Two hundred sixty-five students from four classes at one school of osteopathic medicine were studied to determine the correlation between several frequently used premedical admission criteria to predict performance on the early and current versions of the osteopathic medical licensure examinations. Further analysis evaluated the predictive value of the examination of the National Board of Osteopathic Medical Examiners (NBOME) and its successor, the Comprehensive Osteopathic Medical Licensing Examination–USA (COMLEX–USA) in determining subsequent performance in the ascending levels of the examination sequence, as well as to predict performance in the courses and rotations of osteopathic principles and practice (OPP) and osteopathic manipulative medicine (OMM).

This study also investigated the relationship between the early and later parts of the osteopathic medical licensure examinations. Analysis showed that of all premedical parameters, only the Medical College Aptitude Test score was positively correlated with performance on COMLEX–USA. In addition, a positive correlation was demonstrated between the performance in either Level 1 of COMLEX–USA or in the earlier form of the NBOME examination Part I and subsequent performance on the current Level 2 of COMLEX–USA or the old NBOME Part II. Finally, analysis indicated that COMLEX–USA predicted performance in OPP knowledge and skills.

(Key words: osteopathic medical education, Comprehensive Medical Licensing Examination–USA [COMLEX–USA], Medical College Aptitude Test [MCAT], examination of the National Board of Osteopathic Medical Examiners [NBOME], medical licensure and board certification, validity)

Many studies have been undertaken to evaluate the predictive value of undergraduate measures in determining scholastic performance in allopathic medical schools and performance on allopathic medical licensure examinations. Studies on similar relationships between osteopathic medical licensure examinations and student academic characteristics are rare.

Osteopathic principles and practice (OPP) and osteopathic manipulative treatment (OMT) are unique to osteopathic medical practice. The Comprehensive Osteopathic Medical Licensing Examination–USA (COMLEX–USA) is the only national standardized examination that tests for OPP and OMM. Therefore, it would be expected that osteopathic medical licensure examinations would predict future performance in OPP and OMT knowledge and skills. However, no previous studies have demonstrated the predictive value of scores on medical licensure examinations on performance on the courses or rotations in OPP and osteopathic manipulative medicine (OMM) in the third or fourth year of osteopathic medical school training.

COMLEX–USA is a three-part examination series. However, it has never been demonstrated how the performance on the older Part I is related to performance on Part II of the old version or Level 2 of the current version of COMLEX–USA. In addition, the osteopathic medical licensure examinations were reformulated in the late 1990s. It is not clear how COMLEX–USA, the current version of the osteopathic medical licensure examinations, differs from its previous version—Parts I and II of the examination of the National Board of Osteopathic Medical Examiners (NBOME)—in its relationship with predmission characteristics and the ability to predict OPP and OMM course and rotation performances in medical school.

To address those issues, this study was undertaken with a threefold purpose:

- To examine the relationships between the performance on osteopathic medical licensure examinations and students’ premedical academic characteristics.
To investigate the predictive value of osteopathic medical licensure examination scores for the OPF and OMM course and rotation performances.

To analyze the relationship between the performance on the first and second part of COMLEX–USA and the relationship of performance across NBOME Parts I and II and COMLEX–USA Levels 1 and 2.

Review of literature

Many studies have looked at the correlation between time-honored premedical standardized tests and medical licensure examinations. Likewise, studies have been undertaken to determine the ability of an examination to predict some future outcome.

Elam and Johnson\(^5\) observed that results from Part I of the examination of the National Board of Medical Examiners (NBME) and Step 1 of the United States Medical Licensing Examination (USMLE) correlated moderately well with preadmission academic measures, as well as medical school performance. In 1996, Roth et al\(^6\) examined the USMLE scores of 480 students and sought to evaluate the predictive value of the Medical College Admission Test (MCAT), Scholastic Aptitude Test (SAT), and nonscience grade-point averages (NSGPAs) and science grade-point averages (SGPAs), as well as overall grade-point average (GPA) in performance on the USMLE. They note that of all the parameters used, the verbal section of the SAT best predicted performance. The MCAT data were excluded from the final regression analysis because pre-1991 MCAT scores were thought not to be useful in predicting performance of students entering medical school. The report of Roth et al\(^6\) indicates that the SAT score is strongly related to the premedical GPA and that the verbal score in the SAT is helpful in defining candidates capable of dealing with complex scientific concepts (eg, problem solving).

