

Kelli Glaser*, DO, Vanessa Pazdernik, MS, Denise Sackett, DO and Valerie Sheridan, DO

Effect of a required online graded curriculum in the clerkship years on medical student national standardized examination performance

<https://doi.org/10.1515/jom-2020-0298>

Received November 19, 2020; accepted April 9, 2021;

published online June 4, 2021

Abstract

Context: Many medical schools have a distributed model for clinical clerkship education, challenging our ability to determine student gaps during clinical education. With the graduating class of 2017, A.T. Still University's School of Osteopathic Medicine in Arizona (ATSU-SOMA) began requiring additional online curricula for all clerkship courses.

Objectives: To determine whether third year and fourth year students receiving ATSU-SOMA's online curricula during core clerkships performed better overall on national standardized examinations than students from previous years who had not received the curricula, and whether scores from online coursework correlated with outcomes on standardized examinations as possible early predictors of success.

Methods: This retrospective cohort study analyzed existing data (demographics and assessments) from ATSU-SOMA classes of 2017–2020 (curriculum group) and 2014–2016 (precurriculum group). The effect of the curriculum on national standardized examinations (Comprehensive Osteopathic Medical Achievement Test [COMAT] and Comprehensive Osteopathic Medical Licensing Examination of the United States [COMLEX-USA]) was estimated using augmented inverse probability weighting (AIPW).

*Corresponding author: Kelli Glaser, DO, Department of Clinical Science Education, A.T. Still University School of Osteopathic Medicine, 5850 E. Still Circle, Mesa, AZ, 85207-3618, USA, E-mail: kglaser@atsu.edu

Vanessa Pazdernik, MS, Department of Research Support, A.T. Still University, Kirksville, MO, USA

Denise Sackett, DO, Department of Clinical Science Education, A.T. Still University School of Osteopathic Medicine, Mesa, AZ, USA

Valerie Sheridan, DO, Department of Clinical Education, A.T. Still University School of Osteopathic Medicine, Mesa, AZ, USA

Correlations between assignment scores and national standardized examinations were estimated using linear regression models.

Results: The curriculum group had 405 students with a mean (standard deviation [SD]) age of 25.7 (± 3.1) years. Two hundred and fifteen (53.1%) students in the curriculum group were female and 190 (46.9%) were male. The precurriculum group had 308 students (mean \pm SD age, 26.4 \pm 4.2 years; 157 [51.0%] male; 151 [49.0%] female). The online curriculum group had higher COMAT clinical subject exam scores in obstetrics and gynecology, osteopathic principles and practice (OPP), psychiatry, and surgery (all $p \leq 0.04$), as well as higher COMLEX-USA Level 2-Cognitive Evaluation (CE) family medicine and OPP subscores (both $p \leq 0.03$). The curriculum group had a 9.4 point increase in mean total COMLEX-USA Level 2-CE score ($p = 0.08$). No effect was found for the curriculum overall on COMAT mean or COMLEX-USA Level 2-Performance Evaluation scores (all $p \geq 0.11$). Total coursework scores in each core clerkship, excluding pediatrics, were correlated with COMAT mean score (all adjusted $p \leq 0.03$). Mean scores for five of the seven assignment types in core clerkships, excluding evidence based medicine types, were positively correlated with COMAT mean scores (all adjusted $p \leq 0.049$). All assignment types correlated with COMLEX-USA Level 2-CE total score (all adjusted $p \leq 0.04$), except interprofessional education (IPE).

Conclusions: Results from this study of 713 students from ATSU-SOMA suggested that our online curriculum supplemented clinic based learning during clerkship courses and improved student outcomes on national standardized examinations.

Keywords: clerkship; COMAT; COMLEX-USA; curriculum; medical education; performance.

Anecdotally, many medical schools have a distributed model for clinical clerkship education, challenging our

ability to determine student gaps during clinical education. Although online curricula for clerkship education have become more common, the study of their influence on medical school outcomes is limited. One prior study [1] of online curricula from 2016 reporting on five narrated PowerPoint modules viewed online by 107 students in an emergency medicine (EM) clerkship found poor participation rates (22.45% reviewed all five modules and 36.4% viewed none), and used faculty derived pretests and posttests to measure student performance. Another study from 2015 [2] reporting the results of blended online learning in a surgery clerkship for 129 students at Johns Hopkins University also showed no statistically significant improvement in outcomes on the National Board of Medical Examiners (NBME) surgery subject exam, and analysis of scores for one term of the study year revealed statistically significant lower scores in the blended online learning group. Those investigators recommended long term follow up to verify the accuracy of their results [2]. Another study [3] published in 2015 evaluated the effect of two infectious disease modules in a pediatrics clerkship for 67 students, and results showed no statistically significant improvement in NBME pediatric shelf exam scores [3]. Yet another [4] evaluated the effect of student selected study materials during clerkships on NBME performance. Some studies have evaluated the effect of other factors on student performance during clinical education, such as preceptor quality [5], preceptor evaluations [6–8], clerkship order [9], and work hours [9]. To our knowledge, only two studies [10, 11] have investigated impact of an online curriculum during clerkships on student performance on National Board of Osteopathic Medical Examiners (NBOME) examinations. One [10] specifically looked at 78 students who opted to participate in a blended online learning supplement to their pediatrics clerkship, and results revealed no statistically significant improvement in their performance on the Comprehensive Osteopathic Medical Achievement Test (COMAT) Pediatrics examination. One of our own prior studies [11] evaluated required online coursework scores for 105 students in a family medicine (FM) clerkship and showed positive correlations with outcomes on National Board of Osteopathic Medical Examiners Comprehensive Osteopathic Medical Licensing Examination of the United States (COMLEX-USA) Level 2-Cognitive Evaluation (CE) examination and the COMAT FM examination [11].

Beginning with the graduating class of 2017 during their third year, A.T. Still University's School of Osteopathic Medicine in Arizona (ATSU-SOMA) began requiring online coursework for all core clerkships for third and fourth year students to standardize and supplement learning from preceptors and local hospitals, with the goal of enhanced learning of different skills and competencies

throughout the clerkship curriculum. Weekly graded assignments were added to supplement standard evaluation methods, such as clinical rotation evaluation, clinical encounter logs, and COMAT clinical subject exams (CSEs). The online curriculum varied, as course directors were given flexibility with its design. Most clerkships had 8–12 hours of online work per week for 4 weeks. This commonly included readings from various sources, procedure videos, medical websites, “gamified” cases, practice questions, and various other activities. Graded assignments created by clerkship directors were used to assess student learning, and the source, type, and frequency of assignments for any given course were also left to the discretion of the clerkship director. Assessments included traditional multiple choice questions and various written assessments that are further detailed in our methods description. Assessments typically occurred on a weekly basis and constituted 20% of the students' final course grades. Annual course reviews were conducted by a subgroup of the curriculum committee to evaluate course structure, materials, and assignments for quality.

