

RESEARCH ARTICLE

Open Access



Predictors of medical school clerkship performance: a multispecialty longitudinal analysis of standardized examination scores and clinical assessments

Petra M. Casey^{1*}, Brian A. Palmer², Geoffrey B. Thompson³, Torrey A. Laack⁴, Matthew R. Thomas⁵, Martha F. Hartz⁶, Jani R. Jensen⁷, Benjamin J. Sandefur⁴, Julie E. Hammack⁸, Jerry W. Swanson⁸, Robert D. Sheeler⁹ and Joseph P. Grande¹⁰

Abstract

Background: Evidence suggests that poor performance on standardized tests before and early in medical school is associated with poor performance on standardized tests later in medical school and beyond. This study aimed to explore relationships between standardized examination scores (before and during medical school) with test and clinical performance across all core clinical clerkships.

Methods: We evaluated characteristics of 435 students at Mayo Medical School (MMS) who matriculated 2000–2009 and for whom undergraduate grade point average, medical college aptitude test (MCAT), medical school standardized tests (United States Medical Licensing Examination [USMLE] 1 and 2; National Board of Medical Examiners [NBME] subject examination), and faculty assessments were available. We assessed the correlation between scores and assessments and determined USMLE 1 cutoffs predictive of poor performance (≤ 10 th percentile) on the NBME examinations. We also compared the mean faculty assessment scores of MMS students vs visiting students, and for the NBME, we determined the percentage of MMS students who scored at or below the tenth percentile of first-time national examinees.

Results: MCAT scores correlated robustly with USMLE 1 and 2, and USMLE 1 and 2 independently predicted NBME scores in all clerkships. USMLE 1 cutoffs corresponding to poor NBME performance ranged from 220 to 223. USMLE 1 scores were similar among MMS and visiting students. For most academic years and clerkships, NBME scores were similar for MMS students vs all first-time examinees.

Conclusions: MCAT, USMLE 1 and 2, and subsequent clinical performance parameters were correlated with NBME scores across all core clerkships. Even more interestingly, faculty assessments correlated with NBME scores, affirming patient care as examination preparation. USMLE 1 scores identified students at risk of poor performance on NBME subject examinations, facilitating and supporting implementation of remediation before the clinical years. MMS students were representative of medical students across the nation.

Keywords: Assessment, Clerkship, Clinical rotation, Standardized scores

* Correspondence: casey.petra@mayo.edu

¹Division of Gynecology, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA

Full list of author information is available at the end of the article



Background

Standardized examinations are ubiquitous throughout medical education and are designed to objectively measure performance. Unsurprisingly, performance on a given standardized test tends to predict performance on subsequent tests. The Medical College Admission Test (MCAT) is widely used to select students with a higher likelihood of success in medical school, and it is considered more valid than letters of recommendation and grade point averages (GPAs) [1, 2]; MCAT performance also predicts performance on the United States Medical Licensing Examination Step 1 (USMLE 1) [3]. USMLE 1, administered after the first 2 years of medical school, assesses understanding and the ability to apply important concepts in basic science to the practice of medicine. It covers traditional disciplines such as anatomy, behavioral sciences, biochemistry, biostatistics and epidemiology, microbiology, pathology, pharmacology, and physiology, as well as interdisciplinary areas such as genetics, aging, immunology, nutrition, and molecular and cell biology [4]. The National Board of Medical Examiners (NBME) comprehensive basic science subject examination also can help identify students at risk for failing USMLE 1 [5, 6]. MCAT and USMLE 1 scores are also predictive of NBME obstetrics and gynecology (OB/GYN) subject examination scores [7].

Further in medical education, several studies link poor test results with lower performance during the clinical years, and moderate-to-strong correlations were identified between NBME subject examination scores and USMLE performance [8, 9]. For example, students failing either the USMLE 1 or NBME OB/GYN subject examination were more likely to fail USMLE 2 [10]. USMLE 2 (Clinical Knowledge and Clinical Skills), taken during the fourth year of medical school, assesses application of medical knowledge, skills, and understanding of clinical science that are essential for providing patient care under supervision. It includes questions regarding the immune system, blood and lymphoreticular systems, behavioral health, nervous system and special senses, skin and subcutaneous tissue, musculoskeletal system, cardiovascular system, respiratory system, gastrointestinal system, renal and urinary systems, pregnancy, childbirth and the puerperium, female reproductive system and breast, male reproductive system, endocrine system, multisystem processes and disorders, biostatistics, and epidemiology/population health, as well as interpretation of the medical literature [11, 12]. Poor performance on the third-year NBME surgery subject examination was strongly correlated with the second-year pathology NBME subject examination scores [13, 14]. Lastly, examinees with low USMLE scores also had a higher risk of failing part 1 of the American Board of Orthopaedic Surgery Certifying Examination [15]. Poor standardized

