This report discusses results of a study that sought to determine if medical students perform equally on undergraduate and postgraduate academic measures regardless of the clinical site of their clerkship. A review of the literature revealed that accreditation requirements and sound education practices require that clinical sites be comparable with regard to medical education experience and student performance. The current study examined performance comparability for two clerkships (medicine and surgery) at a variety of sites using both academic measures (final clerkship percent score, National Board of Medical Examiners (NBME) Part 2 subtest, and NBME Part 2 total) and postgraduate academic measures (NBME Part 3 total and supervisor rating of first-year residents). It was found that virtually no differences in academic performance were found among sites for the two clerkships. The findings are discussed in the context of methods for monitoring comparability of clerkship sites. Contains 32 references. (GLR)
Comparability of clerkship sites in a community-based medical school: Evaluation of undergraduate and postgraduate measures
R. J., Markert, H.V. Barnes, M.M. Dunn, K. Goldenberg, J.V. Hennessey

Abstract

Ronald J. Markert
Wright State University
School of Medicine
P.O. Box 927
Dayton, OH 45401-0927
Comparability of clerkship sites in a community-based medical school: Evaluation of undergraduate and postgraduate measures
R. J. Markert, H.V. Barnes, M.M. Dunn, K. Goldenberg, J.V. Hennessey
Abstract

Accreditation requirements and sound education practices require that clinical sites be comparable with regard to medical education experience and student performance. Past studies have generally found this to be the case for undergraduate academic measures in individual clerkships at medical schools with university hospitals. The current study examined performance comparability for two clerkships (medicine and surgery) in a community-based medical school using both medical school academic measures (final clerkship percent score, NBME Part 2 subtest, and NBME Part 2 total) and postgraduate academic measures (NBME Part 3 total and supervisor rating of first-year resident). Virtually no differences in academic performance were found among sites for the two clerkships. The findings are discussed in the context of methods for monitoring comparability of clerkship sites.

Purpose
Changes in health care delivery systems over the last two decades have prompted medical schools which previously conducted clinical education exclusively in a university-based hospital to extend students' clinical education experiences into community sites. The newer, community-based medical schools have used this regional model since their inception. Thus, for most medical schools this development has raised the program evaluation issue of comparability among clinical sites. The research hypothesis for the study was: Do medical students perform equally on undergraduate and postgraduate academic measures regardless of the clinical site of their clerkship?

Review of Literature
Early reports described curricular variability among clerkship sites and the need for comparable experiences in medical student education.\textsuperscript{1-3} Garrard and Verby\textsuperscript{4} found that medical students in a seven-month rural clinical experience had three times more patient encounters; saw different types of and were assigned more responsibility for clinical problems; did more procedures; and saw more female patients and more young patients but fewer age 41-65 year old patients than did their university-based peers. Friedman et al.\textsuperscript{5} compared education experiences in an obstetrics/gynecology clerkship and found that individual community hospitals differed as much among themselves as from the university hospital. The university hospital tended to provide a more academic clinical experience and less active participation in patient care. Bornsztein and Julian\textsuperscript{6} verified the percent of required clinical activities completed by students at nine obstetrics/gynecology clerkship sites and found that the mean percent completion rate varied from 71.5% to 98.3%. The completion rates for full-time faculty sites (86.6%) and private practice sites (84.2%) were similar. Strand et al.\textsuperscript{7} noted differences in number of deliveries and student ratings of teaching effectiveness among eight obstetrics/gynecology clerkship sites. Calhoun et al.\textsuperscript{8} examined a surgery clerkship at five different sites and found no differences in the total time allocated to patient care or education activities, but there were site variations for some of the 12 specific patient care or education activities. Rosen et al.\textsuperscript{9} analyzed student logs to compare a university hospital to five community hospitals with regard to diagnostic categories. Students saw more hematology/oncology and cardiovascular problems at the
university hospital but more neurologic and endocrine problems at the community hospital. Taylor et al. examined the activities and types of clinical problems seen at seven pediatric clerkship sites ranging from exclusively inpatient experience to a near total outpatient experience. Not surprisingly, the university and related hospital training locations had more structured teaching, interaction with residents, and more severely ill patients. In contrast, community outpatient settings offered greater ambulatory patient opportunities, greater variety of clinical problems, and more of the common pediatric illnesses. Miller et al. documented differences among six psychiatry clerkship sites in a variety of clinical education activities but reported that students saw no noteworthy advantages among the learning experiences at different sites. In Great Britain Wakeford found that peripheral (i.e., community) hospitals involved students more in practical procedures but did not differ from a teaching (i.e., university) hospital in experience with common medical conditions. More encouragement and more informal teaching occurred at peripheral hospitals, and students were generally more satisfied with their education experiences at those sites.