Because of preceptor experience or practice preferences, some clerkships may have had limited education on or utilization of osteopathic principles and practice (OPP) or osteopathic skills, causing reduced exposure to osteopathic manipulative medicine (OMM). Therefore, ATSU-SOMA also offered an asynchronous online OPP course that ran concurrently with other clerkships throughout the third and fourth year. This course had no daily clinic based component with a preceptor, but required osteopathic manipulative treatment (OMT) practice activity at least once each semester.

The purpose of the current study was to determine whether third and fourth year students receiving the online curriculum during core clerkships performed better overall on national standardized examinations than students from previous years who had not received this online curriculum, and whether scores from the online coursework correlated with outcomes on national standardized examinations as possible early predictors of success.

Methods

Study design and participants

This retrospective cohort study analyzed existing data collected during the normal course of education at ATSU-SOMA. Records of 718 osteopathic medical students from the classes of 2014–2020 were included. Data were grouped by year of graduation, which determined participation in the online curricula (curriculum group, classes of 2017–2020; $n=308$) or nonparticipation (precurriculum group, classes of 2014–2016; $n=405$); data were excluded for five students who

appeared in both groups. All data were stored on secure servers and deidentified before analysis. The A.T. Still University-Arizona Institutional Review Board considered the study exempt.

Outcomes and predictor variables

Outcomes scores on national standardized examinations were used as benchmarks of success: COMAT CSEs, COMAT mean for all seven required clerkships, Comprehensive Osteopathic Medical Licensing Examination of the United States (COMLEX-USA) Level 2-CE, and COMLEX-USA Level 2-Performance Evaluation (PE). To evaluate outcomes relative to national performance level, we analyzed nationally standardized COMAT CSE scores, taking the difference between COMAT CSE scores from our cohorts and the mean score for the United States relative to standard deviation (SD). National statistics were obtained from COMAT national performance annual reports shared with our institution and applied to scores with exam dates that were within the respective reporting timeline.

Online coursework started in the 2015–2016 academic year, so the class of 2016 had minor exposure during their fourth year, possibly affecting COMAT scores for the EM clerkship. We included this cohort in the precurriculum group because this was our only data for the precurriculum group for the COMAT EM; this examination did not exist prior to 2015. The OPP course was analyzed separately because it had an existing online curriculum without the daily patient care component of other clerkships.

To compare groups, we collected the following data for potential confounding variables: demographics (matriculation age, sex, race, ethnicity), preadmission factors (undergraduate science and nonscience grade point average [GPA], Medical College Admission Test [MCAT] Biological and total scores), preclinical curriculum performance (GPA at end of second year, number of times COMLEX-USA Level 1 was taken before passing, COMLEX-USA Level 1 examination date relative to matriculation, COMLEX-USA Level 1 scores [total and subject specific]), and years to graduate relative to matriculation.

To determine whether scores from online coursework and individual assignments were correlated with outcomes on standardized examinations, we collected scores from several types of graded assignments and aggregate coursework scores for each clerkship. The seven assignment types were case analysis; article analysis; evidence based medicine (EBM) literature search; EBM research question; interprofessional education (IPE) reflection notes; multiple choice quiz (MCQ); and subjective, objective, assessment, and plan (SOAP) notes. For OPP courses, there were four assignment types: MCQ, scholarly activity, OMT Practice, and video.

Statistical analysis

Using fixed sample sizes and estimates from previous studies [11, 12] with R^2 of 0.27 for group averaged R^2 between each outcome and its covariates, we estimated the smallest differences in outcomes between groups with 80% power at 0.05 significance level to detect a difference: a 1.6 point difference in means of COMAT CSE scores, a 20.6 point difference in means of COMLEX-USA Level 2-CE scores, and a 6% change in passing rate for COMLEX-USA Level 2-PE.

To determine effect of the online curricula during core clerkships, we used a doubly robust estimator (augmented inverse probability weighting [AIPW]) for estimating average treatment (curricular) effect

[13]. We used this approach because propensity score is a validated statistical method for inferring causation in observational trials without randomization [13–18]. To infer causation, we assumed no unmeasured confounders existed between group and national standardized examination scores. Doubly robust estimation uses inverse probability weighting (IPW) using the propensity score and group specific regressions, allowing the estimator to remain consistent even if models (propensity score model or two group specific regression models) are not correctly specified.

For selection of covariates for the propensity score model, we considered confounding variables before the third year, except years to graduate, to avoid bias in our curricular effect estimate. Although years to graduate may have been affected by participation in the online curricula, other factors such as time for additional experiences for career success or personal leaves of absence may have affected that outcome. Thus, including time to graduate in the model balanced groups with respect to this variable. American Indian, Black, Pacific Islander, multiracial, and unknown race categories were balanced between the cohorts in aggregate because of small individual counts. We refined our variables using COMLEX-USA Level 1 total score as an outcome proxy to remove unrelated covariates and avoid increasing variance of the estimated curricular effect [14, 19, 20]. Ethnicity was unrelated to COMLEX-USA Level 1 and excluded as a covariate in the model. Because of an ATSU-SOMA administrative decision, minimum passing score on the Comprehensive Osteopathic Medical Self-Assessment Exam to attempt the COMLEX-USA Level 1 examination was higher for students in the curriculum group. This change may have caused higher COMLEX-USA Level 1 scores on the first attempt and delayed examination dates for the curriculum group, so we included time to first examination relative to matriculation when adjusting for scores. Twenty covariates were used for the propensity score model.

Depending on the variable, data were missing for at most 13 (1.8%) students in age, ethnicity, MCAT, undergraduate GPA, GPA at end of second year, and COMLEX-USA Level 1 and were imputed using chained equations with predictive mean matching [21]. Twenty two students had MCAT total and Biological scores from the new examination starting April 2015. New scores were converted to old by equating percentile ranks of historical data using linear interpolation from examinations administered January 2012 through September 2014 and May 2016 through April 2017 [22]. The highest score for all attempts was used for analysis. For each covariate, standardized differences between groups using nonimputed data with and without IPW were calculated. The propensity score method was considered adequate to control for potential confounding when IPW standardized Differences were within ± 0.10 . With AIPW, outcome models were fitted separately for each group and controlled for student GPA from first and second years and COMLEX-USA Level 1 total score. With low power to detect a difference between groups of less than 6% in passing rates for COMLEX Level 2-PE, we also assessed the Pearson correlation between this outcome with COMLEX Level 2-CE total score.

