2022

The Effects of COVID-19 on Staying Alive NBME Performance and Exam Testing Conditions

Anh Phuong Nguyen
Wright State University - Main Campus, nguyen.204@wright.edu

The effects of COVID-19 on Staying Alive NBME performance and exam testing conditions

Anh Phuong Nguyen

Dr. Irina Overman, Assistant Professor of Geriatrics & Internal Medicine

Medical Education Research

Scholarship in Medicine Final Report

☐ **By checking this box, I indicate that my mentor has read and reviewed my draft proposal prior to submission**

Abstract

Objective: This study will study the impact of COVID-19 on Wright State University BSOM class of 2023's performance on the SA NBME exam and whether there was a difference in the scores between the group who took the test at school compared to at home. *Methods:* independent samples t-test was used to analyze SA NBME exam scores from BSOM class of 2022 and class of 2023 (n = 231) and also from students in the class of 2023 who took the SA NBME from home and at school. *Results:* The class of 2023's performance on the SA NBME exam was comparable to the year before. Members of the class of 2023 who took this test at school did better than those who took it at home.

Keywords: COVID-19, online testing, online proctoring, medical student, medical education

Introduction/Literature Review

COVID-19 started as a local transmission in the region of Wuhan, China in late December 2019 and by February 2020, it was declared a pandemic by the World Health Organization[1]. As the virus continued to spread, aggressive measures such as closing down public spaces like restaurants and schools were taken to limit the contagion. By mid-March 2020, numerous medical schools in the U.S had suspended all clinical rotations and medical students were asked to stay home and maintain proper social distancing[2]. At Wright State University Boonshoft School of Medicine (BSOM), a public Midwestern allopathic medical school in the U.S, for the class of 2023, COVID-19 necessitated the shift from in-person classes to online classes with instructors and students communicating in live, real-time virtual classrooms and chat forums for the module Staying Alive (SA). This form of communication is known as synchronous learning[3] and many features of the BSOM curriculum such as live polling and quizzes and small groups discussion were preserved through the use of this learning delivery method. However, synchronous learning has also been noted to have a few drawbacks. Technical difficulties such as poor audio or poor connections often cause frustrations in students and instructors[4,5] and impede the delivery of learning material[6]. The final exam for this course which consisted of 150 questions selected from a question bank provided by the National Board of Medical Examiners (NBME) and was to be taken at an on-campus facility. However, given that many BSOM students were out of state and had decided to shelter in place at home, students were offered two options: 1. take the SA final exam remotely with online proctoring using webcams; 2. take it at school with in-person proctoring (while following procedures and policy regarding social distancing). As more and more students are taking their exams online as the result of an increase in online classes and universities, the need for online proctoring also increases[7]. Although some forms of proctoring during online exams are needed to maintain academic integrity, online proctoring has been found to affect students' performance, especially those with high trait test anxiety[8]. Given the disruptions of COVID-19 on students' daily routines, study and test taking schedules, and its negative impacts on people's overall well-being[9], the purposes of this study were to explore: 1. the impact of COVID-19 on the overall NBME performance in Staying Alive this year compared to the previous year; 2. whether there is a difference in the scores of students who took it remotely with online proctoring (at-home group) versus on campus with in-person proctoring (at-school group). We expected to find no difference in SA NBME performance between the class of 2022 versus 2023 and the at-home versus at-school group given a multitude of factors.

At the time of this writing, July 2020, even though social distancing measures are cautiously being relaxed, the total number of COVID-19 cases in the U.S are still rising[10] and predictions about what will happen are still changing by days, if not hours. Such unpredictability and uncertainty are challenging to medical students and administrations alike as delays in medical education are costly and yet it is the health and safety of students and staff that are at stake. Therefore, going forward, knowing how medical students perform under such circumstances and

different testing conditions is critical to making adjustments to the curriculum and teaching methods in medical education.

