Association Between Performance on COMLEX-USA and the American College of Osteopathic Family Physicians In-Service Examination

Kimberly M. Hudson, MEd
Gail Feinberg, DO, MEd
Laura Hempstead, DO
Christopher Zipp, DO
John R. Gimpel, DO, MEd
Yi Wang, MS

ABSTRACT

Background The primary goal of residency programs is to select and educate qualified candidates to become competent physicians. Program directors often use performance on licensure examinations to evaluate the ability of candidates during the resident application process. The American College of Osteopathic Family Physicians (ACOFP) administers an in-service examination (ISE) to residents annually. There are few prior studies of the relationship between the Comprehensive Osteopathic Medical Licensing Examination of the United States of America (COMLEX-USA) series and formative assessments of residents in training.

Objective We explored the relationship between performance on COMLEX-USA and the ACOFP in-service examination to offer support on the use of licensing examinations in resident selection.

Methods In 2016, performance data from the COMLEX-USA and the ISE were matched for 3 resident cohorts (2011–2013, inclusive; N = 1384). Correlations were calculated to examine the relationship between COMLEX-USA and ISE scores. Multiple linear regression models were used to determine if performance on COMLEX-USA significantly predicted third-year ISE (ISE-3) scores.

Results Findings indicated that correlations among performance on COMLEX-USA and ISE were statistically significant (all $P < .001$), and there was strong intercorrelation between COMLEX-USA Level 3 and ISE-1 performance ($r = 0.57$, $P < .001$). Performance on the COMLEX-USA Levels 1 and 2–Cognitive Examination significantly predicted performance on the ISE-3 ($F(2,1381) = 228.8$, $P < .001$).

Conclusions The results support using COMLEX-USA as a part of resident selection in family medicine. Additionally, program directors may use performance on COMLEX-USA to predict success on the ISE-3.

Introduction

The Comprehensive Osteopathic Medical Licensing Examination of the United States of America (COMLEX-USA) is the national series of licensure assessments for osteopathic physicians. The COMLEX-USA series assesses osteopathic medical knowledge and clinical skills for the practice of osteopathic medicine.[1] It consists of 4 examinations in the COMLEX-USA series: Level 1, Level 2–Cognitive Evaluation (CE), Level 2–Performance Evaluation (PE), and Level 3. Levels 1, 2-CE, and 3 are predominantly multiple-choice examinations, and Level 2-PE is a standardized patient-based examination.[1] All levels follow the same blueprint: they are designed to measure competency across the same 9 domains of patient presentations and 6 domains of physician tasks.[1] Levels 1, 2-CE, and 2-PE scores are widely used by osteopathic medical schools for graduation purposes, and by programs for resident selection, as scores provide reliable evidence of residents' medical knowledge and application. Numerical scores on Level 2-PE are not reported; therefore, the predictive validity of these scores is not applicable to the current study.

Consistent with the mission of the American College of Osteopathic Family Physicians (ACOFP) "to promote excellence in osteopathic family medicine through quality education, visionary leadership, and responsible advocacy,"[2] it partners with the NBOME to offer an annual in-service examination (ISE) to evaluate the progress of each resident and provide feedback for continuous professional development. The NBOME develops, processes, and scores the ISE as a service to ACOFP. Scores on the ISE are used by residency program faculty to determine areas of academic strengths and deficiencies for individual residents and to plot a curriculum to ensure each resident meets the minimum osteopathic medical knowledge benchmarks and milestones in the program.

In evaluating the relationship between performance on COMLEX-USA and ISE, the timing of both examinations is a factor. Medical students take levels

DOI: http://dx.doi.org/10.4300/JGME-D-17-00997.1

1, 2-CE, and 2-PE during their second, third, and fourth years of osteopathic medical school. COMLEX-USA Level 3 is typically administered at the end of the first osteopathic residency year (OGME-1) or the beginning of the second year (OGME-2). It is required for progression to the third year of residency (OGME-3). The FIGURE displays the sequence of COMLEX-USA and ISE administration.

