronous learning activities were purely passive (e.g., online didactics), most offered some level of interactivity. Common mechanisms used to foster virtual engagement within the video-conferencing platforms were discussion, chat, question and answer (Q & A) sessions, breakout rooms, the whiteboard annotate feature, and polling. Other platforms (e.g., WhatsApp, Google documents, Kahoot!) were used in conjunction with video-conferencing tools to allow for further connectivity via instant messaging, collaborative workspaces, or gamification. In addition, teachers used real-time feedback / coaching, facilitated debriefs, and reflective exercises during synchronous sessions to further connect with learners (Figure 2, mechanisms to foster engagement and Appendix 2, column techniques used to increase virtual engagement).

Asynchronous educational offerings used learning management systems or shared online storage sites to house educational materials, such as lectures, videos, and readings. Learners were able to access materials and submit assignments in a flexible, self-directed manner that best suited their needs. Sud et al. (2020) described an ophthalmology curriculum consisting of annotated PowerPoints and multiple choice questions. Mahima et al. (2020) and Singal et al. (2020) both reported on online anatomy curricula, with the former requiring submission of assignments. Sud et al. (2020) and Iqbal et al. (2020) used applications (e.g., Telegram and WhatsApp) that
allowed for instant messaging and sharing of files. This allowed communication among participants and crowdsourcing of resources, promoting interactions, even as learners accessed course content at different times.

Educational offerings that incorporated synchronous and asynchronous components harnessed both the flexibility and autonomy of asynchronous content, while providing opportunities for real-time engagement, interactivity and discourse in synchronous formats. These educational offerings incorporated a combination of instructional methods, most commonly asynchronous lectures with synchronous small groups, to allow students to engage in critical thinking and apply their knowledge and skills.

**Instructional methods**

The educational method most commonly used was the lecture or ‘didactic’ (n=40) with or without interactive components (Figure 2, type of instruction and Appendix 2, column instructional methods). Podcasts, webinars and narrated video-conferences were included under the broader category of ‘didactics’, as all these modalities involved a teacher (or a peer) giving an educational talk to students with an aim of transmitting knowledge. The majority of developments that included didactics combined them with other instructional methods. For example, educational offerings that incorporated telesimulation, clinical skills or SP exercises often began with an introductory didactic (Martinez et al. 2020; Rossasco et al. 2020; Sa-Couto and Nicolau 2020; Newcomb et al. 2021).
Small groups (n=26) were the second most commonly utilized instructional method. Small groups provided opportunities for discussions with faculty and peers. Several small groups utilized the format to specifically connect students considering a specialty with faculty in that discipline (Burns and Wenger 2020; Geha and Dhaliwal 2020; Spaletra et al. 2020; Steehler et al. 2021; Tan et al. 2020; Thum Dicesare et al. 2020). Four developments utilized online small groups for anatomy tutorials (Naidoo et al. 2020; Parker et al. 2020; Srinivasan 2020; Wish-Baratz et al. 2020). Three offerings utilized small groups to foster interaction among students from different health professions (Jones et al. 2020; Rutledge et al. 2020; Robertson et al. 2021). A few combined small groups with simulation or SP activities (Martinez et al. 2020; Mohos et al. 2020) as a mechanism to foster reflection and debriefing.

Clinical skills / telesimulation / procedural skills sessions (n=12) were the third most commonly utilized instructional method. Several offerings incorporated standardized patients (SPs) remotely: Rucker et al. (2020) detailed the use of Zoom breakout rooms to replicate formerly in-person SP activities; Mohos et al. (2020) and Sudhir et al. (2020) described communication skills using SPs; and Martinez et al. (2020) outlined telemedicine interactions with SPs. Newcomb et al. (2021) used role-plays to teach communication skills. Telesimulation was also successfully employed via available technological tools in a few studies, including by Prasad et al. (2020) for maternal and neonatal emergencies and Sa-Couto and Nicolau (2020) for a variety of emergent clinical scenarios. One exemplary article, Rosasco et al. (2020), compared in-person behavioral health screenings of SPs by an interprofessional group of learners to a telesimulation experience with a telepresence robot. Schlégl et al. (2020) and Co and Chu (2020) described teaching surgical skills remotely through the innovative use of cameras and video-conferencing.
technology. Proper positioning of devices allowed teachers to demonstrate skills, then assess learner performance of the skills while providing feedback.

A handful of studies described the implementation of team-based learning (TBL, n=3) and problem-based learning (PBL, n=2) via video-conferencing platforms. For TBL, Vollbrecht et al. (2020) and Gaber et al. (2020) described utilizing the breakout room features on Microsoft Teams and Zoom, respectively, to achieve the TBL format, while Jumat et al. (2020) described creating new video-conferencing sessions and use of an instant messaging platform to manage side conversations. Different programs utilized different approaches to conduct the individual readiness assurance tests (iRATs), group readiness assurance tests (gRAT), and modified TeamLEAD readiness assurance processes. Jumat et al. (2020) noted that currently no all-encompassing video-conferencing platform exists to facilitate TBL, and the ideal platform would need to facilitate both breakout rooms and administration of individual and group assessments.

