Relationships Between Clinical Rotation Subscores, COMLEX-USA Examination Results, and School-Based Performance Measures

Michael K. Cope, PhD; Helen H. Baker, PhD, MBA; Robert W. Foster, DO; and Craig S. Boisvert, DO

At the West Virginia School of Osteopathic Medicine (WVSOM) in Lewisburg, an electronic rating form was created to assist preceptors in evaluating student performance for third- and fourth-year clinical rotations. Multiple preceptors, often in rural locations, rated the clinical performance of 70 students in the WVSOM graduating class of 2005. The current study analyzes these ratings. Using Cronbach α, subscore reliability coefficients were obtained for all rotations: clinical knowledge, 0.80; data collection, 0.59; application of knowledge, 0.65; and professionalism, 0.78. For the three required family medicine rotations, which were almost always supervised by osteopathic physicians, reliability of the rating for osteopathic principles and practice (OPP) was 0.44. Intercorrelations among these five subscores ranged from 0.46 to 0.94, all statistically significant at the .01 level. Ratings for the five subscores were compared with 19 measures of student performance in other parts of the WVSOM curriculum, many correlations were significant at the .01 level. Clinical knowledge correlated 0.59 with year 2 grade point average (GPA), 0.51 with years 1 and 2 OPP GPA, 0.50 with Comprehensive Osteopathic Medical Licensing Examination USA Level 2 Cognitive Evaluation, and 0.45 with years 1 and 2 physical diagnosis GPA. Application of knowledge correlated 0.55 with year 2 GPA and 0.50 with the organization score on the college’s year 3 objective structured clinical evaluation. Professionalism correlated 0.51 with year 2 GPA and 0.49 with OPP years 1 and 2 written examination score. The average preceptor rating using the new electronic form was 92.6, compared with 96.8 when the previous paper-based form was used for the WVSOM class of 1998 (change significant at .05 level). These moderate correlations provide some support for the validity of the Clinical Education Grade Form.

Conventional wisdom holds it impossible for a broad cross-section of clinical preceptors to evaluate student clinical performance subjectively in a valid and reliable way. Medical schools must develop such a method for clinical teachers to assess student performance on clinical rotations. For the osteopathic medical schools that conduct clinical education at multiple locations, this challenge may seem particularly daunting.

The present study describes a new electronic form, the Clinical Education Grade Form, used by clinical preceptors to grade osteopathic medical student performance on third- and fourth-year rotations at the West Virginia School of Osteopathic Medicine (WVSOM) in Lewisburg. This study also describes procedures used to evaluate the extent to which the Clinical Education Grade Form decreased grade inflation, and the extent to which preceptor ratings were correlated with other measures of academic ability.

The use of evaluation forms and systems in clinical clerkships has a long history that has run parallel with concerns over the validity of those forms and systems. Articles evaluating medical student clerkship performance generally fall into three categories: surveys and suggestions; evidence of validity, reliability, and generalizability; and correlation with other measures of medical student performance.

Surveys and Suggestions

In 1987, Tonesk and Buchanan1 conducted a pilot study of 10 medical schools that described common problems with clerkship evaluation systems. Some problems cited were the evaluators’ unwillingness to record negative evaluations, inadequately defined evaluation criteria, and lack of training for evaluators. In 1990, Magarian and Mazur2 conducted a survey of 101 medical schools on the type of clerkship evaluation system used to assess student performance. The survey found that 68% of schools used a pass-fail grading system, 28% used letter grades, and 4% reported numeric scores. In 1992, Hunt3 described a model evaluation system containing four stages and provided a list of “signs and symptoms” one could consult to detect an evaluation system with problems. Other articles4–7 express a general desire for a valid and reliable clerkship evaluation system that provides adequate and appropriate evaluation and feedback concerning the developing clinical skills of medical students.
Evidence of Validity, Reliability, and Generalizability

