ORIGINAL RESEARCH

Relationships of demographic variables to USMLE physician licensing exam scores: a statistical analysis on five years of medical student data

Jacqueline L Gauer[1]
J Brooks Jackson[2]

[1]Office of Medical Education, Medical School, University of Minnesota, Minneapolis, MN, USA; [2]Department of Laboratory Medicine and Pathology, Medical School, University of Minnesota, Minneapolis, MN, USA

Introduction: The purpose of this study was to determine the associations of the demographic variables of gender, state of legal residency, student age, and undergraduate major with scores on the Medical College Admissions Test (MCAT) and the United States Medical Licensing Exam (USMLE) Step 1 and Step 2 Clinical Knowledge.

Methods: The researchers collected and analyzed exam scores and demographic student data from participants of five graduating classes of students at the University of Minnesota Medical School (N = 1,067).

Results: Significant differences ($p < 0.05$) were found for traditional-aged (defined as < 25 years old at matriculation) versus nontraditional-aged students on USMLE Step 1 scores ($t[1065] = 2.91, p = 0.004$) and USMLE Step 2 scores ($t[1061] = 4.39, p < 0.001$), both in favor of traditional-aged students. Significant differences were found for males versus females on MCAT Composite scores ($t[1063] = 6.53, p < 0.001$) and USMLE Step 1 scores ($t[1065] = 5.14, p < 0.001$), both in favor of males. There were no significant differences between science and nonscience majors or between Minnesota legal residents and nonresidents.

Conclusion: Traditional age and male gender were associated with higher exam scores, although patterns differed between tests, whereas undergraduate major and state of legal residency were not associated with higher exam scores.

Keywords: licensing exams, demographics, gender, age, undergraduate major, state of legal residency

Plain language summary

The purpose of this study was to determine associations of demographic variables of gender, state of legal residency, student age, and undergraduate major with scores on the Medical College Admission Test (MCAT) and the United States Medical Licensing Exam (USMLE) Step 1 and Step 2 Clinical Knowledge (CK), in a sample of 1,067 graduates of the University of Minnesota Medical School. The authors found that male students scored higher than female students on the MCAT and USMLE Step 1, but not on USMLE Step 2 CK. Moreover, students younger than 25 scored higher than students aged 25 or older on USMLE Step 1 and Step 2 CK, but not on the MCAT. The authors found no significant differences between science and nonscience majors or between Minnesota legal residents and nonresidents. These patterns can help medical school administrators select high-performing students, and identify students who may be at risk of low exam performance, in order to provide targeted support.

Introduction

Medical school admissions committees are faced with difficult choices when determining which students to admit to medical school. For example, committees debate whether

Correspondence: J Brooks Jackson
University of Minnesota, MMC 293 C607,
420 Delaware Street, Minneapolis, MN
55455, USA
Tel +1 612 626 4949
Email jacksonb@umn.edu

students entering medical school at a "nontraditional" age are more likely to succeed, due to increased maturity and life experiences, or less likely, due to factors such as increased responsibilities at home. This study aims to explore several demographic variables debated by admissions committees, as they relate to one critical measure of medical school success: scores on physician licensing exams.

In the United States, the United States Medical Licensing Exam (USMLE) is generally required for physician licensure. The USMLE is a three-step examination sponsored by the Federation of State Medical Boards and the National Board of Medical Examiners. Together, Step 1 and Step 2 Clinical Knowledge (CK) assess a physician's ability to apply knowledge and concepts to provide safe and effective patient care.[1] Step 2 Clinical Skills (CS) scores, although also an important indicator, were not included in this analysis because they are reported as pass/fail, as opposed to numeric scores. Many factors affect the quality of care a physician provides, but because passing scores on the USMLE exams are generally required for physician licensure in the United States, they are a critical indicator of medical school success. Furthermore, residency directors typically consider USMLE scores in selection of applicants for their residency programs.[2]

