Ranking Practice Variability in the Medical Student Performance Evaluation

BY

MEGAN BOYSEN OSBORN
B.A. University of San Diego, 2002
M.D., University of California, Irvine, 2007

THESIS

Submitted as partial fulfillment of the requirements for the
Degree of Master of Health Professions Education
in the Graduate College of the
University of Illinois at Chicago, 2015

Chicago, Illinois

Defense Committee:

Ilene B Harris, Chair and Advisor
Ara Tekian
John Christian Fox, University of California, Irvine
ACKNOWLEDGEMENTS

I would like to thank my thesis committee—Drs. Harris, Tekian, and Fox—for their guidance, mentorship, and inspiration during the preparation of my thesis. Their insights molded my thesis into a work of which I am very proud. I would also like to thank the faculty of the Department of Medical Education. Their instruction has been foundational for my early career in medical education. I also greatly appreciate the administrative support of Ms. Janet Settle, who has provided alacritous support throughout my entire journey at the University of Illinois, Chicago.

Mr. James Mattson and Mr. Justin Yanuck, both fourth-year medical students at the University of California, Irvine, worked tirelessly to collect data for this project and Dr. Craig Anderson helped me tremendously with the statistical analysis for this project.

I truly appreciate the support of my husband, Brad, who has enabled me to achieve my life’s aspirations of being a wife, mother, and academic clinician.

MBO
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION ........................................ 1</td>
<td></td>
</tr>
<tr>
<td>A. Background ........................................... 1</td>
<td></td>
</tr>
<tr>
<td>B. Purpose .............................................. 2</td>
<td></td>
</tr>
<tr>
<td>II. METHODS .............................................. 3</td>
<td></td>
</tr>
<tr>
<td>A. Sample ............................................... 3</td>
<td></td>
</tr>
<tr>
<td>B. Questions Answered During Data Collection .......... 3</td>
<td></td>
</tr>
<tr>
<td>C. Process of Data Collection .......................... 4</td>
<td></td>
</tr>
<tr>
<td>D. Data Analysis ........................................ 4</td>
<td></td>
</tr>
<tr>
<td>E. IRB Approval .......................................... 5</td>
<td></td>
</tr>
<tr>
<td>III. RESULTS .............................................. 6</td>
<td></td>
</tr>
<tr>
<td>A. Sample Size .......................................... 6</td>
<td></td>
</tr>
<tr>
<td>B. Overall Ranking Practice ............................ 6</td>
<td></td>
</tr>
<tr>
<td>C. Most Common Ranking Practices ..................... 8</td>
<td></td>
</tr>
<tr>
<td>D. Overall Most Common Terms .......................... 11</td>
<td></td>
</tr>
<tr>
<td>E. Location of Students’ Ranking Descriptions ........ 13</td>
<td></td>
</tr>
<tr>
<td>F. Location of Ranking System Legend .................. 13</td>
<td></td>
</tr>
<tr>
<td>G. Description of Non-ranking Schools .................. 13</td>
<td></td>
</tr>
<tr>
<td>H. Differences from 2013 to 2015 ....................... 14</td>
<td></td>
</tr>
<tr>
<td>IV. DISCUSSION ............................................ 15</td>
<td></td>
</tr>
<tr>
<td>A. Background and Implications of Ranking Variation .. 15</td>
<td></td>
</tr>
<tr>
<td>B. Limitations .......................................... 18</td>
<td></td>
</tr>
</tbody>
</table>
C. Conclusions........................................................................................................... 19

