

Kenneth D. Royal*, PhD, MEd, Christian Meyer, BS, Erik Guercio, MA, Mark Speicher, PhD, Joseph Flamini, MBA, Jeanne M. Sandella, DO, Tsung-Hsun Tsai, PhD and Cynthia A. Searcy, PhD

The predictive validity of MCAT scores and undergraduate GPA for COMLEX-USA licensure exam performance of students enrolled in osteopathic medical schools

<https://doi.org/10.1515/jom-2023-0265>

Received December 5, 2023; accepted May 16, 2024;

published online July 18, 2024

Abstract

Context: Osteopathic (Doctor of Osteopathic Medicine [DO]) medical students account for more than 25 % of all medical students in the United States.

Objectives: This study examined the predictive validity of Medical College Admission Test (MCAT) total scores and cumulative undergraduate grade point averages (UGPAs) for performance on the Comprehensive Osteopathic Medical Licensing Examination of the United States (COMLEX-USA) Level 1 and Level 2-CE (Cognitive Evaluation) licensure examinations administered by the National Board of

Osteopathic Medical Examiners (NBOME). Additionally, the study examined the degree to which MCAT total scores and UGPAs provide comparable prediction of student performance by key sociodemographic variables.

Methods: This study involved a collaborative effort between the Association of American Medical Colleges (AAMC), the American Association of Colleges of Osteopathic Medicine (AACOM) and the NBOME. Data were examined for 39 accredited DO-granting medical schools in the United States during the 2017 application cycle. Researchers utilized three regression models that included MCAT total scores, cumulative UGPA, and combined MCAT total scores and cumulative UGPA to determine predictive validity. Researchers also examined the comparability of prediction for sociodemographic variables by examining the differences between observed and predicted error for both scores and pass/fail success rates.

Results: Medium to large correlations were discernible between MCAT total scores, UGPA, and COMLEX-USA examination outcomes. For both COMLEX-USA Level 1 and Level 2-CE scores and pass/fail outcomes, MCAT scores alone provided superior predictive value to UGPA alone. However, MCAT scores and UGPA utilized in conjunction provided the best predictive value. When predicting both licensure examination scores and pass/fail outcomes by sociodemographic variables, all three models provided comparable predictive accuracy.

Conclusions: Findings from this comprehensive study of DO-granting medical schools provide evidence for the value-added benefit of taking MCAT scores and UGPA into consideration, particularly when these measures are utilized in conjunction. Further, findings provide evidence indicating that individuals from different sociodemographic backgrounds who enter medical school with similar MCAT scores and UGPA perform similarly on licensure examination outcome measures.

Keywords: assessment; evaluation; medical education; medical licensure; osteopathic medical education; validity

***Corresponding author: Kenneth D. Royal**, PhD, MEd, Director of MCAT Research and Data Science, Association of American Medical Colleges, 655 K Street, Washington, DC 20001, USA, E-mail: kroyal@aamc.org. <https://orcid.org/0000-0002-5508-1480>

Christian Meyer, BS, Psychometric Intern, Association of American Medical Colleges, Washington, DC, USA

Erik Guercio, MA, Senior Director of Research, American Association of Colleges of Osteopathic Medicine, Bethesda, MD, USA. <https://orcid.org/0000-0002-8748-9114>

Mark Speicher, PhD, Senior Vice President for Learning, Innovation and Research, American Association of Colleges of Osteopathic Medicine, Bethesda, MD, USA. <https://orcid.org/0000-0002-5273-8299>

Joseph Flamini, MBA, Vice President for Administration/Chief Operating Officer, National Board of Osteopathic Medical Examiners, Chicago, IL, USA

Jeanne M. Sandella, DO, Vice President for Professional Development Initiatives and Communications, National Board of Osteopathic Medical Examiners, Conshohocken, PA, USA

Tsung-Hsun Tsai, PhD, Associate Vice President for Assessment Services and Research, National Board of Osteopathic Medical Examiners, Chicago, IL, USA

Cynthia A. Searcy, PhD, Senior Director of MCAT Science Strategy and Solutions, Association of American Medical Colleges, Washington, DC, USA

Admissions committees utilize holistic review, a process that considers various academic metrics, attributes, and experiences in their appropriate context, to identify applicants that are likely to be successful in medical school [1]. Part of the holistic review process also involves identifying the best combination of students that can help medical schools carry out their individual missions, which typically include educating a diverse body of competent physicians who provide quality care to diverse populations and contributing to research that improves health-related outcomes for all [2]. Holistic review typically involves consideration of several important domains for applicants, including academic metrics, backgrounds, lived experiences, and personal characteristics. A 2021 survey of admissions officers by the Association of American Medical Colleges (AAMC) found that 95 % of Doctor of Medicine (MD) programs reported utilizing elements of holistic review [3], although specific practices (e.g., weighting across and within domains) vary considerably across institutions.