test scores before the clinical years may identify students who would benefit from proactive support to avoid poor performance and thereby increase their ability to successfully complete clerkships. In addition, student assessments do not need to be limited to standardized tests. For example, a faculty-developed pretest given on day 1 of an internal medicine clerkship reliably identified students at risk for poor performance [16].

Curriculum design also affects standardized test scores. Decreasing the duration of an OB/GYN medical student clerkship resulted in lower subject examination scores, especially for students enrolled in the first half of the academic year [17]. Similarly, shortening the psychiatry clerkship length at Mayo Medical School (MMS) negatively affected NBME scores [18]. Certain clerkship characteristics are associated with better student examination performance, the most important being caring for more patients per day [19]. Self-assessment practice examinations can accurately predict actual performance on standardized tests, although some variation in predictive accuracy occurs across different test administration conditions [5, 20].

Medical schools strive to produce inquisitive physicians with a basic foundation of knowledge, technical skills, reasoning ability, and empathy; this provides a sufficient framework for graduate medical education and practice and helps develop skills for self-directed lifelong learning. Factors that predict a student's success in all these areas are difficult to identify [21]. Critics of standardized examinations point out that the tests primarily measure medical knowledge and may be poor predictors of clinical performance and professionalism [22, 23]. Not surprisingly, a survey of students reported that those who scored well on numerically scored standardized tests were more likely to favor their use, whereas those who scored poorly favored pass/fail grading [24, 25].

Although several studies have confirmed correlations between various standardized tests, less is known regarding the relationship between standardized test results and clinical performance. In fact, subjective assessment of surgical knowledge by faculty and residents correlated poorly with objective performance measures, bringing into question whether subjective appraisal of surgical knowledge should be included in the assessment process [26]. Nevertheless, others have reported that grades from an OB/GYN rotation correlated with USMLE scores [27] and NBME subject examination scores [7].

The merits of using either objective standardized testing or subjective clinical assessments can be debated, but both remain common components of medical student assessment. We asked the following questions: 1) How do test scores relate to medical student clinical performance? and 2) How do faculty assessments relate to overall student performance on core clerkships? Our

study focused on standardized testing relationships throughout medical school as they relate to performance on all clinical clerkships. If longitudinal and cross-sectional relationships between standardized testing (both before and during medical school) and clinical performance can be confirmed, poor performers may be identified and supported before their clinical years.

Methods

We included students who matriculated at MMS from 2000 through 2009. We selected 2009 as the last year of matriculation to allow access to complete information across all 4 years of medical school. The study was deemed exempt by the Mayo Clinic Institutional Review Board (protocol 13-003310), and data were de-identified before analysis.

The following information was obtained for each student: gender, age, underrepresented minority status, degree program, undergraduate GPA, MCAT subscores and total score, USMLE 1 and 2 scores, NBME subject examination scores (range, 0–100), and faculty assessment scores (FAS) (range, 1–5, with 5 representing the highest possible score). The subcategories of faculty assessment encompassed domains of clinical knowledge, including history taking, examination skills, and decision making. Further, professional behaviors such as communication, teamwork, and patient-centered approach were assessed. In addition, we obtained FAS and USMLE 1 scores of visiting medical students completing clerkships (in internal medicine, neurology, OB/GYN, pediatrics, surgery, family medicine, psychiatry, and emergency medicine) from August 2006 through August 2012. All data were electronically retrieved from databases maintained by the MMS.

For MMS students, correlations between measures were evaluated graphically and quantified using the Pearson correlation coefficient. Multivariable linear regression models were fit using backward variable elimination to identify sets of independent predictors of USMLE 1, USMLE 2, and each of the NBME subject examinations, respectively. Predictors with a P value $< .05$ were retained in each final model.