Of the articles that described remote PBL, Coiado et al. (2020) used Zoom with Google Docs as a virtual whiteboard, Alkhowailed et al. (2020) used Blackboard with Zoom as a backup, and Rehman and Fatima (2021) used Microsoft Teams with WhatsApp group chat. Authors described both PBL and TBL formats as time- and energy-intensive, requiring increased administrative and faculty support, due to new facilitator responsibilities that included managing participants with regard to muting and unmuting, staying on time, and maintaining engagement (Alkhowailed et al. 2020; Coiado et al. 2020; Jumat et al. 2020; Vollbrecht et al. 2020).

Additional instructional methods such as debates (Lapane and Dube 2020), discussion boards (Jones et al. 2020; Mendez-Reguera and Lopez Cabrera 2020), mentorship/networking sessions
(Thum DiCesare et al. 2020), group assignments (Cuschieri and Calleja Agius 2020; Geha and Dhaliwal 2020; Liang et al. 2020; Rutledge et al. 2020) and question and answer sessions (Clemmons et al. 2020) were less commonly described.

**PICRAT technology integration framework**

The PICRAT Matrix shows the number of developments assigned to each category on a continuum (Figure 2, PICRAT: technology integration framework and Appendix 2, columns PICRAT code and PICRAT intervention label). Studies that described multiple interventions were assigned more than one PICRAT category. Our findings revealed that teachers’ use of technology during the pandemic largely ‘replaced’ existing instructional methods. Within this group, the majority incorporated technology for interactive replacement (IR, n=41), several used technology for passive replacement (PR, n=26), and only a handful leveraged technology for creative replacement (CR, n=3). There were small numbers of exemplary uses of technology that ‘amplified’ traditional practices (PA, n=1; IA, n=5; and CA, n=2), including a multi-view surgical demonstration (Co and Chu 2020), serious games (Moro and Stromberga 2020), point-of-view wearable devices for clinical skills instruction (Wintraub et al. 2020), and student-developed memes (Mendez-Reguera and Lopez Cabrera 2020). Only one study ‘transformed’ learning through technological integration to an extent that could not have been otherwise achieved. This study was Wish-Baratz et al. (2020), which used a HoloAnatomy software suite to implement mixed-reality anatomy dissections.

**Resources explicitly mentioned by authors**
The most commonly listed technological resource was video-conferencing software, including Zoom (n=25), Microsoft Teams (n=5), Google Meet (n=2), WebEx (n=1), DingTalk (n=1), and unspecified applications (n=6) (Appendix 2, column resources). Several highlighted use of features within the software application including breakout rooms (e.g., Gaber et al. 2020; Jumat et al. 2020; Thum DiCesare et al. 2020; Vollbrecht et al. 2020) and virtual whiteboards (Coiado et al. 2020). Others highlighted messaging platforms such as WhatsApp or WeChat (e.g., Alkhowailed et al. 2020; Gaber et al. 2020; Naidoo et al. 2020; Roy et al. 2020; Sud et al. 2020; Zhang et al. 2020; Rehman and Fatima 2021). One utilized social media (Facebook group) to target specific students (Chandrasignhe et al. 2020). Learning management systems (Alkhowailed et al. 2020; Fatani 2020; Jones et al. 2020; Joseph et al. 2020; Kim et al. 2020; Sud et al. 2020; Zhang et al. 2020), cloud sharing via Box file systems (Liang et al. 2020), Google Documents/Classroom (Mahima et al. 2020; Sharma et al. 2020; Singh et al. 2020), Telegram (Iqbal et al. 2020), and other storage software (Rutledge et al. 2020) allowed for easy distribution of learning materials and submission of assignments. Examples of novel software to enhance interactivity and engagement included gamification platforms (Moro and Stromberga 2020; Sa-Couto and Nicolau 2020), interactive pathology slide viewing with live annotating capabilities (Parker et al. 2020), as well as interactive polling (Moro and Stromberga 2020; Srinivasan 2020; Tan et al. 2020; Vollbrecht et al. 2020). Innovative use of technologic hardware in medical education included Go-Pro cameras with different accessories (Wintraub et al. 2020), telepresence robots to enhance clinical skills education in the virtual environment (Rosasco et al. 2020), and mixed-reality headsets to facilitate anatomy education (Wish-Baratz et al. 2020).

The majority of articles did not specifically mention direct financial or human resource costs. Some described use of freely available tools (Iqbal et al. 2020; Mahima et al. 2020; Sa-Couto...
Several publications broadly highlighted the importance of adequate faculty support and training resources for successful implementation (Jumat et al. 2020; Khalil et al. 2020; Kim et al. 2020; Mahima et al. 2020; Mohos et al. 2020; Naidoo et al. 2020; Sa-Couto and Nicolau 2020; Verma et al. 2020; Vollbrecht et al. 2020). A few provided details regarding time investment for implementation (Clemmons et al. 2020; Steehler et al. 2021). Joseph et al. (2020) noted that new faculty appointments were necessary to quickly produce online educational content. Authors occasionally mentioned other human resources in the form of standardized patients (Martinez et al. 2020; Mohos et al. 2020; Rosasco et al. 2020; Rucker et al. 2020; Shahrviní et al. 2020; Sudhir et al. 2020; Newcomb et al. 2021), administrative support (Jumat et al. 2020; Zhang et al. 2020), technologic support (Coiado et al. 2020; Garg et al. 2020; Jeong et al. 2020; Singh et al. 2020), and creators/facilitators of faculty development tools and sessions (Co and Chu 2020; Fatani 2020; Naidoo et al. 2020; Singh et al. 2020).