Regarding the issue of academic readiness in medical school admissions, research has long indicated that two of the most widely utilized academic metrics are scores from the Medical College Admission Test (MCAT), which measures an applicant's knowledge in foundational concepts and reasoning skills necessary for medical school, and undergraduate grade point average (UGPA), which provides a comprehensive measure of the applicant's academic achievement during their undergraduate education [4].

A recent study by Hanson and colleagues [5] examined the predictive validity of MCAT scores and UGPA for United States Medical Licensing Examination (USMLE) Step 1 and Step 2 Clinical Knowledge (CK) examinations. Although the study reported multiple, important findings, the primary outcomes were that: (1) MCAT scores and UGPA utilized together provided the best prediction for Step 1 and Step 2 CK examination performance; and (2) a combination of MCAT scores and UGPA also provided comparable prediction for Step 1 and Step 2 CK examination performance by race and ethnicity, sex, and socioeconomic status (SES). These findings provide strong evidence that MCAT scores and UGPA play an important role in helping admissions committees match individuals' premedical preparation with future medical school curricula and learning environments.

The Hanson et al. [5] study was limited to evaluating the predictive validity of MCAT scores for students in US-based MD-granting medical schools. The other type of physician training program in the United States, osteopathic medical (Doctor of Osteopathic Medicine [DO]) schools, account for a growing portion of the total physician workforce and medical student body. According to the American Osteopathic Association (AOA), the field of osteopathic medicine has experienced tremendous growth over the last decade, with an 81 % increase in the number of osteopathic physicians

($n=141,759$) currently in the workforce [6]. Further, DO physicians currently account for approximately 11 % of the entire physician workforce. With respect to medical education, there are currently over 36,500 students enrolled in a DO program in the United States, accounting for more than 25 % of all medical students in the United States [6, 7].

DO students follow a training model that focuses on osteopathic principles and practice and osteopathic manipulative treatment [8]. Despite these key differences, both MD and DO students follow a similar educational pathway and often train alongside one another in internship and residency programs. Further, both MD and DO physicians obtain a license to practice medicine. Whereas students in MD programs seek licensure by passing the USMLE examination administered by the National Board of Medical Examiners (NBME), students in DO programs seek licensure by passing the Comprehensive Osteopathic Medical Licensing Examination (COMLEX-USA) administered by the National Board of Osteopathic Medical Examiners (NBOME).

Like MD-granting medical schools, admissions committees in DO-granting medical schools also consider MCAT scores and UGPA as two important academic metrics in their holistic review process. However, there are several important differences between MD and DO students before they matriculate to medical school that may make the two student populations inherently different. Since 2011, there has been a discernible trend in primary care that has seen an increasing number of DO students and a decreasing number of MD students pursue a career in primary care specialties (family medicine, general internal medicine, or general pediatrics) [9]. Recent data indicate that these trends will continue, as 32 % of all DO students plan to specialize in primary care compared to 18 % of MD students, and 53 % of DO students plan to practice in underserved areas upon graduation compared to 35 % of MD students [10, 11].

Further, because institutions are committed to identifying applicants who will best fit their curricular environment and help advance their educational mission, this may lead to differences in other metrics, such as MCAT scores and UGPAs, which may in turn impact relationships with future measures (e.g., USMLE and COMLEX examinations). Because DO students must take a different licensing examination than the MD students presented in the Hanson et al. [5] study, it is worth replicating this study with the DO student population for direct and meaningful comparison.

Thus, the purpose of this study was to examine the predictive validity of MCAT scores and UGPA in relation to DO student performance on COMLEX-USA licensure examinations. We attempted to accomplish this aim in the most comprehensive manner possible by examining data across all DO medical schools. The research questions guiding this study were: (1) do MCAT total scores and UGPA predict student performance on the COMLEX-USA Level 1 and Level

2-CE (Cognitive Evaluation) examinations?; and (2) do MCAT total scores and UGPA provide comparable prediction of student COMLEX-USA Level 1 and Level 2-CE performance by sex, race and ethnicity, and SES?

Methods

Data sources and participants

This study was part of a collaborative effort between the AAMC, the American Association of Colleges of Osteopathic Medicine (AACOM) and the NBOME. Each participant's most recent MCAT record was provided by the AAMC; participant background characteristics and matriculation information were provided by AACOM; and the NBOME provided COMLEX-USA examination results for students who consented for their results to be utilized for research purposes.

The authors analyzed data from individuals who matriculated to 1 of 39 accredited DO-granting medical schools in the United States during the 2017 application cycle. To qualify for inclusion in the study, participants must have reported UGPA and scores from the current MCAT examination on their application. In the sample, a total of 17,470 individuals applied to at least one DO-granting medical school, and 5,919 (33.9 %) individuals matriculated.

Based on the 5,919 individuals who matriculated, 52.4 % ($n=3,100$) were male and 47.5 % ($n=2,811$) were female. A total of 683 (12.4 %) matriculants self-reported a race or ethnicity that was categorized as underrepresented in medicine (UIM). This group of students consisted of individuals who identified as Black or African American; Hispanic, Latino or Spanish Origin; American Indian or Alaska Native; or Native Hawaiian or other Pacific Islander. Further, 1,295 (23.4 %) matriculants' parents did not attain a bachelor's degree.

