ORIGINAL RESEARCH

A New Normal: Assessment Outcomes and Recommendations for Virtual Versus In-Person Curricula in Post-COVID-19 Times

Eileen Cowan1 · Beth Altschafl2 · Julie Foertsch2 · Denise Barnes2 · Michael Lasarev3 · Elaine Pelley4

Accepted: 8 March 2022 / Published online: 19 March 2022
© The Author(s) under exclusive licence to International Association of Medical Science Educators 2022

Abstract
Introduction  COVID-19 response efforts that began in March 2020 prompted an urgent need to transition medical education from an in-person to a virtual format. Our aim is to provide evaluation of a virtual platform for a fully integrated curriculum to provide future guidance in teaching methods.

Materials and Methods  We used summative assessments and course evaluations from pre- and post-transition from in-person to virtual delivery of educational content to measure the impacts of this transition on student performance and perceptions. Additionally, we surveyed students about their in-person versus online educational preferences.

Results  There were no statistically significant differences in student knowledge acquisition as assessed by weighted averages of summative assessments when comparing an in-person to a virtual educational platform. While the transition to virtual learning was initially well-received by students, our studied cohorts gave lower scores for the overall learning experience after prolonged virtual learning (p < 0.001). Students had a strong preference that anatomy and other group sessions should be delivered in-person. There was no strong preference whether other learning modalities should be given in-person or virtually.

Conclusions  Although student knowledge acquisition remained stable on a virtual platform, the student learning experience varied. We recommend that when returning to a new normal after COVID-19 restrictions are lifted, sessions that require 3-dimensional or group learning should remain in-person, while other educational activities may be offered on a virtual platform and that, whenever virtual learning is employed, attention be paid to ensuring ongoing social and academic engagement between learners and faculty.

Keywords  Medical student · Assessment · COVID-19 · Distance education · Virtual · Burnout

Introduction

The World Health Organization declared a state of emergency due to the COVID-19 pandemic on March 11, 2020 [1]. In order to keep students safe and on track to graduate, there was an urgent need to shift medical education curricula, including learning activities and assessments, from an in-person to a virtual format [2]. Enacting this change created many challenges, from the restructuring of lectures, labs, and group activities, to streamlining test administration, to managing potential increases in student anxiety [3]. This sudden transition provided a “natural experiment” for exploring the effectiveness of various learning modalities when moved online. We felt it critical to examine the resulting effects on student acquisition of knowledge and learning experience as medical schools and universities will benefit from this education to mitigate the negatives and accentuate the positives in any setting where online learning occurs. Our goals are to evaluate an integrated preclinical medical school curriculum that was converted from an in-person to a virtual platform and to inform best practices in a post-pandemic era.

1 Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
2 Academic Affairs, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
3 Department of Biostatistics and Medical Informatics, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
4 Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
Materials and Methods

The Context

The ForWard curriculum at the University of Wisconsin School of Medicine and Public Health (UWSMPH) is a medical doctorate program that is divided into three phases spanning four years [4]. Phase 1 compresses the traditional 2-year preclinical curriculum into six sequential, integrated thematic blocks spanning 18 months. Subsequently, Phase 2 dedicates one year to required integrated clinical rotations, while Phase 3 is composed of 18 months of specialty electives, acting internships, and career exploration opportunities. The data presented here focuses on outcomes from Phase 1. The first block, Patients, Professionalism, and Public Health (PPP), is a 4-week introduction to medical school. Following PPP, the M1 academic year continues with three large Phase 1 blocks that integrate basic science with core clinical conditions centered around specific organ systems, starting with Body in Balance (BIB), followed by Food, Fasting, and Fitness (FFF), and Human Family Tree (HFT). Invaders and Defense (I&D) begins the M2 academic year, followed by Mind and Motion (M&M), which completes the Phase 1 curriculum. Each block in this study had a class size of approximately 175 students, with slight variations due to the number of students progressing through each block. Table 1 shows the name, sequence, duration, and general scope of content covered in each of the six integrated thematic blocks that comprise Phase 1 of the UWSMPH ForWard curriculum.