Shen and Comrey\(^7\) studied the relationship among medical students’ cognitive abilities, personal traits, and medical school performance using GPAs, MCAT scores, and the Comrey Personality Scales. The results of the study indicated that the MCAT score was a strong predictor of performance in medical school and on NBME examinations. However, in clinical studies, the correlation dropped sharply.

In 1996, Swanson et al\(^8\) reported that of all the premedical parameters used for medical school admission, the MCAT score is the most accurate barometer in predicting USMLE Step 1 performance. Hodgson and Silver\(^9\) reviewed a study conducted by Cunningham and Norman in 1997. Swanson et al\(^8\) indicate that MCAT scores correlated moderately well with the NBME Step 1 and clerkship performance in medical school.

Silver and Hodgson\(^10\) observed that MCAT scores and undergraduate GPAs (UGPAs) were good indicators of NBME Part I performance. These scores were not predictive of clinical performance as judged by grades in clinical subjects. They also noted that students in the upper quarter of their under-graduate class did significantly better on NBME Part I than those in the lowest quarter of their class.\(^11\)

The study conducted by Swanson et al\(^8\) in 1996 explored the notion that an early level of a medical licensure examination could predict performance on the next level. The authors demonstrated strong correlation between performance on Part I of the NBME examination and performance on Parts II and III. From these findings, they suggested that results of the NBME scores might be useful in selecting resident applicants. Vancouver et al\(^12\) also noted that SGPA and MCAT scores are predictive of medical school performance as measured by the NBME Part I examination.

By and large, most of the studies to date imply that MCAT scores correlate well with performance on medical licensure examinations. The GPAs—whether total, SGPA, or NSGPA—appear less valuable in predicting overall performance in a medical licensure examination and medical school performance, particularly when judged by performance in the clinical years.\(^11\)

With regard to the relationship between the lower and higher levels of medical licensure examinations, most studies indicate that the lower level of a medical licensure examination is a predictor of performance in higher levels of the same examination.\(^13\) However, some general disagreement exists as to whether MCAT scores or NBME Part I scores predict performance in clinical courses. For example, Elam and Johnson\(^5\) demonstrated moderate correlation. Shen and Comrey\(^7\) report little correlation with clinical studies.

Baker et al\(^14\) in a study conducted at the West Virginia School of Osteopathic Medicine in Lewisburg, observed that scores in COMLEX–USA Level 1 correlate significantly with GPAs achieved in the first 2 years, as well as individual course performance. In their study, Baker et al found no significant correlation between preadmission GPAs and MCAT scores with performance on COMLEX–USA Level 1.

In another study, Baker et al\(^15\) sought to determine the relationship of performance on COMLEX–USA Level 1 with academic performance at 18 of the then 19 AOA-accredited colleges of osteopathic medicine. The study involved 2146 students taking COMLEX–USA Level 1. The study demonstrated that academic performance in the first 2 years of osteopathic medical school is strongly associated with performance on COMLEX–USA Level 1.

Results of a study to determine whether different curricula have an impact on student performance on COMLEX–USA Level 1 demonstrated that curriculum type explains only a 1% variance in student scores.\(^16\)

Methods

Subjects and instruments

Four classes of osteopathic medical students from a single osteopathic medical college participated in this study. All classes were in the 1990s. The sizes of classes 1, 2, 3, and 4 were 57, 63, 68, and 77 students, respectively.
Three types of student academic information were collected from the school database:
- preadmission grades and test scores,
- osteopathic medical school course grades, and
- examination scores on COMLEX–USA Levels 1 and 2 or the NBOME Parts I and II scores.

COMLEX–USA or NBOME scores used in this study were the scores for the whole examination. COMLEX–USA is a three-test medical licensing examination. Each level has approximately 800 multiple-choice items. Level 1, given to osteopathic medical students at the end of the second year of medical school, measures the students’ basic science knowledge relevant to the most common or vital medical problems defined by the COMLEX–USA blueprint. Level 2, given to osteopathic medical students at the end of medical school, measures students’ knowledge of the concepts and principles involved in all steps of medical problem solving.

Parts I and II of the NBOME were the predecessor of COMLEX–USA Levels 1 and 2. Level 2 replaced the Part II examination in 1997, and Level 1 replaced the Part I examination in 1998. Level 3 of COMLEX–USA was the first section to be converted to the new format, replacing NBOME Part III in 1995. All COMLEX–USA and NBOME examinations report standard scores. Level 1 is scaled with a mean of 500 (SD 72). Level 2 is scaled with a mean of 500 (SD 85). All NBOME examinations were on a scale with a mean of 500 (SD 100).