For students in the curriculum group (classes of 2017–2020), we estimated relationships between outcomes with online coursework total scores using Pearson correlation coefficient. Third year core clerkship assignment scores except those for OPP were unavailable for the class of 2017 because of data loss due to a change in the software used for clerkship assignment grading; fourth year core clerkship scores were unavailable for the class of 2020 because of the study timeline. The EM and surgery clerkships had a shift in curriculum assessment methods during the study period and were analyzed by

year. After assignments were categorized into one of the seven types for core clerkships or one of the four types for the OPP course, linear regression models were applied to estimate relationships between mean assignment scores and outcomes. For COMAT CSEs and COMLEX-USA Level 2-CE subscores with multiple outcomes for different clerkships, we regressed on clerkship specific pairs, allowed the intercept to vary by clerkship, and included a random effect for student for correlations of scores from the same student. Sensitivity analyses were performed by repeating analyses after excluding incompletes (zero score); results were reported when significance changed. Further, for both sets of analyses we use Bonferroni step-down adjustment for multiple comparisons within an outcome. Data are presented as frequency and percentage, mean and associated SD or 95% confidence interval (CI), and correlation or regression coefficient. SAS version 9.4 software (SAS Institute, Inc., Cary, NC) was used to conduct analyses. A value of $p < 0.05$ was considered statistically significant.

Results

Data from 718 osteopathic medical students from the classes of 2014–2020 were included. For the curriculum assessment, we excluded five students who matriculated in 2012 and graduated in 2017 because scores indicated partial participation in both groups. The curriculum group had 405 students with a mean age (\pm SD) of age 25.7 (\pm 3.1) years. Two hundred and fifteen (53.1%) students in the curriculum group were female and 190 (46.9%) were male. The precurriculum group had 308 students (mean \pm SD age, 26.4 \pm 4.2 years); 157 (51.0%) were male and 151 (49.0%) were female (Table 1). Standardized differences without IPW were found for age, race, undergraduate nonscience GPA, MCAT scores, number of times the COMLEX-USA Level 1 was taken before passing, subject specific COMLEX-USA Level 1 scores, and time to graduate (all unadjusted $p \leq 0.03$; Figure). IPW standardized differences ranged from -0.10 to 0.09 . The standardized difference between cohorts for Native Hawaiian or Pacific Islander race was slightly higher for the curriculum group (0.14), but again, groups were balanced among American Indian, Black, Pacific Islander, Multiracial, and unknown races in aggregate because of small individual counts (standardized difference = 0.02). Ethnicity was not used in the propensity model; the difference between groups was negligible using IPW ($p = 0.10$).

The curriculum group had higher COMAT CSE scores in obstetrics and gynecology (OB), OPP, psychiatry, and surgery than the precurriculum group (all $p \leq 0.04$; Table 2). For nationally standardized COMAT CSEs, the curriculum group had higher OPP and lower EM and pediatrics scores (both $p \leq 0.03$) and higher COMLEX-USA Level 2-CE FM and OPP subscores (both $p \leq 0.03$). The curriculum group had

higher scores on COMLEX Level 2-CE total score with a 9.4 point (95% CI, from -1.3 to 20.0) increase in mean total score ($p = 0.08$). Students who passed the COMLEX-USA Level 2-PE also had a mean score 39 points higher (95% CI, from 11.8 to 66.3) on the COMLEX-USA Level 2-CE examination ($p = 0.005$).

For relationships between coursework scores and outcomes, data from 410 students were analyzed; however, the sample size depended on the specific course and exam, ranging from 102 students to 406 students. Correlations between clerkship coursework total mean scores and national examination outcomes are reported in Table 3. For six clerkships (excluding pediatrics) and the OPP course, total mean coursework score was correlated with COMAT mean score (all adjusted $p \leq 0.03$) and with respective mean COMAT CSE scores for FM, IM, OBGYN, and OPP (all adjusted $p \leq 0.006$). Total coursework score correlated with COMLEX-USA Level 2-CE total score for FM, EM, IM, Psychiatry and OPP (all adjusted $p \leq 0.03$), and FM and psychiatry total coursework score correlated with respective COMLEX-USA Level 2-CE subscores (both adjusted $p \leq 0.02$).

Regression coefficients between outcomes and clerkship assignment types are reported in Table 4. The assignment types that positively correlated with COMAT mean score included case analysis, article analysis, IPE reflection note, SOAP Note, MCQ, and OPP scholarly (all adjusted $p \leq 0.049$). Assignment type MCQ and OPP MCQ were positively correlated with respective COMAT CSE scores (both adjusted $p \leq 0.007$). All assignment types, except IPE and OMT practice and video for OPP, were positively correlated with COMLEX-USA Level 2-CE total score (all adjusted $p \leq 0.04$).

There were exceptions found for aggregated assignment types failing to show clear correlations with a national exam outcome where an assignment in a single course correlated with a national exam outcome. All assignments for the studied courses and their correlations with national examination outcomes can be found in the Supplemental Material.

Discussion

We investigated whether third year and fourth year students receiving online curricula during core clerkships performed better overall on national standardized examinations than students from previous years and whether scores from the online coursework correlated with outcomes on national standardized examinations. Our results indicated the required online curricula, designed to supplement learning in the clinical environment for core

Table 1: Estimates and standardized differences in student characteristics between precurriculum and curriculum groups.

Characteristic	Precurriculum (n=308)		Curriculum (n=405)		Standardized difference without IPW	Unadjusted p-Value	Standardized difference with IPW	Adjusted p-Value
	No.	Mean \pm SD or n (%)	No.	Mean \pm SD or n (%)				
Demographic								
Age, years	307	26.4 \pm 4.2	404	25.7 \pm 3.1	-0.20	0.01	0.01	0.88
Sex	308		405		-0.08	0.28	-0.01	0.96
Female		151 (49.0)		215 (53.1)				
Male		157 (51.0)		190 (46.9)				
Ethnicity		20 (6.6)		34 (8.6)	0.08	0.32	0.14	0.10
Race	308		405			0.01		0.79
Asian		83 (27.0)		141 (34.8)	0.17		0.05	
White		179 (58.1)		191 (47.2)	-0.22		-0.06	
American Indian or Alaska native		1 (0.3)		3 (0.7)	0.06		0.07	
Black		3 (1.0)		8 (2.0)	0.08		0.02	
Native Hawaiian or other Pacific Islander		0 (0.0)		1 (0.3)	0.07		0.14	
Multiracial		20 (6.5)		26 (6.4)	0.00		-0.05	
Race unknown		22 (7.1)		35 (8.6)	0.06		0.03	
Preadmission^a								
Undergraduate science GPA	297	-0.03 \pm 0.3	404	0 \pm 0.3	0.10	0.19	-0.07	0.48
Undergraduate nonscience GPA	297	-0.04 \pm 0.2	405	0 \pm 0.2	0.17	0.03	-0.05	0.59
MCAT biology	297	-0.2 \pm 1.1	403	0 \pm 1.2	0.19	0.01	-0.08	0.44
MCAT total	297	-0.8 \pm 2.6	403	0 \pm 2.7	0.28	<0.001	-0.03	0.75
Preclinical curriculum								
SOMA GPA ^a	304	0.1 \pm 3.8	405	0 \pm 3.8	-0.03	0.67	-0.09	0.35
No. COMLEX-1 takes	303	1.1 \pm 0.4	405	1.0 \pm 0.2	-0.22	0.006	0.06	0.56
1		276 (91.1)		389 (96.1)				
2		20 (6.6)		14 (3.5)				
3+		7 (1.3)		2 (0.5)				
COMLEX-1 date, y	303	2.0 \pm 0.2	405	2.0 \pm 0.1	-0.12	0.16	0.09	0.40
<2		274 (90.4)		362 (89.4)				
2 < 2.5		19 (6.3)		39 (9.6)				
2.5 or more		10 (3.3)		4 (1.0)				
COMLEX-1 scores ^a	303		405					
Total		-30.6 \pm 78.1		0 \pm 80.9	0.38	<0.001	-0.07	0.55
Anatomy		-45.5 \pm 110.5		0 \pm 101.5	0.42	<0.001	-0.07	0.63
Behavioural science		-11.2 \pm 168.2		0 \pm 172.9	0.07	0.38	-0.02	0.79
Biochemistry		-12.9 \pm 125.9		0 \pm 132.8	0.10	0.19	-0.04	0.64
Microbiology		-49.6 \pm 106.5		0 \pm 120	0.42	<0.001	-0.02	0.86