Methods

Design

We conducted a retrospective study on BSOM class of 2023's SA NBME performance compared to the class of 2022's and whether there was a difference in SA NBME scores between at-home and at-school groups. This study was conducted within the auspices of the Office of Medical Education as a curriculum evaluation study and therefore exempted by Wright State University's Institutional Review Board.

Staying Alive NBME examination

The exam consisted of 150 multiple choice questions and was 3 hours and 45 minutes long. The questions were selected by BSOM faculty from the NBME Customized Exam Services (CES) to test students' understanding of physiology, pharmacology, system pathology, microbiology, general principles, clinical diagnoses, and gross anatomy in the cardiovascular, renal, and respiratory organ systems. The NBME exam that the class of 2022 and 2023 took were deemed equivalent in difficulty by NBME CES, with both having the anticipated mean of 80% correct.

For students who opted to take the exam remotely, NBME Instructions with Web Conferencing using Webex were used (see Appendix A).

Data Collection

Demographic data such as gender, race, and underrepresented in medicine (URM) status, along with MCAT score and percentile, undergraduate GPA, and undergraduate science GPA were collected from matriculates' American Medical College Application Service (AMCAS) and BSOM secondary applications. Students who self-reported as American Indian or Alaska Native, black or African American, Hispanic/Latino, Native Hawaiian, or other Pacific Islander are considered URM by BSOM.

SA NBME scores from BSOM class of 2022 and 2023 were used to establish the study population. If students had to retake the exam, only scores from their first exam were included. Exclusion criteria included students who were not eligible to sit for the exam and those who had to take it at a later date due to personal reasons.

Anonymous data files with no student identification number were obtained from the Office of Medical Education. Scores from the other three exams that BSOM class of 2022 and 2023 took in their first year which were Origins NBME, Human Architect 1 (HA1), and Host and Defense (HD) were also collected. Scores from the Objective Structured Clinical Examination 1(OSCE1) were also analyzed for completeness (see appendix B for module descriptions).

The data also included whether students took the SA NBME exam remotely with online proctoring or at school with in-person proctoring.

Data Analysis

Data analysis was conducted using IBM SPSS Statistics software with the assistance of Dr. Adrienne Stolfi, an assistant professor at BSOM.

For research purpose 1 (RP 1): what is the impact of COVID on the overall NBME performance in SA this year (class of 2023 cohort) compared to last year (class of 2022 cohort), prior to running an independent sample t-test, we wanted to see if any of the categorical demographic variables including gender, race, URM status was a confounding variable. A cross-tabulation of those variables was conducted between the 2022 and 2023 cohorts and was followed by a chi-square test. It was determined that they were not confounding variables. Next, using Levene's test for equality of variances followed by an independent samples test, we analyzed the continuous data on MCAT total scores, MCAT percentiles, undergrad GPA, undergrad science GPA, Origins NBME exam, HA1 final exam, HD final exam, and SA NBME exam scores.

For research purpose 2 (RP 2): is there a difference in SA NBME scores between at-home versus at-school groups, a similar approach was taken where a cross-tabulation of gender, race, and URM were run between the two groups and followed by a chi-square test. Levene's test for equality of variances followed by an independent samples test was used to analyze the performance between the at-home group and at-school group on OSCE1, Origins, HA1, HD, and SA final exams.

Results

RP 1

Data from 231 students were analyzed and the descriptive characteristics of the study population were displayed in Tables 1 and 2. Pearson's Chi-Square test showed no significant differences ($p > 0.05$) between the two cohorts with regard to gender, race, and URM status (table 1). Independent samples t-tests were used to compare MCAT total scores and percentiles, undergrad GPA, and science GPA between the two cohorts. There was no statistically significant difference ($p > 0.05$) between the two groups (table 2). The Independent samples t-test was used again to compare the performance of the two cohorts on four major exams that took place during the first year at BSOM (table 3). Levene's test showed equal variances for all variables. There was no significant statistical difference between the class of 2022 and the class of 2023's performance on all four exams ($p>0.05$), including SA NBME, the exam of interest, since this took place during COVID-19.