Because the transition from the NBOME examination to COMLEX–USA examinations occurred in the late 1990s, some classes took only the NBOME examination or COMLEX–USA; others took both examinations. Classes 1 and 2 took NBOME Parts I and 2. Class 3 took NBOME Part I and COMLEX–USA Level 2. Class 4 took COMLEX–USA Levels 1 and 2. If a student took the same examination more than once, only the scores of the first were used for this study.

Preadmission information included premedical school overall GPA, premedical SGPA, premedical NSGPA, and MCAT scores. The MCAT scores that were used were taken from the candidates’ preadmission entry records. The MCAT score used for analysis in our study is the mean score compiled by averaging each candidate’s performance on all the numerical parts of the MCAT examination. The letter-graded language score was not taken into account in this study. All GPAs were on a scale of 0 to 4.

**Table 1**

**Performance on Examination of the National Board of Osteopathic Medical Examiners Parts I and II* and Preadmission Grade-Point Averages**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>Mean (SD)</td>
<td>No.</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Part I*</td>
<td>43</td>
<td>638.35 (85.67)</td>
<td>54</td>
<td>584.48 (82.28)</td>
</tr>
<tr>
<td>Part II*</td>
<td>43</td>
<td>602.07 (88.81)</td>
<td>52</td>
<td>566.58 (104.27)</td>
</tr>
<tr>
<td>GPA†</td>
<td>49</td>
<td>3.29 (0.47)</td>
<td>60</td>
<td>3.32 (0.26)</td>
</tr>
<tr>
<td>SGPA‡</td>
<td>49</td>
<td>3.18 (0.33)</td>
<td>60</td>
<td>3.22 (0.32)</td>
</tr>
<tr>
<td>NSGPA§</td>
<td>49</td>
<td>3.34 (0.36)</td>
<td>60</td>
<td>3.40 (0.30)</td>
</tr>
</tbody>
</table>

* Predecessor of the Comprehensive Osteopathic Medical Licensing Examination–USA (COMLEX–USA) Levels 1 and 2.
† GPA = Grade-point average.
‡ SGPA = Premedical science grade-point average.
§ NSGPA = Premedical nonscience grade-point average.

**Analysis**

All the analyses in this study are class based. Although analysis based on the total group by combining all classes would have yielded a larger sample size and therefore be more likely to yield significant results, the class-based analysis enabled us to look at patterns of the results among classes. If there were a clearly consistent pattern of the relationships across
the four classes based on small sample sizes, the findings would be more convincing.

The relationship between premedical academic characteristics and the COMLEX–USA/NBOME examinations were studied by the Pearson correlation. The relationship between COMLEX–USA Level 1/NBOME Part I and COMLEX Level 2/NBOME Part II was subjected to the same analysis.

The course or rotation grades of OPP were categorical. Therefore, their relationship with COMLEX–USA or the NBOME examination was explored by the two-tailed t test if there were actually only two categories of performance levels (pass, high pass, etc, as no numerical grades were given to OMM students) or by one-way analysis of variance (ANOVA) if there were more than two categories of performance levels (α level = 0.05; df, N –1, or 264). In those analyses, COMLEX–USA Level 1 and NBOME Part I or COMLEX Level 2 and NBOME Part II were the dependent variables and the categorical grades were the independent variables. The purpose of these analyses was to investigate whether the scores on the NBOME Part I examination predicted subsequent performance on Part II or COMLEX–USA Level 2. Moreover, COMLEX–USA Level 2 differentiated the performance of students in the OMM course. If students who scored high on the NBOME examinations also received high grades on osteopathic medical courses, or vice versa, the predictive value or the positive relationship could be demonstrated.

To avoid samples that were too small within a grade category, categories with small numbers were combined with the immediate next category on the scale for some classes.

- For class 2:
  - for OPP scores, “low pass” and “pass” were combined into “pass,” and “honor” and “high pass” were combined into “high pass”;
  - for OMM-1 scores, “honor” and “high pass” were combined into “high pass”;
  - for OMM-2, “low pass” and “pass” were combined into “pass.”

- For class 3:
  - for OPP scores, “honor” and “high pass” were combined into “high pass”;
  - for OMM-1, “pass” and “high pass” were combined into “pass.”

- For class 4:
  - for OMM-1 scores, “pass” and “high pass” were combined into “high pass.”