Many articles have also addressed the validity, reliability, and generalizability of data gathered from existing clerkship evaluation systems.7-12 While some studies7,8 have supported the validity of rating forms used in these evaluation systems, many researchers argue that the topic of form reliability has not been adequately addressed in the literature. For example, Levine and colleagues9 determined that one way to increase form reliability would be to have multiple preceptors rate each student. Reznick and coauthors10 used interrater agreement coefficients to look at relative and absolute reliabilities and determined that only five raters per student were required to obtain the commonly recommended reliability of 0.80. Carlone and coinvestigators11 determined that a minimum of seven raters per student were required to obtain a reliability of 0.80. One study12 noted that “raters usually have had limited opportunities to observe a student's skills before completing the evaluation, affecting the reliability of the assessment.” Researchers13 further posited that three raters were required to obtain rating reliability.

Correlations With Other Performance Measures

In the process of demonstrating the validity of clerkship evaluation forms, many authors13-17 have shown correlations between these evaluation scores. Correlations with other academic measures or correlations with measures from outside agencies such as the National Board of Osteopathic Medical Examiners (NBOME) or the National Board of Medical Examiners (NBME) have also been demonstrated.13-17 Lazaro and colleagues also reported the relative lack of correlation between clinical performance and written examinations whether they were generated internally or by national agencies. Metheny also found a weak relationship between clinical evaluations and NBME Subject Examination scores. Campos-Outcalt and colleagues15 compared ratings of clinical performance and problem-solving skills with preadmission measures (ie, undergraduate GPA and Medical College Admission Test [MCAT] score) and also with NBME Parts I and II subtests and total score values. Their study found no correlation between clinical evaluations and problem-solving skills and any preadmission measures. In addition, their study found mild correlations between clinical evaluations and NBME Parts I and II scores. They found no correlation between problem-solving skills and NBME Part I, and only a moderate correlation between problem-solving skills and NBME Part II. Although Silver and Hodgson16 reported no correlation between admissions measures (ie, undergraduate GPA and MCAT score) and clinical performance, they found a good correlation between admissions measures and NBME Part I performance. Callahan and coauthors17 reported a high (0.01) correlation between family medicine, internal medicine, and pediatrics clerkship evaluations and USMLE Step 3. Ferguson and Kreiter18 reported a high correlation between preclinical and clinical training measures such as objective structured clinical evaluations (OSCEs), case-based reports, multiple-choice clinical skills examinations, and clinical preceptor evaluation forms. They also noted that basic science grades, undergraduate GPA, and MCAT scores did not correlate with clinical preceptor evaluation forms.

To determine form validity, we looked at clerkship evaluation systems and forms and their correlation to institutional and national measures of both preclinical and clinical performance. The Clinical Education Grade Form was created in Microsoft Word 98 for Windows (Microsoft Corporation, Redmond, Wash). It was studied previously19 and researchers found that student grades were distributed more accurately (ie, more even distribution and lower mean) using the Clinical Education Grade Form than WVSOM’s previous paper-based form.

Preceptor Evaluations

Prior to 2002, WVSOM used a paper-based clinical education grading form that had several descriptive items that required individual numeric ratings from clinical preceptors. However, the final letter grade was intended as a global assessment and was assigned without reference to the numeric ratings. Although students were satisfied with the results achieved using the old form, WVSOM administration was concerned about its validity and the potential for grade inflation. For the WVSOM graduating class of 1998, 81% of the grades assigned by preceptors were A’s (which was assigned a numeric grade of 99); 7%, B (89); and 2%, C (79) or F (65), resulting in a mean GPA of 96.8 (almost perfect) for phase 3, the clinical years of osteopathic medical school. The same students received an average of 85.7 for phase 1 (courses taught through April of year 1) and 89.5 for phase 2 (organ systems courses and related clinical courses taught in the last part of year 1 and throughout year 2), indicating significant grade inflation during phase 3. Further, for this graduating class, the correlation between phase 3 GPA and performance on the Comprehensive Osteopathic Medical Licensing Examination USA Level 2 Cognitive Evaluation (COMLEX-USA Level 2-CE) written examination was only 0.16, leading to even more questions about the form’s validity.

