Advances in Medical Education and Practice

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Predictive Validity of Preclerkship Performance Metrics on USMLE Step 2 CK Outcomes in the Step 1 Pass/Fail Era

Karl Aharonian, Maggie Sanders, Tal Schlesinger, Victoria Winter, Edward Simanton

Kirk Kerkorian School of Medicine, Las Vegas, NV, USA

Correspondence: Karl Aharonian, Email aharonia@unlv.nevada.edu

Purpose: With the transition of USMLE Step 1 to pass/fail, Step 2 CK has become a critical factor in residency selection. This study evaluates the predictive value of preclinical academic metrics for Step 2 CK outcomes to help students assess their metrics for residency before starting clerkships.

Methods: We analyzed data from 58 students at the Kirk Kerkorian School of Medicine at UNLV. Academic metrics included MCAT scores, NBME-style Phase 1 exam performance, and CBSE scores. Pearson correlation coefficients were calculated between each predictor and Step 2 CK scores. Binary logistic regression was used, with Step 2 CK upper quartile (score >255) as the dependent variable and thresholds for each predictor as independent variables.

Results: MCAT scores showed a moderate positive correlation with Step 2 CK ($r=0.37$, $p=0.004$). Students scoring >510 on the MCAT were significantly more likely to achieve >255 on Step 2 CK. Phase 1 NBME exam performance had the strongest correlation ($r=0.67$, $p<0.0001$), with students scoring above the national mean on >75% of exams more likely to achieve >255. While CBSE scores showed a similar correlation ($r=0.67$, $p<0.0001$), logistic regression analysis revealed that their predictive ability was less consistent for students in the middle performance range, indicating variability in classification utility.

Conclusion: Preclinical NBME exam performance emerged as the strongest predictor of Step 2 CK success in our cohort, with MCAT and CBSE scores offering additional, though less robust, predictive value. These findings highlight the importance of early academic guidance to optimize residency preparedness and provide actionable insights for students assessing their readiness within the new Step 1 pass/fail framework. Further multi-institutional research is warranted to validate and generalize these results.

Keywords: USMLE step 2 CK, step 1 pass/fail, NBME, MCAT, CBSE, phase 1 examinations, medical student performance, residency, medical education, competitiveness in residency, step 2 CK predictors, preclerkship, preclinical

Introduction

The United States Medical Licensing Examination (USMLE) is a three-step examination designed to assess a medical student's ability to apply knowledge, concepts, and principles essential for patient care and clinical practice. Step 1 traditionally focused on foundational medical sciences, Step 2 Clinical Knowledge (CK) on clinical science application, and Step 3 on managing real-life scenarios as a licensed physician. Step 1 and Step 2 examinations are offered during medical school while Step 3 is offered typically during first year as a resident.[1] These exams are strategically integrated into the medical curricula to evaluate and guide student progression through their training.

Since the USMLE Step 1 examination transitioned to a Pass/Fail grading format, the landscape of medical education has adapted significantly. One of the most significant challenges following this change is how to best advise students in this new paradigm. Thus far, the primary focus of this guidance has been predicting a passing score for the USMLE Step 1 based on practice test scores. Leading medical institutions, such as Emory School of Medicine, have released an official guide suggesting that students should take the exam once they reach a 99% likelihood of passing, which corresponds to achieving a score of approximately 68% on practice exams.[2]

Received: 21 November 2024
Accepted: 22 January 2025
Published: 25 February 2025

However, for students aspiring to enter competitive specialties, a crucial question arises regarding how this will affect their performance on USMLE Step 2. A recent study indicated that 70.5% of students believe they will need to devote more time to preparing for USMLE Step 2 to remain competitive now that USMLE Step 1 is Pass/Fail.[3] Studies have shown that this concern is well-founded and supported by evidence. One survey of over 2000 residency programs found that 81% of program directors anticipate USMLE Step 2 CK becoming a more critical factor in determining applicant matches following the transition of Step 1 into pass/fail format.[4]