Lastly, MMS students were compared with other medical students in 2 ways. First, FAS and USMLE 1 scores were each compared between MMS matriculating students and visiting medical students who concurrently completed the same clerkships. Scores were compared using the 2-sample t test. Second, using national data on NBME first-time examinees, we stratified students by clerkship and determined the overall percentage of MMS matriculating students who scored at or below the 10th national percentile for that academic year on the NBME subject examination from 2003–2012 [28].

Analyses were performed using SAS version 9.2 (SAS Institute Inc). All calculated P values were 2-sided, and P values less than .05 were considered statistically significant.

Results

We identified 435 students who matriculated in 2000 through 2009 and completed USMLE 1. Of the 435 students, 219 (50.3 %) were men; 65 (14.9 %) were underrepresented minorities. The majority ($n = 356$ [81.8 %]) were in the standard MD program, the rest were in joint-degree programs for MD-PhD ($n = 53$ [12.2 %]), MD-DDS ($n = 19$ [4.4 %]), or MD-PhD-MS ($n = 7$ [1.6 %]) degrees. The results are reported for all programs combined.

Table 1 summarizes the observed correlation of student characteristics, undergraduate GPA, and MCAT scores with USMLE 1 scores. The strongest predictor of USMLE 1 was the MCAT total score ($r = 0.50$; $R^2 = 25$ %). In a multivariable analysis using backward variable elimination, the following variables were identified as independent predictors of USMLE 1 (P values $< .05$) with an overall R^2 of 35 %: GPA, MCAT–biological science, MCAT–physical science, and gender. As previously described [29], USMLE 1 scores were slightly higher for men than women. Table 1 also summarizes the observed correlations with USMLE 2 clinical knowledge scores (available for 324 students). Overall, the strongest predictor of the USMLE 2 score was the USMLE 1 score ($r = 0.77$; $R^2 = 59$ %). In a multivariable analysis using backward variable elimination, the following variables were identified as independent predictors of USMLE 2, with an overall $R^2 = 61.6$ %: USMLE 1, gender, and MCAT–total.

Table 2 summarizes correlations of the test measures and student characteristics with NBME subject examinations, based on 407 students with at least 1 NBME subject examination score. The variables that independently correlated with each NBME score (based on a multivariable analysis) are shown. Both USMLE 1 and 2 independently correlated with NBME scores across clerkships.

Of the 435 students in the cohort, FAS were available for 222 students (available for 27/211 students [12.8 %] who matriculated in 2000–2004 and 195/224 students [87.1 %] who matriculated in 2005–2009). Table 3 summarizes correlations between FAS and the MCAT, USMLE, and NBME scores for each clerkship. FAS correlated well with USMLE 1 and 2 across almost all clerkships. NBME and FAS also correlated well across all clerkships. FAS was not correlated with MCAT scores.

Given the high correlation between USMLE 1 and NBME subject examination scores, we sought to determine a

Table 1 Correlation of test measures and student characteristics with USMLE 1 and 2

Test measure	Analysis of USMLE 1		Analysis of USMLE 2	
	r ^a	P value ^b	r ^a	P value ^b
GPA	0.39	<.001	0.35	<.001
MCAT				
Total	0.50	<.001	0.38	<.001
Biological science	0.49	<.001	0.37	<.001
Physical science	0.47	<.001	0.36	<.001
Verbal	0.21	<.001	0.19	<.001
Writing	0.16	<.001	0.11	.06
USMLE 1	0.77	<.001
Student Characteristic	USMLE 1, mean (SD)	P Value ^c	USMLE 2, mean (SD)	P Value ^c
Gender		<.001		.55
Male	236.3 (20.9)		236.1 (23.6)	
Female	228.6 (20.5)		237.6 (22.4)	
Underrepresented minority		<.001		<.001
No	235.0 (19.7)		239.1 (21.2)	
Yes	218.3 (23.3)		224.2 (28.5)	

Abbreviations: GPA Grade point average, MCAT Medical college admission test, USMLE 1 and 2 United States medical licensing examination steps 1 and 2

^ar denotes the Pearson correlation coefficient

^bP values indicate whether the correlation is significantly different from 0

^cP values were based on a 2-sample t test

USMLE 1 cutoff that was predictive of poor performance on the NBME examinations, defined as a score at or below the 10th percentile. For each clerkship, we selected the USMLE 1 threshold at which both sensitivity and specificity were simultaneously maximized, to minimize the false-negative and false-positive rates. As summarized in Table 4, depending on the NBME clerkship, a USMLE 1 score ranging from 220 to 223 was predictive of poor performance on NBME subject examinations, with 77 % to 83 % sensitivity and specificity.