REFERENCES........................................................................................................... 21

VITA............................................................................................................................... 23
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. DOES THE MSPE CONTAIN COMPARATIVE PERFORMANCE (RANK) FOR THE STUDENT?</td>
<td>7</td>
</tr>
<tr>
<td>II. TYPE OF RANK USED AMONG RANKING SCHOOLS</td>
<td>9</td>
</tr>
<tr>
<td>III. MOST COMMON DESCRIPTORS FOR EACH PERCENTILE GROUP AMONG CATEGORICAL RANKING GROUPS</td>
<td>10</td>
</tr>
<tr>
<td>IV. OVERALL, MOST COMMON TERMS AMONG SCHOOLS THAT USE CATEGORICAL RANKING SYSTEMS</td>
<td>12</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>AAMC</td>
<td>Association of American Medical Colleges (AAMC)</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>ERAS</td>
<td>Electronic Residency Application Service</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile Range</td>
</tr>
<tr>
<td>MSPE</td>
<td>Medical Student Performance Evaluation</td>
</tr>
<tr>
<td>PD</td>
<td>Program Director</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
</tbody>
</table>
SUMMARY

The Medical Student Performance Evaluation (MSPE) is an assessment of a student’s performance during medical school. There is great variability in the composition and format of the MSPE from school to school.

We sought to appraise the variability in ranking systems in MSPE among US medical schools. We reviewed MSPE documents from 125 (95%) US allopathic medical schools over the course of two academic years, 2012-2013 and 2014-2015. We recorded whether the medical school had a ranking system, the type of ranking system used, the size and description of student categories, and the location of the ranking statement and category legend.

Seventy-four percent (74%) of medical schools provide comparative performance (ranks) for their students. Of these, the vast majority (87%) group their students into categories or quantiles, but there is wide variation in the number and size of these groups. The most common descriptors for the top, second, third, and lowest groups were “outstanding”, “excellent”, “very good”, and “good”. We found the student’s rank and description of ranking system in variable locations in the MSPE.

There is extensive variation among ranking systems within MSPEs. Program directors (PDs) may find it tedious to locate and translate a student’s rank and may have difficulty making inter-school comparisons.
I. INTRODUCTION

A. Background

The Medical Student Performance Evaluation (MSPE), formerly called the “Dean’s letter”, is an important component in a medical student’s application for residency training. The guidelines for preparing the document were revised in 2002 by the Association of American Medical Colleges (AAMC), in an attempt to promote consistency in the letters among medical schools.\(^1\) Three years later, Shea reported that there was still tremendous variability from institution to institution.\(^2\) The MSPE summary statement is one such area of variability.

The 2002 guidelines state that the summary paragraph should include “a summative assessment…of the student’s comparative performance in medical school, relative to his peers, including any information about any school-specific categories used in differentiating among levels of student performance.”\(^3\) Notwithstanding this recommendation, analyses of MSPEs demonstrate variable ranking systems among schools.\(^2-5\) Some medical schools provide a numerical rank for their students, while others group their students into quartiles or quintiles. Many medical schools group their students into categories, with descriptors such as “outstanding” or “very good”. Several schools do not describe or use a ranking system.

Categorical ranking systems vary widely in the terminology used and in the size of the category groups. For example, many schools use the word “excellent” to describe high-achieving students, while others use the term "excellent" to describe average students.\(^5\)
“Good”, on the other hand, is typically used to describe students in the bottom 50% of their graduating class. The AAMC recommends that Appendix D of the MSPE should provide a legend for these categorical ranking systems, but this explanation of ranking categories is commonly found elsewhere in the MSPE.

B. **Purpose**

While some authors have described the variable use of certain descriptive ranking terms within the MSPE, no authors have quantified the variability in ranking practices among US medical schools. Our purpose, based on a review of MSPEs from 95% of US medical school, is to: 1) report the proportion of medical schools that use a defined ranking system; 2) explore the variability among ranking systems and the most common language used among medical schools; and 3) appraise the summary paragraphs of MSPEs from schools that do not have formal ranking systems to identify statements suggestive of rank.
II. METHODS

A. **Sample**

We electronically extracted the MSPE document from each application to our emergency medicine residency program from two application cycles: 2012-2013 and 2014-2015. We did not have MSPEs from the 2013-2014 cycle available to us. We included any applicant from a United States (US) allopathic medical school, excluding Puerto Rico. We excluded any MSPEs for students who did not graduate in 2013 or 2015. For each application year, we selected one MSPE document per school for review, from the student whose application was listed first alphabetically in the electronic residency application service (ERAS).