Historically, USMLE Step 1 score was the best predictor of student performance on Step 2. A systematic review and meta-analysis confirmed this correlation, providing robust evidence supporting Step 1 performance as a useful predictor of Step 2 scores.[5] A study conducted at our institution, the Kirk Kerkorian School of Medicine at the University of Nevada, Las Vegas, revealed that USMLE Step 1 had the strongest correlation with performance on Step 2 ($r = 0.752$, $p < 0.001$).[6] This study also found that the Internal Medicine shelf exam also held an important predictive value (r=) for Step 2 scores.[7] This study concluded that Step 2 CK scores could be effectively predicted using both Step 1 and shelf exam performance.

However, following the shift of Step 1 to pass/fail, there is now a noticeable gap in the performance metrics able to predict and optimize outcomes on Step 2 CK.

The absence of Step 1 numerical scores creates uncertainty for both students and educators. Previously, a strong Step 1 score could guide students and programs in setting realistic expectations and goals for Step 2 preparation. Now, with Step 1 being Pass/Fail rather than a numerical score, the focus has shifted to finding new benchmarks and strategies to ensure success on Step 2 CK. Research is still evolving to identify these benchmarks, and current recommendations are largely based on historical data and expert consensus rather than robust empirical evidence.

Our study aims to address this gap in evidence by analyzing the performance of the Class of 2025, the first cohort at the Kirk Kerkorian School of Medicine to take Step 2 CK following the transition of Step 1 to Pass/Fail. By examining their scores and correlating them with various preclerkship metrics—including MCAT, CBSE, and the average scores of Phase 1 NBME exams—we aim to identify the factors that best predict Step 2 CK performance. This study will offer valuable insights to help students and educators navigate the evolving landscape of USMLE assessments.

Methods

The Institutional Review Board (IRB) approval for this study was granted by the University of Nevada, Las Vegas School of Medicine IRB, under protocol number IRB#UNLV-2022-293. The data used in the study is private and was accessed with permission from the Kirk Kerkorian School of Medicine at the University of Nevada, Las Vegas, specifically authorized by the Office of Program Evaluation. The study utilized previously de-identified academic performance data provided by the Office of Program Evaluation. Based on the IRB's review of the study protocol, no consent was required.

The Class of 2025 cohort at the Kirk Kerkorian School of Medicine (KSOM) provided a total of 58 subjects. All members of the class were included in the study. This is the only class in KSOM that has taken Step 2 after the formal transition of Step 1 to pass/fail (P/F).

Three main areas of performance were explored in relation to USMLE Step 2 CK scores:

1. Premedical Performance Measures: This initially included the average MCAT score and GPA. GPA was excluded due to significant variation caused by different institutional policies and the selection of major/minor.
2. Preclinical Performance Variables: Two variables were used to assess preclinical performance: performance on NBME system block exams and NBME Comprehensive Basic Science Exam (CBSE) scores.

Phase 1 of our medical school curriculum consists of organ system blocks lasting 3 to 6 weeks, each concluding with an NBME-style examination administered in a standardized environment following NBME protocols. Questions are chosen from the NBME Customized Assessment Service (CAS). The questions in the CAS exams are composed from a database of retired USMLE Step 1 questions. The school's faculty chooses questions for specific system and concepts.

The exams administered at our school and included in this study consist of Foundational Medical Science (FMS) exams and organ block exams. FMS exams include basic science knowledge such as biochemistry, cell physiology, immunology, microbiology, and immunology. The organ systems tested include Cardiology, Pulmonology, Renal, Hematology/Oncology,

Gastrointestinal, Reproductive, Endocrinology, Neurology, Psychiatry, Dermatology, and Musculoskeletal systems. Student grades are adjusted based national difficulty as reported by the NBME. For the purpose of this study, student performance on each exam was coded as 1=performance equal to or greater than the national mean and 0=performance below the national mean. The "%national mean P1" variable represents the mean of all exams based on the binary coding. In essence this is the percentage of exams in which each student equalled or outperformed the national average. This was done to use national standards which would be more useful to students rather than local averages that would not be generalizable.