Clerkships were completed by 755 visiting medical students from August 2006 through August 2012. The mean USMLE 1 was identical for the 435 MMS students and 207 visiting students for whom data were available ($P = .99$). The number of faculty assessments performed varied for each clerkship; however, mean FAS differed between the 2 groups only for OB/GYN (mean difference, 0.17; $P = .05$), surgery (mean difference, 0.17; $P < .001$), and family medicine (mean difference, 0.56; $P = .04$).

To benchmark MMS student performance on NBME subject examinations against the national pool of first-time examinees, we determined the percentage of MMS students from each clerkship who scored at or below the 10th national percentile for that academic year on the NBME subject examination from 2003 through 2012. The overall percentages were 5.5 % (95 % CI, 3.0-8.0 %)

for OB/GYN, 7.1 % (95 % CI, 4.2-10.0 %) for both surgery and neurology, 7.7 % (95 % CI, 4.8-10.7 %) for internal medicine, 9.4 % (95 % CI, 6.0-12.9 %) for pediatrics, and 11.7 % (95 % CI, 8.1-15.3 %) for psychiatry.

Discussion

Medical student performance during the clinical years is predictable across clerkships

We examined the feasibility of predicting performance across core clerkships by examining relationships among examination scores and clinical performance parameters before and during medical school. Clinical performance encompassed FAS and various other aspects of learning that were specific to each clerkship, including demonstrated skills, topic presentations, case summaries, and reflections. We included MCAT scores, undergraduate GPA, USMLE 1 and 2 scores, NBME subject examination scores, and FAS. We confirmed several previously reported relationships, including strong correlation of all subsections of the MCAT with USMLE 1. Further, all MCAT subsections except writing were highly correlated with USMLE 2 (Tables 1 and 2). The latter finding is less relevant because the future MCAT format will not include any writing [30]. Of primary interest to our anticipated outcomes, NBME subject examination scores across all clerkships, not reported previously, were highly correlated with USMLE 1 and 2 scores, whereas

Table 2 Correlations of GPA, MCAT, and USMLE scores with NBME subject examination scores for each clerkship

Test measure	Medicine		Neurology		OB/GYN		Pediatrics		Psychiatry		Surgery	
	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b
GPA	0.33	<.001	0.31	<.001	0.29	<.001	0.38	<.001	0.32	<.001	0.35	<.001
MCAT												
Total	0.38	<.001	0.40	<.001	0.33	<.001	0.39	<.001	0.40	<.001	0.45	<.001
Biological science	0.36	<.001	0.38	<.001	0.33	<.001	0.34	<.001	0.35	<.001	0.43	<.001
Physical science	0.31	<.001	0.33	<.001	0.29	<.001	0.32	<.001	0.31	<.001	0.39	<.001
Verbal	0.23	<.001	0.26	<.001	0.17	<.001	0.27	<.001	0.30	<.001	0.26	<.001
Writing	0.10	.05	0.13	.01	0.09	.07	0.14	.006	0.21	<.001	0.12	.23
USMLE 1	0.69	<.001	0.69	<.001	0.65	<.001	0.68	<.001	0.62	<.001	0.67	<.001
USMLE 2	0.77	<.001	0.75	<.001	0.74	<.001	0.77	<.001	0.69	<.001	0.70	<.001
Student Characteristic	Mean (SD)	P Value ^c	Mean (SD)	P Value ^c	Mean (SD)	P Value ^c	Mean (SD)	P Value ^c	Mean (SD)	P Value ^c	Mean (SD)	P Value ^c
Gender		.77		.54		.08		.26		.02		.26
Male	78.4 (8.3)		75.2 (8.2)		74.5 (8.8)		76.4 (9.1)		77.2 (9.0)		75.6 (9.5)	
Female	78.2 (8.0)		75.7 (8.2)		76.0 (8.3)		77.4 (9.3)		79.2 (8.9)		74.6 (8.9)	
Underrepresented minority		<.001		<.001		<.001		<.001		<.001		<.001
No	79.1 (7.8)		76.2 (7.7)		76.0 (8.4)		77.9 (8.7)		79.0 (8.7)		75.9 (9.1)	
Yes	73.9 (8.6)		70.9 (9.1)		70.9 (8.2)		71.6 (10.1)		73.7 (9.7)		70.3 (8.4)	
Independent predictors from multivariable models	USMLE 2		USMLE 2		USMLE 2		USMLE 2		USMLE 2		USMLE 2	
	USMLE 1		USMLE 1		USMLE 1		USMLE 1		USMLE 1		USMLE 1	
	MCAT, Physical		MCAT, Physical		Gender		Gender		Gender		MCAT, biological science	
	MCAT, Total		MCAT, Total				MCAT, Verbal		MCAT, Physical			
									MCAT, Total			
	(R ² = 60.6 %)		(R ² = 59.0 %)		(R ² = 58.4 %)		(R ² = 63.0 %)		(R ² = 53.4 %)		(R ² = 52.6 %)	