Learning activities in the ForWard curriculum include individualized experiences (coaching and competency (C&C) reviews, office hours), small-group learning sessions (patient-centered education (PaCE) cases, anatomy dissections), medium-sized group sessions (case-based learning (CBL) sessions), and whole class sessions (team-based learning (TBL) sessions, lectures). These educational activities had to be quickly converted to a virtual format during COVID-19 [5]. Table 2 briefly describes various learning modalities as well as how they were structured in-person versus in the virtual format. With the exception of two online history-taking sessions early in the pandemic, Phase 1 clinical skills teaching remained in-person with standardized patients and appropriate COVID-19 precautions (barrier masks, face shields, and room capacity limitations) throughout the time period presented in this paper.

Data Collection and Analysis

Knowledge Acquisition Outcomes

We used student performance on summative assessments to measure the impact of the transition to virtual learning on student acquisition of knowledge. Weighted averages of
Table 2 Description of various learning modalities used in Phase 1 of the University of Wisconsin School of Medicine and Public Health ForWard curriculum

| Learning modality                  | Description                                                                 | In-person                                                      | Virtual                                                                 |
|------------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------------------|-------------------------------------------------------------------------|
| Anatomy labs                        | Hands-on gross dissections (in-person); image labeling (virtual)            | ~8 students per human cadaver                                 | ~8 students per image in each virtual breakout room                     |
| Case-based learning (CBL) sessions | Group work involving clinical cases                                          | Classroom work with ~6 students per table, followed by a group debrief with ~36 students and 1–2 facilitators | 4–10 students per virtual breakout room, followed by group debrief in virtual main room ranging from 36 students to whole class with 1–5 facilitators |
| Team-based learning (TBL) sessions | Group work involving individual and team-based quizzes and clinical cases   | Classroom work with ~8 students per table, followed by group debrief with whole class (~175 students) with 1–2 facilitators | 4–10 students per virtual breakout room, followed by group debrief in virtual main room ranging from 36 students to whole class with 1–5 facilitators |
| Patient-centered education (PaCE) cases | Student inquiry-based creation of learning objectives surrounding patient cases | 8 students in classroom with 1 facilitator                     | 8 students in virtual classroom with 1 facilitator                      |
| Lectures                            | Didactic content delivery with interspersed knowledge check questions/polling | All students in 1 lecture hall, raise hand to ask questions    | Pre-recorded lectures or 1 virtual lecture room with chat feature enabled to allow for real-time student questions |
| Office hours/question and answer (Q&A) sessions | Informal Q&A sessions between faculty and interested students | 1 faculty member and several students face-to-face             | 1 faculty member and several students via a virtual meeting             |
| Coaching and competency (C&C) reviews | Meetings between students and their faculty coaches to discuss student progress and goals | 1 faculty member and 1 student face-to-face                    | 1 faculty member and 1 student via a virtual meeting                   |
the summative assessments in each block were compared between different graduating cohorts, with the earlier cohort experiencing the curriculum in-person and the subsequent cohort experiencing the curriculum virtually. Statistical analyses were performed between iterations of a given block and not between blocks. Means and standard deviations of the weighted averages were determined for each block and the 95% confidence intervals (CIs) for the difference in means were calculated between cohorts for the same block. Weighted averages were based on the number of exams per block (4 in BIB; 3 in FFF, I&D, and M&M; 2 in HFT) and exams were additionally weighted based on the amount of content covered. Each block’s exams was weighted the same for each of the two cohorts that were compared. Based on previous observations of stability of exam performance year-to-year, differences between means of less than 5% points were predefined as being not educationally significant, while larger differences (5 points or more) were considered indicative of a notable change in class performance. Second-generation p-values ($p_b$) were used to identify changes that met this minimum threshold for a meaningful difference [6]. A second-generation p-value is the proportion of the 95% CI for the difference overlapping the predefined region of indifference. Having $p_b=0$ signals that differences best supported by the data do not overlap with the region of indifference (i.e., all differences are meaningful in size), while $p_b=1$ implies data only support differences of no real importance due to complete overlap of the CI with the region of indifference. Values between these two extremes are inconclusive due to data simultaneously supporting both unimportant and meaningful differences, with values closer to zero signifying less overlap and therefore greater likelihood of meaningful effects, and values closer to one signifying greater overlap with the region of indifference and signaling greater support for nugatory effects. Statistical analyses were done using R (v. 4.1.0) [7]. PPP assessment data were not included as it is a 4-week introduction to medical school that does not assess academic (BIB and FFF) and courses taken after the transition to online learning via the following survey question: “In a COVID-free Phase 1, where all options are available, how would you prefer the following learning activities to be delivered?” This question was answered on a 5-point scale, where 1 = “online strongly preferred,” 2 = “online somewhat preferred,” 3 = “neutral,” 4 = “in-person somewhat preferred,” and 5 = “in-person strongly preferred.” For purposes of analysis, student preference was further summarized as “online preferred” (strongly or somewhat), “neutral,” or “in-person preferred” (strongly or somewhat) and described using frequencies and percentages. Students were also asked to provide open-ended narrative feedback. The learning activities in question included anatomy labs, CBL sessions, TBL sessions, PaCE cases, lectures, office hours/ question and answer (Q&A) sessions, and C&C reviews. We did not ask student preference regarding clinical skills teaching as curricular leaders deemed in-person teaching the de facto better approach, independent of student preference. Among students who held an opinion, the ratio of percentages for those who preferred an in-person activity relative to those who preferred an online activity was computed, together with a supporting 95% CI for the ratio. We predefined a meaningful and definitive direction of preference as a ratio of at least 2:1, equating to at least 66% of students preferring one modality over the other. Second-generation p-values ($p_b$) were calculated to assess the proportion of ratios best supported by our data that overlap with a region of no genuine preference (spanning from 0.5:1 up to 2:1).