Table 1: (continued)

Characteristic	Precurriculum (n=308)		Curriculum (n=405)		Standardized difference without IPW	Unadjusted p-Value	Standardized difference with IPW	Adjusted p-Value
	No.	Mean ± SD or n (%)	No.	Mean ± SD or n (%)				
OPP		-48.1 ± 94.4		0 ± 106.3	0.46	<0.001	-0.08	0.47
Pathology		-12.1 ± 119.8		0 ± 114.2	0.10	0.17	-0.10	0.47
Pharmacology		-31 ± 110.7		0 ± 125.7	0.26	<0.001	-0.04	0.75
Physiology		-20.3 ± 116.3		0 ± 111.6	0.18	0.02	-0.06	0.54
Graduation								
Time to graduate, y	308	4.00 ± 0.42	405	3.91 ± 0.25	-0.27	<0.001	0.05	0.70
<4		275 (89.3)		384 (94.8)				
4 < 4.5		7 (2.3)		6 (1.5)				
4.5 < 5		18 (5.8)		12 (3.0)				
5 or more		8 (2.6)		3 (0.7)				

A standardized difference is the difference between two estimates divided by the pooled SD. It reflects the size of the difference in the estimates relative to the SD. Standardized differences were calculated as curriculum minus precurriculum; therefore, negative values indicate the curriculum group was relatively lower than the precurriculum group and likewise positive values indicate the curriculum group was relatively higher than the precurriculum group. The propensity score method was considered adequate to control for potential confounding when IPW standardized differences were within ±0.10. Age was age at matriculation. American Indian, Black, Pacific Islander, and multiracial patients as well as those with unknown races were balanced between the cohorts in aggregate because of small individual counts. COMLEX-1 date was the COMLEX-USA Level 1 examination date relative to matriculation. All variables were used in the propensity score model, except ethnicity (Hispanic/Latino). ^aMean data for preadmission variables, SOMA GPA, COMLEX-1 scores are reported after subtracting the means of the curriculum group. COMLEX-1, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 1; GPA, grade point average; IPW, inverse probability weighting; MCAT, Medical College Admission Test; OPP, osteopathic principles and practice; SD, standard deviation; SOMA, School of Osteopathic Medicine in Arizona.

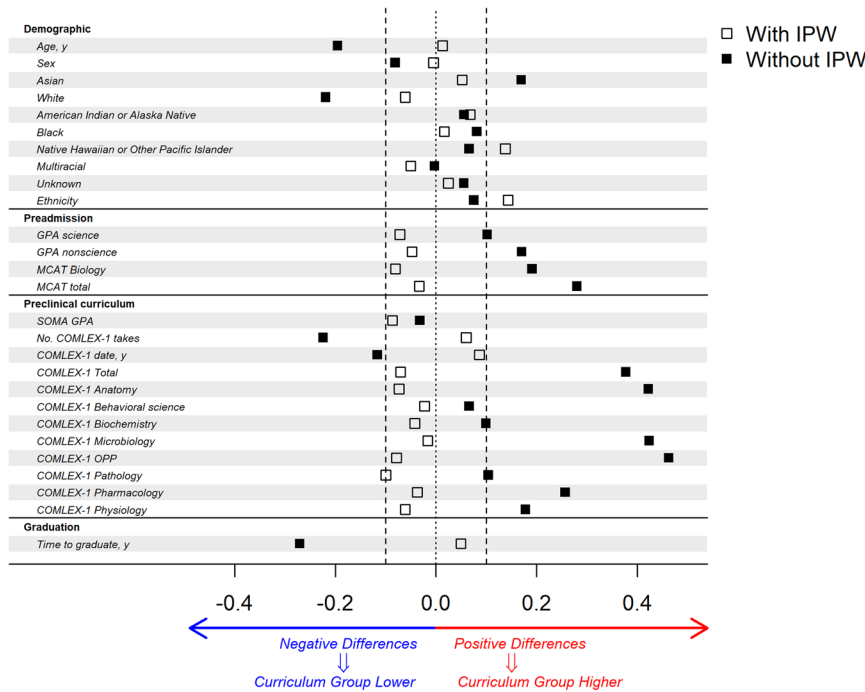


Figure 1: Standardized differences in student characteristics between the curriculum and precurriculum groups with and without inverse probability weighting (IPW) using the propensity score. A standardized difference is the difference between two estimates divided by the pooled standard deviation (SD); it reflects the size of the difference in the estimates relative to the SD. Standardized differences were calculated as curriculum minus precurriculum; therefore, negative values indicate the curriculum group was relatively lower than the precurriculum group and likewise positive values indicate the curriculum group was relatively higher than the precurriculum group. The propensity score method was considered adequate to control for potential confounding when IPW standardized differences were within ± 0.10 . Age was age at matriculation. COMLEX-1 date was the COMLEX-USA Level 1 examination date relative to matriculation. All variables were used in the propensity score model, except ethnicity (Hispanic/Latino). COMLEX-1, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 1; GPA, grade point average; MCAT, Medical College Admission Test; OPP, osteopathic principles and practice; SOMA, School of Osteopathic Medicine in Arizona.

clerkships, improved student performance on national standardized examinations. We also found correlations between total coursework scores and specific assignment types and outcomes on standardized examinations, which may be early predictors of success.

When analyzing COMAT using AIPW, the online curricula had a positive effect on scores in OPP, OB, psychiatry, and surgery. These clerkships, except for the OPP course, exclusively used MCQ assessments. When considering MCQ assessments across all clerkships, this positive correlation suggested that MCQ assessments may effectively predict COMAT CSE scores. Since MCQs are frequently used on standardized examinations, these assessments allowed students to practice, suggesting a possible causal relationship. Total coursework scores for two core clerkships and the OPP course were positively correlated with COMAT CSE scores. Some courses used multiple assessment types that correlated with COMAT CSE scores in at least one course and included case analysis, EBM literature search, EBM research

question, FM SOAP notes, IPE reflection notes, OPP scholarly activity, and MCQs.