Abbreviations: GPA Grade point average, MCAT Medical college admission test, NBME National board of medical examiners, OB/GYN Obstetrics and gynecology, USMLE 1 and 2 United States medical licensing examination steps 1 and 2

^ar denotes the Pearson correlation coefficient

^bP values indicate whether the correlation is significantly different from 0

^cP values were based on a 2-sample t test

Table 3 Correlation of Clerkship Faculty Assessment Score with MCAT, USMLE, and NBME Scores

Clerkship	MCAT, total		USMLE 1		USMLE 2 clinical knowledge		NBME	
	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b	r ^a	P value ^b
Medicine	0.08	.26	0.26	<.001	0.14	.06	0.30	<.001
Neurology	-0.05	.49	0.20	.007	0.20	.009	0.25	<.001
OB/GYN	-0.02	.75	0.15	.04	0.12	.11	0.22	.002
Pediatrics	-0.08	.30	0.13	.07	0.20	.009	0.18	.01
Psychiatry	0.03	.68	0.19	.01	0.22	.003	0.30	<.001
Surgery	-0.02	.77	0.18	.02	0.11	.15	0.16	.03
Family medicine	0.20	.009	0.37	<.001	0.26	.001	...	^c
Emergency medicine	0.06	.37	0.20	.005	0.16	.03	...	^c
Internal medicine subspecialties	0.12	.12	0.22	.003	0.15	.04	...	^c

Abbreviations: MCAT Medical college admission test, NBME National board of medical examiners, OB/GYN Obstetrics and gynecology, USMLE 1 and 2, United States medical licensing examination steps 1 and 2

^a r denotes the Pearson correlation coefficient

^b P values indicate whether the correlation is significantly different from 0

^c NBME subject examination scores were not available for some clerkships

undergraduate GPA and MCAT had less robust correlations.

The next, less expected finding is educationally relevant and reflects the benefits of thoughtful, high-quality faculty feedback. When considering students from all core clerkships for whom NBME subject examination scores were available, FAS robustly correlated with NBME examination scores, more so than they did with MCAT or USMLE 1 and 2 scores, although for most clerkships, FAS and USMLE scores also were correlated (Table 3). At our institution, students frequently describe tension between patient care responsibilities during clerkships and the need to master material tested on the NBME subject examinations (“Do we study for the test or take care of patients?”). To the extent that faculty assessments are based on observations of students taking care of patients, we are now able to support our answer to this

question with evidence—taking care of patients is excellent preparation for the test.

Although the MCAT total score had high predictive value for USMLE 1, USMLE 2, and NBME examination scores, it was not well correlated with core clerkship FAS. In fact, our data suggest the absence of a relationship between MCAT score and FAS (Table 3), even though MCAT and FAS correlated well with performance on the corresponding NBME examination. This finding may warrant further inquiry into which aspects of the MCAT predict success in clinical medicine.

We investigated the predictive value of standardized tests in identifying students who may need additional help or support during the clinical years (Table 4). We chose the 10th percentile as a cutoff because it is identical to the minimum passing NBME subject examination score in core clerkships at MMS. We outline, with maximal sensitivity and specificity, the USMLE 1 score below which we expect to see an NBME subject examination score in the lowest 10th percentile on each clerkship. The cutoff scores varied by clerkship, but the USMLE 1 may facilitate identification of at-risk students before clerkships begin, which potentially allows proactive enhanced support before and during the clerkship, rather than retroactive remediation after failure of the clerkship. The nature of such remediation is beyond the scope of this work, but determining parameters most relevant for identifying students potentially at risk is an educationally sound first step.

