Predicting United States Medical Licensing Examination Step 2 Clinical Knowledge Scores from Previous Academic Performance Measures within a Longitudinal Interleaved Curriculum

Rachel A. Kracaw 1, Wynona Dizon 1, Sabrina Antonio 1, Edward Simanton 1

1. Medical Education, University of Nevada, Las Vegas School of Medicine, Las Vegas, USA

Corresponding author: Rachel A. Kracaw, sigurdso@unlv.nevada.edu

Abstract

Background

United States Medical Licensing Examination (USMLE) Step 1 is a common metric looked at by residency programs to determine invitations for candidates to interview. However, USMLE Step 2 Clinical Knowledge (CK) has also been an important factor for selecting applicants to interview and plays a significant role during applicant selection. This study aims to identify academic performance measures that correlate with USMLE Step 2 CK scores and to develop a model to predict USMLE Step 2 CK scores using previous academic measures from the first two cohorts in the longitudinal interleaved clerkship (LInC) at the Kirk Kerkorian School of Medicine at the University of Nevada, Las Vegas (KSOM).

Setting

The KSOM is a newly accredited US allopathic medical school that accepted its first class in 2017. At KSOM, a LInC model is used in the primary clinical year. In this model, rotations are two weeks in duration before moving on to the next specialty. Students complete the National Board of Medical Examiners (NBME) subject examinations in all six specialties in one week at the midpoint and the end of the LInC. Students who passed an exam at the midpoint can opt out of that exam at the end as the higher of the two exam scores is recorded. However, most students choose to take all the exams again to improve their scores and prepare for USMLE Step 2 CK.

Methodology

Academic performance measures were gathered from the class of 2021 and 2022 (n = 101) including undergraduate grade point average (GPA), undergraduate science GPA, medical college admission test score, USMLE Step 1 score, NBME clinical subject exam scores, and USMLE Step 2 CK scores. Pearson correlations were run between the performance variables and USMLE Step 2 CK scores to measure influence variables individually, then a regression model measured impacts of variables together.

Results

All variables except undergraduate science GPA significantly correlated with USMLE Step 2 CK score. USMLE Step 1 had the strongest correlation (r = 0.752, p < 0.001). The regression model had an R of 0.859 with the internal medicine subject exam showing the highest beta coefficient (0.327, p < 0.001).

Conclusions

This study determined that USMLE Step 2 CK scores can be effectively predicted using available performance measures. With USMLE Step 1 becoming pass/fail in January 2022, the importance of USMLE Step 2 CK as a screening tool in the residency application process will likely increase. This study was conducted within a LInC curriculum and may have limited value in the prediction of scores within other clinical year curricula.

Introduction

The United States Medical Licensing Examination (USMLE) has three exams consisting of Step 1, Step 2 Clinical Knowledge (CK), and Step 3. USMLE Step 1 tests students’ knowledge of science and preclinical medicine, whereas USMLE Step 2 CK focuses more on clinical medicine. Traditionally, USMLE Step 1 has been the most impactful measure looked at by residency programs to determine invitations for candidates to interview according to the National Resident Matching Program (NRMP) data in 2020 [1]. USMLE Step 1 scores have been correlated with resident performance and ability to pass medical licensing board
found in Table 1. Pearson correlations for all of the academic performance measures with USMLE Step 2 CK scores can be

Results

Pearson correlations for all of the academic performance measures with USMLE Step 2 CK scores can be found in Table 1.
### TABLE 1: Results of academic performance measures' correlation with USMLE Step 2 CK scores.

* Statistically significant with an alpha of 0.05.

| Academic performance measure                      | R-value | P-value  |
|---------------------------------------------------|---------|----------|
| Family Medicine Subject Exam                      | 0.627   | <0.001*  |
| Medicine Subject Exam                              | 0.745   | <0.001*  |
| Obstetrics and Gynecology Subject Exam             | 0.616   | <0.001*  |
| Pediatric Subject Exam                             | 0.672   | <0.001*  |
| Psychology Subject Exam                            | 0.530   | <0.001*  |
| Surgery Subject Exam                               | 0.617   | <0.001*  |
| USMLE Step 1                                       | 0.752   | <0.001*  |
| Basic Sciences Exam Mean                           | 0.604   | <0.001*  |
| MCAT                                               | 0.276   | 0.005*   |
| Undergraduate GPA                                  | 0.233   | 0.019*   |
| Undergraduate Science GPA                          | 0.196   | 0.050    |