For the validity analyses, matriculants must have completed either the COMLEX-USA (Level 1 or Level 2-CE) examination, or both, during their current program of study. The final sample frame consisted of 5,208 students for the COMLEX-USA Level 1 examination and 4,795 students for the COMLEX-USA Level 2 CE examination. All students with COMLEX-USA Level 2-CE examination results had COMLEX-USA Level 1 examination results. Of the 5,919 matriculant records provided by AACOM, 711 (12 %) could not be matched to a COMLEX-USA Level 1 examination record and 1,124 (19 %) could not be matched to either a COMLEX-USA Level 1 or 2 CE examination record.

Selection variables and outcome measures

The authors investigated two selection variables: MCAT total scores and cumulative UGPA. These variables were chosen because they typically are regarded as two important variables utilized by admissions committees [4, 12, 13]. The MCAT total score, ranging from 472–528, is the sum of scores from the 4 sections of the examination (Biological and Biochemical Foundations of Living Systems; Chemical and Physical Foundations of Biological Systems; Psychological, Social, and Biological Foundations of Behavior; and Critical Analysis and Reasoning Skills). We studied the most recent MCAT total scores.

To determine if MCAT total scores and UGPA provide comparable prediction of students' performance on COMLEX-USA Level 1 and Level 2-CE examinations by sex, race and ethnicity, and SES, we divided students into two groups for each demographic variable. Students self-reported their sex as either male or female. Race and ethnicity were categorized into UIM or non-UIM based on the aforementioned definitions. Because parental educational attainment is strongly associated with family income and wealth, parental educational attainment was utilized as a proxy for SES [14]. Students for whom neither parent held a bachelor's degree were categorized as lower-SES, and students for whom at least one parent held a bachelor's degree were categorized as upper-SES.

The outcome variables investigated in this study were student scores based on their first attempt at the COMLEX-USA Level 1 and COMLEX-USA Level 2-CE licensure examinations. According to the NBOME, the COMLEX-USA Level 1 assesses "foundational biomedical sciences with other areas of medical knowledge relevant to clinical problem-solving and the promotion of health maintenance" [15]. The COMLEX-USA Level 2-CE assesses "application of knowledge in clinical and foundational biomedical sciences with other physician competences related to the clinical care of patients and promoting health in supervised clinical settings" [16]. Both COMLEX-USA examinations produce a scaled score and a pass/fail outcome measure. The scaled score range for both examinations is 9–999. This study was deemed exempt by the American Institutes for Research in Behavioral Sciences Institutional Review Board (Protocol # EX00591).

Data analysis and procedures

The authors utilized three linear regression models to determine if MCAT total scores and UGPA predict performance on students' first attempt of COMLEX-USA Level 1 and

Level 2-CE examinations. The first model examined the predictive value of MCAT scores alone; the second model examined the predictive value of UGPA alone; and the third model examined the predictive value of MCAT scores and UGPA combined. Due to the many nuances and variations in medical schools (e.g., mission, applicant pools, curricula, etc.), all analyses were estimated within each medical school and then summarized across schools. Because COMLEX-USA scores can only be obtained for matriculants, the bivariate UGPA and MCAT score distribution of the analytical sample is restricted relative to the population. We utilized the simultaneous multivariate correction method given in Sackett and Yang [17] to adjust the correlation in the restricted sample to estimate the correlation between the predictors and COMLEX-USA scores. This procedure alleviated the underestimation bias associated with utilizing only the matriculant sample. For each model, correlations across schools were summarized with median and interquartile range (IQR).

The authors utilized a logistic regression on pass/fail outcomes. Schools with 100 % pass rates on either the COMLEX-USA Level 1 or Level 2-CE examinations were excluded from analysis because relations between variables cannot be estimated when one or both is constant in a sample. Thus, a logistic regression was performed on 35 schools for the Level 1 analysis and on 31 schools for the Level 2-CE analysis. We examined the area under the receiver operating curve (AUC) to determine how well each model differentiated students. We then summarized the AUC for each model across schools with the median and IQR.

For the comparable prediction analysis, we calculated the aggregate prediction error within each group and then summarized the values across schools. For scores, this was computed as the average of the difference between the predicted and observed scores. For pass/fail indicators, this was determined by dividing the observed success rate by the predicted success rate. Standardized effect sizes (Cohen's *d* and Cohen's *h*, respectively) were then calculated to discern the practical significance of any observed differences [18]. We performed all statistical analyses utilizing R statistical computing programming language version 4.2.3 for Windows (R Core Team).