Given the small class size at MMS, we evaluated the generalizability of these data beyond comparison of standardized scores. Because FAS in this study are specific to our institution, comparison of MMS students with visiting students showed that USMLE 1 scores were identical, confirming a similar level of knowledge. FAS are

Table 4 USMLE 1 cutoff scores predictive of NBME subject examination scores below the 10th percentile for each core clerkship

Clerkship	NBME subject examination 10th percentile ^a	USMLE 1 cutoff score ^b	Sensitivity, %	Specificity, %
Internal medicine	68	≤223	77.1	77.9
Neurology	65	≤222	77.8	79.7
OB/GYN	65	≤222	81.3	80.4
Pediatrics	65	≤222	80.0	79.9
Psychiatry	67	≤223	78.7	77.9
Surgery	63	≤220	82.6	83.0

Abbreviations: NBME National board of medical examiners, OB/GYN Obstetrics and gynecology, USMLE 1 United States medical licensing examination step 1

^aTen percent of 435 students scored at or below this threshold

^bCutoff scores were selected to maximize sensitivity and specificity

inherently subjective, and they varied among MMS students in some clerkships (family medicine, OB/GYN, pediatrics, and surgery) but not in others (internal medicine, neurology, and emergency medicine). Given this variability among clerkships and the relatively small sample sizes for individual clerkships (but not for the overall comparison), we do not believe that these relatively small differences are educationally meaningful, particularly given the absence of an overall trend. We believe that the MMS cohort adequately represents US medical students. Therefore, whereas the calculation of a USMLE 1 cutoff predictive of lower clerkship performance is institution-specific, the method we used could be easily replicated at other medical schools using available data. These predictors also may provide meaningful guidance to the admissions process and curriculum design.

Strengths

The key strengths of this study include the longitudinal and cross-sectional correlations between standardized scores and clinical performance parameters across all core clerkships. Unlike earlier studies that focused on single clinical clerkships [7, 15, 18, 19, 27, 31, 32], we analyzed scores and clinical assessments across all required clerkships. These findings may allow us to identify students with need of assistance in improving performance in any of the core clerkships. We also identified clinical acumen in terms of FAS as a vital contributor to overall clinical performance. This finding is both novel and educationally meaningful. Lastly, we generalized MMS data through comparison with clinical assessments of visiting medical students from across the United States and comparison with national NBME standards for the subject examinations pertaining to each clerkship.

Limitations

Limitations include a retrospective design at a single institution over a long study period. Some variation in scores may have been introduced by a curriculum change in 2006, although we included similarly sized groups of students enrolled before and after the change. Further, not all scores were available for all students in all clerkships during the study interval.

Conclusions

In summary, we described academic assessment of MMS students over a 10-year interval. We confirmed a number of well-established findings regarding the predictive value of MCAT, USMLE 1 and 2, and NBME examinations in individual clerkships while expanding to report trends across all core clerkships. We also showed that both standardized testing and FAS were highly predictive

of NBME subject examination scores, whereas MCAT scores did not correlate well with FAS in the clinical years. We confirmed that USMLE 1 sensitively and specifically identified students at risk for poor performance in clerkships, before clerkships began. Our findings can facilitate identification of students at risk for poor clinical performance who would potentially benefit from proactive remediation. Because the MMS cohort was comparable in terms of the primary outcomes of interest (NBME and FAS) to other US medical students, we believe that our findings are generalizable and that the methodology can be replicated to determine institution-specific performance cutoffs.

Consent to publish

No potentially identifying details are included in the text. Consent to publish is not applicable for this manuscript.

Availability of data and materials

The dataset supporting the conclusions of this article is included within the article.

Abbreviations

FAS: faculty assessment scores; GPA: grade point average; MCAT: Medical college admission test; MMS: Mayo medical school; NBME: National board of medical examiners; OB/GYN: obstetrics and gynecology; USMLE 1 and 2: United States medical licensing examination steps 1 and 2.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors participated in the design and implementation of the study, as well as provided substantive edits to the manuscript. Specific contributions were as follows: PMC and BAP provided oversight for the data acquisition, analysis, and interpretation and drafted the initial manuscript. GBT, TAL, MRT, MFH, JRJ, BJS, JEH, JWS, and RDS provided substantive edits and comments to the evolving data and manuscript drafts. JPG contributed the initial idea toward this project and provided substantive edits and comments to the evolving data and manuscript drafts. He served as mentor to the committee's effort at educational scholarship and to the committee's chair (P.M.C.). All authors read and approved the final manuscript.