Data analysis was used to evaluate the extent to whether the Clinical Education Grade Form decreased grade inflation and correlated with other measures of academic success, including OSCE performance. Other than those found in the OSCE and COMLEX-USA examinations, the current investigation does not include analysis of written tests or case studies required during phase 3 for the class of 2005. This omission is due in part to changes in process at WVSOM with regard to how those components are assigned.
Methods
Approval was obtained from WVSOM’s institutional review board to conduct this retrospective review using existing educational data. Subjects for the study were 70 students who were in the class of 2005 as of April 2004 and who took WVSOM’s year 3 OSCE as well as the COMLEX-USA Level 2-CE (written examination) and COMLEX-USA Level 2 Performance Evaluation (COMLEX-USA Level 2-PE), a clinical skills examination.

Clinical Rotations
Preceptors are asked to evaluate student performance on clinical rotations. Five performance factors are addressed (Figure 1):

- clinical knowledge (subscore 1)
- data collection, history, and physical examination (subscore 2)
- application of knowledge to clinical care/clinical judgment (subscore 3)
- osteopathic principles and practice (OPP) (subscore 4)
- professionalism, responsibility, dependability, and reliability (subscore 5)

Possible ratings were “failure,” “needs improvement” (61-68), “adequate” (73-77), “good” (81-89), “excellent” (92-96), “truly exceptional” (100), or “not observed.” A brief description of the behaviors associated with each rating is included in the Clinical Education Grade Form. All ratings were converted to numeric scores using the values listed. The subscore values for all core clerkships, selective clerkships, and elective clerkships were averaged for use in analysis. Figure 2 lists the various clinical rotations available at WVSOM and describes their respective durations at the time of the study.

At WVSOM, clinical rotation grades are determined predominantly by ratings from clinical preceptors who have not previously taught the student in a more formal educational environment. These preceptors spend extensive time in one-on-one interaction with the students and thus have enough direct contact with students to provide a thorough evaluation.

For a few rotations (predominately electives), students had the option of spending 50% of the rotation with one preceptor and the other 50% with a different preceptor (9.1% of total rotations, 2.5% of required and selective rotations, and 36% of elective rotations). In these instances, the subscore values assigned by each preceptor were averaged to determine the rotation grade.

Student Performance on COMLEX-USA
The National Board of Osteopathic Medical Examiners administers a series of licensing examinations designed to evaluate student competency at different stages in the educational process. At the end of year 2, students take COMLEX-USA Level 1; during year 4, students take COMLEX-USA Level 2-CE. In the spring of 2005, COMLEX-USA Level 2-PE was introduced. All members of WVSOM’s class of 2005 were required to pass COMLEX-USA Level 1 and COMLEX-USA Level 2-CE in order to graduate. All members of WVSOM’s class of 2005 were required to participate in, but not necessarily pass, COMLEX-USA Level 2-PE. Numeric scores were available for both the Level 1 and Level 2-CE examinations, and pass-fail status was established for the Level 2-PE examination. The scores for the COMLEX-USA Level 2-PE examination were coded as “1” for pass and “0” for fail.

School-Based Performance Measures

- Preclinical Grades—All members of the WVSOM class of 2005 participated in an organ systems–based curriculum. The first 2 years of the curriculum are divided into two phases. The first phase of the year 1 curriculum is composed of a variety of basic science courses designed to lay the foundation for the study of clinical medicine in the next phase. Also included in the coursework for phase 1 are several clinical components including physical diagnosis and OPP. Phase 1 runs for the entire first semester and 3 months of the second semester of students’ first year at WVSOM. Phase 2 of the curriculum, primarily organ system courses, takes up the remainder of the second semester of year 1 and all of year 2. Course grades are given based on a 100-point scale with a minimum passing grade of 70. When a student failed a course and was required to repeat it, the first (failing) numeric score was used for the current analysis. The following variables were calculated from raw numeric course grades:
  - phase 1 GPA
  - year 1 GPA
  - phase 2 GPA
  - year 2 GPA
  - GPA for the two physical diagnosis courses
  - GPA for the four OPP courses