The variables most correlated included USMLE Step 1 (r = 0.752, p < 0.001), the medicine NBME subject exam (r = 0.745, p < 0.001), and the pediatric NBME subject exam. The least correlated variables included MCAT (r = 0.276, p = 0.005) and undergraduate GPA (r = 0.233, p = 0.019). The only variable not significantly correlated with USMLE Step 2 CK was undergraduate science GPA (r = 0.196, p = 0.050).

For the regression model predicting USMLE Step 2 CK scores, the Pearson coefficient was reported as an R-value of 0.859 (p < 0.001). The standardized and unstandardized coefficients along with their respective p-values are shown in Table 2.
### Table 2: Standardized and unstandardized coefficients for the predictive model to predict USMLE Step 2 CK scores.

*Statistically significant with an alpha of 0.05.

| Academic performance measure          | Standardized coefficient | Standard error | Unstandardized coefficient (beta) | P-value |
|--------------------------------------|--------------------------|----------------|-----------------------------------|---------|
| Family Medicine                      | 0.349                    | 0.219          | 0.134                             | 0.114   |
| Medicine                             | 0.634                    | 0.175          | 0.327                             | <0.001* |
| Obstetrics and Gynecology            | 0.086                    | 0.203          | 0.036                             | 0.672   |
| Pediatrics                           | 0.355                    | 0.179          | 0.171                             | 0.051   |
| Psychiatry                           | 0.077                    | 0.235          | 0.023                             | 0.744   |
| Surgery                              | -0.117                   | 0.170          | -0.060                            | 0.493   |
| Basic Sciences Exam mean             | 32.721                   | 24.083         | 0.114                             | 0.178   |
| USMLE Step 1                         | 0.281                    | 0.084          | 0.321                             | 0.001*  |
| MCAT                                 | -0.168                   | 0.181          | -0.060                            | 0.356   |
| Undergraduate GPA                    | 17.549                   | 8.388          | 0.390                             | 0.039*  |
| Undergraduate Science GPA            | -16.038                  | 6.847          | -0.432                            | 0.021*  |
| Constant                             | 123.897                  | 85.680         | -                                 | 0.152   |

The variables with significant beta coefficients included NBME medicine subject exam (b = 0.327, p < 0.001), USMLE Step 1 score (b = 0.321, p = 0.001), undergraduate science GPA (b = -0.432, p = 0.021), and undergraduate science GPA (b = 0.390, p = 0.039).

The percentage error and difference for the predictive regression model are shown in Table 3, and Figure 1 shows a graph of the actual versus the predicted USMLE Step 2 CK score.

| Statistic                              | Average | Minimum | Maximum | Standard deviation |
|----------------------------------------|---------|---------|---------|--------------------|
| Difference between actual versus predicted USMLE Step 2 CK | 0       | -26     | 21      | 2.100              |
| Percentage error                       | 2.199%  | 0%      | 12.617% | 7.240%             |

### Table 3: Averages, minimums, maximums, and standard deviations for the prediction model percentage error and the difference between actual versus predicted.

USMLE: United States Medical Licensing Examination; CK: Clinical Knowledge
Discussion

This study determined that USMLE Step 2 CK scores can be effectively predicted using available performance measures. The variables with the strongest correlations were USMLE Step 1, the NBME medicine subject examination, and the NBME pediatrics subject examination. The lowest correlated measures were MCAT scores and undergraduate GPA. Undergraduate science GPA was the only variable that was not significantly correlated with USMLE Step 2 CK. All but one variable was significantly correlated, which shows these measures are predictive. Comparable trends have been found in studies that examine similar academic performance measures correlated with USMLE Step 1, which was expected, given that USMLE Step 1 and USMLE Step 2 CK also correlate well [3-7,10,12-19]. However, these correlations could also be explained by students’ academic aptitudes, as demonstrated previously [17]. A similar correlation showed that poor performance on standardized exams before and early in medical school is associated with poor performance on other exams later in school and beyond [3].