Authors' information

Most authors are clerkship directors of their respective areas. P.M.C. is the Director of Clinical Education, Mayo Medical School. J.P.G. served as Senior Associate Dean of Academic Affairs at Mayo Medical School. J.W.S. is Director of Program Evaluation at Mayo Medical School.

Acknowledgment

The authors gratefully acknowledge Ms Bonnie J. Denzer for gathering and formatting an enormous quantity of data and Ms Amy L. Weaver for expert statistical analysis, table preparation, and manuscript review. The authors gratefully acknowledge the Mayo Clinic College of Medicine for providing funding for this study. No other funding was received for this work. The Mayo Clinic College of Medicine had no role in study design; in the collection, analysis, and interpretation of data; writing of the manuscript; and in the decision to submit the manuscript for publication. Oral presentation at the Association for Medical Education in Europe (AMEE) meeting, Prague, Czech Republic, August 24-28, 2013.

Author details

¹Division of Gynecology, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ²Department of Psychiatry and Psychology, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ³Division of Subspecialty General Surgery, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁴Department of Emergency Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁵Department of Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁶Department of Pediatrics and Adolescent Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁷Division of Reproductive Endocrinology and Infertility, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁸Department of Neurology, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ⁹Department of Family Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ¹⁰Department of Anatomic Pathology and Laboratory Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA.