Results

When the data referable to the premedical school variables were used, MCAT scores had moderate but significant correlations with performance on all levels of COMLEX–USA and all parts of the NBOME examination at the conventional P level of 0.05 for all classes. At the same time, none of the classes had a significant correlation between COMLEX–USA and NBOME scores and overall GPA, SGPA, and NSGPA. A P value of .05 was used as the significance level to all tests applied in this study. Table 3 lists the correlation between
MCAT and COMLEX–USA and NBOME examination scores. Table 3 also demonstrates the correlation between COMLEX–USA Level 1 and NBOME Part I and COMLEX–USA Level 2 and NBOME Part II. Classes 1 and 2 took both NBOME Part I and Part II examinations. The correlation between NBOME Parts I and II for those two classes was high and significant.

Class 3 took NBOME Part I and COMLEX–USA Level 2. The correlation between NBOME Part I and COMLEX–USA Level 2 for that class was significant. Although substantial, the size of the correlation was not as high as the correlation among NBOME Parts I and II, as classes 1 and 2 demonstrated. This phenomenon was expected because both COMLEX–USA and the NBOME examination test medical knowledge, so they should correlate well. Also, COMLEX–USA and the NBOME examination had substantially different designs; therefore, it was understandable that the correlation between two examinations from different systems was not as high as that between the two parts of the same system.

The correlation between COMLEX–USA Levels 1 and 2 for class 4 was high and significant. It was slightly lower than that between NBOME Parts I and II for class 1 but substantially higher than that for class 2. The information obtained from this aspect of this study was not decisive enough to draw conclusions regarding the differences of the relationship between the NBOME examination and COMLEX–USA. Nevertheless, it appeared that the relationship between COMLEX–USA Levels 1 and 2 was not dramatically different from that between NBOME Parts I and II.

The predictive value of the COMLEX–USA Level 1 and NBOME Part I examination scores for performance in OPP and OMM courses was demonstrated by the analyses of the mean differences of the COMLEX–USA Level 1 and the NBOME Part I scores among the students in different course and rotation grade categories in undergraduate years 3 and 4. As Table 4 indicates, students with high grades in OPP and OMM courses and rotations were more likely to have had high COMLEX–USA Level 1 or NBOME Part I scores. The differences for the OMM-1 scores for classes 2 and 3 were not significant, but the pattern of the differences was the same: students with higher scores on COMLEX–USA Level 1 or NBOME Part I tended to have higher grades in all OPP, OMM-1, and OMM-2 courses. Clearly, COMLEX–USA Level 1 and NBOME Part I scores are good predictors of third- and fourth-year OMM and OPP course performance.

Similarly, Table 5 demonstrates that students with higher COMLEX–USA Level 2 or NBOME Part II scores did significantly better in all OPP, OMM-1, and OMM-2 courses for all classes; exceptions are the performance of class 3 in OPP and OMM-1 and the performance of class 1 in OMM-1. For the stu-
dent who took COMLEX–USA Level 2 or NBOMe Part II tests before the OPP and OMM courses in this study, clearly, their examination scores were clearly good predictors for their course performance. For those who took COMLEX–USA Level 2 or NBOMe Part II after some of the OPP and OMM courses and rotations, examination scores could not be called predictors. However, the high and positive correlation between the results of the two types of tests was clear and significant.

Comment
Our study demonstrates several interesting areas of correlation that have been substantiated in the literature.

First, of all the premedical parameters used to evaluate future performance in a medical licensure examination, only the MCAT score is an effective barometer of future performance on COMLEX–USA. The GPA did not predict performance on COMLEX–USA.

Second, performance on COMLEX–USA Level 1 accurately predicts performance on Level 2 of COMLEX–USA, and a significant relationship exists between the NBOMe examination and COMLEX–USA in their first two components of the examination.

Third, perhaps the unique feature, is that COMLEX–USA Level 1 scores predict performance in the clinical courses of OPP and OMM.

Although the total number of students in the cohort that was studied was relatively small, our study parallels results obtained in previous studies. Additionally, our study strongly implies that COMLEX–USA can predict candidate performance in the next highest level of COMLEX–USA as well as performance in the clinical courses unique to osteopathic medicine.

Further study may be necessary as larger groups of students complete the COMLEX–USA sequence to evaluate the examination’s predictive value of performance in other undergraduate domains, as well as postgraduate parameters such as in-service examinations, intern and resident examinations, and eventually board certification examinations. Ultimately, other studies may address relationships to actual practice outcomes.

References