B. **Questions Answered During Data Collection**

We answered five questions for each school's MSPE document: 1) Does the school use a defined ranking system? 2) What type of ranking system is utilized? 3) Into how many categories are the students divided? What are the most common category descriptors for each group? What percentage of students is in each category? 4) Where is the student’s rank provided in the MSPE document? Where is the legend for the ranking system? 5) Do non-ranking schools use similar language to schools that rank?

For the last question, we reviewed the summary statements of the schools that did not rank their students. We searched for language that could be confused for a ranking statement, particularly the sentence: “[Student name] is a [descriptor] candidate for residency training/graduate medical education,” which was a phrase that appeared in
many MSPEs we reviewed.

C. **Process of Data Collection**

A non-blinded, trained, single reviewer (JM) extracted and recorded the data on a data abstraction form for the 2013 applications. A second reviewer (MBO) reviewed the MSPEs for all schools without a defined ranking system to ensure that none had been overlooked. We (JM and MBO) held periodic meetings and resolved any inconsistencies between raters through discussions to reach consensus. The second reviewer (MBO) then reviewed the 2015 documents and abstracted data onto a data abstraction form. A third reviewer (JY) re-reviewed all 2015 documents of schools that did not appear to have a formal ranking system, to verify that there was indeed no ranking system. This reviewer also checked these documents for the fifth study question.

D. **Data Analysis**

We calculated descriptive statistics for data related to each of the study questions. Specifically, we calculated the percentage of schools using ranking systems and the type of ranking system used. We described common terminology among medical schools. For any schools that changed their ranking practice (e.g., changing from non-ranking to ranking or changing the number/description of groups) between 2013 and 2015, we use the 2015 practice for our data analysis (11% of schools). For any schools that changed the percentage of students in each category group, we averaged the group size between years. We performed a Wilcoxon signed rank test to compare the percentage of students in the first, second, third, and last category between 2013 and 2015 to ensure that there was no
statistically significant difference between the group sizes in 2013 and 2015.

E. **IRB Approval**

The University of California, Irvine and University of Illinois, Chicago Human Subjects Institutional Review Boards approved this study.
III. RESULTS

A. **Sample Size**

In 2015, 132 US allopathic medical schools had graduating classes, while there were 128 in 2013. In each of the 2012-2013 and 2014-2015 application seasons, we received approximately 650 applications to our emergency medicine residency program. Of these applications, approximately 550 per year were from United States allopathic medical students. From this pool, we had applicants (and MSPEs) from 125 of the 132 (95%) US allopathic medical schools. We had data for both years (2013 and 2015) for 90 (72%) of these medical schools. For the remainder, we had 2015 data for 12 (10%) and 2013 data for 23 schools (18%).

B. **Overall Ranking Practice**

Of the 125 schools’ MSPEs we examined, 93 (74%) provided ranks for the medical students (Table 1).
TABLE I

DOES THE MEDICAL STUDENT PERFORMANCE EVALUATION CONTAIN COMPARATIVE PERFORMANCE (RANK) FOR THE STUDENT? (N=125)

<table>
<thead>
<tr>
<th>Does school provide comparative performance (rank)?</th>
<th>Number (% of total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>93 (74%)</td>
</tr>
<tr>
<td>No</td>
<td>32 (26%)</td>
</tr>
</tbody>
</table>
C. **Most Common Ranking Practices**

Of these, 59 (63%) schools used named category groups, such as “outstanding” for their top group and “good” for their lowest group. Twenty-two (n=22, 24%) schools broke the class into segments such as tertiles (thirds), quartiles (fourths) or quintiles (fifths), without other category descriptors. The remaining schools (n=12, 13%) used a variety of descriptive methods, such as a numerical class rank or cumulative grade average compared to a mean (Table 2).