- WVSOM’s Third-Year OSCE—To prepare students for the COMLEX-USA Level 2-PE and to provide an institutional measure of student development in clinical skills, WVSOM developed a year 3 OSCE. The graduating class of 2005 was the first WVSOM class to take this new examination. The OSCE was a 12-station examination given to all third-year students. The test was administered during 3 half-day sessions in April 2004. Testing dates were consecutive and accommodated schedule conflicts. At each of the 12 stations, students had 13 minutes to take a patient history and perform a focused physical examination. They were then given 9 minutes to complete the SOAP (subjective, objective, assessment, and plan) Note Form. Because this was the first time an OSCE was used at WVSOM, all students received a
**Clinical Education Grade Form**

**West Virginia School of Osteopathic Medicine**
400 North Lee Street, Lewisburg, WV 24901, Phone: (304) 647-6278, Fax: (304) 647-6258

Student: ________________________________
Service: ________________________________
 dates: ________________________________
Hospital, Clinic, or Office: ________________________________
Location: ________________________________
RHEP: Yes No

Please darken the circle [ ] which best describes this student's performance for the student's current level of training.

### 1. Clinical Knowledge

<table>
<thead>
<tr>
<th>Failure/Improvement</th>
<th>Needs Improvement</th>
<th>Adequate</th>
<th>Good</th>
<th>Excellent</th>
<th>Truly Exceptional</th>
<th>Not Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear failure - unacceptable performance, fails to meet minimum standards. Fails rotation.</td>
<td>Major gaps in knowledge base. Fails to consider important issues or aspects of common problems.</td>
<td>Adequate knowledge base and understanding. Acceptable but less than expected differential diagnosis and evaluation.</td>
<td>Consistently demonstrates expected, good knowledge of disease processes and related concepts. Entirely satisfactory knowledge of differential diagnosis, diagnostic criteria, and evaluation of conditions.</td>
<td>Excellent knowledge of disease process and differential diagnosis, far beyond what would be expected.</td>
<td>Truly exceptional - highest 1% on above.</td>
<td>Not observed or unable to rate.</td>
</tr>
</tbody>
</table>

### 2. Data Collection, History, and Physical Examination

<table>
<thead>
<tr>
<th>Failure/Improvement</th>
<th>Needs Improvement</th>
<th>Adequate</th>
<th>Good</th>
<th>Excellent</th>
<th>Truly Exceptional</th>
<th>Not Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear failure - unacceptable performance, fails to meet minimum standards. Fails rotation.</td>
<td>Misses important information or misinterprets findings. Omits relevant parts of the exam.</td>
<td>Adequate history and physical exam; usually includes most relevant information. Some unexpected errors or misinterpretation.</td>
<td>Consistently obtains expected, appropriate patient history, including bio-psycho-social. Competent physical exam. Few errors in examination. Performance entirely satisfactory.</td>
<td>Excellent history and examination technique, far beyond what is expected at this level.</td>
<td>Truly exceptional - highest 1% on above.</td>
<td>Not observed or unable to rate.</td>
</tr>
</tbody>
</table>

### 3. Application of Knowledge to Clinical Care/ Clinical Judgment

<table>
<thead>
<tr>
<th>Failure/Improvement</th>
<th>Needs Improvement</th>
<th>Adequate</th>
<th>Good</th>
<th>Excellent</th>
<th>Truly Exceptional</th>
<th>Not Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear failure - unacceptable performance, fails to meet minimum standards. Fails rotation.</td>
<td>Unable to synthesize information. Interpretations not consistent with clinical data. Unable to prioritize tasks and develop.</td>
<td>Adequate application of medical knowledge. Although some unexpected errors, generally adequate clinical judgment for this level of training.</td>
<td>Consistently demonstrates good application of medical knowledge. Usually interprets findings correctly, develops reasonable plan, consistent with finding. Performance entirely satisfactory.</td>
<td>Exceptionally skilled in integration of knowledge and interpretation of diagnostic results. Uses logic well above expected level.</td>
<td>Truly exceptional - highest 1% on above.</td>
<td>Not observed or unable to rate.</td>
</tr>
</tbody>
</table>