For the predictive model, the average percentage error and the r-value for the Pearson coefficient model both indicate that this model is highly predictive of student performance on USMLE Step 2 CK. Therefore, using previous academic performance measures can help determine performance on USMLE Step 2 CK. It is important to start predicting USMLE Step 2 CK scores given its rising importance in the residency application process [8,9]. With the introduction of a pass/fail USMLE Step 1 beginning January 2022, USMLE Step 2 CK will likely become more widely used during the residency application process [8,9]. Before this transition to pass/fail, USMLE Step 1 was used heavily as a cutoff for residency program interview invitations [1]. Because of this correlation, residency programs are projected to use USMLE Step 2 CK as they have used USMLE Step 1 in the past [8,9]. This exam will likely become the new deciding factor when determining to interview residency candidates due to USMLE Step 1 becoming pass/fail and the fact that USMLE Step 2 CK focuses more on clinical knowledge which is more applicable to residency performance [8,9].

Several limitations exist in this study. First, all participants were students from the charter and second class at KSOM. Being a new school could have influenced test scores and other academic measures because the curriculum is new and less developed compared to more established institutions. Second, the students involved in the study participated in a LInC model for their third year of medical school. Recent studies have shown that a longitudinal clerkship model can increase standardized test scores compared to the traditional clerkship model [20]. This is likely because students benefit from having to continuously study all specialty subjects throughout their clerkship year rather than studying one specialty at a time. Another benefit to the LInC model is that it removes the bias from the order in which the NBME subject exams were taken because KSOM administers all these exams at once rather than sequentially after each respective rotation. Some studies show that the order in which NBME subject exams are taken can affect scores on subsequent subject exams [21]. For example, the family medicine exam contains much medicine and obstetrics and gynecology content. Thus, if the family medicine exam is completed after having taken and completing medicine and obstetrics and gynecology, students will feel more prepared [21]. The LInC model also allows the NBME subject exams to be taken twice with the highest of the two scores being reported, thus, giving students more exposure to these exams, which may influence the model’s accuracy for other programs. Taking all six NBME subject exams in a one-week period could affect the predictive model and the actual USMLE Step 2 CK scores.

FIGURE 1: Actual versus predicted USMLE Step 2 CK scores.

USMLE: United States Medical Licensing Examination; CK: Clinical Knowledge

Actual vs Predicted USMLE Step 2 CK Scores

- Actual USMLE Step 2 CK Score vs Predicted USMLE Step 2 CK Score
- R² = 0.738

2021 Kracaw et al. Cureus 13(9): e18143. DOI 10.7759/cureus.18143

5 of 7
scores because students are either fatigued from so many exams or they have the opportunity to build their mental endurance to be able to sit for the USMLE Step 2 CK exam. Third, this study used USMLE Step 1 as a preclinical variable in the predictive model. As previously stated, USMLE Step 1 will become a pass/fail exam with a non-numeric score in January 2022. It cannot be used as a variable in future predictive models. Lastly, the number of students included in the study was small and the correlations could likely be strengthened if more students were included.

Interestingly, performance in medical school is a much better predictor of USMLE results than pre-admission variables [11]. In future studies, it would be beneficial to include more pre-clinical and clinical predictors rather than pre-admission variables. Kleshinski et al. found an unconventional correlation between age and USMLE Step 2 CK performance that has not been demonstrated elsewhere in the literature [7]. This study did not include student age in the analysis but could in future iterations. Lastly, a predictive model exists for USMLE Step 1 using measures such as study resources, study methods, length of study, financial need, and ethnicity, which could also be useful in future studies [5].

Conclusions
This study correlated academic performance measures with USMLE Step 2 CK scores to help predict students’ performance on USMLE Step 2 CK. In the future, adding in other measures such as age, financial need, and ethnicity could help improve the predictive model. For future iterations, this model could be more realistic if it did not include USMLE Step 1 scores, given the transition to pass/fail scoring system.