Theoretical underpinnings explicitly mentioned by authors

When evaluating underpinning theory in text, we were flexible and applied a broad definition of ‘theory’ including established theories, frameworks, principles, models, concepts, and approaches (Mann et al. 2011). Despite this, less than half of papers (n=21, 37.5%) explicitly described the use of ‘theories’ supporting the development (Appendix 2, column explicit 'theories' underpinning development). Eight papers mentioned ‘grand’ overarching theories as described by Laksov et al. (2017), e.g. community of inquiry, social learning theory, cognitive learning theory, active learning theory, and cognitive apprenticeship (Fatani 2020; Geha and Dhaliwal 2020; Jumat et al. 2020; Naidoo et al. 2020; Prasad et al. 2020; Rucker et al. 2020;
Shahrvin et al. 2020; Tan et al. 2020). The remaining authors used ‘mid–range’ theories or recognised approaches, e.g. reflective learning, gamification, etc. Text in papers often cited instructional approaches (e.g., TBL, PBL), rather than explicitly describing the theories upon which these approaches were based. Mention of key words (e.g. ‘flipped classroom’) left undefined often resulted in a lack of clarity over whether the underpinning pedagogy had been embedded. The pivoting of education from face-to-face teaching to online models was at times colloquially called ‘flipping the classroom’, contributing to this ambiguity. Sometimes more than one theory and approach was applicable and adopted (Fatani 2020; Geha and Dhaliwal 2020; Jumat et al. 2020; Naidoo et al. 2020; Tan et al. 2020). Surprisingly few articles included mention of instructional design or technology-enhanced learning: studies described the affordances of the technology rather than grounding the work in theory or approaches specifically applicable to online or digital learning. Rutledge et al. (2020) was an exception, describing a 4-P framework (Planning, Preparing, Providing, and Performance) which is similar to the ADDIE (Analysis, Design, Development, Implementation and Evaluation) framework (Allen 2006) used in instructional design.

**Summary of Kirkpatrick’s Outcomes**

Virtually all studies (n=54, 96.4%) assessed Kirkpatrick’s level 1 (reaction, satisfaction), but fewer than half (n=23, 41.1%) assessed levels 2a or 2b (change in attitudes or change in knowledge or skills). No study described Kirkpatrick’s levels 3 or 4 (Figure 2, *Kirkpatrick’s outcomes and Appendix 2, columns Kirkpatrick outcome and results*). Student perceptions of online learning were neutral to favorable in most studies (n=47, 83.9%), although many questions interrogated newly online material as an abstract concept of acceptable or not
acceptable, or a desirable or undesirable addition to traditional learning, rather than a comparison of on-line versus in-person education. Among the papers reporting negative impressions (n=9) from the students, most involved anatomy, pathology, physical examination skills or surgical skills. Only five papers commented on faculty perceptions, which showed a lower rate of positivity than student opinions. In the studies documenting outcomes 2a or 2b, a few (n=3) showed positive impacts on student performance, and each of these studies was assessing student familiarity and facility with new online or telepresence technology. A similarly small number of studies (n=3) showed a negative effect on student performance when comparing online teaching to traditional in-person alternatives. Most studies (n=17) reported neutral or unclear impacts on student performance, often simply assessing whether student knowledge of a topic was higher at the end of an intervention than at the beginning, but not making a comparison between two cohorts with different interventions.

Quality assessment / risk of bias

Risk of bias in study methodology

Methodological rigor was assessed using the MERSQI, which revealed lower scores across most domains (Figure 2, Risk of Bias in Methodology and Appendix 2, column risk of bias in study methodology (MERSQI)). Most studies were determined to be of low to moderate quality, but a few exhibited exceptional rigor across multiple domains (Fatani 2020; Kim et al. 2020; Naidoo et al. 2020; Parker et al. 2020; Rosasco et al. 2020; Rutledge et al. 2020; Newcomb et al. 2021; Rehman and Fatima 2021). Since we applied the MERSQI to all study methodologies, not just experimental, quasi-experimental or observational studies, there were
some gaps (i.e., not applicable (N/A)) in various domains of scoring. Additionally, brevity in reporting of some studies (e.g., letters to the editor) limited description of some items and missing items received a score of 0. Thus, we decided not to report total MERSQI scores.