RP 2

Descriptive characteristics of at-home (n = 64) and at-school (n= 49) group were outlined in table 4 and 5. Pearson's Chi-Square test was used to see if there were significant differences between the two cohorts in terms of gender, race, and URM status. It was determined that there was no significant difference ($p > 0.05$) with regard to those variables. Based on what are usually considered innate academic ability parameters such as MCAT scores and percentiles, undergrad GPA, and science GPA, Levene's test found equal variance for those variables between the at-home and at-school groups. The results from the independent sample t-test found that there was no significant difference ($p > 0.05$) between the at-home and at-school groups based on innate academic ability (table 5). However, the at-school group performed consistently better on Origins, HA1, HD, and SA final exams. The two-tailed p-value was significantly different ($p < 0.05$) across all four exams (table 6) with the at-school group scoring higher. The two groups performed similarly on the OSCE 1 exam ($p >0.05$) (table 7).

Discussion

At the time of this writing, July 2020, there have been 10 million confirmed COVID-19 cases globally and 2.56 million in the United States[10]. In March 2020, when the number of COVID-19 first started increasing in the U.S, given the severity of the situation and the need for social distancing, classes at BSOM were moved from in-person to an online platform without much advance notice. Students, therefore, had very little time to prepare and had to learn to quickly adapt to a new learning environment. Nevertheless, despite all the challenges that COVID-19 presented, we found that BSOM class of 2023's performance on SA NBME was comparable to the previous year's. There might be multiple reasons why such results were seen. First, even though there was a period of getting used to classes being held virtually, eventually all of the major aspects of the curriculum in the first two years at BSOM such as peer instruction or team-based learning (TBL) were replicated in the online format. Peer instruction is a didactic method where students come to class and get quizzed on the materials they are responsible for that day. The answers are then discussed and debated in small, randomly assigned groups. During COVID-19, small group discussions were done via Webex and the virtual lecture hall was done via Blackboard. Both Webex and Blackboard are online platforms where students and instructors can communicate via webcam and chat boxes. Therefore, there were very minimal losses of teaching materials during this period. Second, given that exam performance is highly correlated with effective study strategies[11,12] such as deep learning and time management, as long as students were able to keep using the same strategies that had been working for them, it is, therefore, reasonable to expect that they would perform well even in the face of a pandemic. Third, an unintended consequence of social distancing was that all in-person activities such as extracurricular activities and clinical medical education were put on hold. Students could have used this extra time to study and this might have negated the effects of COVID-19 on the learning process. And lastly, medical students are known to have a high level of traits like perseverance[13] which not only allow them to perform well in medical school but also in challenging times such as the COVID-19 pandemic.

For the second research purpose, contrary to our hypothesis, the at-school group was found to perform better on the SA NBME exam than the at-home group. Given that academic performance during medical school has been shown to be proportionally correlated with entrance MCAT scores[14], it is reasonable to assume that the at-school group did better on the SA NBME exam because they had higher MCAT scores. However, this did not turn out to be the case as the at-home and at-school groups were comparable in their MCAT scores and percentiles. On all of the other descriptive categories that we have data on such as gender, race, and URM status, the two groups were comparable as well (Tables 1 and 2). They also performed similarly on the OSCE 1 exam which tests what clinical skills medical students can perform rather than their recall of knowledge. Another reasonable explanation for this puzzling result is perhaps the at-school group did better because their selected testing environment was more conducive to taking tests. However, this explanation is irreconcilable with the finding that the at-school group also did better on the other three exams which took place pre-COVID-19 and hence were administered in person for both groups. This means that testing conditions did not seem to play a role in students' exam performance. Given those puzzling results, we suspected that there might be an unknown third variable that was yet to be uncovered or controlled for. One candidate for such a variable is the study strategy. As previously mentioned, study strategy is highly correlated to exam performance. Therefore, if the students in the at-school group utilized study strategies that were more effective than those used by the at-home group, this could potentially explain those results. Future studies could follow up with a qualitative survey on what study strategy students in each group used to study for the SA NBME. Another follow-up survey that could be sent out is to look at the reasons why students decided to take the exam at school versus at home. Numerous reasons could be included for students to choose from such as being out of state at the time of the exam, having too many distractions at home, or preferring to take tests at school as they had been in the past. Although this study found that testing conditions did not seem to play a role in student's performance, if COVID-19 necessitated the use of online testing again, the same video proctoring protocol should be used to ensure academic integrity, since past research has shown that academic integrity was compromised when exams were left unproctored[15,16]