Following best practices in assessment, it is important to evaluate and provide empirical evidence supporting the relationship between performance on examinations that intend to measure similar constructs.[3] A meta-analysis conducted on the relationship between performance on the United States Medical Licensing Examination (USMLE) and ISE showed there was a strong positive correlation.[4] Findings from research on the association between performance on the USMLE and the internal medicine ISE also have shown significant, positive correlations.[5] Yet, few studies have investigated the relationship between performance on the COMLEX-USA and specialty ISEs. One study on the relationship in performance on the COMLEX-USA, American Board of Family Medicine (ABFM) ISE, and ABFM certification examination showed positive, significant associations between performance on COMLEX-USA levels 1, 2-CE, and 3; ABFP ISE; and ABFP certification examination.[6] Our study builds on existing research by examining the association between COMLEX-USA scores and the ISEs administered in family medicine osteopathic residency programs. We sought to evaluate the predictive validity of COMLEX-USA scores on performance on ACOFP ISE, focusing on the following research questions:

What was known and gap
Program directors have used licensure examinations for resident selection. More research is needed on whether these tests are a valid metric in a holistic selection process.

What is new
A study assessed the correlations between the osteopathic licensing examination and the osteopathic in-service examination.

Limitations
Single specialty study; relationship between examination performance and performance in practice needs further study.

Bottom line
There is a relationship between the examinations, and the osteopathic licensure examination is predictive of performance on the final practice-focused segment of the in-service examination.

1. What is the relationship between performance on COMLEX-USA levels 1, 2-CE, and 3 and performance on the ISE-1, -2, and -3?
2. How well does performance on COMLEX-USA levels 1 and 2-CE predict ISE-3 performance?

Methods

Participants

In 2016, we compiled scores from the ACOFP ISE for the 2011, 2012, and 2013 entering resident cohorts. The completion years of these cohorts were 2014, 2015, and 2016, respectively. We were able to obtain ACOFP data for 1720 residents who were administered the ISE each program year (OGME-1, OGME-2, and OGME-3) between 2011 and 2015.

Levels 1, 2-CE, and 3 COMLEX-USA data between 1990 and 2015 were used to create a longitudinal

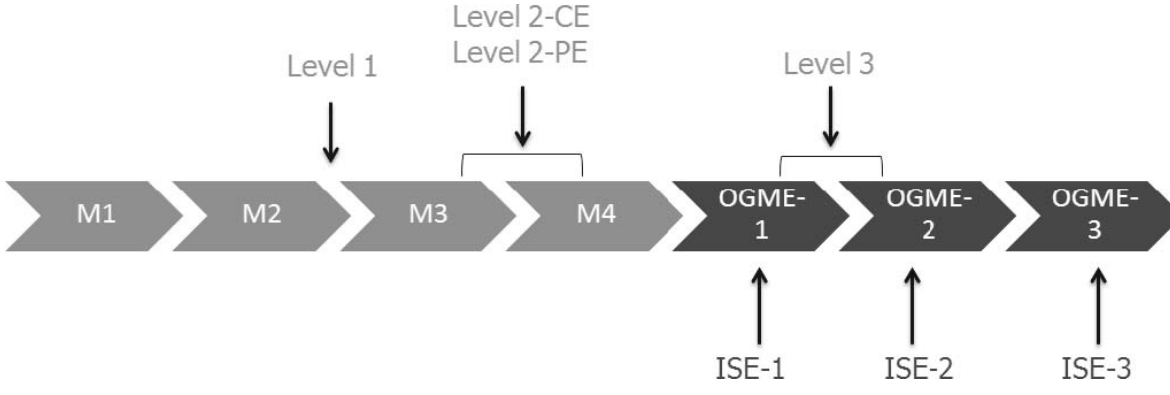

FIGURE
Sequence of Examination Administration of COMLEX-USA Series and ACOFP In-Service Examination[a]

[a] M1 denotes the first year in osteopathic medical school, and OGME-1 denotes the first year in an ACOFP-affiliated residency program. ISE-1 represents the examination administered during the first year of residency.

Abbreviations: COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination of the United States of America; ACOFP, American College of Osteopathic Family Physicians; ISE, in-service examination.