Studies in a number of specialties have examined medical student performance in a multi-site clerkship: four in family practice, three in internal medicine, four in obstetrics/ gynecology, two in pediatrics, two in psychiatry, and five in surgery. Rabinowitz reported on the variability in final examination performance (modified essay questions) and student grades at six family medicine clerkship sites. Hobbs et al. found that family medicine clerks did not differ in final percent score in a comparison between residency teaching locations (6 sites) and private group practices (5 sites).

Cahoun et al. found that time allotted to different internal medicine clerkship activities and level of responsibility assigned to the student varied among four hospitals, but these differences did not correlate with National Board of Medical Examiners (NBME) Part 2 medicine subtest scores. Ramsey et al. studied internal medicine clerks and found no site differences for a variety of objective measures. Whalen and Moses reported no differences for an internal medicine clerkship in a comparison between two medical center locations and a combination of four community hospitals for subjective and final grades or an NBME medicine shelf exam.

Irby et al. in a study of six obstetrics/gynecology clerkship sites found differences in the grades for four within-clerkship criteria (a clerkship posttest, presentation of an obstetrics/gynecology topic to peers, clinical performance, and patient write-ups). However, sites did not differ on oral examination scores or two external measures (ob/gyn portion of a third-year comprehensive exam and NBME Part 2 ob/gyn subtest). In a later analysis of the same obstetrics/gynecology clerkship (but with eight rather than six sites) investigated by Irby et al., Strand et al. reported variability among sites in oral exam performance but minimal site differences in written exam performance. Talbert et al. in a study of a university hospital and four community hospitals found that obstetrics/gynecology clerks did not differ by site on NBME Part 2 or faculty clinical ratings over a four year study period. Also, first postgraduate year performance of students entering obstetrics/gynecology residencies did not differ when analyzed by obstetrics/gynecology clerkship site. Bornsztain and Julian found no differences in cognitive examination performance among nine obstetrics/ gynecology clerkship sites.

Joorabchi et al. investigated a pediatric clerkship in three hospital settings (university, county, private) and found differences in education activities, types and number of patients, and time allocation but found no differences in problem-solving and higher cognitive level examinations. Similarly, in Israel Tamir et al. found that a university hospital and two nonuniversity hospitals differed in pediatric clinical experiences, but students did not score
differently on their final pediatric examination. However, students at nonuniversity hospitals were more satisfied with the quality of their training.

Puryear and Miller\(^2\)\(^4\) described how clerks at one site of a psychiatry clerkship (emergency room), where students historically scored lower on a comprehensive final exam, improved their exam performance through special attention from residents. Miller et al.\(^1\)\(^1\) reported small but statistically significant differences among six psychiatry clerkship sites on written work, clinical work, and written examination.