Students were most commonly divided into four groups (Table 2). Of the 59 schools that used named category groups, the term used to describe the top group of students was most commonly “outstanding” (n=33, 56%); the second highest group was most commonly described as “excellent” (n=30, 51%); the third group was described as “very good” (n=32, 54%), while the lowest group was described as “good” (n=28, 47%) (Table 3). On average, the top group contained 21% of students (range 3-38.5% of students; interquartile range (IQR) 15-25%); the second group contained 30% of students (range 10-75% of students, IQR 20-33%); the third group contained 30% of students (range 0-61% of students; IQR 22-37%); and the lowest group contained 11% of students (range 0-33% of students; IQR 2-19.5%). Less common descriptors for each of the four groups are listed in Table 3. We found 42 unique terms or phrases to describe student category groups. When the negative terms “marginal”, “below average”, or “recommended/presented with reservation” were used for the lowest ranked students, these groups included, on average, 1% of students (range 0-2%).
### TABLE 2
TYPE OF RANK USED AMONG RANKING SCHOOLS (N=93 SCHOOLS)

<table>
<thead>
<tr>
<th>Type of ranking practice</th>
<th>Number of schools with this ranking practice&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three groups (categories or tertiles)</td>
<td>7 (8%)</td>
</tr>
<tr>
<td>Four groups (categories or quartiles)</td>
<td>41 (44%)</td>
</tr>
<tr>
<td>Five groups (categories or quintiles)</td>
<td>29 (31%)</td>
</tr>
<tr>
<td>Six or seven groups (categories)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Numerical (e.g., 23/100)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Alternative ranking method&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6 (6%)</td>
</tr>
<tr>
<td>Not well-defined&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2 (2%)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Percentage is among schools that had ranking systems (n=93)

<sup>b</sup> Students are grouped by total number of honors grades, grade point average ranges, or a cumulative earned percentage compared to a class mean.

<sup>c</sup> It is clear that the school has a ranking system, but it is not well defined. All letters examined mention the student being in the “top 10%” or “lower 1/3” and use terms such as “exceptional” or “good” candidate for residency. However, there is no definition of this ranking system in the MSPE document.
**TABLE 3**
MOST COMMON DESCRIPTORS FOR EACH PERCENTILE GROUP AMONG SCHOOLS THAT USE NAMED CATEGORY GROUPS (N=59 SCHOOLS)

<table>
<thead>
<tr>
<th></th>
<th>Highest group</th>
<th>Second group</th>
<th>Third group</th>
<th>Lowest group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentile</strong></td>
<td>21% of students</td>
<td>30% of students</td>
<td>30% of students</td>
<td>11% of students</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>(range: 3-38.5% of students, IQR 15-25%)</td>
<td>(range: 10-75% of students, IQR 20-33%)</td>
<td>(range: 0-61% of students, IQR 22-37%)</td>
<td>(range: 0-33% of students, IQR 2-19.5%)</td>
</tr>
<tr>
<td><strong>Most common descriptors</strong></td>
<td>Outstanding 33 (56%)</td>
<td>Excellent 30 (51%)</td>
<td>Very good 32 (54%)</td>
<td>Good 28 (47%)</td>
</tr>
<tr>
<td></td>
<td>Exceptional 9 (15%)</td>
<td>Outstanding 17 (23%)</td>
<td>Excellent 20 (34%)</td>
<td>Satisfactory 7 (12%)</td>
</tr>
<tr>
<td></td>
<td>Superior 6 (10%)</td>
<td>Superior 5 (8%)</td>
<td></td>
<td>Very good 6 (10%)</td>
</tr>
<tr>
<td></td>
<td>Distinguished 4 (7%)</td>
<td></td>
<td></td>
<td>Recommended 6 (10%)</td>
</tr>
<tr>
<td><strong>Less common descriptors (used 1-2 times)</strong></td>
<td>Superb, exemplary, elite, enthusiastically recommended, enthusiastically and highly recommended, recommended enthusiastically, highest recommendation</td>
<td>Highly recommend, high recommendation, strongly recommend, recommend strongly, with great enthusiasm, excellent/outstanding, and exemplary</td>
<td>When the two terms above were not used, the third highest group was commonly the lowest group, so the majority of these terms are included in the last column</td>
<td>Adequate, capable, acceptable, competent, marginal, well-qualified, good/fair, recommended/presented with reservation, or below average</td>
</tr>
</tbody>
</table>