TABLE 1
COMLEX-USA and ISE Mean (Standard Deviation) by Cohort

	2011	2012	2013	Total
Count	392	487	505	1384
Level 1	465.70 (66.31)	455.35 (70.99)	464.13 (76.03)	461.49 (71.7)
Level 2-CE	458.65 (81.52)	460.07 (82.38)	463.17 (84.65)	460.80 (82.94)
Level 3	510.88 (106.24)	524.82 (105.31)	515.06 (114.46)	517.31 (109.07)
ISE-1	467.45 (90.44)	468.78 (91.07)	465.76 (90.56)	467.30 (90.65)
ISE-2	522.66 (85.38)	523.14 (85.46)	513.89 (89.98)	519.63 (87.16)
ISE-3	551.88 (88.03)	552.75 (90.53)	543.35 (88.53)	549.07 (89.14)

Abbreviations: COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination of the United States of America; ISE, in-service examination; CE, cognitive examination.

review of resident performance. COMLEX-USA data were matched to ACOFP data by residents' first and last names. Residents were then matched based on their first year in an ACOFP-affiliated residency program and their osteopathic medical school graduation year. Residents with 1 or more gap years between graduation and residency were excluded from this study due to the matching criteria.

Outcomes

COMLEX-USA levels 1, 2-CE, and 3 include 350 to 400 multiple-choice questions. Cronbach's alpha measure for reliability (internal consistency) is approximately 0.90 for each of the 3 cognitive examinations in the COMLEX-USA series.[7] Although there are small differences in the number of scored items across years, the ACOFP ISE contains approximately 200 scored multiple-choice questions. The reliability over the indicated time frame ranged from 0.81 to 0.86.[8–10] The standard scores for all examinations were used for analyses.

The study was declared exempt by the Institutional Review Board of the Center for the Advancement of Healthcare Education and Delivery.

Analyses

Descriptive statistics, including frequencies, means, and standard deviations, were used to describe the sample with respect to performance on COMLEX-USA and ISE. One-way analysis of variance (ANOVA) was used to identify statistically significant differences among the means of the 3 cohorts' COMLEX and ISE scores. The Pearson product-moment correlation was used to measure the strength of the linear relationship among performance on levels 1, 2-CE, 3, and ISE-1, -2, and -3. We used Cohen's effect size to interpret effect sizes to measure strength of associations (eg, 0.10 for small effect, 0.3 for medium effect, and 0.5 for large effect),[11] and applied the Bonferroni correction for multiple comparisons.[12] Predictive validity evidence was evaluated using a multiple linear regression model of scores on levels 1 and 2-CE onto ISE-3. Levels 1 and 2-CE are often used in residency selection.[13] Therefore, we used residents' ISE-3 scores as a proxy for successful completion of the residency program. All analyses were done in R version 2.1 (R Foundation for Statistical Computing, Vienna, Austria). We employed a significance level of .05, unless otherwise specified.

TABLE 2
Results From Analysis of Variance

Examination	DF	F	Sig
Level 1	(1, 1384)	2.2592	.1330
Level 2-CE	(1, 1384)	2.4785	.1156
Level 3	(1, 1384)	0.0958	.7570
ISE-1	(1, 1384)	0.0221	.8818
ISE-2	(1, 1384)	0.6792	.4100
ISE-3	(1, 1384)	0.1819	.6698

Abbreviations: CE, cognitive examination; ISE, in-service examination.

Results

Of the original 1720 residents in the sample, 1384 (80%) residents were successfully matched with COMLEX-USA performance across the 3 cohorts. Six residents (0.3%) were excluded from the analysis due to duplicated records of name and graduation year. One hundred thirty-nine residents who did not begin residency immediately following graduation were excluded from the analysis. Results from independent sample *t* tests indicated these residents scored significantly different on the COMLEX-USA and ISE ($P < .001$; see online supplemental material).

The match rate by cohort ranged from 76% to 84%. TABLE 1 displays the mean and standard deviations of COMLEX-USA standard scores and ISE standard scores by cohort. One-way ANOVA results indicated no significant differences among COMLEX-USA or ISE scores across the 3 cohorts ($P > .05$; see TABLE 2), which provide rationale for analyzing all cohorts together.