Previous research has begun exploring the relationships of various demographic variables with performance on the USMLE. For example, one study found that, after controlling for pre-matriculation measures, men outperformed women slightly on Step 1, and that undergraduate science grade point averages were more associated with Step 1 performance for women than for men.[3] However, the same research team found that, in a very large sample of examinees, women outperformed men in most content areas on the USMLE Step 2 CK exam, and that the gender-related difference increased when controlling for Step 1 scores.[4,5] Moreover, previous research has shown an association between age and test scores on the USMLE Step 2 exam. One study examining the records of 171 medical students found that age was negatively correlated with USMLE Step 2 scores.[6] The same study found no differences in Step 2 scores based on gender or race. The authors suggested that their findings could inform the identification of medical students at risk of poor USMLE Step 2 performance. Another factor that medical school admissions committees debate is whether majoring in science or mathematics in undergraduate school affects eventual medical school success. Previous research has found little to no statistical effect of undergraduate major on medical school success (as measured by course grades,

USMLE scores, and residency program evaluations), with the caveat that preselection bias may have influenced the results.[7]

The medical school under examination in the current study is housed in a land-grant state university. Various stakeholders apply pressure to admit either legal residents of the state, as their tax dollars support the institution, or nonresidents, as they generally pay more in tuition. As we could not find previous research that has explored the effects of state of legal residency on medical school performance, this study examined whether state of legal residency was associated with USMLE scores, in order to help guide these admission decisions.

Some may argue that certain demographic groups may enter medical school with higher test scores or test-taking ability, thus affecting the groups' relationships to USMLE scores. Previous research has shown that high scores on the Medical College Admissions Test (MCAT), widely used for selection and screening of medical applicants in the United States and Canada, are predictive of high scores on licensing exams.[8–11] Therefore, it was important for this study to explore, furthermore, whether differences in USMLE scores could be explained by differences in scores on the MCAT, and to look for patterns in test scores over time.

The primary objective of this study was to determine the association of various demographic variables such as gender, traditional or nontraditional student age, Minnesota residency status, and science or nonscience undergraduate major with USMLE Step 1 and Step 2 CK scores. These variables have been prominently debated within the admissions committee at our institution, as well as nationally. This study attempted to replicate previous findings for certain demographic variables (such as gender), introduce new variables that have not been explored previously in the literature (such as state of legal residency), and look for patterns in tests across time to improve our understanding of the relationships of these demographic variables to USMLE scores. We hope that these findings can both help guide admission decisions, and also help identify students who may be at risk of lower USMLE scores and would benefit from targeted exam preparation support.

Methods
Institutional approval
Ethical approval for this research was granted by the Institutional Review Board at the University of Minnesota on March 19, 2015 (reference number: 1503E66021). As the study was completed on deidentified institutional data, the requirement for consent to participate was waived by the board.

Participants

Participants of this study included all students pursuing MD (N = 1,062) and MD/PhD (N = 5) degrees at the University of Minnesota who both matriculated between 2007 and 2011 and graduated between 2011 and 2015 (N = 1,067). Of the included students, 546 (51.2%) were male and 521 (48.8%) female. Age at matriculation ranged from 19 to 42 years (mean [M] = 23.7 years, standard deviation [SD] = 2.6 years). At the University of Minnesota, medical students matriculate at either the Twin Cities or Duluth campus. They complete the first two years of the degree (basic science courses) at their campus of matriculation and, then, all students complete the second two years of the degree (clinical clerkships) through the Twin Cities campus. Of the students in this study, 282 (26.4%) matriculated at the Duluth campus and 785 (73.6%) matriculated at the Twin Cities campus. Two students did not have MCAT scores and four students did not have Step 2 CK scores in the record, and their data are not included in results involving scores from those exams. Seventeen students did not have a state of legal residency status in the record, and their data are not included in results for analyses including that variable.

Sources of data

Academic and demographic data were collected from student records held by the Office of Medical Education in the Academic Health Center at the University of Minnesota. Demographic data including date of matriculation and graduation date were retrieved from two complementary University and Medical School tracking databases. Data pertaining to participants' undergraduate majors, age at matriculation, gender, and state of legal residency were retrieved from the University of Minnesota's access to primary application data through the American Medical College Application Service (AMCAS). For the purpose of this study, students were categorized as "nontraditional-aged" if they were 25 years of age or older at the time of matriculation. Data on participants' USMLE Step 1 and Step 2 CK scores were provided to the Medical School by the National Board of Medical Examiners upon students' completion of each exam.

Analyses

Using SPSS Statistics v.22 (IBM: Armonk, NY, USA), we conducted independent-samples t-tests for undergraduate science vs nonscience majors, traditional-aged vs nontraditional-aged students, gender, and Minnesota legal residency status to explore group differences in scores on the MCAT, USMLE Step 1, and USMLE Step 2 CK exams. We calculated effect sizes (Cohen's d) for each comparison using the online calculator available on the Social Science Statistics website (www.socscistatistics.com/effectsize. Accessed November 21, 2017).