When evaluating effect of the online curricula on COMAT scores relative to national performance level, the curriculum group performed worse on EM and pediatrics COMAT examinations. A possible explanation is that national means for pediatric COMAT scores were high in 2015–2016 when the first curriculum group cohort took the examination. After excluding this class from sensitivity analyses, results were unchanged. The COMAT national performance reports showed that the national mean during the time when the majority of the curriculum group took these examinations was higher (0.2–3.0 points) than when the precurriculum group took them, except IM (–0.2 points). This upward trend may have lessened the estimated effect of the online curriculum, especially since these negative effects in pediatrics and EM were not seen in raw scores. Perhaps these courses inadequately prepared students for the COMAT due to their absent or limited use of MCQ

Table 2: Average curriculum effect on scores of national standardized examinations between the curriculum group (n=405) and pre-curriculum group (n=308) using AIPW.

Outcome	Estimate	95% CI	p-Value
COMAT			
Mean	0.7	-0.2 to 1.6	0.11
Standardized	-0.07	-0.2 to 0.01	0.09
EM ^a	-1.8	-3.9 to 0.3	0.10
Standardized	-0.2	-0.4 to -0.02	0.03
FM	0.7	-0.8 to 2.1	0.37
Standardized	-0.01	-0.2 to 0.1	0.86
IM	-0.3	-1.6 to 0.9	0.60
Standardized	-0.02	-0.2 to 0.1	0.76
OBGYN	2.0	0.4–3.6	0.01
Standardized	0.01	-0.1 to 0.2	0.89
OPP ^b	1.8	0.2–3.4	0.02
Standardized	0.2	0.04–0.3	0.01
Pediatrics	-0.7	-2.4 to 1.0	0.44
Standardized	-0.4	-0.5 to -0.2	<0.001
Psychiatry	1.8	0.2–3.3	0.02
Standardized	-0.03	-0.2 to 0.1	0.72
Surgery	1.4	0.1–2.7	0.04
Standardized	0.00	-0.1 to 0.1	0.97
COMLEX-2 CE^c			
Total	9.4	-1.3 to 20.0	0.08
EM	11.8	-8.6 to 32.1	0.26
FM	18.9	1.6–36.2	0.03
IM	-6.2	-24.8 to 12.4	0.51
OBGYN	-9.2	-39.6 to 21.2	0.55
OPP	27.4	6.7–48.1	0.01
Pediatrics	-3.1	-22.7 to 16.4	0.75
Psychiatry	11.5	-38.0 to 61.0	0.65
Surgery	-15.1	-47.6 to 17.4	0.36
COMLEX-2 PE			
	0.0	-0.05 to 0.05	0.93

The COMAT mean was the mean score for all seven required clerkships (EM, FM, IM, OBGYN, pediatrics, psychiatry, and general surgery). The COMAT standardized was the difference between the student score and the national mean score divided by the national SD based on data from the annual COMAT national performance report. ^aNo scores were available for 207 students in the precurriculum group. ^bMissing one score from the precurriculum group. ^cMissing one score from the curriculum group and five scores from the precurriculum group. AIPW, augmented inverse probability weighting; CI, confidence interval; COMAT, Comprehensive Osteopathic Medical Achievement Test; COMLEX-2 CE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2-Cognitive Evaluation; COMLEX-2 PE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2-Performance Evaluation; EM, emergency medicine; FM, family medicine; IM, internal medicine; OBGYN, obstetrics and gynecology; OPP, osteopathic principles and practice; SD, standard deviation. The bold values signify the statistically significant results.

assignments. Students mostly take COMAT examinations in the third year; the COMAT CSE in EM is typically taken during students' fourth year. According to COMAT national performance reports that ATSU-SOMA receives

from NBOME, fewer students take COMAT CSE in EM than other COMAT CSEs. It is unclear how this factor affects results.

Our results showed that the curriculum group had a 9.4 point increase in mean total COMLEX-USA Level 2-CE score; an effect ranging from a 1.3 point decrease to a 20.0 point increase was reasonably compatible with our data, suggesting the potential for a positive effect although this outcome was not statistically significant. The NBOME discourages use of COMLEX-USA examination subscores for interpreting discipline performance [23], so using scores to identify strengths and weaknesses in a curriculum is limited by variety in number of examination items and overlap of assessment among disciplines. However, the curriculum group had significantly higher COMLEX-USA Level 2-CE FM and OPP subscores; FM total scores were strongly correlated with COMLEX-USA Level 2-CE total and FM subscore, particularly for FM II, which had the strongest correlation among coursework total scores. The 8 week FM clerkship at ATSU-SOMA is actually two courses (FM I and FM II), and each 4 week portion positively correlated with COMLEX-USA Level 2-CE total mean score. This FM subscore finding may result from broad coverage of topics from many disciplines during the FM clerkship. Curricular content from other clerkships may have also improved this subscore. This finding is supported by results from our previous research [11], which showed that some FM clerkship curriculum assignment scores were positively correlated with COMLEX-USA Level 2-CE. When comparing the FM clerkship with other clerkships, FM clerkships had a greater variety of assignment types (four), three of which were positively correlated with performance on COMLEX-USA Level 2-CE. Thus, effectiveness of the online curriculum may be partly explained by variety of assignment types.

Our findings showed a strong positive effect of the online curriculum for the COMLEX-USA Level 2-CE OPP mean subscore, which increased 27 points. This effect may be caused by added assessments during the OPP course but could be related to other clerkships integrating OPP concepts and assessments into their curricula. Ultimately, continuing to teach OPP during the clerkship years may improve performance on COMLEX-USA Level 2-CE.

In a previous study, objective structured clinical encounter scores and their constituent SOAP note scores correlated with performance on COMLEX-USA Level 2-PE [24]. In our current study, 302 students completed an aggregate of 1,011 SOAP note assignments graded by the clerkship directors as part of the clerkship curriculum. Student performance on these correlated, although perhaps weakly, with likelihood of passing COMLEX-USA Level 2-PE.

Table 3: Correlations between outcomes on national standardized examinations and clerkship coursework total scores.