**Student Experience Outcomes**

We collected student course evaluation data at the end of each block of the Phase 1 curriculum. This allowed us to measure the impact of the transition from the in-person to the virtual platform on students’ overall satisfaction of the learning experience. Students rated the statement “Overall, this course provided a good learning experience” on a 7-point Likert scale that ranged from 1 = “strongly disagree” to 7 = “strongly agree” at the conclusion of each block. We examined student experience in each cohort. For the Class of 2023, this included courses taken in-person before the start of the COVID-19 pandemic (BIB and FFF) and courses taken after the transition to a virtual platform in March of 2020 (HFT, I&D, and M&M).

The Class of 2024 started Phase 1 in the fall of 2020 and thus experienced their entire M1 year virtually (PPP, BIB, FFF, and HFT). PPP data were not included, as this block differs in both content and assessment strategies from the other blocks. Average student-reported experience for each course was compared between time-adjacent cohorts using two-sample t-tests. We had no a priori way to define what size shift in satisfaction would be meaningful, so traditional p-values and calculated measures of effect size (Cohen’s $d$) were used to guide interpretation of results. These analyses were also done using R.

**Survey on Student Learning Preferences**

At the completion of Phase 1, the graduating Class of 2023, who had experienced their first three Phase 1 blocks in-person and their final three Phase 1 blocks on a virtual platform, were asked their preference for in-person versus online learning via the following survey question: “In a COVID-free Phase 1, where all options are available, how would you prefer the following learning activities to be delivered?” This question was answered on a 5-point scale, where 1 = “online strongly preferred,” 2 = “online somewhat preferred,” 3 = “neutral,” 4 = “in-person somewhat preferred,” and 5 = “in-person strongly preferred.” For purposes of analysis, student preference was further summarized as “online preferred” (strongly or somewhat), “neutral,” or “in-person preferred” (strongly or somewhat) and described using frequencies and percentages. Students were also asked to provide open-ended narrative feedback. The learning activities in question included anatomy labs, CBL sessions, TBL sessions, PaCE cases, lectures, office hours/ question and answer (Q&A) sessions, and C&C reviews. We did not ask student preference regarding clinical skills teaching as curricular leaders deemed in-person teaching the de facto better approach, independent of student preference. Among students who held an opinion, the ratio of percentages for those who preferred an in-person activity relative to those who preferred an online activity was computed, together with a supporting 95% CI for the ratio. We predefined a meaningful and definitive direction of preference as a ratio of at least 2:1, equating to at least 66% of students preferring one modality over the other. Second-generation p-values ($p_b$) were calculated to assess the proportion of ratios best supported by our data that overlap with a region of no genuine preference (spanning from 0.5:1 up to 2:1).