Furthermore, using SPSS Statistics v.22 (IBM: Armonk, NY, USA), we calculated ordinary least squares (OLS) regression coefficients and confidence intervals (CIs) for each demographic factor, for both USMLE Step 1 and Step 2 CK scores, both with and without the MCAT score as a covariate.

Results

Means and SDs for each of the demographic comparison groups can be found in Table 1. Initial inspection of q-q plots of MCAT, USMLE Step 1, and USMLE Step 2 scores showed no noticeable concerns with regard to the normality of the data.

The results of our t-test analyses are shown in Table 2. We found significant differences at the $p < 0.05$ (2-tailed) level for two of the compared categories. Significant differences were found for traditional-aged vs nontraditional-aged students on USMLE Step 1 scores [$t(1065) = 2.91$, $p = 0.004$] and USMLE Step 2 scores [$t(1061) = 4.39$, $p < 0.001$], both in favor of traditional-aged students. Significant differences were found for males vs females on MCAT Composite scores [$t(1063) = 6.53$, $p < 0.001$] and USMLE Step 1 scores [$t(1065) = 5.14$, $p < 0.001$], both in favor of males. There were no significant differences between science and nonscience majors or between Minnesota legal residents and nonresidents. Cohen's d (a measure of effect size) ranged from 0.00 to 0.40 for these comparisons, indicating negligible to moderately small effects.

The pattern that male students scored significantly higher on both the MCAT and Step 1 exams indicates the possibility that the gender difference in USMLE Step 1 scores could be explained by the gender difference in MCAT scores. We explored that possibility using multiple linear regression and found that, even after correcting for MCAT score, gender was associated with Step 1 score in the regression model, reducing the mean squared error of the model from 60,432.26 to 31,413.58.

Table 3 shows OLS multiple linear regression coefficients and 95% CIs for the variables under analysis. Reflecting the findings of the t-tests, age was significantly associated ($p < 0.01$) with both Step 1 and Step 2 CK scores, with and without MCAT scores as a covariate, and gender was significantly associated ($p < 0.01$) with Step 1 (but not Step 2 CK scores), both with and without MCAT scores as a covariate. Undergraduate major and state of legal residency were not significantly associated with USMLE scores in the regression analysis.

Table 1 Means and SDs of demographic comparison groups for MCAT, USMLE Step 1, and USMLE Step 2 CK scores, for students pursuing MD (N = 1,062) and MD/PhD (N = 5) degrees at the University of Minnesota who matriculated between 2007 and 2011 and graduated between 2011 and 2015

Demographic comparison	Group	N	MCAT		Step 1		Step 2 CK	
			Mean	SD	Mean	SD	Mean	SD
Undergraduate major	Science	850	31.51	3.40	226.89	19.09	238.33	16.52
	Nonscience	217	31.34	3.35	224.92	18.60	237.16	16.43
Age[a]	Traditional	831	31.53	3.37	227.39	19.12	239.26	16.38
	Nontraditional	236	31.29	3.45	223.33	18.29	233.95	16.30
Gender	Male	546	32.13	3.41	229.38	19.27	238.05	17.20
	Female	521	30.80	3.23	223.46	18.25	238.13	15.75
State of legal residency[b]	MN	898	31.48	3.31	226.68	19.25	238.28	16.53
	Not MN	152	31.56	3.80	224.99	17.84	237.53	16.33
	Total	1067	31.48	3.39	226.49	19.00	238.09	16.50

Notes: [a]"Traditional-aged" was defined as <25 years old and "Nontraditional-aged" as ≥25 years old. [b]Seventeen participants did not have a state of legal residency in the record. Their data are not included in analyses using that variable.
Abbreviations: MCAT, Medical College Admissions Test; MN, Minnesota; SD, standard deviation; CK, Clinical Knowledge; USMLE, United States Medical Licensing Exam.