An analysis of the categories revealed several patterns in the data (Table 2). Single group cross-sectional study designs were the most common (n=44, 78.6%). Ten (17.8%) used a single group pre-post design, one (1.8%) utilized a nonrandomized two group design and one (1.8%) reported on a randomized control trial. Fifty-five studies (98.2%) sampled only one institution and only one study (1.8%) sampled three or more institutions. Sampling response rates were distributed with 21 (37.5%), 13 (23.2%), and 18 (32.2%) studies with response rates of <50% or not described, 50-74% and > 75%, respectively. For three studies (5.4%) a sampling response rate was deemed N/A. Type of data presented focused on assessment by study participants for 45 (80.3%), whereas eight (14.3%) presented objective data, and three (5.4%) were not described. Validity evidence was not described in 39 studies (69.6%) representing the weakest domain on methodological assessment. Data analysis sophistication was low and mostly descriptive, with only 11 studies (19.6%) providing any tests of statistical inference. Outcomes in 40 studies (71.4%) focused on satisfaction/attitude/perception, whereas 16 (28.6%) noted changes in knowledge/skills.

[Insert Table 2 near here]

Risk of bias in study reporting
We visually portrayed the risk of bias in study reporting (**Figure 2, Risk of Bias in Reporting** and **Appendix 2, column risk of bias in study reporting**). Reporting quality was low to average across studies and correlated with both article length and type. Short (1-2 page) letters to the editor exhibited the highest risk of bias, followed by brief reports, then articles. Five studies (Jumat et al. 2020; Lapane and Dube 2020; Naidoo et al. 2020; Rutledge et al. 2020; Rehman and Fatima 2021) were determined as low risk of bias in >4 domains. The domains identified at highest risk of bias were underpinning and content with 38 (67.9%) and 30 (53.4%) studies not reporting (i.e., red), respectively. Resource, setting and educational methods were more often found to be at moderate risk of bias (i.e., amber), with the majority of studies providing at least some details.

**Thematic analysis of lessons learned**

We completed a thematic analysis of the lessons learned as reported by study authors (**Appendix 2, column lessons learned**). The following themes were identified:

*Added challenges / considerations in the online environment*

Pivoting to online remote learning posed significant challenges to the already daunting task of creating engaged learning in a classroom. Studies of most developments, especially those with synchronous learning, highlighted absence of non-verbal cues and suboptimal social interaction within virtual experiences as threats to engaged learning (Shahrvini et al. 2020; Zhang et al. 2020). The unfamiliar nature of virtual experiences and technical glitches added barriers and challenges for some learners, especially those who were easily distracted or had reticent tendencies. Some learners lacked the focus and self-discipline necessary to remain actively
engaged during sessions (Coiado et al. 2020; Shahrvini et al. 2020; Zhang et al. 2020). Without extra effort from instructors to facilitate active learning, learners easily reverted to passive participation (Coido et al 2020; Tan et al. 2020). For instructors, the burdens of creating emergent remote learning, the steep learning curve in catching up with technological advances and the need to adjust their own teaching perspectives posed ‘extraneous cognitive load’ to deliver high quality teaching (Liang et al. 2020; Verma et al. 2020; Steehler et al. 2021). ‘Digital / video-conferencing fatigue’ was evident, and compromised learning (Shahvini et al. 2020; Tan et al 2020). The ‘big elephant’ in the room (i.e., the unique stressors imposed by the pandemic, including sickness, uncertainty, stay-at-home orders, lack of childcare, and the ‘blurring’ of work-life balance) compounded the difficulties of adjusting to remote education. Educators warned others to be mindful about these unique situations and set reasonable expectations of learners (Liang et al. 2020; Vollbrecht et al. 2020).

Best practices for online classrooms

All developments used a videoconferencing platform as a primary approach to creating virtual experiences aimed at replacing in-person classrooms. An overarching theme emerged that ‘instructor presence and an interactive style’ significantly enhanced teaching quality and learner satisfaction (Fatani 2020). Learners valued extra efforts from the instructors (Mendez-Reguera and Lopez Cabrera 2020). Some simple adjustments in teaching practices (e.g., asking students questions more frequently with significant pauses (>10 s), anticipating everything to take extra time online, and/or including quizzes and polling to check learners’ understanding of content) significantly augmented learning (Vollbrecht et al. 2020). Other best practices recommended by
authors included providing instructions or guides for efficient navigation and use of the video-conferencing platforms, being readily available for clarification and troubleshooting using an app for instant communications, or holding ‘virtual office hours’ (Cuschieri and Calleja Agius 2020; Mahima et al. 2020; Vollbrecht et al. 2020). Giving private and timely feedback to less engaged learners via the chat function also enhanced participation (Coiado et al. 2020).

*Using technology to enhance remote engagement*

Chat functionality available in most video-conference platforms created unique opportunities for students to participate in classroom discussions and ask questions without disrupting class flow. This allowed some students to more easily engage with content and lectures through sharing of related references and knowledge and working more with other students (Garg et al. 2020). Remote teaching encouraged faculty to innovate their teaching and integrate tools they may have been reluctant to try before. These included games (Moro and Stromberga 2020), case-based scenarios, and flipped classrooms. Students appreciated digital slide annotation (Parker et al. 2020) and polling tools such as Poll Everywhere or Zoom to increase engagement and gauge understanding (Srinivassan et al. 2020; Tan et al. 2020; Vollbrecht et al. 2020). Augmented reality or mixed-reality resources such as HoloAnatomy were mentioned in several studies but were generally considered costly with high support needs (Wish-Baratz et al. 2020). In the absence of available in-person clinical skills instruction, mixed-reality was considered an acceptable substitute short-term (Wintraub et al. 2020).
Incorporating asynchronous learning to increase flexibility