Received: 9 December 2015 Accepted: 26 April 2016

Published online: 27 April 2016

References

- Hoffman KI. The USMLE, the NBME subject examinations, and assessment of individual academic achievement. *Acad Med.* 1993;68(10):740–7.
- Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med.* 2005;80(10):910–7.
- Basco Jr WT, Way DP, Gilbert GE, Hudson A. Undergraduate institutional MCAT scores as predictors of USMLE step 1 performance. *Acad Med.* 2002;77(10 Suppl):S13–6.
- USMLE step 1: content description and general information. [Internet] Federation of the State Medical Boards of the United States, Inc (FSMB) and the National Board of Medical Examiners® (NBME®): c2015. [updated 2015 Feb; cited 2016 Mar 15]. Available from: http://www.usmle.org/pdfs/step-1/2015content_step1.pdf. Accessed 15 March 2016.
- Morrison CA, Ross LP, Fogle T, Butler A, Miller J, Dillon GF. Relationship between performance on the NBME Comprehensive Basic Sciences Self-Assessment and USMLE Step 1 for U.S. and Canadian medical school students. *Acad Med.* 2010;85(10 Suppl):S98–101.
- Glew RH, Ripkey DR, Swanson DB. Relationship between students' performances on the NBME Comprehensive Basic Science Examination and the USMLE Step 1: a longitudinal investigation at one school. *Acad Med.* 1997;72(12):1097–102.
- Ogunyemi D, De Taylor-Harris S. NBME Obstetrics and Gynecology clerkship final examination scores: predictive value of standardized tests and demographic factors. *J Reprod Med.* 2004;49(12):978–82.
- Zahn CM, Saguil A, Artino Jr AR, Dong T, Ming G, Servey JT, et al. Correlation of National Board of Medical Examiners scores with United States Medical Licensing Examination Step 1 and Step 2 scores. *Acad Med.* 2012;87(10):1348–54.
- Elam CL, Johnson MM. NBME Part I versus USMLE Step 1: predicting scores based on preadmission and medical school performances. *Acad Med.* 1994;69(2):155.
- Myles TD, Henderson RC. Medical licensure examination scores: relationship to obstetrics and gynecology examination scores. *Obstet Gynecol.* 2002;100(5 Pt 1):955–8.
- Step 2 ck (clinical knowledge): content outline and specifications. [Internet] Federation of the State Medical Boards of the United States, Inc. (FSMB) and the National Board of Medical Examiners® (NBME®): c2016. [cited 2016 Mar 15]. Available from: <http://www.usmle.org/step-2-ck/#contentoutlines>. Accessed 15 March 2016.
- Step 2 clinical skills (cs): content description and general information. [Internet] Federation of State Medical Boards of the United States, Inc., and the National Board of Medical Examiners® (NBME®): c2015. [updated 2015 Feb; cited 2016 Mar 15]. Available from: <http://www.usmle.org/pdfs/step-2-cs/cs-info-manual.pdf>. Accessed 15 March 2016.
- Kozar RA, Kao LS, Miller CC, Schenarts KD. Preclinical predictors of surgery NBME exam performance. *J Surg Res.* 2007;140(2):204–7.
- Williams PC, Epps AC, McCammon S. The strategic impact of a changing curriculum and learning environment on medical students' academic performance. *J Natl Med Assoc.* 2011;103(9-10):802–10.
- Swanson DB, Sawhill A, Holtzman KZ, Bucak SD, Morrison C, Hurwitz S, et al. Relationship between performance on part I of the American Board of Orthopaedic Surgery Certifying Examination and Scores on USMLE Steps 1 and 2. *Acad Med.* 2009;84(10 Suppl):S21–4.
- Denton GD, Durning SJ, Wimmer AP, Pangaro LN, Hemmer PA. Is a faculty developed pretest equivalent to pre-third year GPA or USMLE step 1 as a predictor of third-year internal medicine clerkship outcomes? *Teach Learn Med.* 2004;16(4):329–32.
- Edwards RK, Davis JD, Kellner KR. Effect of obstetrics-gynecology clerkship duration on medical student examination performance. *Obstet Gynecol.* 2000;95(1):160–2.
- Bostwick JM, Alexander C. Shorter psychiatry clerkship length is associated with lower NBME psychiatry shelf exam performance. *Acad Psychiatry.* 2012;36(3):174–6.
- Griffith 3rd CH, Wilson JF, Haist SA, Albritton TA, Bogner BA, Cohen SJ, et al. Internal medicine clerkship characteristics associated with enhanced student examination performance. *Acad Med.* 2009;84(7):895–901.
- Sawhill A, Butler A, Ripkey D, Swanson DB, Subhiyah R, Thelman J, et al. Using the NBME self-assessments to project performance on USMLE Step 1 and Step 2: impact of test administration conditions. *Acad Med.* 2004;79(10 Suppl):S55–7.
- Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ.* 2002;324(7343):952–7.
- Reed DA, West CP, Mueller PS, Ficalora RD, Engstler GJ, Beckman TJ. Behaviors of highly professional resident physicians. *JAMA.* 2008;300(11):1326–33.
- Dong T, Saguil A, Artino Jr AR, Gilliland WR, Waechter DM, Lopreato J, et al. Relationship between OSCE scores and other typical medical school performance indicators: a 5-year cohort study. *Mil Med.* 2012;177(9 Suppl):44–6.
- Lewis CE, Hiatt JR, Wilkerson L, Tillou A, Parker NH, Hines OJ. Numerical versus pass/fail scoring on the USMLE: what do medical students and residents want and why? *J Grad Med Educ.* 2011;3(1):59–66.
- Weiss PM, Koller CA, Hess LW, Wasser T. How do medical student self-assessments compare with their final clerkship grades? *Med Teach.* 2005;27(5):445–9.
- Awad SS, Liscum KR, Aoki N, Awad SH, Berger DH. Does the subjective evaluation of medical student surgical knowledge correlate with written and oral exam performance? *J Surg Res.* 2002;104(1):36–9.
- Ogunyemi D, Taylor-Harris D. Factors that correlate with the U.S. Medical Licensure Examination Step-2 scores in a diverse medical student population. *J Natl Med Assoc.* 2005;97(9):1258–62.
- National Board of Medical Examiners (NBME). Subject examination program. Obstetrics and gynecology subjection examination: score interpretation guide [Internet] [cited 2014 Oct 7]. Available from: http://www.nbme.org/pdf/samplescorereports/clinical_sci_score_report.pdf. Accessed 15 March 2016.
- Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of examinee gender and USMLE step 1 performance. *Acad Med.* 2008;83(10 Suppl):S58–62.
- Association of American Medical Colleges (AAMC). Medical College Admission Test (MCAT). What will be different about the MCAT exam in 2014-January 2015? [Internet] Washington (DC) [cited 2014 Oct 7]. Available from: <https://www.aamc.org/students/applying/mcat/faq/313936/differentexam.html>. Accessed 15 March 2016.
- 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(4):175–81.
- Ripkey DR, Case SM, Swanson DB. Predicting performances on the NBME Surgery Subject Test and USMLE Step 2: the effects of surgery clerkship timing and length. *Acad Med.* 1997;72(10 Suppl 1):S31–3.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