TABLE 3
Pearson Correlation Coefficients of COMLEX-USA and ISE Scores

Examination	Level 1	Level 2-CE	Level 3	ISE-1	ISE-2	ISE-3
Level 1	1.00					
Level 2-CE	0.64[a]	1.00				
Level 3	0.59[a]	0.62[a]	1.00			
ISE-1	0.45[a]	0.46[a]	0.57[a]	1.00		
ISE-2	0.45[a]	0.48[a]	0.53[a]	0.65[a]	1.00	
ISE-3	0.42[a]	0.47[a]	0.52[a]	0.59[a]	0.70[a]	1.00

[a] Denotes significant correlation coefficient ($P < .001$).
Abbreviations: COMLEX-USA, Comprehensive Osteopathic Medical Licensing Examination of the United States of America; ISE, in-service examination; CE, cognitive examination.

Association of Performance on COMLEX-USA and ISE

The results for our examination of the relationship among COMLEX-USA scores and ISE scores indicated a significant correlation among performance on COMLEX-USA levels 1, 2-CE, and 3 and performance on all levels of the ISE (TABLE 3).

The strongest correlations were found within each examination group. For example, the strongest associations are between performance on the ISE-2 and ISE-3 ($r = 0.70$, $P < .001$) and ISE-1 and ISE-2 ($r = 0.65$, $P < .001$). Similarly, there was a significant strong correlation between COMLEX-USA Level 1 and Level 2-CE ($r = 0.64$, $P < .001$). Correlations greater than 0.5 indicate a large effect.[10]

Across examinations, the strongest intercorrelation was found between performance on ISE-1 and COMLEX-USA Level 3 ($r = 0.57$, $P < .001$). This is likely due in part to the similar timing of these 2 examinations, as more than 95% (1326 of 1384) of residents were administered COMLEX-USA Level 3 after the ISE-1. The strong intercorrelations may also be related to the similarity in test blueprints for the 2 examinations.

The intercorrelations indicated that performance on levels 1, 2-CE, and 3 accounted for 18%, 22%, and 27% of the variance on ISE-3 scores, respectively. In addition, first- and second-year performance on ISE accounted for 35% and 49% of the variance in ISE-3 scores, respectively. While COMLEX-USA Level 1 and ISE-3 have the weakest relationship, that relationship is significant ($r = 0.42$, $P < .001$).

TABLE 4
Summary of Multiple Regression Analysis for Performance on ISE-3 (N = 1384)

Variable	β	SE (β)	*t*	*P* Value
Intercept	262.409	14.08	18.631	< .001
Level 1	0.252	0.43	6.665	< .001
Level 2-CE	0.370	0.03	11.332	< .001

Abbreviations: ISE, in-service examination; CE, cognitive examination.

Predictive Validity of COMLEX-USA on ISE-3 Scores

Results from a multiple linear regression indicated that performance on COMLEX-USA levels 1 and 2-CE significantly predicted ISE-3 scores ($F(2,1381) = 228.8$, $P < .001$). Performance on levels 1 and 2-CE accounted for 25% of the variance in ISE-3 scores. TABLE 4 displays the results of the multiple regression analysis, including the coefficient (weight) of each variable, the standard error, and the significance associated with the t-statistic.

Discussion

Our findings support the utility of COMLEX-USA scores as a part of the process in selecting qualified residents for matriculation into postgraduate residency programs. The use of ISE-3 as a proxy for successful completion of residency programs was based on prior work showing that residents who earn high scores on ISE-3 are considered to have the requisite knowledge and skills to be osteopathic family physicians. Examining thresholds of resident knowledge is a plausible topic for future research.

The significant correlations among ISE scores were expected due to the consistent test blueprint. Additionally, results showed a significant strong correlation between ISE-1 and Level 3 scores, which may partially be explained by closeness in timing between the examinations. The significant correlations provide evidence of the congruent validity of the COMLEX-USA series and the ISE. It is important to review associations between examinations that purport to measure similar constructs as this evidence supports the congruent validity of both assessment programs.

Limitations of our study include that it was restricted to family medicine residents who were administered the examination for 3 consecutive years from 2011 through 2015. Research with different groups of residents, date ranges, and specialties is warranted to validate the consistency of results. Study of residents who did not begin their residency

program immediately after graduation and residents who did not take the ISE for 3 consecutive years also is of interest. Finally, evaluating the relationship between performance on COMLEX-USA, ISE, and board certification and recertification examinations as well as quality indicators in practice is also warranted.