Before the pandemic, many instructors prioritized in-class synchronous learning over online asynchronous learning. Some used asynchronous learning only in small parts as pre-session assignments. In many studies, however, students valued the flexibility and self-paced nature of online lectures and the ability to complete assignments asynchronously at their convenience. (Shahrviní et al. 2020; Vala et al. 2020; Zhang et al. 2020). In one study of telehealth education, faculty learned after gaining more experience that some previously in-class modules could be successfully implemented and offered completely online in the future (Rutledge et al. 2020). Given instructors could not replace all learning components on a virtual platform, authors noted that it would be wise to adapt, ‘think outside the box’ and put effort to flip or transform the classroom whenever possible (Gaber et al. 2020; Vollbrecht et al. 2020).

Additional resources required for remote learning

Considering resources required to transition to remote teaching, time was of the utmost importance. Both instructors and learners needed additional time for preparation and practice using technology (Burns and Wenger 2020; Zhang et al. 2020). Everything took longer and occurred at a slower pace in the virtual environment than in-person, particularly with procedure- and exam-related activities (Burns and Wenger et al. 2020; Coiado et al. 2020; Vollbrecht et al. 2020). Many authors highlighted the reliance on technology resources and the need for a reliable internet connection for teaching, learning, and communicating. Most developments anchored on video-conferencing software (free or premium) and learning management software for course content sharing and organization. Many studies leveraged readily available and free platforms
such as communication apps (WhatsApp, Slack, etc.) to provide an easy and efficient way for students to connect with instructors and each other (Gaber et al. 2020; Geha and Dhaliwal 2020; Vollbrecht et al. 2020). In addition, technical and administrative support staff needs increased with the rapid transition as some instructors did not have the required technology literacy (Jumat et al. 2020). In synchronous sessions, having additional assistants or instructors (to monitor chat for questions, transition to breakout rooms, etc.) could contribute to a successful session (Vollbrecht et al. 2020).

Exemplary development strategies for online learning

Some noteworthy approaches to the process of educational development were identified by study authors. A medical school in the United Kingdom (Joseph et al. 2020) chose to empower their faculty to innovate online learning based on their own interpretation of educational principles, rather than any top-down directive. The authors attributed the rapid and successful transition to ‘imaginative, committed and creative’ faculty who embraced the power of learning technology. Students enjoyed a broad range of learning technologies and various instructional formats, as well as, being involved in short-loop feedback for rapid iterations of the sessions. An interprofessional group of faculty educators who transitioned their existing interprofessional curriculum on discharge planning to a completely virtual experience found themselves modelling the interprofessional education core competencies during the development process, and attributed their success to their own interprofessional teamwork (Robertson et al. 2020). Eight neurosurgery residency programs in the US modeled an exemplary collaboration to organize the first live, cross-institutional virtual training camp to deliver standardized neurosurgical
educational content to medical students during the pandemic. Collectively, they afforded the diverse content and availability of experts from numerous programs creating an accessible educational venue to a large number of students with decreased cost (Thum DiCesare et al. 2020).

Discussion

Summary of main results

Amidst the stress of the COVID-19 pandemic, educators rapidly pivoted classroom, clinical skill and laboratory learning online to safely continue education for medical students. Their experience, summarized in this review, can provide considerable guidance for educators. While the majority of developments represented the transition of existing offerings online, new innovations were also reported. Educators made use of both synchronous and asynchronous learning formats to promote flexibility and interactivity, and described a myriad of ways to foster virtual engagement in both formats. Interactive didactics and small groups were the most common instructional methods, though educators also utilized telesimulation, group assignments and a variety of other formats to provide opportunities for discourse, critical thinking, and application of knowledge and skills. Technology, including video-conferencing platforms and their embedded features, learning management systems, instant messaging applications, and other software and hardware, allowed teachers to ‘replace’ or even ‘amplify’ many previously in person activities. However, the potential for technology to ‘transform’ learning remained largely unrealized. Authors described many challenges inherent to the online learning environment that
must be considered as we transition to a ‘new normal’: ‘video-conferencing fatigue’ threatened engaged learning, social interactions were suboptimal, and faculty instructional burden was high. Authors also highlighted best practices for remote teaching and described how to leverage technology to enhance engagement, incorporate asynchronous learning to promote flexibility, and apply exemplary development strategies for online learning.

**Quality of the evidence base**

As indicated by both the MERSQI and the RAG risk of bias in reporting tool, the overall quality of the evidence base one year into the pandemic was modest at best. We assert the findings are reflective of a pervasive problem with the primary literature in medical education, a problem that has been repeatedly observed by authors of evidence syntheses. This problem needs to be urgently addressed by authors and editors alike.

We want to highlight two interrelated yet distinct areas for improvement of the primary literature - study methodology and study reporting. Most studies exhibited high risk of bias across both quality assessment tools. Rutledge et al. (2020), Rehman and Fatima (2021), and Naidoo et al. (2020) were notable exceptions, being exemplary in both methodologic rigor and reporting. As such, they may serve as a guide for other educators and scholars.