(continued)
The Clinical Education Grade Form is an electronic form created in Microsoft Word 98 for Windows (Microsoft Corporation, Redmond, Wash) developed at West Virginia School of Osteopathic Medicine (WVSOM) in Lewisburg. It was developed by WVSOM for use mainly by off-site preceptors, often in rural locations, who rate the clinical performance of WVSOM students completing clinical rotations. The form has been recreated in hard copy format and proofread for publication in JAOSA—The Journal of the American Osteopathic Association. Abbreviations: DME, director of medical education; OMT, osteopathic manipulative treatment; OPP, osteopathic principles and practice; RHEP, West Virginia Rural Health Education Partnerships program.

### 4. Osteopathic Principles and Practice

<table>
<thead>
<tr>
<th>Failure</th>
<th>Needs Improvement</th>
<th>Adequate</th>
<th>Good</th>
<th>Excellent</th>
<th>Truly Exceptional</th>
<th>Not Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○ 61 68</td>
<td>○ 73 77</td>
<td>○ 81 85 89</td>
<td>○ 92 96</td>
<td>○ 100</td>
<td>○</td>
</tr>
<tr>
<td>Clear failure—unacceptable performance, fails to meet minimum standards. Fails rotation.</td>
<td>Unable to recognize and apply basic osteopathic principles and practice. Very uncomfortable with osteopathic manipulative treatment.</td>
<td>Adequate application of OPP and OMT. Although some unexpected errors, generally adequate performance for this level of training.</td>
<td>Consistently demonstrates entirely satisfactory application of osteopathic principles. Good grasp of osteopathic diagnosis, OMT, and documentation. Good skills, appropriate for level of training. Performance entirely satisfactory.</td>
<td>Exceptionally skilled in OMT and OPP, far beyond expected level. Teacher and role model for appropriate osteopathic care.</td>
<td>Truly exceptional—highest 1% on above. (Detailed justification must be given below.)</td>
<td>Not observed or unable to rate.</td>
</tr>
</tbody>
</table>

### 5. Professionalism, Responsibility, Dependability, and Reliability

<table>
<thead>
<tr>
<th>Failure</th>
<th>Needs Improvement</th>
<th>Adequate</th>
<th>Good</th>
<th>Excellent</th>
<th>Truly Exceptional</th>
<th>Not Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○ 61 68</td>
<td>○ 72 77</td>
<td>○ 81 85 89</td>
<td>○ 92 96</td>
<td>○ 100</td>
<td>○</td>
</tr>
<tr>
<td>Clear failure—unacceptable performance, fails to meet minimum standards. Fails rotation.</td>
<td>Some behavior inappropriate. Exhibits one or more of the following: dishonnest, disconsidered, manipulative, abusive of others; sloppy.</td>
<td>Adequate but less than expected performance. A few questionable absences or lapses in performance. May have missed some required rounds or lectures.</td>
<td>Ethical judgement, behavior and appearance consistently appropriate. Entirely satisfactory attendance at required conferences. Completed expected service. Dependable, punctual; fulfilled obligations. Worked well with other healthcare team members.</td>
<td>Maturity and professional judgment far exceed expected. Exceptionally conscientious. Anticipated and fulfilled needs far beyond basic rotation.</td>
<td>Truly exceptional—highest 1% on above. (Detailed justification must be given below.)</td>
<td>Not observed or unable to rate.</td>
</tr>
</tbody>
</table>

**Overall strengths:**

**Overall weaknesses/areas for improvement:**

- Were rotation objectives accomplished? ○ Yes ○ No
- Was evaluation discussed with student? ○ Yes ○ No
- If RHEP rotation - was feedback incorporated from the RHEP Site Coordinator? ○ Yes ○ No
- Signature of Evaluator: ○ DO ○ MD ○ Other
- Name: ____________________________ Date: ________________