**Results**

**Knowledge Acquisition Outcomes**

There were no statistically significant differences in student knowledge acquisition, as measured by weighted
averages of summative assessments, between the earlier in-person cohort and the later virtual cohort for each block (Table 3). Variation in n is due to the variation in number of students progressing through each block. All second-generation p-values equaled to 1, indicating no significant differences, due to the 95% CIs for the difference being fully nested in the interval from −5 to 5% points. In fact, in all comparisons, no block experienced a greater than 2% point difference in the calculated weighted average of the block’s summative assessments between the earlier in-person cohort and the later virtual cohort (Fig. 1).

### Student Experience Outcomes

Table 4 shows student responses in each Phase 1 block to the evaluation item, “Overall, this block provided a good learning experience” and the p-value and effect size (d) of the independent t-test comparison between cohorts are shown. Variation in n is due to the variation in number of students completing the survey for each cohort. For the Class of 2023, the first online version of a block was HFT, followed by I&D, and then M&M. This cohort gave mean ratings for these online versions of HFT, I&D, and M&M that were significantly higher, unchanged, and significantly lower,
respectively, than mean ratings for the in-person version of these blocks given by the previous year’s cohort. The Class of 2024, who had all of the M1 Phase 1 blocks on a virtual platform, gave their first large online block, BIB, a mean rating which was unchanged from the prior year and gave their second large online block, FFF, a significantly higher mean rating than the previous year’s in-person cohort. However, this same cohort rated their third large online block of their M1 year, HFT, substantially lower than the prior year’s cohort had, despite both iterations being delivered on a virtual platform and few changes to the structure, management, and content of the block.

Survey on Student Learning Preferences

The graduating Class of 2023, who had experienced their first three Phase 1 blocks in-person and their final three Phase 1 blocks on a virtual platform, were asked their preference for online versus virtual learning and the results are shown in Table 5. Of the students who had a preference, there was a strong preference (by at least a ratio of 2:1) for anatomy labs and CBL sessions to be taught in-person versus online, as evidenced by the 95% CI for each of these learning modalities fully residing outside the predefined region of indifference (Fig. 2). A definite conclusion was also found for lectures, where no preference in favor of either delivery method was identified, as indicated by the 95% CI completely nested within the region of indifference. Due to nearly complete overlap with the region of indifference, there was no genuine preference for whether students preferred PaCE cases to be online or in-person. Other modalities, specifically TBL sessions, office hours/Q&A sessions, and C&C reviews, had 95% CIs for the ratio that partially overlapped with the region of indifference by a non-trivial amount and therefore failed to show a convincing preference for either modality (Fig. 2). Student narrative feedback included some comments suggesting periods of depression and isolation with online learning.

Discussion

The need to convert medical education curricula to an online platform during COVID-19 was a worldwide phenomenon, with reports of an increase in the number of students spending >15 h per week on virtual platforms post-pandemic.
versus pre-pandemic [8]. Given ever-increasing advancements in educational technology, it is reasonable to assume that some aspects of virtual learning are here to stay in a post-pandemic world, making it essential to evaluate student acquisition of knowledge and experience with this format.