Results

Overall summary of MCAT scores, UGPA, and COMLEX-USA performance outcomes

The mean MCAT total score for matriculants ($n=5,208$) was 503.0 ($SD=5.5$). The mean cumulative UGPA for matriculants

was 3.55 ($SD=0.26$). With respect to matriculants' performance on COMLEX-USA examinations, the mean COMLEX-USA Level 1 score was 533.1 ($SD=87.0$), with 4,941 (94.9 %) passing on their first attempt. The mean COMLEX-USA Level 2 score was 573.5 ($SD=95.4$), with 4,677 (97.5 %) passing on their first attempt. A statistical summary of the overall results by subgroup is presented in Table 1.

Predictive value of MCAT total scores and UGPA for COMLEX-USA examination performance

For COMLEX-USA Level 1 scores, MCAT scores alone provided superior predictive value to UGPA alone, yielding a median (IQR) correlation of 0.47 (0.42–0.51) compared to 0.28 (0.22–0.33) for UGPA alone. The best predictive value was provided by MCAT scores and UGPA combined, which yielded a median (IQR) correlation of 0.51 (0.45–0.55). A similar pattern was found for COMLEX-USA Level 2-CE scores. MCAT scores alone provided superior predictive value to UGPA alone, yielding a median (IQR) correlation of 0.39 (0.36–0.47) compared to 0.25 (0.23–0.32). MCAT scores and UGPA combined provided the best predictive value, yielding a median (IQR) correlation of 0.45 (0.40–0.49).

Regarding the prediction of success and failure on COMLEX-USA Level 1, MCAT scores alone provided superior classification accuracy to UGPA alone, each yielding a median (IQR) AUC of 65 % (57 %–72 %) and 57 % (55 %–64 %), respectively (Table 2). MCAT scores and UGPA combined provided the best classification accuracy, with a median (IQR) AUC of 70 % (64–75 %). For COMLEX-USA Level 2, MCAT scores alone provided superior classification accuracy than UGPA alone, each yielding a median (IQR) AUC of 71 % (59 %–78 %) and 65 % (62 %–70 %), respectively. MCAT scores and UGPA combined provided the best classification accuracy, with a median (IQR) AUC of 75 % (70–81 %).

Comparable prediction of COMLEX-USA examination scores by sociodemographic characteristics

Overall, participants belonging to both under- and well-represented sociodemographic groups did not score meaningfully higher or lower on the COMLEX-USA Level 1 or 2-CE examinations than what would be predicted for any other student with the same MCAT total score and UGPA (Table 3). Notably, for all comparisons, Cohen's *d* effect size estimates

Table 1: Descriptive statistics for the full study cohort and associated subgroups.

Variable	Full sample	Race and ethnicity		Parental education		Sex	
		Non-UIM	UIM	Upper-SES	Lower-SES	Male	Female
Total student count	5,208	4,262	583	3,741	1,141	2,713	2,487
Count (%) took COMLEX-USA level 1	5,208 (100.00 %)	4,262 (100.00 %)	583 (100.00 %)	3,741 (100.00 %)	1,141 (100.00 %)	2,713 (100.00 %)	2,487 (100.00 %)
Count (%) passed COMLEX-USA level 1	4,941 (94.87 %)	4,064 (95.35 %)	534 (91.60 %)	3,561 (95.19 %)	1,077 (94.39 %)	2,596 (95.69 %)	2,338 (94.01 %)
Count (%) took COMLEX-USA level 2-CE	4,795 (92.07 %)	3,955 (92.80 %)	510 (87.48 %)	3,456 (92.38 %)	1,050 (92.02 %)	2,518 (92.81 %)	2,270 (91.27 %)
Count (%) passed COMLEX-USA level 2-CE	4,677 (97.54 %)	3,868 (97.80 %)	489 (95.88 %)	3,375 (97.66 %)	1,021 (97.24 %)	2,454 (97.46 %)	2,218 (97.71 %)
Mean (SD) MCAT total score	502.98 (5.48)	503.30 (5.37)	499.82 (5.23)	503.29 (5.47)	501.95 (5.28)	503.70 (5.53)	502.20 (5.31)
Mean (SD) undergraduate total GPA	3.55 (0.26)	3.56 (0.25)	3.46 (0.28)	3.55 (0.26)	3.54 (0.26)	3.53 (0.26)	3.56 (0.26)
Mean (SD) COMLEX-USA level 1 score	533.06 (86.95)	536.80 (87.46)	506.97 (80.76)	535.86 (87.23)	526.83 (84.54)	545.19 (88.92)	519.97 (82.77)
Mean (SD) COMLEX-USA level 2-CE score	573.51 (95.35)	577.08 (95.26)	548.92 (93.86)	576.73 (95.68)	565.07 (92.41)	577.35 (98.19)	569.51 (91.77)

COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination; COMLEX-USA Level 2 CE, Comprehensive Osteopathic Medical Licensing Examination Level 2 Cognitive Evaluation; IQR, interquartile range; MCAT, Medical College Admissions Test; SD, standard deviation; UGPA, undergraduate grade point average.

Table 2: Correlations between COMLEX-USA examination scores and pass/fail by selection variable.