Abbreviations: Interquartile range (IQR)

a: 0% is included in the range because schools provided that category group, but listed 0% of students in that group.
D. **Overall Most Common Terms**

The most common terms used overall to describe student performance, regardless of class position, were “excellent” (n=44, 75%), “very good” (n=44, 75%), “outstanding” (n=42, 71%), and “good” (n=36, 61%) (Table 4). The term “excellent” was used among schools to describe students ranging from the 1st percentile to the 95th percentile; the term "outstanding" was used to describe students who were in the 33rd to 99th percentile; the term "very good" was used to describe students who were in the 1st to 69th percentile; and the term "good" was used to describe students who were in the 1st to 57th percentile.
TABLE 4
OVERALL, MOST COMMON TERMS AMONG SCHOOLS THAT USE NAMED CATEGORY GROUPS (N=59 SCHOOLS)

<table>
<thead>
<tr>
<th>Most common MSPE terms</th>
<th>Number of schools using descriptor and range of student’s percentile rank for this descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>44 (75%) of schools used this term</td>
</tr>
<tr>
<td></td>
<td>Range of students in this category: 1st to 95th percentile</td>
</tr>
<tr>
<td>Very good</td>
<td>44 (75%) of schools used this term</td>
</tr>
<tr>
<td></td>
<td>Range: 1st to 69th percentile</td>
</tr>
<tr>
<td>Outstanding</td>
<td>42 (71%) of schools used this term</td>
</tr>
<tr>
<td></td>
<td>Range: 33rd to 99th percentile</td>
</tr>
<tr>
<td>Good</td>
<td>36 (61%) of schools used this term</td>
</tr>
<tr>
<td></td>
<td>Range: 1st to 57th percentile</td>
</tr>
</tbody>
</table>
E. **Location of Students’ Ranking Descriptions**

Of the schools with formal ranking systems, there was variability in where the reader could locate the student’s rank. The majority (n=76, 82%) of medical schools identified the individual student’s rank in the summary section. Other locations included the appendices (n=13, 14%) or another section within the MSPE (n = 4, 4%). Many schools (n=49, 53%) made an effort to highlight the rank by bolding, capitalizing, or underlining the rank.

F. **Location of Ranking System Legend**

The location for the ranking system legend was also varied. Thirty-four (n=34, 37%) schools described their ranking system in appendix D, as suggested by the AAMC. Thirty-eight (n=38, 30%) described their ranking system in another appendix, including the medical school information page. Other locations were on a cover letter (n=6, 6%) or within the body of the MSPE (n=5, 5%). Ten (n=10, 11%) did not fully describe their ranking system anywhere, but a ranking system was inferred by their giving a numerical or quantile rank to their student at some point in the MSPE.

G. **Description of Non-ranking Schools**

Thirty-two (n=32, 26%) schools did not rank their students. Of these schools, 15 (47%) included a statement somewhere in their MSPE that they “do not rank” their students. There were eight (n=8, 25%) schools whose summary statement contained the sentence: “[The student] is an [outstanding/excellent/superior/above average/very good] candidate for residency/post-graduate training.” Of these eight schools, five had
statements of “we do not rank our students.” Other statements related to ranking included “performance has been strong in comparison to his peers” and “his progress throughout medical school has been EXCELLENT.”