Importantly, our knowledge acquisition outcomes data demonstrated that student acquisition of knowledge remained stable despite conversion from an in-person to a virtual platform, as indicated by second-generation \( p \)-values \( (p_\delta) \) show the proportion of overlap between the 95% CI and the shaded region of indifference. Importantly, our knowledge acquisition outcomes data demonstrated that student acquisition of knowledge remained stable despite conversion from an in-person to a virtual platform, as indicated by second-generation \( p \)-values \( (p_\delta) \) show the proportion of overlap between the 95% CI and the shaded region of indifference. CBL = case-based learning sessions, TBL = team-based learning sessions, PaCE = patient-centered education cases, Q&A = question and answer sessions, C&C = coaching and competency reviews, \( n = 174 \)

Although students performed well on assessments in all virtual blocks, our student experience outcomes data revealed that their experience block-to-block varied. Overall, the first block of online learning was received well, with ratings that were either stable or higher compared to ratings for the same block given in-person the year prior. Some students reported to instructors that the online platform improved scheduling flexibility and that they appreciated the decreased travel time and the ability to self-pace asynchronous activities. In fact, the Kim et al. study found that a majority of students at their medical school wanted to maintain an online curriculum after COVID-19 [12]. It is worth noting, however, that the medical school described does not have a fully integrated curriculum, maintained some aspects of in-person learning, and the study did not follow students as they continued the virtual learning experience over time. We found that for each cohort at our school, after stable or higher scores given for the first two large blocks, the third large virtual block received significantly lower scores compared to the year prior. For the HFT block, this is a particularly interesting occurrence as the student cohort for whom it was their first virtual block rated the learning experience with a statistically significant higher score than the year prior, but the following cohort, for which it was their third sequential virtual block, rated it with a statistically significant lower score. Only minimal changes to the virtual curriculum were made to the HFT block between the two iterations, as per normal yearly quality assurance processes.
Similar to our findings, there are several studies that found student perception of virtual learning to be overall positive, at least in the beginning stages [8, 12, 13]. However, multifactorial barriers exist to online learning, including, but not limited to, family interruptions, poor internet connection [8, 14], and a decline in mental health and an increase in cynicism [15]. These could explain our findings that students gave the learning experience lower scores with prolonged virtual learning. Potentially supporting this, some of our students commented to instructors or provided written narrative feedback that they had “Zoom fatigue,” found it difficult to make meaningful connections with their peers and faculty mentors, and missed the social aspects of studying together. To mitigate this, whenever a virtual approach is selected, some students suggested a “cameras on” policy and informal social exchanges between students and faculty to help facilitate ongoing social connection and engagement. Additionally, the physical strain of being at a computer for long periods of time is well-documented in the literature, with double vision, blurred vision, and musculoskeletal pain contributing to poor experiences [16]. It is also possible that after multiple prior online blocks, students may have higher expectations for the online educational experience and for peer and instructor interaction and engagement.

Although COVID-19 forced medical schools to convert educational experiences to an online curriculum, in the post-COVID-19 era, in-person activities will resume. However, some form of virtual learning is here to stay. Our survey on student learning preference data indicates that students have a strong preference for hands-on activities, such as anatomy, to remain in-person. Students also appear to perceive the benefit of the in-person format in CBL sessions, which involve peer teaching and small-group instructor interaction, which is difficult to replicate online. Alternatively, students had no true preference for whether other activities, such as office hours and C&C sessions, were conducted online, which could be because this approach allows more scheduling flexibility for both faculty and students.

While data about assessment and student experience outcomes are demonstrated here, less tangible outcomes, such as how online delivery of educational material may impact attitudes, values, and professional identity formation of physicians-in-training, may be of equal importance and warrant further study. In addition, our study does not allow for a perfectly controlled comparison of a virtual to an in-person curriculum. Limitations include differences in class cohorts, changing COVID-19 regulations both in and out of school, and continuous quality improvement initiatives leading to a stable, but not fully static, curriculum. However, we feel that that our findings can help inform best practices.