Clerkship course	Total score ^a	COMAT mean			COMAT CSE			COMLEX-2 CE total			COMLEX-2 CE subject			COMLEX-2 PE							
		No.	r	p-Value	Adj No.	r	p-Value	No.	r	p-Value	No.	r	p-Value	No.	r	p-Value					
Third year core clerkship courses with mandatory coursework																					
Family medicine		303	0.28	<0.001	0.002	0.23	<0.001	0.002	0.302	0.18	0.001	0.02	0.302	0.13	0.03	0.37	0.303	0.10	0.08	>0.99	
Total FM I		303	0.36	<0.001	0.002	0.303	0.26	<0.001	0.002	0.302	0.35	<0.001	0.002	0.302	0.29	<0.001	0.002	0.303	0.08	0.16	>0.99
Internal medicine		303	0.07	0.24	0.96	0.303	0.02	0.70	0.85	0.302	-0.02	0.79	>0.99	0.302	-0.04	0.47	>0.99	0.303	-0.06	0.30	>0.99
Total IM I		303	0.17	0.003	0.03	0.303	0.20	<0.001	0.006	0.302	0.18	0.002	0.03	0.302	0.16	0.004	0.07	0.303	-0.006	0.92	>0.99
OBGYN		298	0.18	0.002	0.02	298	0.21	<0.001	0.006	297	0.08	0.17	>0.99	297	0.08	0.19	>0.99	298	0.07	0.22	>0.99
Total (Completed)		275	0.11	0.07	0.53	275	0.12	0.04	0.36	274	0.05	0.46	>0.99	274	0.13	0.04	0.48	275	0.05	0.45	>0.99
Pediatrics		302	0.12	0.04	0.41	302	0.13	0.02	0.21	301	0.06	0.31	>0.99	301	0.09	0.13	>0.99	302	0.07	0.26	>0.99
Psychiatry		303	0.21	<0.001	0.003	303	0.16	0.005	0.07	302	0.26	<0.001	0.002	302	0.19	<0.001	0.02	303	0.03	0.65	>0.99
Surgery		104	0.15	0.13	0.75	104	0.17	0.08	0.54	104	0.15	0.13	>0.99	104	0.08	0.43	>0.99	104	0.05	0.63	>0.99
Total (2018)		197	0.22	0.002	0.03	197	0.15	0.04	0.36	196	0.17	0.02	0.17	196	0.09	0.23	>0.99	197	0.001	0.99	>0.99
Total (2019, 2020)																					
Fourth year core clerkship courses with mandatory coursework																					
Emergency medicine		102	0.09	0.39	>0.99	102	0.16	0.11	0.55	102	0.08	0.42	>0.99	102	-0.02	0.84	>0.99	102	-0.15	0.13	>0.99
Total (2017)		209	0.27	<0.001	0.002	209	0.12	0.09	0.55	209	0.28	<0.001	0.002	209	0.16	0.02	0.27	209	0.08	0.25	>0.99
Total (2018, 2019)																					
Third year and fourth year courses with mandatory coursework without daily clinical requirement																					
OPP		405	0.16	0.001	0.02	405	0.19	<0.001	0.003	404	0.16	0.001	0.02	404	0.14	0.004	0.07	405	0.15	0.003	0.05
Third year sem 1 total		406	0.09	0.06	0.53	406	0.13	0.008	0.10	405	0.10	0.04	0.36	405	0.06	0.25	>0.99	406	0.06	0.23	>0.99
Third year sem 2 total		405	0.15	0.002	0.02	405	0.13	0.007	0.09	404	0.13	0.008	0.11	404	0.06	0.20	>0.99	405	0.11	0.02	0.43
(Completed)		314	-0.01	0.83	>0.99	314	-0.05	0.42	0.85	314	0.008	0.89	>0.99	314	-0.05	0.42	>0.99	314	0.03	0.58	>0.99
Fourth year sem 1 total		307	0.09	0.13	0.75	307	0.08	0.14	0.55	307	0.12	0.03	0.31	307	0.11	0.06	0.71	307	-0.02	0.76	>0.99
(Completed)		306	0.04	0.54	>0.99	306	0.09	0.12	0.55	306	0.06	0.28	>0.99	306	0.03	0.62	>0.99	306	0.007	0.91	>0.99
Fourth year sem 2 total		304	0.11	0.06	0.53	304	0.13	0.02	0.23	304	0.14	0.01	0.15	304	0.11	0.06	0.71	304	0.04	0.47	>0.99
(Completed)																					

Bonferroni step-down adjusted p where the number of hypotheses within each outcome is $m=19$. Reported r is the Pearson correlation coefficient. COMAT mean was the mean score for all seven required clerkships (emergency medicine, FM, IM, OBGYN, pediatrics, psychiatry, and general surgery). Sensitivity analyses excluded incompletes (zero scores), and results are reported when statistical significance changed. Third year core clerkship assignment scores, except for OPP, were unavailable for the class of 2017; fourth year core clerkship scores were unavailable for the class of 2020. ^aData are reported as a student's total score or the score from a single MCQ representing the coursework total score. Data are reported separately when content depended on the course. Adj, adjusted; COMAT, Comprehensive Osteopathic Medical Achievement Test; COMLEX-2 CE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2-Cognitive Evaluation; COMLEX-2 PE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2-Performance Evaluation; CSE, clinical subject exams; FM, family medicine; IM, internal medicine; MCO, multiple choice quiz; NA, not applicable; OBGYN, obstetrics and gynecology, OPP, osteopathic principles and practice; sem, semester. The bold values signify the statistically significant results.

Table 4: Linear regression coefficients from modelling outcomes on national standardized examinations and clerkship coursework scores by assignment type.