Table 2 Results of independent-samples *t*-tests comparing various demographic factors on MCAT and USMLE scores, for students pursuing MD (N = 1,062) and MD/PhD (N = 5) degrees at the University of Minnesota who matriculated between 2007 and 2011 and graduated between 2011 and 2015

Comparison	Exam	t	p (two-tailed)	Cohen's d
Science vs nonscience major	MCAT	0.68	0.50	0.05
	Step 1	1.37	0.17	0.10
	Step 2 CK	0.93	0.35	0.07
Traditional vs nontraditional age[a]	MCAT	0.96	0.35	0.07
	Step 1	2.91	0.004**	0.22
	Step 2 CK	4.39	<0.001**	0.32
Male vs female	MCAT	6.53	<0.001**	0.40
	Step 1	5.15	<0.001**	0.32
	Step 2 CK	−0.08	0.94	0.00
MN resident vs nonresident[b]	MCAT	−0.23	0.82	0.02
	Step 1	1.01	0.31	0.09
	Step 2 CK	0.52	0.61	0.05

Notes: **$p < 0.01$. [a]"Traditional-aged" was defined as <25 years old and "Nontraditional-aged" as ≥25 years old. [b]Seventeen participants did not have a state of legal residency in the record. Their data are not included in analyses using that variable.
Abbreviations: MCAT, Medical College Admissions Test; MN, Minnesota; CK, Clinical Knowledge; USMLE, United States Medical Licensing Exam.

Discussion

Overall, the results of this study show that several demographic variables may be associated with performance on the USMLE Step 1 and USMLE Step 2 CK exams.

In comparisons between groups, traditional-aged vs nontraditional-aged and male vs female comparisons showed significant differences for some exams. The patterns for each comparison were different. Nontraditional-aged students scored significantly lower than traditional-aged students on Step 1 and Step 2 CK but not on MCAT, indicating a possible decline in test performance, relative to the traditional-aged students, as they progressed from MCAT (and admission) through Step 1 and, then, Step 2 CK. Reasons for this decline are not clear, but might be related to an increased likelihood for older students to experience external life events and responsibilities (children, ill parents, etc.) that could decrease time and concentrated effort available for exam preparation.

In a converse pattern to that of the age results, male students scored significantly higher than female students on the MCAT and Step 1 exams, but not on the Step 2 CK exam, indicating that performance between genders evened out over time. The Step 1 and Step 2 CK gender results are consistent with previously reported data.[2] The gender data reveals an interesting pattern, wherein the male students in this sample entered medical school with higher test scores, but the female students caught up to their male counterparts over time. The reason for this pattern is unclear, but a multiple linear regression model indicates that gender is associated with Step 1 score even after correcting for MCAT score.

No significant differences were found in this dataset for comparisons of science vs nonscience majors and Minnesota residents vs nonresidents.

Results of the OLS regression analysis support those from the *t*-tests. When combined into models including all predictor variables included in this study, age was significantly associated with both USMLE Step 1 and Step 2 CK scores, gender was significantly associated with Step 1 scores, and undergraduate major and state of legal residency were not significantly associated with either. These patterns held even when controlling for MCAT score by introducing it to the model as a covariate.

Table 3 OLS regression coefficients[a] and 95% CIs for demographic comparison variables, for models predicting USMLE Step 1 and USMLE Step 2 CK scores, with and without MCAT scores as a covariate, for students pursuing MD (N = 1,062) and MD/PhD (N = 5) degrees at the University of Minnesota who matriculated between 2007 and 2011 and graduated between 2011 and 2015

Demographic variable	Without MCAT covariate		With MCAT covariate	
	Step 1	Step 2	Step 1	Step 2
Undergraduate science major	0.015 (−2.15, 3.54)	0.007 (−2.20, 2.79)	0.013 (−2.05, 3.25)	0.005 (−2.17, 2.55)
Nontraditional age[b]	−0.105** (−7.58, −2.03)	−0.134*** (−7.74, −2.88)	−0.085** (−6.46, −1.30)	−0.115*** (−6.87, −2.26)
Gender	0.170*** (4.20, 8.78)	0.016 (−1.49, 2.53)	0.096** (1.507, 5.85)	0.051 (−3.63, 0.25)
MN legal resident[c]	0.019 (−2.22, 4.25)	0.007 (−2.51, 3.17)	0.026 (−1.60, 4.43)	0.010 (−2.20, 3.18)

Notes: **$p < 0.01$, ***$p < 0.001$. [a]Values shown are standardized beta coefficients (95% CI). [b]"Traditional-aged" was defined as <25 years old and "Nontraditional-aged" as ≥25 years old. [c]Seventeen participants did not have a state of legal residency in the record. Their data are not included in analyses using that variable.
Abbreviations: CI, confidence interval; MCAT, Medical College Admissions Test; MN, Minnesota; OLS, ordinary least squares; CK, Clinical Knowledge; USMLE, United States Medical Licensing Exam.