Additional Information
Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. UNLV Biomedical IRB issued approval 1020906-1. Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue. Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: Payment/services info: All authors have declared that no financial support was received from any organization for the submitted work. Financial relationships: All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. Other relationships: All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. National Resident Matching Program (NRMP): results of the 2020 National Program Director Survey. (2020). Accessed: June, 27, 2021; https://nk0rmp.reports@tmgkiminstadn.com/wp-content/uploads/2020/08/2020- PD-Survey.pdf.
2. Gauer JL, Jackson JB: The association between United States Medical Licensing Examination scores and clinical performance in medical students. Adv Med Educ Pract. 2019, 10:209-16. 10.2147/AMEP.S19201
3. Casey PM, Palmer BA, Thompson GB, et al.: Predictors of medical school clerkship performance: a multivariable longitudinal analysis of standardized examination scores and clinical assessments. BMC Med Educ. 2016, 16:128. 10.1186/s12909-016-0652-y
4. Guoit 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, 9:945-9. 10.2147/AMEP.S180786
5. Giordano C, Hutchinson D, Peppler R: A predictive model for USMLE Step 1 scores. Cureus. 2016, 8:e769. 10.7759/cureus.769
6. Basco WT Jr, Way DP, Gilbert GE, Hudson A: Undergraduate institutional MCAT scores as predictors of USMLE step 1 performance. Acad Med. 2002, 77:S13-6. 10.1097/00001888-200201000-00005
7. Kleshinski I, Khuder SA, Shapiro II, Gold JP: Impact of pre-admission variables on USMLE step 1 and step 2 performance. Adv Health Sci Educ Theory Pract. 2009, 14:69-78. 10.1007/s10459-007-9087-x
8. 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, 382:2389-92. 10.1056/NEJMp2006148
9. Huq S, Khalafallah AM, Botros D, Jimenez AE, Lam S, Huang J, Mukherjee D: Perceived impact of USMLE Step 1 pass/fail scoring change on neurosurgery: program director survey. J Neurosurg. 2020, 155:928-35. 10.3171/2020.4.INN2074
10. 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, 8:385-91. 10.2147/AMEP.S158557
11. Gohara S, Shapiro J, Jacob A, et al.: Joining the conversation: Predictors of success on the United States Medical Licensing Examinations (USMLE). Learn Assist Rev. 2011, 16:12-20.
12. Ghaffari-Rafi A, Lee RE, Fang R, Miles JD: Multivariable analysis of factors associated with USMLE scores across U.S. medical schools. BMC Med Educ. 2019, 19:154. 10.1186/s12909-019-1605-z
13. Julian ER: Validity of the Medical College Admission Test for predicting medical school performance. J Acad Med. 2005, 80:910-7. 10.1097/00001888-200510000-00010
14. Jones RF, Thomae-Forges M: Validity of the MCAT in predicting performance in the first two years of medical school. J Med Educ. 1984, 59:455-64. 10.1007/00001888-198406000-00010
15. Silver B, Hodgson CS: Evaluating GPAs and MCAT scores as predictors of NBME I and clerkship

2021 Kraczew et al. Cureus 13(9): e18143. DOI 10.7759/cureus.18143
performances based on students’ data from one undergraduate institution. Acad Med. 1997, 72:394-6. 10.1097/00001888-199705000-00022

16. Donnon T, Paolucci EO, Violato C: The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research. Acad Med. 2007, 82:160-6. 10.1097/01.ACM.0000249878.25186.b7

17. Zahn CM, Sagui A, Artino AR Jr, 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, 87:1348-54. 10.1097/ACM.0b013e31826a13bd

18. Veloski JJ, Callahan CA, Xu G, Hojat M, Nash DB: Prediction of students’ performances on licensing examinations using age, race, sex, undergraduate GPAs, and MCAT scores. Acad Med. 2000, 75:528-30. 10.1097/00001888-200010001-00009

19. 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, 97:1258-62.

20. Latessa R, Beaty N, Royal K, Colvin G, Pathman DE, Heck J: Academic outcomes of a community-based longitudinal integrated clerkships program. Med Teach. 2015, 37:862-7. 10.3109/0142159X.2015.1009020

21. Reteguiz JA, Crosson J: Clerkship order and performance on family medicine and internal medicine National Board of Medical Examiners Exams. Fam Med. 2002, 34:604-8.