Study limitation

This study was conducted at one medical school in the US. Moreover, given the recency of the COVID-19 pandemic, studies conducted and published and the literature on the impact of COVID-19 on medical education are scant.

Conclusion

This study examined the SA NBME performance of BSOM class of 2023 compared to that of BSOM class of 2022. Results showed that there was no difference in NBME scores between the two cohorts when controlling for gender, race, URM status, MCAT scores and percentiles, undergrad GPA, and science GPA. This points to how adaptable the curriculum at BSOM is and how adaptable and resilient medical students are. This study also looked at whether there was a difference in the SA NBME score between those who took the exam at home and those who took it at school. This self-selected grouping was necessitated by the social distancing that COVID-19

created. Based on the results, we found that after controlling for gender, race, URM status, MCAT score and percentiles, undergrad GPA, and science GPA, the at-school group not only did better on SA NBME but also the other three exams during the first year of the BSOM curriculum. There is a need for additional research to be done to uncover why such puzzling results existed and also how medical students will continue to perform during this troubling time.

References

Accessed

Tables and Figures

Table 1. Demographics of the class of 2022 and 2023

		Class of 2022	**Class of 2023**	**Total population**
Sample size		119	113	231
Gender	Female	56.3% n = 67	61.9% n = 70	59.1% n = 137
	Male	43.7% n = 52	38.1% n = 43	40.9% n = 95
Race	White, non-Hispanic	60.7% n = 71	49.1% n = 52	55.2% n = 123
	Black, non-Hispanic	16.2% n = 19	22.6% n = 24	19.3% n = 43
	Asian/Pac Islander	19.7% n = 23	21.7% n = 23	20.6% n = 46
	Other	3.4% n = 4	6.6% n = 7	4.9% n = 11
Underrepresented in medicine	Qualified	18.5% n = 22	24.8% n = 28	21.6% n = 50
	Not qualified	81.5% n = 97	75.2% n = 85	78.4% n = 182

Table 2. MCAT total scores and percentile, undergrad GPA and science GPA for the class of 2022 and 2023

	Class of 2022 (*n* = 119)	**Class of 2023 (*n* = 113)**	**p value**
MCAT total scores	**Mean = 506.42** SD = 5.130 Range = 495 - 517	**Mean = 506.54** SD = 5.734 Range = 495 - 522	$p > 0.05$ $p = 0.867$
MCAT percentile	**Mean = 70.61** SD = 15.630 Range = 33 - 95	**Mean = 69.42** SD = 16.684 Range = 32 - 99	$p > 0.05$ $p = 0.573$
Undergrad GPA	**Mean = 3.621** SD = 0.280	**Mean = 3.630** SD = 0.302	$p > 0.05$ $p = 0.825$

	Range = 2.54 – 4.00	Range =2.48 – 4.00	
Undergrad science GPA	**Mean = 3.525** SD = 0.327 Range = 2.08 – 4.00	**Mean = 3.535** SD = 0.364 Range = 2.23 – 4.00	p > 0.05 p = 0.811

Table 3. Performance of class of 2022 and 2023 on four major exams during the first year