Conclusion

The COVID-19 pandemic thrust the world into a state of social distancing that necessitated adapting in-person educational experiences to virtual formats. Our data suggest that virtual learning approaches can be implemented without fear of negatively impacting student performance on assessments. Although there were aspects of virtual learning that were perceived positively by students, students had clear preferences for certain learning modalities to occur in-person. We found that, while initially well-received, prolonged online learning was associated with lower ratings in the student-reported learning experience in both of our studied cohorts. Although we do not know the exact cause for this phenomenon, some individual student feedback suggested that fatigue, isolation, and burnout can occur with online learning, which may contribute to these findings. Individual students also recommended social exchanges between faculty and students to promote connection and engagement during virtual learning. Thus, it is our recommendation that the new normal of medical education curricula post-COVID-19 should consider keeping certain educational experiences in-person, including courses that either require 3-dimensional learning such as anatomy, or courses that require group learning and peer teaching, and that heightened attention be given to encouraging social and academic engagement with faculty and peers whenever virtual learning is being employed.

Author Contribution All authors contributed to the study conception and design. Material preparation and data collection were performed by Eileen Cowan, Beth Altschaffl, Julie Foertsch, Denise Barnes, and Elaine Pelley. Analysis was performed by Julie Foertsch and Michael Lasarev. The first draft of the manuscript was written by Eileen Cowan, and all authors commented on previous versions of the manuscript. All authors read and approved of the final manuscript.

Availability of Data and Material All data generated or analyzed during this study are included in this published article.

Code Availability All applicable software applications or custom codes used during this study are included in this published article.

Declarations

Ethics Approval Not applicable. Based on the UW-Madison Health Sciences Institutional Review Board’s Quality Improvement/Program Evaluation Decision Tree, this project constitutes Program Evaluation and does not fit the federal definition of research. Further IRB review was not required.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.
References

1. AJMC_Staff. A timeline of COVID-19 developments in 2020. 2021 [cited 2021 05 Aug]; Available from: https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020.
2. Rose S. Medical student education in the time of COVID-19. JAMA. 2020;323(21):2131–2.
3. Chandratte S. Medical students and COVID-19: challenges and supportive strategies. J Med Educ Curric Dev. 2020;7:2382120520935059.
4. Pelley EM, et al. University of Wisconsin School of Medicine and Public Health. Acad Med. 2020;95(9S A Snapshot of Medical Student Education in the United States and Canada: Reports From 145 Schools):S559-S562.
5. Hueston WJ, Petty EM. The impact of the COVID-19 pandemic on medical student education in Wisconsin. WMJ. 2020;119(2):80–2.
6. Blume JD, et al. Second-generation p-values: improved rigor, reproducibility, & transparency in statistical analyses. PLoS One. 2018;13(3):e0188299.
7. R_Development_Core_Team. R: a language and environment for statistical computing. 2021. Available from: http://cran.r-project.org.
8. Dost S, et al. Perceptions of medical students towards online teaching during the COVID-19 pandemic: a national cross-sectional survey of 2721 UK medical students. BMJ Open. 2020;10(11):e042378.
9. Almarzooq ZI, Lopes M, Kochar A. Virtual learning during the COVID-19 pandemic: a disruptive technology in graduate medical education. J Am Coll Cardiol. 2020;75(20):2635–8.
10. Prigoff J, Hunter M, Nowygrod R. Medical student assessment in the time of COVID-19. J Surg Educ. 2021;78(2):370–4.
11. Jones VA, et al. Rescuing medical education in times of COVID-19. Clin Dermatol. 2021;39(1):33–40.
12. Kim JW, et al. How medical education survives and evolves during COVID-19: our experience and future direction. PLoS One. 2020;15(12):e0243958.
13. Khalil R, et al. The sudden transition to synchronized online learning during the COVID-19 pandemic in Saudi Arabia: a qualitative study exploring medical students’ perspectives. BMC Med Educ. 2020;20(1):285.
14. Baticulon RE, et al. Barriers to online learning in the time of COVID-19: a national survey of medical students in the Philippines. Med Sci Educ. 2021:1–12.
15. Zis P, et al. Medical studies during the COVID-19 Pandemic: the impact of digital learning on medical students’ burnout and mental health. Int J Environ Res Public Health. 2021;18(1).
16. Gowrisankaran S, Sheedy JE. Computer vision syndrome: a review. Work. 2015;52(2):303–14.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.