This review and the companion piece on postgraduate medical education (PGME) by Khamees et al. (2021) are the first reviews with the complementary use of the MERSQI and the risk of bias in reporting tool for a holistic approach to assessing primary studies. Prior reviews (e.g.,
Gordon et al. 2020) and the concurrent review by Grafton-Clarke et al. (2021) recognized the importance of assessing both methodologic quality and study reporting but attempted to use the Cochrane Risk of bias tool or ROBINS-I (Sterne et al. 2016) for study methodology. That approach could only be meaningfully applied to a handful of studies due to challenges with the range of methodologies employed. While the MERSQI is better suited to accommodate such a range of methodologies, it has limitations. The domains that constitute the MERSQI are most conducive for assessing quality of experimental, quasi-experimental or observational studies, yet we extended its application to other study types in this review. Thus, we elected to report MERSQI subscores while highlighting gaps in the data such that readers might evaluate the overall quality of the evidence using a constructivist / interpretivist approach, in lieu of presenting total scores which align with a more post-positivist approach.

Through MERSQI assessment, we derived assertions for future work: 1) While a single group cross-sectional study design was appropriate during the early phase of the pandemic, more rigorous study designs are now needed; 2) Multi-institution sampling was more prevalent in the PGME review (56.9 vs. 1.8%). UGME should follow PGME’s lead and work to break down current institutional silos, focusing more on shared problems and solutions, instead of exclusively focusing on local issues; 3) Validity evidence for evaluation instrument scores was practically non-existent. Given valid outcome measures are critical to decision-making that directly impacts learners, such a finding is unacceptable and should be urgently rectified; 4) Finally, the predominance of studies focused on satisfaction / reaction limits the range of conclusions that can be drawn concerning educational effectiveness. Educators should attempt to determine efficacy evidence through multi-level evaluation.
The risk of bias in reporting tool has highlighted the need for improved reporting across all domains to facilitate replication of developments. While inadequate reporting does not necessarily indicate methodological weakness, it tampers the strength and utility of evidence. Resources, setting, and educational methods were more often reported on than underpinning or content, however, there were still notable gaps. For example, many articles were given an amber for reporting on technology resources, yet few reported on costs, in terms of financial and human resources, leaving institutions that fund education unclear on needs moving forward. This is not a new finding (Gordon et al. 2013), but in the context of the current global and rapid shifts in teaching practice, represents a distinct limitation to dissemination and replication of potentially useful works. We hypothesize that content reporting in particular may have suffered from restrictions placed on article lengths. Some of the most innovative articles in our sample were brief reports, yet the sparse details make it difficult to build on the evidence base. Thus, we encourage authors and journals to utilize creative means for providing content via supplemental digital appendices or links to online repositories.

The lack of underpinning theories needs utmost attention. The use of theories helps educators make informed decisions about design, development and implementation. Without explicit descriptions of theories or conceptual frameworks, authors in this review failed to justify the reasons for, and provide rigor to the implementation and evaluation of their intervention(s), limiting transferability of their work. The last decade has seen growing awareness in the Technology Enhanced Learning (TEL) sector of educational theories (especially constructivist and social constructivist theories) (Millwood 2013). Given the plethora of readily applicable
frameworks (e.g., cognitive load theory by Sweller et al. (2011) and multimedia learning principles by Mayer (2005)), we expected more employment of these frameworks within instructional design. During the pandemic, however, many educators took a tools-based approach (focusing on the affordances of technology to pivot a course), and/or a materials-based approach (using the same materials to teach the course regardless of the format), rather than a pedagogy-centered approach that considers educational purpose, desired learning outcomes and context (Rapanta et al. 2020). Effective online learning results from careful, systematic design and planning, yet the time-pressures exerted by the unforeseen and emergent nature of the pandemic dramatically impaired the feasibility of such an approach (Branch and Dousay 2015; Hodges et al. 2020). Instructional design approaches that leverage technology to enhance learning may take more time to be operationalized. Future faculty development efforts should encourage incorporation of underpinning theory and evidence-informed practices into program development.

Comparison to prior reviews, literature and the other reviews in the series

In comparison to the rapid review by Gordon et al. (2020) and the scoping review by Daniel et al. (2021), which investigated the impact of COVID-19 on medical education broadly, this review offered a narrower focus on the pivot to online learning in UGME. This allowed for more in-depth reporting on educational formats, instructional methods, and technology utilized; assessment of quality to determine risk of bias; and robust thematic analyses of lessons learned.

Two prior works investigated the pivot to online learning in UGME. Gaur et al. (2020) conducted a literature review that examined the challenges and opportunities faced by medical
schools in implementing remote learning for *preclinical teaching* during COVID-19. Their review was neither rigorous nor systematic, though they did identify several parallel themes. Wilcha (2020) also conducted a brief qualitative review of the application and effectiveness of virtual teaching. This review only spanned May-June 2020. Thus, our review represents the most methodologically rigorous and comprehensive systematic review to date (spanning a full year since COVID began).