Scores			
Outcome	Predictor	Median correlation	IQR correlation
COMLEX-USA level 1	MCAT only	0.47	0.42–0.51
	UGPA only	0.28	0.22–0.33
	MCAT and UGPA	0.51	0.45–0.55
COMLEX-USA level 2 CE	MCAT only	0.39	0.36–0.47
	UGPA only	0.25	0.23–0.32
	MCAT and UGPA	0.45	0.40–0.49
Pass/fail Outcome	Predictor	Median %	IQR %
COMLEX-USA level 1	MCAT only	65 %	57–72 %
	UGPA only	57 %	55–64 %
	MCAT and UGPA	70 %	64–75 %
	UGPA		
COMLEX-USA level 2 CE	MCAT only	71 %	59–78 %
	UGPA only	65 %	62–70 %
	MCAT and UGPA	75 %	70–81 %
	UGPA		

COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination; COMLEX-USA Level 2 CE, Comprehensive Osteopathic Medical Licensing Examination Level 2 Cognitive Evaluation; IQR, interquartile range; MCAT, Medical College Admissions Test; UGPA, undergraduate grade point average.

ranged from 0.00 to 0.20 in magnitude, indicating negligible practical differences between the predicted and observed scores [18]. Although practically equivalent, predictions based on MCAT scores alone deviated less from the observed scores than those based on UGPA alone; and the predictions based on MCAT scores and UGPA together provided the most comparable prediction.

Comparable prediction of passing the COMLEX-USA examinations by sociodemographic characteristics

Overall, participants belonging to both under- and well-represented sociodemographic groups passed both COMLEX-USA Level 1 and Level 2-CE examinations at similar rates to what would be predicted for any other student with the same values on the selection variables (Table 4). The Cohen's *h* effect sizes for all comparisons ranged from 0.00 to 0.06, indicating negligible practical differences between predicted and observed pass rates [18]. As with the prediction of COMLEX-USA examination scores, predictions based on MCAT scores alone deviated less from the observed pass rates than those based on UGPA alone; in addition, the predictions based on MCAT scores and UGPA together provided the most comparable prediction.

Table 3: Comparison of observed and predicted COMLEX-USA score outcomes by sociodemographic characteristics.

Outcome	Variable	Classification	Student count	Observed mean (SD)	MCAT only			UGPA only			MCAT and UGPA		
					Predicted mean (SD)	Difference mean (SD)	Effect size	Predicted mean (SD)	Difference mean (SD)	Effect size	Predicted mean (SD)	Difference mean (SD)	Effect size
COMLEX-USA level 1	Race and ethnicity ^a	Non-UJM	4,262	536.8 (87.5)	535.3 (32.2)	1.5 (81.4)	0.02	534.6 (25.9)	2.2 (83.8)	0.03	535.8 (34.4)	1.0 (80.3)	0.01
		UJM	583	507.0 (80.8)	513.9 (32.4)	-7.0 (75.6)	-0.09	522.1 (27.2)	-15.2 (76.5)	-0.20	509.7 (35.2)	-2.8 (74.7)	-0.04
	Parental education ^b	Upper-SES	3,741	535.9 (87.2)	534.7 (32.9)	1.1 (81.0)	0.01	533.9 (26.3)	2.0 (83.3)	0.02	535.1 (35.2)	0.8 (79.8)	0.01
		Lower-SES	1,141	526.8 (84.5)	527.5 (32.4)	-0.7 (77.3)	-0.01	531.6 (26.6)	-4.8 (80.3)	-0.06	527.2 (35.2)	-0.3 (77.0)	0.00
	Sex	Male	2,713	545.2 (88.9)	537.0 (32.6)	8.2 (83.7)	0.10	533.2 (25.8)	12.0 (85.5)	0.14	536.3 (35.0)	8.9 (82.5)	0.11
COMLEX-USA level 2-CE		Female	2,487	520.0 (82.8)	528.8 (32.8)	-8.8 (75.8)	-0.12	532.9 (26.9)	-13.0 (77.9)	-0.17	529.5 (35.6)	-9.6 (74.7)	-0.13
	Race and ethnicity ^a	Non-UJM	3,955	577.1 (95.3)	575.7 (33.3)	1.4 (89.2)	0.02	574.8 (28.4)	2.3 (91.1)	0.02	576.1 (36.1)	1.0 (88.1)	0.01
		UJM	510	548.9 (93.9)	554.4 (34.8)	-5.5 (88.9)	-0.06	563.6 (27.7)	-14.7 (89.6)	-0.16	551.1 (37.2)	-2.2 (87.8)	-0.02
	Parental education ^b	Upper-SES	3,456	576.7 (95.7)	575.4 (34.0)	1.3 (89.4)	0.01	574.7 (28.5)	2.0 (91.1)	0.02	575.9 (36.9)	0.9 (88.0)	0.01
		Lower-SES	1,050	565.1 (92.4)	567.8 (33.5)	-2.7 (86.4)	-0.03	571.6 (28.4)	-6.5 (89.1)	-0.07	567.2 (36.3)	-2.2 (86.1)	-0.03
	Sex	Male	2,518	577.4 (98.2)	576.9 (33.4)	0.5 (93.3)	0.00	573.2 (27.9)	4.2 (94.8)	0.04	576.2 (36.3)	1.1 (91.9)	0.01
		Female	2,270	569.5 (91.8)	569.8 (34.3)	-0.3 (84.0)	0.00	573.9 (29.1)	-4.4 (86.3)	-0.05	570.5 (37.4)	-1.0 (83.1)	-0.01

COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination; COMLEX-USA Level 2 CE, Comprehensive Osteopathic Medical Licensing Examination Level 2 Cognitive Evaluation; IQR, interquartile range; MCAT, Medical College Admissions Test; SD, standard deviation; SES, socioeconomic status; UGPA, undergraduate grade point average; UJM, underrepresented in medicine. ^aStudents were categorized into UJM or non-UJM groups based on self-reported race/ethnicity. UJM groups consisted of Black or African American; Hispanic, Latino, or Spanish Origin; American Indian or Alaska Native; or Native Hawaiian or other Pacific Islander. Non-UJM groups consisted of White and Asian. ^bParental educational attainment was utilized as a proxy for SES. Students for whom neither parent held a bachelor's degree were categorized as lower-SES, and students for whom at least one parent held a bachelor's degree were categorized as upper-SES.

Table 4: Comparison of observed and predicted COMLEX-USA pass/fail outcomes by sociodemographic characteristics.

Outcome	Variable	Classification	Student count	Observed success rate	MCAT only			UGPA only			MCAT and UGPA		
					Predicted success rate	Difference	Effect size	Predicted success rate	Difference	Effect size	Predicted success rate	Difference	Effect size
COMLEX level 1	Race and ethnicity ^a	Non-UIM	3,917	94.9 %	94.8 %	0.2 %	0.01	94.7 %	0.2 %	0.01	94.8 %	0.1 %	0.00
		UIM	551	91.1 %	91.7 %	-0.6 %	-0.02	92.7 %	-1.6 %	-0.06	91.3 %	-0.2 %	-0.01
	Parental education ^b	Upper-SES	3,474	94.8 %	94.7 %	0.1 %	0.01	94.6 %	0.2 %	0.01	94.7 %	0.1 %	0.00
		Lower-SES	1,046	93.9 %	93.8 %	0.1 %	0.01	94.1 %	-0.3 %	-0.01	93.7 %	0.2 %	0.01
COMLEX level 2-CE	Sex	Male	2,499	95.3 %	94.8 %	0.5 %	0.02	94.6 %	0.7 %	0.03	94.8 %	0.5 %	0.02
		Female	2,309	93.5 %	94.0 %	-0.5 %	-0.02	94.3 %	-0.8 %	-0.03	94.1 %	-0.5 %	-0.02
	Race and ethnicity ^a	Non-UIM	2,957	97.1 %	97.0 %	0.1 %	0.01	96.9 %	0.2 %	0.01	97.1 %	0.1 %	0.00
		UIM	416	95.2 %	95.3 %	-0.1 %	0.00	96.1 %	-0.9 %	-0.05	95.1 %	0.1 %	0.00
	Parental education ^b	Upper-SES	2,623	97.0 %	97.0 %	0.0 %	0.00	97.0 %	0.0 %	0.00	97.0 %	0.0 %	0.00
		Lower-SES	795	96.5 %	96.5 %	0.0 %	0.00	96.6 %	-0.1 %	0.00	96.3 %	0.1 %	0.01
	Sex	Male	1,921	96.8 %	97.1 %	-0.4 %	-0.02	96.9 %	-0.1 %	-0.01	97.1 %	-0.3 %	-0.02
		Female	1,714	97.0 %	96.5 %	0.5 %	0.03	96.8 %	0.2 %	0.01	96.6 %	0.4 %	0.02

COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination; COMLEX-USA Level 2 CE, Comprehensive Osteopathic Medical Licensing Examination Level 2 Cognitive Evaluation; IQR, interquartile range; MCAT, Medical College Admissions Test; SES, socioeconomic status; UGPA, undergraduate grade point average; UIM, underrepresented in medicine. ^aStudents were categorized into UIM or non-UIM groups based on self-reported race/ethnicity. UIM groups consisted of Black or African American; Hispanic, Latino, or Spanish Origin; American Indian or Alaska Native; or Native Hawaiian or other Pacific Islander. Non-UIM groups consisted of White and Asian. ^bParental educational attainment was utilized as a proxy for SES. Students for whom neither parent held a bachelor's degree were categorized as lower-SES, and students for whom at least one parent held a bachelor's degree were categorized as upper-SES.

Discussion

Results of this study largely mirror those of Hanson et al. [5], who examined the predictive validity of MCAT scores and UGPA in MD-granting medical schools. Specifically, MCAT scores alone had greater predictive validity than UGPA alone, but when MCAT scores were utilized in conjunction with UGPA, they offer greater predictive validity than the use of either MCAT scores or UGPA alone. In the context of DO-granting medical schools, MCAT scores are a good predictor of both medical students' scores and pass/fail outcomes on COMLEX-USA licensure examinations, as evidenced by median correlations of a moderate to large magnitude.