We found that a higher percentage (28%, n=9 of 32, 95% confidence interval (CI) 12.5-43.7%) of non-ranking schools were a US News and World Report top 20 medical school than the ranking schools (12%, n=11 of 93, 95% CI 5.2-18.4%, chi-squared test for independence: p=0.03).

H. **Differences from 2013 to 2015**

Five schools that did not rank their students in 2013 ranked their students in 2015. No schools changed from ranking to non-ranking. Five schools in the ranking group decreased their number of groups between 2013 and 2015 and one of these schools changed its descriptor terms. Thus, in the schools for which we had two years of data, 11% had a change in either the ranking practice or in the number of groups utilized in determining ranks. Several schools’ group percentages changed (n=32, 54% of named category group schools). A Wilcoxon signed rank test detected no statistical difference in the percentage of students in the first, second, third, or last category between 2013 and 2015.
IV. DISCUSSION

A. Background and Implications of Ranking Variability

Despite the 2002 AAMC recommendations to standardize MSPEs, there is still considerable variability in their structure and content among US medical schools.\textsuperscript{2,4} Our paper focuses on the ranking systems reported in MSPEs and clearly demonstrates their current variability among US medical schools. Variability in the format of MSPEs may contribute to difficulty by program directors (PDs) in interpreting them. On average, PDs receive 856 applications per year.\textsuperscript{8} And while the recommended length of the MSPE document is two to three pages, it is usually longer.\textsuperscript{1,2} Reading an MSPE for every applicant requires the PD to read thousands of pages each application season. Therefore, it is not surprising that one-third of all PDs do not use the MSPE to make rank-order decisions.\textsuperscript{8} Even among those who use the MSPE to make rank-order decisions, the average PD weighted the MSPE lower than 19 other factors in making rank-order decisions.\textsuperscript{8} Authors of one study found that for any specialty, the MSPE was never a top factor in making selection decisions.\textsuperscript{9}

This lack of importance placed on MSPEs by PDs is a matter of concern, in light of how much time and resources go into creating the MSPE at each institution.\textsuperscript{10,11} The effort involved in the registrar analyzing transcript data, the student summarizing her salient unique characteristics, and the student affairs officer meeting with each student and writing the MSPE should not be minimized. In 1998, student affairs officers estimated an average annual cost of $26,000 per school to produce these letters, although this estimated cost was prior to ERAS and was very roughly estimated.\textsuperscript{12} Assuming a
cumulative inflation rate of 46.4%, the 2015 value of $38,000 would equate to $5.1 million spent annually in the US on production of the MSPEs. Issues in the readability and usability of the MSPE undermine the efforts and resources spent by all parties involved.

The majority of PDs still use the MSPE it in their rank list decisions, even though it is difficult to find and interpret the student performance descriptions. Directors of anesthesiology training programs felt that the student’s academic rank was one of the most important components of the MSPE. Lurie, et al. demonstrated that a higher categorical class rank predicted stronger residency performance. In order to locate a class rank, the PD must identify the sentence containing the student’s rank category, and then locate the legend that describes the ranking system. Both of these items have a variable location and presence within the MSPE. As a result, a PD’s time is inevitably taken away from other important aspects of a student’s application.

Despite the AAMC recommendations, our study demonstrates that a quarter of schools do not provide comparative performance information for their students. For these schools that do not provide comparative performance information, at least one quarter contained a sentence that could suggest the student's rank to an experienced PD. Despite claiming that they have no ranking system, those who write MSPEs are either consciously or unconsciously using the ranking language used by other schools that do provide ranking information.
Each of the most common descriptors used by MSPE writers was associated with a broad range of student percentiles. “Outstanding”, for example, was used, variably among schools to refer to any student in the top two-thirds of their class. “Excellent” was used variably among schools to describe any student, excluding those in the top five percent of their class. While “good” was usually used to refer to the bottom half of students in their class, its use ranged above the median in at least one school.