Note: The Faculty of Record (DME, Preceptor, or Attending) is responsible for determining the ratings on this sheet. Any attempt by a student to pressure an Attending/Preceptor/DME into changing a rating would be unprofessional behavior, and cause for disciplinary action. Any student appeal of a grade may take place only after the student files written notification with the Associate Dean for Clinical Education. The only acceptable reasons for a grade appeal are: a) error such as mistake in computation; b) malicious or discriminatory grading; or c) no discernible rationale for arriving at grade. Grades cannot be appealed because of: a) student disagreement with preceptor requirements; b) student disagreement with grading standards; c) student disagreement with faculty judgment in applying standards, as long as the faculty member has made reasonable effort in good faith; or d) a student’s desire for an “A” for a particular grade.
passing grade for participation, but numeric scores were available for use in the current study. Faculty evaluators provided numeric scores for the history and physical examination portions of the SOAP Note Form, but simulated patients were asked to provide communications ratings. In a postexamination feedback session, students were able to review their completed evaluation forms as well as summary reports for all 12 stations and overall scores for history, physical examination, OPP, SOAP Note Form, and communication components of the examination. In addition to providing detailed written feedback to participants, numeric results were also used to revise OSCE administration procedures. The OSCE data used for the current analysis were: (1) the total history score, (2) the total physical examination score, and (3) the total SOAP Note Form score. In addition, while the OSCE was given, each student was evaluated on three global factors. These factors, communication, organization, and professionalism, were graded either by the simulated patient or by an in-room evaluator.

**Results**

As noted, the scores for COMLEX-USA Level 2-PE were reported to colleges only as a “pass” or “fail” for each student. Therefore, these data are dichotomous. The other variables under study were continuous. Accordingly, appropriate correlations were calculated using SPSS for Windows (version 10.1; SPSS Inc, Chicago, Ill).

Correlations with selected academic performance measures are shown in the Table.

**Reliability of Clinical Education Grade Form Subscores**

The reliability of the subscores from the Clinical Education Grade Form was varied. Subscore 1 had a reliability of 0.80; subscore 2, 0.59; subscore 3, 0.65; and subscore 5, 0.78.

For subscore 4, there were not enough data to determine reliability because all WVSOM students in this graduating class took at least one rotation with a preceptor (sometimes an allopathic physician [MD]) who did not evaluate students for OPP. The clerkships that had the largest proportion of osteopathic physicians (DOs) serving as preceptors, namely the three family medicine clerkships, were isolated from the rest of the dataset and used to determine the reliability of the OPP subscore. All other subscores for these clerkships were also reevaluated for reliability based on this reduced dataset.

The reliability ratings for the family medicine clerkship subscores were as follows: subscore 1, 0.52; subscore 2, 0.53; subscore 3, 0.51; subscore 4, 0.44; and subscore 5, 0.34.

All subscores were correlated, ranging from a low correlation of 0.46 to a high correlation of 0.94. The highest correlation was between subscore 1, clinical knowledge, and subscore 3, application of knowledge.

**Correlations With COMLEX-USA Results**

The correlations of subscores 2 through 5 with COMLEX-USA Level 1 were not statistically significant. Subscore 1 was significantly correlated with student performance on COMLEX-USA Level 1 ($P = .05$). The correlation of subscores 1, 2, 3, and 5 with student performance on COMLEX-USA Level 2-CE was statistically significant ($P = .01$). Subscore 4 was not correlated with student performance on COMLEX-USA Level 2-CE. Subscores 1 through 4 were not correlated with student performance on COMLEX-USA Level 2-PE. However, subscore 5 did have a statistically significant correlation with student performance on COMLEX-USA Level 2-PE ($P = .05$).

**Correlations With Academic Measures**

Subscore 1, subscore 3, and subscore 5 were significantly correlated with all academic measures (GPAs for both preclinical phases and years, as well as for physical diagnosis and OPP courses) ($P = .01$). Subscore 2 was significantly correlated both with phase 2 and year 2 GPAs ($P = .01$) as well as with GPAs from physical diagnosis and OPP courses ($P = .05$). Subscore 2 had no significant correlation with either phase 1 GPA or year 1 GPA. Subscore 4 was correlated with GPAs for both preclinical phases and years as well as OPP courses ($P = .01$). Subscore 4 was significantly correlated with GPA for physical diagnosis courses ($P = .05$).