Although members of every racial and ethnic group attain scores that span the entire MCAT score scale, it is widely acknowledged that the "achievement gap," structural racism, and unequal opportunity in the United States has contributed to group level differences in MCAT scores [19]. When predicting both licensure examination scores and pass/fail outcomes based on race and ethnicity, SES, and sex, all three models (MCAT scores alone, UGPA alone, and MCAT scores and UGPA combined) provided comparable predictive accuracy. More specifically, individuals from different sociodemographic backgrounds who enter medical school with similar MCAT scores and UGPA perform similarly on licensure examination outcome measures. These findings are consistent with the findings from Hanson et al. [5] and Busche et al. [20], who studied the current version of the MCAT examination (introduced in 2015), and Davis et al. [21], who studied the previous version of the MCAT examination (in effect from 1991 to January 2015).

In comparison to the study by Hanson et al. [5] involving MD-granting medical schools, the present study of DO-granting medical schools found slightly lower predictive accuracy for pass/fail outcomes. We believe that this is likely due to very high success rates (94.9 % for COMLEX-USA Level 1 and 97.5 % for COMLEX-USA Level 2-CE) in the dataset that we examined. Additionally, four programs were excluded from the Level 1 analysis and eight programs were excluded from the Level 2-CE analysis because students in each of those programs had 100 % pass rates. Thus, we suspect the lack of variability in the pass/fail variable likely resulted in lower predictive accuracy.

Collectively, these findings provide evidence for the value-added benefit of taking MCAT scores into consideration, particularly when they are utilized in conjunction with other academic variables such as UGPA. Further, the ability to anticipate if students are likely to experience academic difficulties in medical school or face potential

challenges when attaining medical licensure offers a tremendous benefit for medical school leaders. These insights can help inform the forecasting, planning, and programming necessary to ensure adequate personnel, and resources are available to provide the necessary supports to maximize students' success.

With respect to validity, the ability of MCAT scores and UGPA to predict COMLEX-USA licensure examination success with good accuracy lends additional predictive validity evidence for these measures [22–25]. Additional validity evidence is discernible given the fact that the findings from this comprehensive study of DO programs largely mirror those in the comprehensive study of MD programs by Hanson et al. [5]. Finally, the ability of MCAT scores, in particular, to yield predictions that hold equally well across racial and ethnic, sex, and SES groups lends additional validity evidence and further demonstrates fairness in prediction.

This study possesses several limitations. First, because the sample size for some racial and ethnic groups was sparse, we were unable to account for intersectionality across sociodemographic groups (e.g., females who are underrepresented in medicine, males whose parents do not hold a bachelor's degree). Further research with larger subpopulations is warranted. A second limitation is that the results are limited to the 2017 applicant cycle. We do not know how data from other cohorts might vary, nor the effects that a different minimum passing standard (MPS) might have on MCAT scores and UGPA predictive abilities. Further, because the COMLEX Level 1 moved to pass/fail score reporting and numeric scores were eliminated as of 2022, we cannot speak to how this change might affect student preparation and correlations involving MCAT scores and UGPA for future cohorts [26]. Thus, this study should be replicated in the future utilizing data from additional cohorts should the MPS be revised. Finally, this study examined only the MCAT total scores and overall UGPA. It is unknown how well MCAT section scores and UGPA limited to various domains or classification schemas (e.g., natural sciences) might predict COMLEX-USA examination performance.

Conclusions

We investigated the predictive value of MCAT scores and UGPA for DO medical student performance on COMLEX-USA Level 1 and Level 2-CE examinations, and whether these selection variables provided comparable predictive value based on sex, race and ethnicity, and SES. Results indicated the best predictive accuracy resulted when MCAT scores and

UGPA were combined, followed by MCAT scores alone, and then UGPA alone. Additionally, individuals from different sociodemographic backgrounds who enter medical school with similar MCAT scores and UGPA perform similarly on licensure examination outcome measures. Collectively, these findings provide evidence for the value-added benefit of taking MCAT scores and UGPA into consideration, particularly when these measures are utilized in conjunction.

Acknowledgments: The authors would like to acknowledge the contributions of our colleague Tsung-Hsun Tsai, PhD (NBOME), who passed away after the conclusion of this research. He was passionate about collaborative research to help further validity studies on the academic progression of our future osteopathic physician workforce. Also, the authors wish to thank the following AAMC personnel for reviewing earlier drafts of the manuscript: Gabrielle Campbell and Javarro Russell. In addition, we thank the following AAMC personnel for their administrative contributions to this article: Lindsey Topp, Robert Santos, Andrea Carpentieri, Brianna Gunter, Ruhyyih Degeberg, Marie Tummarello, and Marissa DuVon.

Research ethics: This study was deemed exempt by the American Institutes for Research in Behavioral Sciences Institutional Review Board (Protocol #EX00591).