It is important to note that CBSE administered by schools is either USMLE Step 1 performance indicator or as a summative phase 1 exam as described in previous studies.[6] The CBSE is the last required examination taken by students at the end of Phase 1 and prior to taking Step 1 at our school. A minimum pass score on CBSE as specified by faculty is required to pass the last course of Phase 1. It is also used as a measure of Step 1 readiness. All students must pass Step 1 prior to proceeding to Phase 2, clinical rotations. Our school used the CBSE as the primary indicator of USMLE Step 1 performance, as other practice exams/forms taken by students were completed in non-standardized environments with significant variability in how students approached these exams.

Pearson correlation coefficients were calculated between each of these variables and the USMLE Step 2 CK score to determine the relationships between the performance measures and the Step 2 CK scores. Following this, a regression analysis (Method = Enter) was conducted using the same variables to develop a prediction model.

All variables were converted to binary variables with the intention of identifying performance in the upper quartile of Step 2 CK. As such, MCAT scores at or above 510 were compared to performance below 510. Percent national mean Phase 1 scores used a cutoff of 75% or greater. This means students who performed at or above the national average on at least 75% of their phase 1 exams. CBSE scores of 65% or more were compared with those below 65%. Step 2 CK upper quartile was defined as a score equal or greater than 255. Chi square analysis was utilized to determine statistical significance between these binary variables.

Results

MCAT

The association between USMLE Step 2 CK and MCAT scores was assessed using Pearson correlation analysis, which revealed a correlation coefficient of 0.37 ($p = 0.004$), indicating a weak to moderate positive relationship between performance on these two examinations.

Chi square analysis showed statistically significant breakdown of 58% of students who scored above 510 on MCAT had Step 2 CK scores of above 255, while 42% of students had scores below 255. In students who scored below 510 on MCAT, 85% had scored below 255 on Step 2 CK and consequently 15% of students scored above 255.

Table 1 provides an overview of both the overall and subgroup analyses, while Figure 1 displays the correlation graph between MCAT scores and Step 2 CK performance.

Average Score of Phase 1 NBME Exams

The performance of students who scored above the national average on at least 75% of their phase 1 CAS exams showed a strong positive correlation with USMLE Step 2 CK scores, with a Pearson coefficient of 0.67 ($p < 0.0001$). Table 2

Table 1 Between Students Who Scored Above/Below 510 on MCAT and Students Who Scored Above/Below 255 on Step 2 CK

	MCAT below 510	MCAT 510+	Total	Chi-square
Step 2 below 255	23 (85.2%)	13 (41.9%)	36 (62.1%)	X^2=11.5
Step 2 of 255+	4 (14.8%)	18 (58.1%)	22 (37.9%)	df=1
Total	27	31	58	Sig. =0.10 (2-sided)

Notes: As shown in Table 1, students with MCAT of 510 and above were significantly more likely to receive scores equal or above 255 on Step 2 CK.