The correlations for WVSOM’s OSCE were as follows: subscore 1 was significantly correlated with total score, history, SOAP Note Form, organization, and professionalism ($P = .01$), as well as with physical examination ($P = .05$). Subscore 1 was not correlated with communication. Subscore 2 was significantly correlated with organization ($P = .01$), total score, and SOAP Note Form ($P = .05$). Subscore 2 was not correlated with history, physical examination, communication, or profession-

---

**Table**

<table>
<thead>
<tr>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Medicine I</td>
<td>Family Medicine III</td>
</tr>
<tr>
<td>Family Medicine II</td>
<td>Internal Medicine II</td>
</tr>
<tr>
<td>Internal Medicine I</td>
<td>Internal Medicine III</td>
</tr>
<tr>
<td>Pediatrics I</td>
<td>Pediatrics II</td>
</tr>
<tr>
<td>Geriatrics</td>
<td>Surgery II</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>Surgery III</td>
</tr>
<tr>
<td>Surgery I</td>
<td>Emergency Medicine</td>
</tr>
<tr>
<td>Elective I</td>
<td>Obstetrics/Gynecology</td>
</tr>
<tr>
<td>Elective II</td>
<td>Elective III</td>
</tr>
<tr>
<td>Elective II</td>
<td>Elective IV</td>
</tr>
</tbody>
</table>

**Figure 2. With the exceptions of the Family Medicine I, Internal Medicine I, and Family Medicine III rotations, which each last 2 months, all clinical rotations at West Virginia School of Osteopathic Medicine in Lewisburg last 1 month. A total of 22 months are required for students to complete their clinical rotations.**
alism. Subscore 3 was significantly correlated with total score, history, SOAP Note Form, organization, and professionalism ($P=.01$), as well as with physical examination ($P=.05$). Subscore 3 was not correlated with communication. Subscore 4 was significantly correlated with total score and physical examination ($P=.01$). Subscore 4 was not correlated with physical examination, SOAP Note Form, communication, organization, or professionalism. Subscore 5 was significantly correlated with total score, organization, and professionalism ($P=.01$), as well as with physical examination.
than the average rating of 96.8 for the class of 1998. Developed to help preceptors conduct better evaluations of student performance in clinical rotations. The Clinical Education Grade Form had an average preceptor rating of 92.6, somewhat lower than the average rating of 96.8 for the class of 1998 (P=.05), which used the old paper-based form that was discontinued in 2002. While we would like to believe that students excelled on their clinical rotations, we believe it is more likely the preceptor ratings remain inflated with the Clinical Education Grade Form. We will continue to expand faculty development activities to better define WVSOM's necessary expectations in order for students to receive an "A" grade.

With the previous paper-based grading form, the correlation between preceptor ratings and COMLEX-USA Level 2-CE was low (0.16). With the new electronic Clinical Education Grade Form, the correlation between subscore 1 and COMLEX-USA Level 2-CE was 0.50, which seems more appropriate. Correlations with other written and performance examination results seem to suggest the validity of the new instrument. For example, the correlation between subscore 1 and performance on WVSOM's OSCE is 0.43, which supports the validity of both measures.

All subscores were correlated, with the highest correlation (0.94) between subscore 1 and subscore 3. Because these two variables (clinical knowledge and application of knowledge) are logically related, we believe this high correlation is appropriate. We intend to leave both items on the scale because students' final grades are determined by calculating the simple average of each subscore. We think it is appropriate to have both items on the Clinical Education Grade Form.