Schwartz et al.\(^2\)\(^5\) studied surgery clerkship sites (a university hospital and nine affiliated hospitals). Although students were randomly assigned to site and did not differ by site on a surgery pretest, there were some site differences on four surgery subtests (general surgery, orthopedics, urology, and anesthesiology). This finding was not surprising given the many hospital pairwise comparisons across four different subtests. Jacobson et al.\(^2\)\(^6\) compared a university setting with private and county hospitals and found no differences among surgery clerkship students on an NBME shelf examination in surgery, oral examination, or clinical evaluation. McCarthy and King\(^2\)\(^7\) in a comparison of four hospital sites (university, county, private, and Veterans Administration) found no differences on a surgery knowledge examination or achievement of procedural skills. Levitsky et al.\(^2\)\(^8\) studied two years of surgery clerks at a university and five community hospitals and found no differences on the surgery subtest or total score on both a school clinical comprehensive examination and NBME Part 2. Bielewicz et al.\(^2\)\(^9\) examined three different site combinations involving a different mix of experiences during a three-month surgery clerkship and found no differences among sites in written (NBME-developed exam) and oral score performance.

The current study extends previous investigations of comparability of education by examining two multi-site clerkships in a community-based medical school (i.e., no university hospital) as well as both medical school and postgraduate measures of performance.

**Methods**

The internal medicine and general surgery clerkships at Wright State University School of Medicine (WSUSOM) were examined. At the time of the study, the internal medicine clerkship used principally five different site combinations. A student spent six weeks at the first hospital and six weeks at the second hospital. Now the clerkship uses four weeks at the first hospital, four weeks at the second hospital, and four weeks in an ambulatory setting. The surgery clerkship lasts eight weeks (six weeks in general surgery at one site and two weeks in a surgical subspecialty). The current study investigated the six sites in which the general surgery rotation took place and the five principal internal medicine site combinations. Students are assigned randomly to a site in the medicine clerkship, but in the surgery clerkship student preferences for a site are accommodated as much as possible.

Sites (surgery) or site combinations (medicine) were compared on three medical school academic variables (final clerkship percent score, NBME Part 2 subtest, and NBME Part 2 total) and two postgraduate academic variables (NBME Part 3 total and total score on an 18-item measure of performance as a first-year resident).

The final clerkship percent score in surgery is calculated from three components: patient history and physical exam records written by the student and presented to faculty (20% of score), preceptor evaluation of student using clerkship performance form (40%), and objective and essay tests (40%). The final clerkship percent score in medicine also has three elements: preceptor evaluation of the student using clerkship performance form (50%), oral exam based on patients written up by the student (10%), and objective quizzes and final examination (40%). This physician supervisor rating of performance as a first-year resident was developed at the University of Arizona.\(^3\)\(^0\) The internal consistency reliability
(Cronbach's alpha) of the physician supervisor rating form with eight WSUSOM classes is .994.

The Arizona form contains 18 scales: (a) 11 scales yield ratings from one to eight - interpersonal skills with staff, interpersonal skills with patients, fund of medical knowledge, motivation/diligence, verbal and presentation skills, history-taking skills, physical examination skills, differential diagnosis/problem-listing skills, therapeutic planning, procedural skills, and overall evaluation; (b) seven scales yield ratings from one to five - preparation at start of residency, resident's view of preparation, understanding of goals and objectives of residency, progress in learning and development during first year of residency, constructive use of feedback, use of education resources of the residency program, and responsibility for self-directed learning. The maximum score for the physician supervisor rating of postgraduate performance is 123 (11 scales x 8 points + 7 scales x 5 points = 123).

Graduates of four classes (1988 through 1991) were investigated (n = 349). Data were collected for 348 (99.7%) graduates. Site combination for 331 medicine clerkship students was used in the analysis. Seventeen medicine clerkship students had site combinations other than the five principal locations. These site combinations were not included due to their small sample size. Since the Class of 1991 had not completed the first year of residency and the Classes of 1988 through 1990 had to grant permission, fewer graduates had data for the postgraduate academic variables (NBME Part 3 total, n = 163 of 265, 62%; supervisor rating of first postgraduate year, n = 151 of 265, 57%).

Univariate analysis of variance with site or site combination as the independent variable and each of the five academic criteria as dependent measures was used. Tukey's multiple range test was used to follow up significant ANOVA findings. Inferences were made at the .05 level of significance.