https://doi.org/10.2147/AMEP.S505612

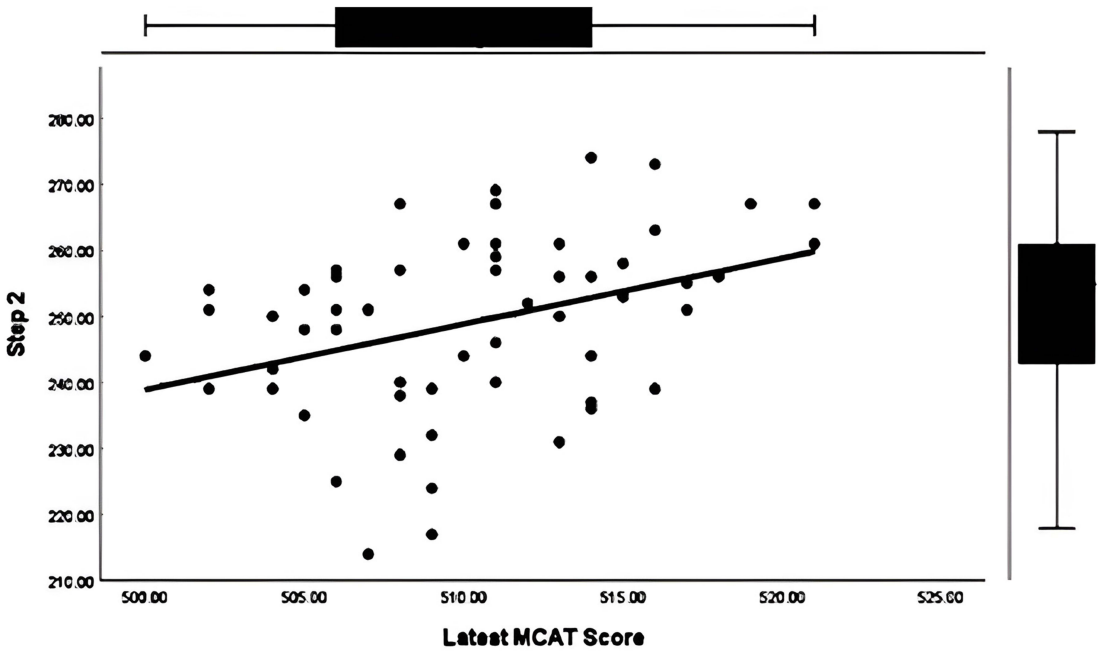

Figure 1 Correlation between latest MCAT score and Step 2 CK score.

presents the overall chi square values for the Phase 1 exam performance while Figure 2 illustrates the trend between the performing above national mean and Step 2 CK performance.

Chi square analysis showed statistically significant breakdown of 51.5% of students who scored above national average on 75% of their exams in phase 1 had Step 2 CK scores of above 255, while 48.5% of students had scores below 255. In students who scored below national average in the majority of their exams, 80% had scored below 255 on Step 2 CK and consequently 20% of students scored above 255.

CBSE

CBSE scores demonstrated a moderate positive correlation with Step 2 CK scores (Pearson=0.67, $p < 0.0001$). Chi square analysis showed statistically significant breakdown of 44% of students who scored above 255 on Step 2 CK had CBSE scores below 65%. Additionally, CBSE scores below 65% were more predictive of lower Step 2 CK performance, with only 24% of students achieving a Step 2 CK score above 255 when their CBSE was under 65%. Figure 3 illustrates the correlation between CBSE and Step 2 CK. Table 3 llustrates the sub-stratified data.

Table 2 Comparison Between Students Who Scored Above/Below National Mean on 75% of Their Phase 1 Exams and Students Who Scored Above/Below 255 on Step 2 CK

	Phase 1 Mean Below 0.75	Phase 1 Mean of 0.75+	Total	Chi-square
Step 2 below 255	20 (80.0%)	16 (48.5%)	36 (62.1%)	X^2=6.00 df=1 Sig. =0.014 (2-sided)
Step 2 of 255+	5 (20.0%)	17 (51.5%)	22 (37.9%)	
Total	25	33	58	

Notes: As shown in Table 2, students with Phase 1 Mean of 0.75 and above were significantly more likely to receive scores equal or above 255 on Step 2 CK.