Preceptors did not always fully complete the new form. Consequently, students did not receive a rating for all subscores. The subscore most frequently left incomplete by preceptors was subscore 4, OPP. We must grant that some rotations are not always appropriate for the demonstration of the clinical skills included in the OPP subscore. In addition, preceptors were either DOs or MDs. Most MDs did not provide a rating for subscore 4. We assume the omission of this rating was often intentional and that most MDs decided not to rate students on this aspect of their performance due to their relative lack of training in OPP. However, some DOs also did not rate students on this item. Because of incomplete results for subscore 4, we were unable to determine the reliability of the OPP subscore. When the dataset was reduced to rotations with a high proportion of preceptors who were DOs (eg, family medicine clerkships), subscore 4 reliability was 0.44. This reliability is lower than expected—and low enough to reduce the possibility that these ratings would highly correlate with academic variables documented earlier in the medical school curriculum (ie, phase 1 or phase 2). Some faculty development in the application of the OPP subscore has already occurred. We intend to work more closely with preceptors to standardize and build confidence in their abilities to evaluate students for OPP.

The current study has several limitations. First, the data were collected at only one rural, osteopathic medical school, with a graduating class composed of 70 students. The ability of schools with other missions and structures to adopt the new electronic Clinical Education Grade Form may be limited; adaptation may be preferable. Also, the current analysis does not consider the impact of written examinations taken at or near the end of rotations. Similarly, it does not include the impact of case studies on rotation grades. Neither does this investigation address the extent to which preceptor ratings may correlate with such measures. Beginning with WVSOM's class of 2005, changes in end-of-rotation examinations have been implemented, so analysis of their impact would have no value for curriculum planning. However, for the continued improvement of the program, eventually it will be necessary to analyze how much end-of-rotation examinations correlate with preceptor ratings.

As noted, changes to the electronic Clinical Education Grade Form have been ongoing and other changes are underway. First, as initially implemented, the numeric values associated with the grade scale were intentionally removed in the hope that preceptors would focus on the characteristics described (eg, "consistently demonstrates expected, good knowledge") rather than the associated grade value (eg, "85," a mid-B grade). However, discussions with students at graduation exit interviews revealed that some students wrote these numbers on the forms in the hope of receiving a better grade. Student-altered forms are a possible source of error variation (ie, variability not related to the skills and abilities we wish to measure). Therefore, to prevent some students from having an unfair advantage, we placed numeric grade values on the electronic evaluation form.

Second, the Clinical Education Grade Form was developed before an emphasis on "core competencies" was implemented at the predoctoral level for colleges of osteopathic medicine, and before WVSOM's faculty expanded the application of contemporary models for teaching and evaluating communication skills. The new electronic evaluation form incorporates taking a patient's medical history as one of the examples for subscore 2, and the item "worked well with other healthcare team members" as an example for subscore 5. The Clinical Education Grade Form is also being revised to
make interpersonal communication a separate subscore: interpersonal communication and professionalism. Accordingly, some descriptors will be modified to emphasize communication skills. Other ways of measuring the core competencies (eg, basic knowledge of OPP and osteopathic manipulative treatment) will also be pursued.²⁰

Finally, it is clear that WVSOM must continue to expand faculty-development activities for clinical preceptors, specifically with regard to faculty expectations for WVSOM students. At its midwinter continuing medical education conferences, WVSOM has offered special sessions regarding contemporary practices in physical examination, communication skills, writing physician progress notes, and osteopathic manipulative medicine. In February 2005, session participants were asked to take part in a three-station OSCE. In February 2006, they were asked to critique three aspects of student performance: a video of osteopathic manipulative treatment, a role play of communication skills, and sample written progress notes. A major objective of these programs was to help preceptors become more confident with WVSOM’s current expectations for students’ clinical performance. Similar programs are being planned, as well as additional training for preceptors that specifically focus on the new electronic evaluation form.

Conclusion
The new electronic Clinical Education Grade Form decreased grade inflation for clinical rotations at WVSOM but did not eliminate it completely. Some subscores demonstrated acceptable reliability and were moderately correlated with other measures of academic ability. These moderate correlations provide some support for the validity of the new electronic form. Continued faculty development will be undertaken at WVSOM regarding the form. A special emphasis on standardization and improving preceptor confidence in using the OPP subscale is planned. Furthermore, the Clinical Education Grade Form will be further revised to focus more on the core competencies now required at colleges of osteopathic medicine.

References