This review occurred in parallel with two other reviews: Khamees et al. (2021) focused on the ‘classroom’ to online pivot in PGME and Grafton-Clarke et al. (2021) focused on the workplace-based clinical learning pivot across the continuum of medical education. In this triad of reviews, we observed that UGME educators tended to face inward and focus on local needs, with few examples of multi-institutional collaborations. Educators in PGME, however, more commonly collaborated across institutions in an effort to provide regional, national and even global solutions to shared problems (Khamees et al. 2021). National specialty-specific organizations in PGME facilitated partnerships and contributed to education delivery more heavily than national UGME organizations. These findings highlight an opportunity for UGME educators to follow the lead of those in PGME in leveraging multi-institutional perspectives and resources to lessen the burden on individual educators and institutions.

Another striking trend noted across the triad of reviews was that the highest impact education journals were largely unrepresented. They clearly contributed to the international dialogue by publishing numerous perspectives, but they published a paucity of research articles. A notable exception was *Medical Education*, though these were almost exclusively brief reports. Several
possible explanations exist. For one, the rapidly composed manuscripts may not have met the level of rigor required of the highest impact journals with regard to methodology and detail of reporting. The highest impact journals may also involve a more rigorous review process that generates more rounds of revisions, such that manuscripts detailing relevant educational activities may have remained under review as of December 21, 2020. Moving forward, we urge authors and editorial boards to prioritize publishing high quality research studies, including studies representing extensions of previous pilots that may now have more robust evaluation data to build upon the existing evidence base.

To our knowledge, the UGME and PGME reviews were the first to apply the PICRAT technology integration framework (Kimmons et al. 2020) to examine the extent to which learners engaged with and teachers utilized technology. The results provide a cautionary tale that technology should only be a means to an end (i.e., enhanced learning). Decisions concerning the level of learner engagement (i.e., passive - interactive - creative) were mostly dictated by the subject matter and learning objectives. While we applaud the number of developments that were ‘interactive’, very few approaches were classified as ‘creative’ across these reviews. This represents an area of opportunity to enhance learner engagement in the future. Teachers most commonly applied technology to ‘replace’ formerly in-person activities. However, UGME educators used technology to ‘amplify’ traditional teaching practices to a far greater extent than in PGME (Khamees et al. 2020). UGME involves a higher proportion of classroom-based, clinical-skills, and laboratory activities compared to PGME, where the majority of learning occurs through clinical service and direct patient care. This may have driven creativity in UGME, as they had a larger number and wider variety of activities to deliver online. We would
encourage educational designers to utilize the PICRAT as a tool for planning technology integration to transform practice in the future, as the rush to find solutions is abating and the time for truly thoughtful design is upon us.

**Strengths and Limitations**

This review had many strengths. Similar to the last review from BEME on COVID-19, we completed the work on a rapid timeline without compromising methodological rigor, in large part due to the benefits of a large and relatively experienced team that had already developed content expertise from the prior reviews. The author group represented an international collaboration of medical students, residents, fellows and faculty with expertise in systematic reviews, medical education, online learning, and educational theory. The narrower focus on UGME ‘classroom’ pivots allowed us to complete a thematic analysis for lessons learned, incorporate a technology integration framework novel to systematic reviews (the PICRAT), and conduct quality assessments using both the MERSQI for study methodology and the RAG risk-of-bias reporting tool.

Our review also has inherent limitations. We adopted independent coding and consensus review, however operationalizing the various tools (e.g., PICRAT, MERSQI, RAG) was at times challenging, due to the inherent variability of the core material. The most notable limitation relates to the relatively short time taken for its completion. While this permits a rapid dissemination of important and relevant updates in the field, we acknowledge the ongoing evolution of the medical education landscape. Original research manuscripts that incorporate
more elements of theory and provide more detailed descriptions of educational offerings may be forthcoming.

**Practical recommendations for educators moving forward**

This review includes publications from the first year following the onset of the COVID-19 pandemic, during which educators rapidly adapted to minimize disruptions for learners. Thus, we must evaluate the educational offerings presented in this review within the context of ‘emergency remote learning’ (ERL). The standards for successful ERL differ from planned online learning, in that the expectation for ERL is to provide an adequate, rather than equal or superior, educational experience relative to the standard learning plan (Vollbrecht et al. 2020).

“The primary objective in these circumstances is not to recreate a robust educational ecosystem but rather to provide temporary access to instruction... in a manner that is quick to set up and is reliably available during an emergency or crisis” (Hodges et al. 2020).

The educational offerings summarized in this review undoubtedly fulfilled an urgent need to continue learning, but most were not intended as permanent replacements for face-to-face learning. While Kirkpatrick’s outcomes demonstrated that ERL developments were palatable during a crisis and that some knowledge and skill development continued, comparative studies that look at longitudinal outcomes to guide decisions about ‘what’s next’ are currently lacking.

As the initial shock of the COVID-19 pandemic wanes and online learning becomes increasingly more accepted as ‘the norm,’ the medical education community must envisage the post-pandemic future; this entails upholding a higher standard of online learning, one that leverages technology to optimize learning, rather than simply maintaining it. Moving forward, we
encourage educators to more fully explore technology’s potential to transform learning. We also encourage educators to urgently engage in studies that provide answers to the questions of what is desirable, sustainable, and effective long term.