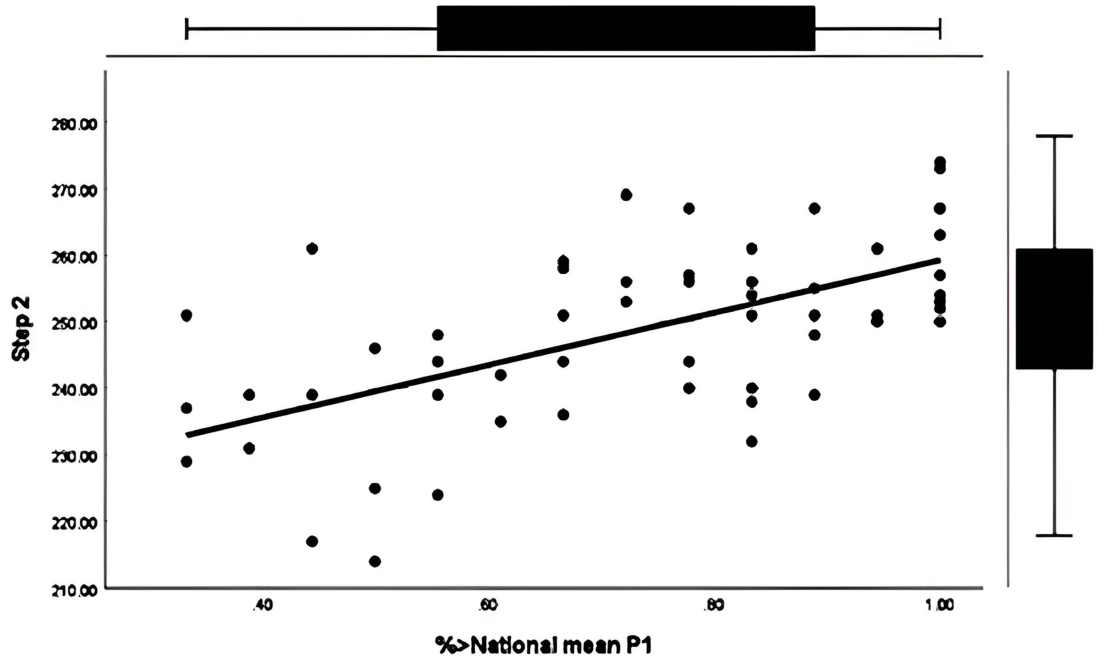

Figure 2 Correlation between percentile above national mean on Phase 1 examinations and Step 2 CK score.

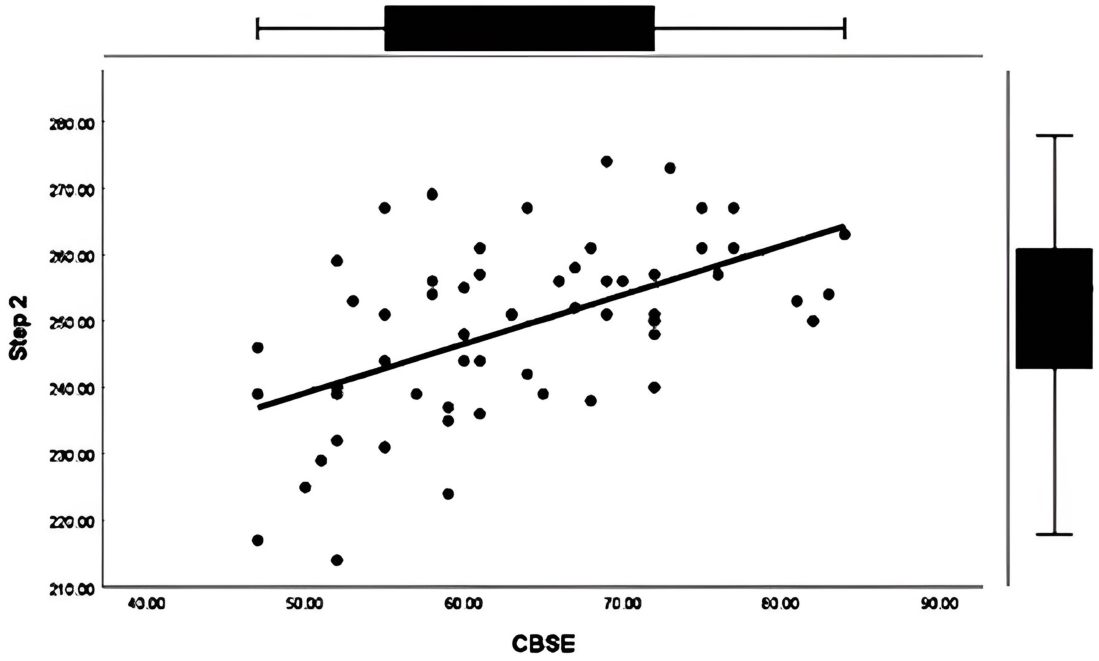

Figure 3 Correlation between percentile CBSE and Step 2 CK scores.