Assignment type	Clerkship			COMAT mean			COMAT CSE			COMPLEX-2 CE total			COMPLEX-2 CE subscore			COMPLEX-2 PE				
	No. ^a	b	p-Value	Adj p-Value	No. ^a	b	p-Value	Adj p-Value	No. ^a	b	p-Value	Adj p-Value	No. ^a	b	p-Value	Adj p-Value	No. ^a	b	p-Value	Adj p-Value
Case analysis	1,212, 405	1.4 (0.6, 2.2)	<0.001	0.006	1,212, 405	63 (8, 118)	0.03	0.20	1,209, 404	18 (7, 29)	0.001	0.02	1,209, 404	4.4 (-2.4, 11.3)	0.21	>0.99	1,212, 405	1.1 (0.8, 1.7)	0.58	>0.99
Article analysis	1,611, 306	2.8 (1.4, 4.2)	<0.001	0.001	1,611, 306	-26 (-80, 28)	0.35	>0.99	1,606, 305	42 (23, 61)	<0.001	0.001	1,606, 305	4.8 (-3.4, 12.9)	0.25	>0.99	1,611, 306	1.4 (0.8, 2.6)	0.27	>0.99
EBM: literature search	208, 208	1.2 (0.3, 2.1)	0.01	0.07	208, 208	132 (28, 236)	0.01	0.13	208, 208	19 (6, 32)	0.004	0.03	208, 208	12.8 (-4.3, 29.9)	0.14	>0.99	208, 208	1.3 (0.95, 1.7)	0.10	>0.99
(Completed)	207, 207	1.2 (0.1, 2.3)	0.04	0.14	207, 207	81 (-48, 210)	0.22	>0.99	207, 207	22 (6, 38)	0.009	0.04	207, 207	10.0 (-11.4, 31.4)	0.36	>0.99	207, 207	0.9 (0.5, 1.8)	0.83	>0.99
EBM: research question	209, 209	1.0 (0.2, 1.8)	0.02	0.09	209, 209	105 (11, 200)	0.03	0.20	209, 209	19 (7, 31)	0.002	0.02	209, 209	17.5 (2.1, 33.0)	0.03	0.31	209, 209	1.2 (0.9, 1.5)	0.31	>0.99
(Completed)	207, 207	1.3 (0.0, 2.6)	0.04	0.14	207, 207	191 (44, 337)	0.01	0.12	207, 207	11 (-8, 29)	0.25	0.92	207, 207	3.9 (-20.1, 27.9)	0.75	>0.99	207, 207	1.4 (1.0, 2.2)	0.05	0.57
IPE reflection note	614, 306	0.9 (0.3, 1.5)	0.006	0.049	614, 306	45 (10, 80)	0.01	0.13	613, 305	5 (-3, 14)	0.23	0.92	613, 305	0.0 (-5.2, 5.2)	>0.99	>0.99	614, 306	1.1 (0.8, 1.5)	0.55	>0.99
MCQ	4,111, 303	2.6 (1.8, 3.4)	<0.001	0.001	4,111, 303	44 (19, 69)	<0.001	0.007	4,096, 302	30 (19, 41)	<0.001	0.001	4,096, 302	6.8 (1.8, 11.7)	0.007	0.09	4,111, 303	1.3 (0.9, 2.0)	0.16	>0.99
SOAP note	1,018, 303	1.6 (0.7, 2.5)	<0.001	0.004	1,018, 303	21 (-29, 71)	0.42	>0.99	1,016, 302	24 (13, 36)	<0.001	0.001	1,016, 302	2.6 (-3.9, 9.1)	0.43	>0.99	1,018, 303	1.2 (0.9, 1.7)	0.19	>0.99
(Completed)	1,011, 302	1.6 (0.5, 2.8)	0.006	0.049	1,011, 302	21 (-39, 80)	0.50	>0.99	1,009, 301	25 (9, 41)	0.002	0.02	1,009, 301	2.3 (-5.4, 10.1)	0.55	>0.99	1,011, 302	1.8 (1.1, 3.0)	0.02	0.25
MCQ	404, 404	1.3 (0.3, 2.2)	0.008	0.05	404, 404	202 (88, 316)	<0.001	0.007	403, 403	20 (7, 33)	0.003	0.02	403, 403	24.9 (7.8, 42.1)	0.005	0.06	404, 404	1.6 (1.1, 2.4)	0.02	0.25

Table 4: (continued)

Assignment type	Clerkship	COMAT mean			COMAT CSE			COMLEX-2 CE total			COMLEX-2 CE subscore			COMLEX-2 PE							
		No. ^a	b	p- Value	Adj p- Value	No. ^a	b	p- Value	Adj p- Value	No. ^a	b	p- Value	Adj p- Value	No. ^a	b	p- Value	Adj p- Value				
Scholarly	OPP	1,420,	2.3(1.0,	<0.001	0.006	1,420,	214(54,	0.009	0.11	1,418,	29(11,	0.002	0.02	1,418,	23.3	0.06	0.64	1,420,	1.9	0.02	0.22
		409	3.5)			409	374)			408	48)			408	(-0.8,			409	(1.1,		
OMT practice	OPP	1,422,	1.1	0.15	0.31	1,422,	181(-8,	0.06	0.36	1,420,	6(-16,	0.60	>0.99	1,420,	3.0	0.83	>0.99	1,422,	1.7	0.10	>0.99
		409	(-0.4,			409	371)			408	28)			408	(-25.6,			409	(0.9,		
Video	OPP	1,228,	0.2	0.76	0.76	1,228,	-59	0.45	>0.99	1,226,	2(-16,	0.85	>0.99	1,226,	-9.1	0.44	>0.99	1,228,	1.1	0.78	>0.99
		409	(-1.0,			409	-213,			408	19)			408	(-32.2,			409	(0.6,		
			1.4)				94)							14.0)							1.8)

Bonferroni step-down adjusted p where the number of hypotheses within each outcome is $m=14$. Unstandardized regression coefficients (β) and associated 95% confidence intervals represent the predicted change in outcome due to a 10% increase in percentage points in mean coursework score for that assignment type. Change is reported as points for COMAT and COMLEX-2 CE and as odds of passing for COMLEX-2 PE. COMAT mean was the mean score for all seven required clerkships (EM, FM, IM, OBGYN, pediatrics, psychiatry, general surgery). Sensitivity analyses excluded incomplete assignments (zero scores), and results are reported when statistical significance changed. Third year core clerkship assignment scores, except for OPP, were unavailable for the class of 2017; fourth year core clerkship scores were unavailable for the class of 2020. ^aNo. is reported as the number of assignments, number of students-Adj, adjusted; COMAT, Comprehensive Osteopathic Medical Achievement Test; COMLEX-2 CE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2 Cognitive Evaluation; COMLEX-2 PE, Comprehensive Osteopathic Medical Licensing Examination of the United States Level 2 Performance Evaluation; CSE, clinical subject exams; EBM, evidence-based medicine; EM, emergency medicine; FM, family medicine; IM, internal medicine; IPE, interprofessional education; MCQ, multiple choice quiz; OBGYN, obstetrics and gynecology; OMT, osteopathic manipulative treatment; OPP, osteopathic principles and practice; SOAP, subjective objective assessment and plan. The bold values signify the statistically significant results.

SOAP notes allow students to consider experiential knowledge and put it into written form, which requires substantial analytical skills. They also allow students to practice this skill assessed on COMLEX-USA Level 2-PE. Perhaps writing SOAP notes, whether for an objective structured clinical encounter or online curriculum, and the resulting critical feedback were beneficial and provided early identification of those who would struggle with this examination. Albeit also somewhat weakly, student performance on scholarly activity coursework during OPP was also positively correlated with passing COMLEX-USA Level 2-PE. Perhaps coursework refreshed knowledge of OMM essential for integration of osteopathic principles into patient encounters of COMLEX-USA Level 2-PE.

We found no specific type of assignment consistently correlated with performance on all national standardized examinations. However, all types of assignments were positively correlated with performance on at least one national standardized examination, except OMT Practice and video. MCQ assessments were most frequently positively correlated but were also the most frequently used assessment. These assignments test foundational knowledge while providing practice with this type of test item. An explanation for observed results for case and article analyses may be that they required critical thinking that enhanced student thought processes. The EBM assignments evaluating student ability to conduct a literature search or identify a research question may also engage critical thinking and decision making skills. Although IPE assignments teach aspects of patient care and collaboration essential for the practice of medicine and are not designed to enhance examination performance, they still correlated with COMAT mean score. Regardless of correlations found, we cannot definitively conclude how assessment type positively impacted mean national standardized examination scores because the curriculum group was exposed to all assessments; future studies could evaluate individual curricular aspects further. Our findings for coursework assignments may help clerkship directors improve learner preparedness by choosing more valid and reliable assessments. Ultimately, a variety of assignment types may better prepare students for national standardized examinations.