As the dust settles, the findings of this review can provide insights into which aspects of online learning will likely persist in a post-pandemic world:

- Didactics achieve reasonable levels of learner engagement and thus might persist online.
- Small groups promote active and engaged learning virtually, but fostering community and connection amongst faculty and peers is more challenging, thus the choice of format should be aligned with the local context and program objectives.
- Clinical skills (most notably, physical exam skills), procedural skills, and laboratory practices (e.g., anatomy dissections) are the most challenging to teach remotely, and should be prioritized to return to face-to-face instruction as soon as possible.

We close by offering the following practical recommendations to educators as they develop and report on online educational developments in UGME:

- Leverage available expertise by forming a robust team of educational experts, instructional designers, faculty developers and other stakeholders.
- Utilize technology as an affordance for contextualized educational development, as a means to an end, to enhance learning.
- Design interventions on a foundation of theory and incorporate evidence-informed educational practices.
- Describe the content, setting and educational methods robustly to promote transferability.
- Report the development in detail, including time, cost, human and material resources, to allow for replicability by others.

**Conclusions:**

UGME educators rose to the challenges presented by the COVID-19 pandemic and rapidly pivoted traditionally face-to-face classroom activities to the online environment. The use of synchronous and asynchronous formats encouraged both virtual engagement and interactivity, while providing opportunities for more flexible, self-directed learning. Although technology’s potential to transform learning is not yet fully realized, this review summarized a number of novel solutions that can form the foundation for future learning in a post-pandemic world. As we transition from emergency remote learning and publications aimed at rapid dissemination, educators must underpin developments with theory, focus on improving study methodology, evaluate additional outcomes, and provide details across all elements to support replication.

**Acknowledgments:**

The authors would like to express their appreciation to Whitney Townsend for her expertise developing the search strategy and assistance running the database search.

**Disclosure statement:**

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.
Notes on contributors:

Jennifer N. Stojan, MD, MHPE, is an Associate Professor of Internal Medicine and Pediatrics at the University of Michigan Medical School in Ann Arbor, Michigan, USA.
Mary Haas, MD, MHPE, is Assistant Residency Program Director and Clinical Instructor, Department of Emergency Medicine, University of Michigan Medical School, Ann Arbor, MI.

Satid Thammasitboon, MD, MHPE, is a Director of the Center for Research, Innovation and Scholarship in Medical Education (CRIS), Co-Director of a BEME International Collaborating Centre, and Associate Professor of Pediatric Critical Care Medicine at Texas Children's Hospital, Houston, Texas, USA.

Lina Lander, ScD is an Associate Professor, Family Medicine and Public Health and Associate Dean, Education Technology, Innovation and Assessment at the University of California, San Diego, School of Medicine, La Jolla, California, USA.

Sean J. Evans, MD is a Clinical Professor of Neurosciences and Associate Dean for Undergraduate Medical Education at the University of California, San Diego, School of Medicine, La Jolla, California, USA.

Cameron Pawlik, BS, is a first-year medical student at the University of Michigan Medical School, Ann Arbor, Michigan, USA.

Teresa Pawlikowska, MB BS, PhD FRCPI, is Foundation Director of The Health Professions Education Centre (HPEC) which is also a BEME International Collaborating Centre, RCSI University of Medicine and Health Sciences, Dublin, Ireland.
Madelyn Lew, MD, is an Associate Professor of Pathology and Director of Medical School Pathology Education Curriculum at the University of Michigan Medical School in Ann Arbor, Michigan, USA.

Deena Khamees, MD, is a Clinical Lecturer and Assistant Program Director of Emergency Medicine, Michigan Medicine, Ann Arbor, Michigan, USA.

William Peterson, MD, is a Director of Emergency Medicine Residency Preparatory Course, Assistant Clerkship Director, and an Assistant Professor of Emergency Medicine, Department of Emergency Medicine, University of Michigan Medical School, Ann Arbor, MI.

Ahmad Hider, MPhil, is a second-year medical student at the University of Michigan Medical School, Ann Arbor, Michigan, USA.

Ciaran Grafton-Clarke, MBChB, is a Clinical Education Fellow at University Hospitals of Leicester, UK, and Honorary Fellow at the University of Leicester, UK.

Hussein Uraiby, MBChB, PGCert, FHEA is a Specialty Trainee in Histopathology and a Clinical Education Fellow at the University Hospitals of Leicester NHS Trust.

Morris Gordon, MBChB, PHD, MMed, is Cochrane Coordinating Editor, Chair of the BEME Executive Committee, and a Professor of Evidence Synthesis and Systematic Review, University of Central Lancashire, Preston, UK.

Michelle Daniel, MD, MHPE, is Chair of the BEME Review Committee, Associate Editor for Medical Teacher, and Vice Dean for Medical Education and Clinical Professor of Emergency
Medicine at the University of California, San Diego School of Medicine, La Jolla, California, USA.

References:


https://doi.org/10.37506/ijphrd.v11i10.11100.


https://doi.org/10.22454/PRiMER.2020.275245.