Discussion

With the recent change of the USMLE Step 1 exam to a pass/fail score system, increasing emphasis is being placed on medical students' scores on the USMLE Step 2 exam for acceptance into residency. The majority of currently available studies evaluate predictors of Step 1 scores, with very few studies evaluating factors that influence performance on Step 2.

Table 3 Comparison Between Students Who Scored Above/Below 65% on CBSE and Students Who Scored Above/Below 255 on Step 2 CK

	CBSE Below 65	CBSE of 65+	Total	Chi-square
Step 2 below 255	25 (75.8%)	11 (44.0%)	36 (62.1%)	X^2=6.09 df=1 Sig. =0.014 (2-sided)
Step 2 of 255+	8 (24.2%)	14 (56.0%)	22 (37.9%)	
Total	33	25	58	

Notes: As shown in Table 3, students with CBSE 65 and above were significantly more likely to receive scores equal or above 255 on Step 2 CK.

Our study analyzed the Kirk Kerkorian School of Medicine at UNLV class of 2025's performance on Step 2 to understand what factors in a medical student's educational history may be predictive of success on Step 2. Our study was a first of its kind now to offer guidelines based on their pre-clinical performance for medical students post USMLE Step 1 pass/fail transition. We focused on student scores on the standardized exams MCAT, CBSE, and Phase I NBME subject exams.

MCAT

In our cohort, 58.1% of students in the class of 2025 who had an MCAT score of 510 or higher ended up scoring higher than 255 on the USMLE Step 2 exam ($\chi^2 = 11.5$, df = 1, p = 0.010, statistically significant). MCAT score has previously been associated with high performance on the USMLE Step 1 exam.[8–11] Thus, it is unsurprising that MCAT is also a predictor of success on the USMLE Step 2 exam. Prior studies have found that MCAT is a stronger predictor of Step 1 versus Step 2 scores due to the basic science nature of Step 1.[7] However with the new shift of the Step 1 exam to pass/fail, more thorough analysis is needed to truly evaluate the correlation between MCAT performance and Step 2 scores.

CBSE

Our statistical analysis found that even though CBSE scores were statistically correlated with USMLE Step 2 Performance, they were not good predictors of the Step 2 USMLE exam score. This conclusion stems from the fact that 44% of students who scored 255 on USMLE Step 2 did not have high CBSE scores. We hypothesize that this lack of meaningful correlation is likely due to a multitude of factors. At UNLV, students take the CBSE on the same date, however at different points of their studying for USMLE Step 1. Additionally, the basic science nature of Step 1, and therefore CBSE, makes the content of the exam itself different from Step 2 and therefore a weaker predictor. Finally, a plausible explanation could be the shift of importance from Step 1 scores to Step 2 scores.

However, one conclusion can be drawn is that the CBSE scores serve as better predictors of the outliers. This finding aligned with studies such as Guiot et al which showed that scoring below 66% had a much higher chance of not performing well on Step 1 and Step 2 CK.[12] In our results, it can be seen that students who scored much higher on CBSE were likely to score higher on USMLE Step 2 and vice versa. For an average performing student, CBSE serves as a poor predictor of performance.

Phase I NBME CAS Exams

The performance on Phase 1 exams was the strongest predictor of USMLE Step 2 score based on the positive correlation of 0.67. This can be explained by the content of these exams which test the fundamental physiology and pathology of various disciplines emphasizing high yield material from each subject. These findings align with previous studies such as by Bird et al, which showed that NBME style examinations during the preclinical years were a significant predictor of performance on Step 2 CK.[13]

Although these exams served as a strong predictor in our cohort, future studies need to be done to assess if this will be replicated in other medical institutions with different exam structures.