The results from the multiple linear regression support the predictive validity of COMLEX-USA on successful completion of ISE-3 (measured by ISE-3 performance). If program directors view scores on ISE-3 as a proxy for successful completion in the residency program, the findings of this research support the utility of COMLEX-USA levels 1 and 2-CE scores as a key parameter for selecting qualified residents for osteopathic family medicine residency programs. Despite the evidence to support the secondary use of scores on licensure examinations for predicting successful performance in residency programs, the National Board of Osteopathic Medical Examiners advocates for holistic review of residency applicant qualifications.

Conclusion

The evidence presented in this study supports the concurrent and predictive validity of performance on COMLEX-USA and performance on the ACOFP ISE. Performance on these 2 examination series is significantly correlated, and can be used to glean information about applicants to osteopathic residency programs. This study contributes to the evidence that supports the use of COMLEX-USA as part of holistic application review for residents entering into osteopathic family medicine residency programs.

References

1. National Board of Osteopathic Medical Examiners Inc. COMLEX-USA Bulletin of Information. https://www.nbome.org/exams-assessments/comlex-usa/new-bulletin. Accessed September 12, 2018.
2. American College of Osteopathic Family Physicians. 2016–2018 Strategic Plan of the American College of Family Physicians. https://www.acofp.org/ACOFPIMIS/acofporg/PDFs/about/strategic_plan_2016.pdf. Accessed September 12, 2018.
3. American Educational Research Association; American Psychological Association; National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.
4. Kenny S, McInnes M, Singh V. Associations between residency selection strategies and doctor performance: a meta-analysis. *Med Educ*. 2013;47(8):790–800.
5. McDonald F, Zeger S, Kolars J. Associations between United States Medical Licensing Examination (USMLE) and Internal Medicine In-Training Examination (IM-ITE) scores. *J Gen Intern Med*. 2008;23(7):1016–1019.
6. O'Neill T, Peabody M, Song H. The predictive validity of the National Board of Osteopathic Medical Examiners' COMLEX-USA Examinations with regard to outcomes on American Board of Family Medicine Examinations. *Acad Med*. 2016;91(11):1568–1575.
7. Cronbach LJ. *Essentials of Psychological Testing*. Harper and Row; 1970.
8. American College of Osteopathic Family Physicians. Resident in-service examination summary report. Chicago, IL: ACOFP; 2011.
9. American College of Osteopathic Family Physicians. Resident in-service examination summary report. Chicago, IL: ACOFP; 2012.
10. American College of Osteopathic Family Physicians. Resident in-service examination summary report. Chicago, IL: ACOFP; 2013.
11. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
12. Keppel G, Wickens TD. *Design and Analysis: A Researcher's Handbook*. 4th ed. Upper Saddle River, NJ: Pearson Prentice Hall; 2004.
13. Sandella JM, Gimpel JR, Smith LL, et al. The use of COMLEX-USA and USMLE for residency applicant selection. *J Grad Med Educ*. 2016;8(3):358–363.

Kimberly M. Hudson, MEd, is a Research Scientist, National Board of Osteopathic Medical Examiners (NBOME); **Gail Feinberg, DO, MEd,** is Professor and Chair, Primary Care, Touro University of California College of Osteopathic Medicine; **Laura Hempstead, DO,** is Accreditation Field Staff, Accreditation Council for Graduate Medical Education; **Christopher Zipp, DO,** is Chair, Department of Family Medicine, Morristown Medical Center; **John R. Gimpel, DO, MEd,** is President, NBOME; and **Yi Wang, MS,** is Principal Research Associate, NBOME.

Funding: The authors report no external funding for this study.

Conflict of interest: The authors declare they have no competing interests.

The authors would like to thank NBOME colleagues, Edward Tsai, PhD, and Sandra Waters, MEM, and American College of Osteopathic Family Physicians colleagues, Bina Mehta and Judy Schwellenbach, for their support in this research study.

Corresponding author: Kimberly M. Hudson, MEd, National Board of Osteopathic Medical Examiners, 8765 West Higgins Road, Suite 200, Chicago, IL 60631, 773.714.0622 ext 266, khudson@nbome.org

Received December 7, 2017; revisions received March 23, 2018, and May 21, 2018; accepted June 14, 2018.