	Class of 2022	**Class of 2023**	**p value**
Origins 1 and 2 NBME	**Mean = 83.017** SD = 6.987 Range = 64 - 96	**Mean = 83.062** SD = 7.989 Range = 62 - 98	p > 0.05 p = 0.963
Human Architecture 1 final exam	**Mean = 83.076** SD = 7.030 Range = 64 - 96	**Mean = 84.770** SD = 6.573 Range = 67 - 98	p > 0.05 p = 0.60
Host and Defense final exam	**Mean = 80.782** SD = 9.974 Range = 49 - 99	**Mean = 81.566** SD = 10.525 Range = 53 - 100	p > 0.05 p = 0.560
Staying Alive NBME	**Mean = 80.496** SD = 8.501 Range = 50 - 98	**Mean = 82.221** SD = 7.361 Range = 59 - 96	p > 0.05 p = 0.101

Table 4. Class of 2023 At-home vs at-school group demographics

		At school	**At home**	**P value**
Sample size		49	64	
Gender	Female	53.1% n = 26	68.8% n = 44	p > 0.05 p = 0.89
	Male	46.9% n = 23	31.3% n = 20	
Race	White, non-Hispanic	60.4% n = 29	39.7% n = 23	p > 0.05 p = 0.185
	Black, non-Hispanic	16.7% n = 8	27.6% n = 16	
	Asian/Pac Islander	18.8% n = 9	24.1% n = 14	
	Other	4.2% n = 2	8.6% n = 5	
Under-represented in medicine	Qualified	20.4% n = 10	28.1% n = 18	p > 0.05 p = 0.346
	Not qualified	79.6% n = 39	71.9% n = 46	

Table 5. MCAT total scores and percentile, undergrad GPA and science GPA of at-home and at-school group

	At school (*n* = 49)	At home (*n* = 64)	p value
MCAT total scores	**Mean = 507.49** SD = 5.8.99 Range = 495 - 522	**Mean = 505.81** SD = 5.540 Range = 496 - 520	p > 0.05 p = 0.124
MCAT percentile	**Mean = 72.18** SD = 16.800 Range = 32 - 99	**Mean = 67.30** SD = 16.411 Range = 35 - 99	p > 0.05 p = 0.123
Undergrad GPA	**Mean = 3.633** SD = 0.333 Range = 2.48 – 4.00	**Mean = 3.627** SD = 0.279 Range =2.89 – 4.00	p > 0.05 p = 0.908
Undergrad science GPA	**Mean = 3.553** SD = 0.406 Range = 2.23 – 4.00	**Mean = 3.522** SD = 0.331 Range = 2.53 – 4.00	p > 0.05 p = 0.438

Table 6. performance of at-home and at-school group on four major exams

	At school (*n* = 49)	At home (*n* = 64)	p value
Origins 1 and 2 NBME	**Mean = 85.592** SD = 6.946 Range = 62 - 95	**Mean = 83.062** SD = 8.241 Range = 62 - 98	p < 0.05 p = 0.003
Human Architecture 1 final exam	**Mean = 87.122** SD = 5.775 Range = 73 - 98	**Mean = 82.969** SD = 6.618 Range = 67 - 97	p < 0.05 p = 0.001
Host and Defense final exam	**Mean = 84.020** SD = 9.564 Range = 59 - 100	**Mean = 79.688** SD = 10.907 Range = 53 - 98	p < 0.05 p = 0.029
Staying Alive NBME	**Mean = 84.082** SD = 7.718 Range = 59 -96	**Mean = 80.797** SD = 6.796 Range = 63 -91	p < 0.05 p = 0.018

Table 7. performance of at school and at home group on OSCE CM1

	At school (*n* = 49)	At home (*n* = 64)	p value
OSCE1	**Mean = 82.267** SD = 4.820	**Mean = 88.911** SD = 5.928	p > 0.05 p = 0.732

Appendix

Appendix A: online testing protocol

Appendix B: Module descriptions

Module title	Description
OSCE 1	Clinical competency
Origins	Fundamentals of biochemistry, molecular biology, and genetics.
Human Architecture 1	An introduction to concepts of human anatomy, dissection, and embryology.

Host and Defense	Principles of immunology and microbiology and their application to human diseases.