Guidelines for Students

One of the main aims of our study was to provide guidelines for students navigating through the new era of Step 1 being Pass or Fail. Step 2 CK is one of the key screening examinations for residency matches, which will likely serve as the sole metrics after Step 1 Pass/fail transition.[3] From the single cohort analysis of our institution, we propose the following guideline for students. Specifically, it appears that one of the best predictors of performance on USMLE Step 2 CK is performing above average on Phase 1 examinations. Although MCAT served as a significant predictor of success in our cohort, the correlation is weaker than that of Phase 1 exams. Finally, it is clear from the data of this cohort that no single variable should discourage students from aiming for competitive fields as there are a number of students in this cohort with below average pre-clinical scores who performed above 255 on Step 2 CK. Additionally, the importance of other aspects of application such as publications and clinical grades are not investigated in our study as balancing factors in application screening.

Limitations

Limitations of this study include its focus on a single institution and a cohort of only 58 medical students. This may restrict the external validity of the findings for students outside the Kirk Kerkorian School of Medicine.

Additionally, all students in the cohort take the CBSE simultaneously, regardless of their scheduled Step 1 exam dates. This may result in some students taking the CBSE without the same level of preparation they would invest for the Step exams or the MCAT.

In 2020, the MCAT format was altered due to the COVID-19 pandemic. Exams during this period were shortened from 7.5 hours to just over 5 hours, enabling more exams to be administered each day (2020 format). If students took the exam in this shortened format, it may have influenced their scores compared to those who took the original 7.5-hour MCAT.

Furthermore, students pursuing less competitive specialties may not prioritize achieving a top score on the Step 2 exam, potentially aiming for a score closer to the average. However, it is generally expected that medical students will strive to perform at their best on such a significant examination in their careers. It is also important to note that different schools have different timelines for Step 2 CK and our study results may not be generalizable to schools with significantly different schedules of Step 2 CK.

A limitation to expanding this study to other institutions is that not all medical schools utilize NBME-style questions for their preclinical exams. Some institutions rely on professor-written exams tailored to their specific curricula, which may complicate efforts to predict Step 2 performance based on preclinical exam scores at those schools.

Future Directions

Future directions for this research include collecting data from multiple institutions to investigate whether the MCAT, CBSE, and preclinical exam scores correlate with Step 2 exam scores. Given that the average score on preclinical NBME-style exams was the strongest predictor of Step 2 scores at the Kirk Kerkorian School of Medicine, we aim to determine if this trend holds true in larger cohorts.

Additionally, it would be valuable to consider other factors when predicting Step 2 exam scores, such as the resources utilized by medical students, total in-class lecture time, and the format of instruction (eg, lecture vs problem-based learning). Identifying the strategies that correlate with the highest Step 2 exam scores is crucial in the context of the Step 1 pass/fail grading system.

Finally, shelf exams and their sequence can be assessed as another predictor of Step 2 CK scores once students start their clerkships. Previous studies, including one at our institution, have shown evidence for these examinations to correlate with Step 2 CK performance.[6,14–16] Therefore, future studies should explore this relationship in Step 1 pass/fail era.

Conclusion

The transition of USMLE Step 1 to a Pass/Fail format has shifted the focus to Step 2 CK as a critical metric for residency selection, raising questions about the best predictors of success on this exam. Our study analyzed the performance of the Class of 2025 at the Kirk Kerkorian School of Medicine, identifying key preclerkship metrics—MCAT scores, Phase 1

NBME exam performance, and CBSE scores. Among these, Phase 1 NBME exam performance showed the strongest correlation, suggesting that consistent success on foundational exams is a strong indicator of Step 2 CK performance.

The study's limitations, including its single-institution focus and limited sample size, may restrict the generalizability of the findings. Multi-institutional research is warranted to validate these predictors and to further refine strategies for supporting students' success.