Informed consent: All participants in this research provided informed consent for their educational records and socio-demographic data to be used for research purposes.

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 declared.

Research funding: None declared.

Data availability: The raw data may be obtained on request from the corresponding author per AAMC, AACOM, and NBOME approval

References

1. Conrad SS, Addams AN, Young GH. Holistic review in medical school admissions and selection: a strategic, mission-driven response to shifting societal needs. *Acad Med* 2016;91:1472–4.
2. Terregino CA, Saguil A, Price-Johnson T, Anachebe NF, Goodell K. The diversity and success of medical school applicants with scores in the middle third of the MCAT score scale. *Acad Med* 2020;95:344–50.
3. Association of American Medical Colleges. Admissions officers survey; 2021. Unpublished data.
4. Association of American Medical Colleges. Using MCAT data in 2023 medical student selection. Washington, DC: Association of American Medical Colleges; 2022.
5. Hanson JT, Busche K, Elks ML, Jackson-Williams LE, Liotta RA, Miller C, et al. The validity of MCAT scores in predicting students' performance and progress in medical school: results from a multisite study. *Acad Med* 2022;97:1374–84.
6. American Osteopathic Association. Osteopathic medical profession report; 2022. <https://osteopathic.org/wp-content/uploads/2022-AOA-OMP-Report.pdf> [Accessed 22 Mar 2023].
7. American Association of Colleges of Osteopathic Medicine. Quick facts 2022. <https://www.aacom.org/become-a-doctor/about-osteopathic-medicine/quick-facts> [Accessed 22 Mar 2023].
8. Cohen J. The separate osteopathic medical education pathway: isn't it time we got our acts together? Counterpoint. *Acad Med* 2009;84:696.
9. Knight V. American medical students less likely to choose to become primary care doctors. *KFF Health News*. <https://kffhealthnews.org/news/american-medical-students-less-likely-to-choose-to-become-primary-care-doctors/>. [Accessed 27 Sep 2023].
10. American Association of Colleges of Osteopathic Medicine. 2021-2022 Academic year entering student survey summary report. 2021-2022-academic-year-entering-student-survey-summary-report.pdf (aacom.org) [Accessed 27 Sep 2023].
11. Association of American Medical Colleges. Matriculating student questionnaire (MSQ). <https://www.aamc.org/media/57926/download> [Accessed 27 Sep 2023].
12. Agahi F, Speicher MR, Cisek G. Association between undergraduate performance predictors and academic and clinical performance of osteopathic medical students. *J Am Osteopath Assoc* 2018;118:106–14.
13. Steed KS, Kadavakollu S. How to prepare for the medical college admissions test (MCAT): six important tips for pre-medical students from rural areas. *Med Sci Educ* 2019;29:1147–53.
14. Wolla SA, Sullivan J. Education, income, and wealth. *Page One Economics*; 2017. <https://research.stlouisfed.org/publications/page1-econ/2017/01/03/education-income-and-wealth> [Accessed 22 Mar 2023].
15. National Board of Osteopathic Medical Examiners. COMLEX-USA level 1; 2022a. <https://www.nbome.org/assessments/comlex-usa/comlex-usa-level-1> [Accessed 22 Mar 2023].
16. National Board of Osteopathic Medical Examiners. COMLEX-USA level 2-CE; 2022b. <https://www.nbome.org/assessments/comlex-usa/comlex-usa-level-2-ce> [Accessed 22 Mar 2023].
17. Sackett PR, Yang H. Correction for range restriction: an expanded typology. *J Appl Psychol* 2000;85:112–18.
18. Cohen D. Statistical power analysis for the behavioral sciences, 2nd ed Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
19. Lucey CR, Saguil A. The consequences of structural racism on MCAT scores and medical school admissions: the past is prologue. *Acad Med* 2020;95:351–6.
20. Busche K, Elks ML, Hanson JT, Jackson-Williams L, Manuel RS, Parsons WL, et al. The validity of scores from the new MCAT exam in predicting student performance: results from a multisite study. *Acad Med* 2020;95:387–95.
21. Davis D, Dorsey JK, Franks RD, Sackett PR, Searcy CA, Zhao X. Do racial and ethnic group differences in performance on the MCAT exam reflect test bias? *Acad Med* 2013;88:593–602.

22. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Psychological Association; 2014.
23. Kane M. Validation strategies: delineating and validating proposed interpretations and uses of test scores. In: Lane S, Raymond MR, Haladyna TM, editors. Handbook of test development. New York, NY: Routledge; 2015:80–96 pp.
24. Kane M. The argument-based approach to validation. *Sch Psychol Rev* 2013;42:448–57.
25. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol* 1995;50:741–9.
26. National Board of Osteopathic Medical Examiners. COMLEX-USA Level 1 to eliminate numeric scores; 2022c. <https://www.nbome.org/news/comlex-usa-level-1-to-eliminate-numeric-scores> [Accessed 22 Mar 2023].