When considering the implications of these comparisons, it is important to also consider possible selection effects due to the admissions process. The dataset analyzed here includes only students who had already achieved admission to the University of Minnesota medical school; therefore, these same variables may have different predictive values for those in the overall applicant pool.

As with any study of this type, we must be careful of the claims we make from the findings. The nature of the observational data collection method is that we cannot draw causal conclusions about our findings, only make claims of associations. Furthermore, it is highly likely that the variables studied interact with each other and with many other variables, including simple error, to affect outcomes. Across all analyses included in this study, it is important to note that although many of the results were statistically significant, we must also consider the practical significance of, for example, the 4.06 point difference in mean Step 1 scores between traditional- and nontraditional-aged students. Even for the variables showing significant *t*-test results, the maximum effect size (as measured by Cohen's *d*) was 0.40, indicating a moderately small effect. It is clear that the variables studied here are far from perfectly predictive of future exam scores. The Association of American Medical Colleges advocates for a holistic approach to reviewing applications to medical school; therefore, it is important to consider these findings in that context.[12] As this analysis includes data on only one specific institution, readers must consider the context of their own institution when determining how to apply these results.

Additionally, it is important to note that the findings here might not necessarily reflect differences in abilities between groups, but instead reflect biases in the designs of the exams or of medical education itself. Future research could use the findings from this study to explore, for example, a possible gender bias in the USMLE Step 1 exam, particularly by determining why the significant gender differences seen in Step 1 scores are not present in Step 2 CK scores.

Conclusion

In conclusion, this study provides an analysis of the relationships between several demographic variables and licensing exam scores. Results of our analyses indicate that there are definite patterns that can be found in the factors that predict student performance on the USMLE Step 1 and Step 2 CK exams, especially when it comes to gender and age. The precise nature of the relationship of these factors remains unclear, prompting further study.

Acknowledgments

Author contributions

JLG is responsible for data acquisition and analysis, interpretation of findings, and the preparation of this manuscript. JBJ is responsible for the conceptualization and oversight of this study and revision of this manuscript. Both authors contributed toward data analysis, drafting, and revising the paper, and agree to be accountable for all aspects of the work, and reviewed and approved the final manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. National Board of Medical Examiners: United States Medical Licensing Exam Website. Available from: http://www.usmle.org. Accessed August 10, 2015.
2. Berner ES, Brooks CM, Erdmann JB. Use of the USMLE to select residents. *Acad Med*. 1993;68(10):753–759.
3. Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of examinee gender and USMLE Step 1 performance. *Acad Med*. 2008;83(10 Suppl):S58–S62.
4. Cuddy MM, Swanson DB, Dillon GF, Holtman MC, Clauser BE. A multilevel analysis of the relationships between selected examinee characteristics and United States Medical Licensing Examination Step 2 clinical knowledge performance: revisiting old findings and asking new questions. *Acad Med*. 2006;81(Suppl 10):S103–S107.
5. Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of the relationships between examinee gender and United States Medical Licensing Exam (USMLE) Step 2 CK content area performance. *Acad Med*. 2007;82(Suppl 10):S89–S93.
6. 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 Nat Med Assoc*. 2005;97(9):1258–1262.
7. Smith SR. Effect of undergraduate college major on performance in medical school. *Acad med*. 1998;73(9):1006–1008.
8. Callahan CA, Hojat M, Veloski J, Erdmann JB, Gonnella JS. The predictive validity of three versions of the MCAT in relation to performance in medical school, residency, and licensing examinations: a longitudinal study of 36 classes of Jefferson Medical College. *Acad Med*. 2010;85(6):980–987.
9. 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(1):100–106.
10. Gauer JL, Wolff JM, Jackson JB. Do MCAT scores predict USMLE scores? An analysis on 5 years of medical student data. *Med Educ Online*. 2016;21:31795.
11. Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med*. 2005;80(10):910–917.
12. Cohen JJ. Will changes in the MCAT and USMLE ensure that future physicians have what it takes? *JAMA*. 2013;310(21):2253–2254.