These findings provide actionable insights into the predictive validity of preclerkship performance metrics, offering guidance for students aiming to optimize their preparation for Step 2 CK within the evolving Step 1 Pass/Fail framework.

Disclosure

The authors report no conflicts of interest in this work.

References

1. United States Medical Licensing Examination. USMLE bulletin of information. 2025. Availble from: https://www.usmle.org/sites/default/files/2024-11/2025bulletin.pdf. Accessed December 18, 2024.
2. Emory School of Medicine. USMLE step 1 guide for students. Center for Holistic Student Success. Emory University; 2024. Available from: https://med.emory.edu/education/center-for-holistic-student-success/_documents/step-1-guide. Accessed September 18, 2024.
3. Girard AO, Qiu C, Lake IV, Chen J, Lopez CD, Yang R. US medical student perspectives on the impact of a pass/fail USMLE step 1. *J Surg Educ*. 2022;79(2):397–408. doi:10.1016/j.jsurg.2021.09.010
4. Makhoul AT, Pontell ME, Ganesh Kumar N, Drolet BC. Objective measures needed — program directors' perspectives on a pass/fail USMLE step 1. *N Engl J Med*. 2020;382(25):2389–2392. doi:10.1056/NEJMp2006148
5. Jacobparayil A, Ali H, Pomeroy B, Baronia R, Chavez M, Ibrahim Y. Predictors of performance on the United States medical licensing examination step 2 clinical knowledge: a systematic literature review. *Cureus*. 2022;14(2):e22280. doi:10.7759/cureus.22280
6. Kracaw RA, Dizon W, Antonio S, Simanton E. Predicting United States medical licensing examination step 2 clinical knowledge scores from previous academic performance measures within a longitudinal interleaved curriculum. *Cureus*. 2021;13.
7. Wright WS, Baston K. Use of the national board of medical examiners® comprehensive basic science exam: survey results of US medical schools. *Adv Med Educ Pract*. 2017;8:377–383. PMID: 28670149; PMCID: PMC5482400. doi:10.2147/AMEP.S136669
8. Saguil A, Dong T, Gingerich RJ, et al. Does the MCAT predict medical school and PGY-1 performance? *Mil Med*. 2015;180(4 Suppl):4–11. doi:10.7205/MILMED-D-14-00550
9. Hanson JT, Busche K, Elks ML, 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(9):1374–1384. doi:10.1097/ACM.0000000000004754
10. Purkiss J, Plegue M, Grabowski CJ, et al. Examination of medical college admission test scores and us medical licensing examination step 1 and step 2 clinical knowledge scores among students with disabilities. *JAMA Network Open*. 2021;4(5):e2110914. doi:10.1001/jamanetworkopen.2021.10914
11. Seal ZA, Koek W, Sharma R. Correlation of medical college admission test scores and self-assessment materials with the United States medical licensing examination step 1 performance. *Cureus*. 2020;12(4):e7519. doi:10.7759/cureus.7519
12. Guiot HM, Franqui-Rivera H. Predicting performance on the United States medical licensing examination step 1 and step 2 clinical knowledge using results from previous examinations. *Adv Med Educ Pract*. 2018;9:943–949. doi:10.2147/AMEP.S180786
13. Bird JB, Olvet DM, Willey JM, Brenner JM. A generalizable approach to predicting performance on USMLE step 2 CK. *Adv Med Educ Pract*. 2022;13:939–944. doi:10.2147/AMEP.S373300
14. Drake E, Phillips JP, Kovar-Gough I. Exploring preparation for the USMLE step 2 exams to inform best practices. *PRiMER*. 2021;5:26. doi:10.22454/PRiMER.2021.693105
15. 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(5):1200–1205. doi:10.1016/j.jsurg.2018.02.017
16. Monteiro KA, George P, Dollase R, Dumenco L. Predicting United States medical licensure examination step 2 clinical knowledge scores from previous academic indicators. *Adv Med Educ Pract*. 2017;8:385–391. doi:10.2147/AMEP.S138557