Other factors may have influenced our students' performance on national standardized examinations. Impact of the online curriculum is likely smaller than student learning with preceptors and patients during clerkships. Further, students may use other resources to prepare for examinations. In one prior study, data showed an effect between quantity and type of resources used and NBME Clinical Science Surgery examination scores [4]. Future studies should investigate use of external resources

on COMAT CSE performance. When comparing our results with national mean scores, resources were likely available at other colleges of osteopathic medicine that our students did not have, impacting effect of our curriculum relative to national performance. Time spent studying correlates with NBME scores [4]; however, this outcome was previously evaluated at our institution with a single graduating class, and no correlation was found between reported study time and performance on COMAT CSEs or COMLEX Level 2-CE [11]. Study time was not constantly measured during the current study, so we could not reevaluate this factor. In general, causation is significantly difficult to prove in educational research due to the multitude of confounding factors that can influence student learning and assessment success.

Limitations

Some limitations of the current study may be addressed in future research. Although we used 7 years of student data, our study was conducted at a single institution. Since other medical schools may lack an online component in clerkship curricula, results may not be generalizable, and future studies should be conducted at other institutions for better applicability. Because we used existing data, we were unable to control for clinical experiences on preceptor quality and student reported clinical logs volume and variety; clinical experiences occurred concurrently with the online curriculum and controlling for these variables may remove part of the curriculum effect. However, we expect clinical that experiences were consistent between groups. Assignments were typically created by our clerkship directors and were not nationally standardized or accessible to the public. Therefore, grading rubrics for coursework may have varied in scope and application between courses. Future studies should compare clerkships with different correlations of assignment scores with national standardized examination scores to identify tools and processes impacting evaluation of student performance. Such information may contribute to deeper understanding of assessment and early identification of struggling learners. More detailed investigations of successful clerkships and assessments may be necessary to develop additional curricular enhancements for other clerkships.

Conclusions

Our results suggested that the online curricula at ATSU-SOMA supplemented traditional clinic based learning during clerkship years and improved student outcomes on

national standardized examinations. Notably, OPP instruction during the clerkship years was quite influential on national examination scores and should be considered by osteopathic schools as a core offering. Given current trends in medical education that have resulted in more clinical education occurring online, enhancing the clerkship curriculum with additional online learning and evaluating its effectiveness is important to ensure students are learning as much as possible during their third and fourth years of osteopathic medical school.

Research funding: This study was supported by a grant from the American Association of Colleges of Osteopathic Medicine (grant no. 501-705).

Author contributions: All authors provided substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; all authors drafted the article or revised it critically for important intellectual content; all authors gave final approval of the version of the article to be published; and all authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests: None reported.

Ethical approval: The A.T. Still University-Arizona Institutional Review Board considered the study exempt.

References

1. Lew EK, Nordquist EK. Asynchronous learning: student utilization out of sync with their preference. *Med Educ Online* 2016;21:30587.
2. Lindeman BM, Law JK, Lipsett PA, Arbella T, Stem M, Lidor AO. A blended online curriculum in the basic surgery clerkship: a pilot study. *Am J Surg* 2015;209:145–51.
3. Khasawneh R, Simonsen K, Snowden J, Higgins J, Beck G. The effectiveness of e-learning in pediatric medical student education. *Med Educ Online* 2016;21:29516.
4. Volk AS, Rhudy AK, Marturano MN, Ott L, DuCoin C. Best study strategy for the NBME Clinical Science Surgery exam. *J Surg Educ* 2019;76:1539–45.
5. Blue AV, Griffith CH 3rd, Wilson J, Sloan DA, Schwartz RW. Surgical teaching quality makes a difference. *Am J Surg* 1999;177:86–9.
6. Lee KB, Vaishnavi SN, Lau SK, Andriole DA, Jeffe DB. “Making the grade”: noncognitive predictors of medical students’ clinical clerkship grades. *J Natl Med Assoc* 2007;99:1138–50.
7. Plymale MA, French J, Donnelly MB, Iocono J, Pulito AR. Variation in faculty evaluations of clerkship students attributable to surgical service. *J Surg Educ* 2010;67:179–83.
8. Riese A, Rappaport L, Alverson B, Park S, Rockney RM. Clinical performance evaluations of third-year medical students and association with student and evaluator gender. *Acad Med* 2017;92:835–40.
9. Dong T, Copeland A, Gangidine M, Schreiber-Gregory D, Ritter EM, Durning SJ. Factors associated with surgery clerkship performance and subsequent USMLE Step scores. *J Surg Educ* 2018;75:1200–5.
10. Langenau EE, Lee R, Fults M. Blended learning educational format for third-year pediatrics clinical rotation. *J Am Osteopath Assoc* 2017;117:234–43.
11. Glaser K, Sackett D, Pazdernik VK. Success predictors for third-year osteopathic medical students on national standardized examinations: a family medicine clerkship course study. *J Am Osteopath Assoc* 2020;120:253–62.
12. Dixon D. Prediction of osteopathic medical school performance on the basis of MCAT score, GPA, sex, undergraduate major, and undergraduate institution. *J Am Osteopath Assoc* 2012;112:175–81.
13. Funk MJ, Westreich D, Wiesen C, Sturmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol* 2011;173:761–7.
14. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
15. Gunasekara FI, Richardson K, Carter K, Blakely T. Fixed effects analysis of repeated measures data. *Int J Epidemiol* 2014;43:264–9.
16. Shardell M, Ferrucci L. Joint mixed-effects models for causal inference with longitudinal data. *Stat Med* 2018;37:829–46.
17. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci* 2010;25:1–21.
18. Yazdani A, Boerwinkle E. Causal inference in the age of decision medicine. *J Data Min Genom Proteonomics* 2015;6:2–7.
19. Kelcey B. Covariate selection in propensity scores using outcome proxies. *Multivariate Behav Res* 2011;46:453–76.
20. Cuong NV. Which covariates should be controlled in propensity score matching? Evidence from a simulation study. *Stat Neerl* 2013;67:169–80.
21. van Buuren S, Goothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Software* 2011;45:1–67.
22. Percentile ranks for the MCAT exam. Association of American Medical Colleges Web site. Available from: <https://students-residents.aamc.org/advisors/article/percentile-ranks-for-the-mcat-exam/> [Accessed 7 Oct 2020].
23. Scoring principles. National Board of Osteopathic Medical Examiners Web site. Available from: <https://www.nbome.org/exams-assessments/comlex-usa/comlex-usa-level-2-ce/scores-transcripts/scoring-principles/> [Accessed 7 Oct 2020].
24. Wang S, Basehore P. Relationship of clinical skills performance in medical school with COMLEX-USA Level 2-Performance Evaluation. *J Am Osteopath Assoc* 2020;120:245–52.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/jom-2020-0298>).