Results

Table 1 shows that in the medicine clerkship students differed by site combination for one undergraduate measure (clerkship score) but not for two other undergraduate measures (NBME Part 2 medicine subtest and NBME Part 2 total). Also, no site differences were found for the two postgraduate measures (NBME Part 3 total and supervisor rating as a first-year resident). The significant ANOVA for clerkship score was followed by Tukey's multiple range test. One site combination (SC1) differed from one other site combination (SC2). Table 1 shows that in the surgery clerkship there was one measure on which sites varied: surgery clerkship score with two pairwise differences.

Discussion

The Liaison Committee on Medical Education has addressed the importance of equivalence or comparability of clerkship education among sites:

If required clerkships in a single discipline are conducted in several hospitals, every effort must be made to ensure that the students receive equivalent educational experiences. (p. 13)

Our study complements previous studies which analyzed curricular content and education activities and/or examined performance. Unlike other studies the current investigation was conducted at a medical school without a university hospital and involved more than one clerkship. In addition, we used both undergraduate and postgraduate measures to assess equivalence. Kassebaum has argued for the assessment of both types of outcomes in the evaluation of medical education programs. Furthermore, the use of a physician supervisor's rating of a graduated student when he/she is a first-year resident answers to some extent Friedman et al.'s concern that studies use "clinical competence as the ultimate criterion measure." Also, our study used a clerkship that assigned students randomly and one which allowed student preferences.
For three measures of medical school academic performance and two measures of postgraduate academic performance, our internal medicine and general surgery clerkships differed by site only on clerkship score. For the final percent score in the internal medicine clerkship one pairwise comparison of site combinations was different by 3.2%. Similarly, we found two pairwise site differences for surgery clerkship score with two sites higher than a third site by 3.9% and 3.7%. The practical importance of these differences does not seem great, and we can identify no factors which would account for the differences. Our results add to those of other studies which have investigated student performance differences among clerkship sites.6-7, 11, 13-29

Wright State University monitors comparability among sites for its nine required clerkships: emergency medicine, family practice, medicine, neurology, obstetrics/gynecology, orthopedics, pediatrics, psychiatry, and surgery. Nine methods are used to assure comparability among education sites. These are: written clerkship objectives, written performance expectations for the clerkship, written evaluation standards, review of patient logs maintained by students, evaluation of a specific number of patient workups regardless of clerkship site, on-site faculty coordinator at each institution, regular meetings of the on-site coordinators, active participation by students in patient care, and case conferences for all clerkship students at a central location. These steps directed toward comparability among sites may account in part for our lack of performance variability in the two clerkships reported.

### Table 1

<table>
<thead>
<tr>
<th>Internal Medicine Clerkship</th>
<th>F statistic</th>
<th>p value</th>
<th>General Surgery Clerkship</th>
<th>F statistic</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clerkship score</td>
<td>3.70</td>
<td>.006*</td>
<td></td>
<td>3.66</td>
<td>.003**</td>
</tr>
<tr>
<td>NBME Part 2 subtest</td>
<td>0.34</td>
<td>.848</td>
<td>1.13</td>
<td>.343</td>
<td></td>
</tr>
<tr>
<td>NBME Part 2 total</td>
<td>0.25</td>
<td>.907</td>
<td>1.71</td>
<td>.132</td>
<td></td>
</tr>
<tr>
<td>NBME Part 3 total</td>
<td>0.17</td>
<td>.952</td>
<td>1.92</td>
<td>.093</td>
<td></td>
</tr>
<tr>
<td>Supervisor rating of first year resident</td>
<td>0.80</td>
<td>.529</td>
<td>1.59</td>
<td>.166</td>
<td></td>
</tr>
</tbody>
</table>

*Site combination 1 [SC1] (mean = 77.6%) was greater than SC2 (mean 74.4%).

**Site 1 (mean = 83.8%) and Site 2 (mean = 83.6%) were greater than Site 3 (mean = 79.9%).

### References


31. Liaison Committee on Medical Education. Functions and Structure of a Medical School: Standards for Accreditation of Medical Education Programs Leading to the M.D. Degree. Washington, D.C.: Association of American Medical Colleges; Chicago: American Medical Association; 1989.