Some schools did not provide complete information for the size of their ranking groups. For example, five schools did not give sizes or percentiles for any of their ranking groups, while six only provided information for their top group(s). Such variability among use of categories may lead to errors of PD interpretation. One PD may assume that an “outstanding” student was in the school’s top category, when in reality the student was in the bottom half of his class. Furthermore, a school may loosely use the term “outstanding”, when they actually have no ranking system at all.

High-achieving and above average students have the most to lose from variable ranking systems. First, they are indirectly impacted if a PD incorrectly assigns their lower-achieving counterparts to a higher quartile than is deserved (for example, describing a student in the 35th percentile of their class as "outstanding"). Furthermore, PDs may not be able to identify above average students from non-ranking schools, since grade distribution in medical schools is also variable.

It is not clear why some schools choose not to rank their students. In our study, a
larger proportion of non-ranking schools were among the US News and World Report top 20 medical schools than the ranking group.\textsuperscript{7} This may suggest that top medical schools may be of the mindset that “all of our students are outstanding” and do not want to place them into categories. Also, schools that provide ranks most likely have rationales for their particular preference of ranking style. One school may prefer to place a small number of students in the top category and a large number of students in the bottom category, to highlight their outstanding students, without punishing their lower achieving students. Other schools may equally distribute their student groups and provide a more precise description of class rank. Inherently, there is a conflict between the needs of the letter writer and the letter reader.\textsuperscript{16} A student affairs officer may feel compelled to advocate for each student,\textsuperscript{11} while PDs attempt to select the best candidates for their residency program.

A variable use and practice of ranking systems by medical schools makes it difficult for a PD to determine a student’s class rank. Regardless of the ranking method chosen, student affairs officers, students, and PDs would benefit if the MSPEs were written to demonstrate a consistent approach to ranking systems, and certainly in accordance with the AAMC’s mandate in 2002.\textsuperscript{1}

B. Limitations

There were a few notable limitations to this study. First, we only examined one MSPE per academic year and assumed that other MSPEs from the same school would have the same format; however, we minimized this potential issue by analyzing two
different years’ documents. Second, our data abstractors were not blinded to the study hypothesis. Finally, we excluded schools from the group size calculations that did not give sizes for their groups.

C. **Conclusions**

Medical schools had significant variability in their ranking practices. They differed in: presence of a ranking system, type of ranking system, size of ranking groups, description of rank categories, location of the student’s rank in the MSPE, and location in the MSPE of the legend/description of the school’s ranking system. The most common ranking practice was to use named category groups. The most common descriptors for the first, second, third, and last student groups were “outstanding”, “excellent”, “very good”, and “good”, although we identified over 40 different terms used to describe category groups. Each of the top terms used in MSPE student performance descriptions had broad percentile ranges, with each term spanning more than 50 percentile points. While a quarter of schools did not rank their students, many non-ranking schools included language in their summary statements that could suggest rank to an experienced PD.

The variability in ranking practices limits the ability of a PD to interpret the MSPE, make inter-school comparisons, and to separate high-achieving students from low-achieving students. Interpretation of the MSPE is time-consuming and prone to errors. Threats to the usability of the MSPE undermine the tremendous time and resources spent by student affairs officers in the production of the MSPE.
The variability in MSPE ranking systems may contribute to the low perceived value of the MSPE by PDs. Program directors, students, and student affairs officers would benefit from a consistent approach to the ranking of students in the MSPE.
REFERENCES


8. National Resident Matching Program, Data Release and Research Committee: Results of the 2014 NRMP Program Director Survey. National Resident Matching Program,


VITA

NAME: Megan Boysen Osborn (former: Megan Marie Boysen)

EDUCATION:
B.A., Chemistry, University of San Diego, San Diego, California, 2002
M.D., University of California, Irvine, Irvine, CA, 2007
M.H.P.E, University of Illinois, Chicago, Chicago, Illinois 2015

TRAINING
Emergency Medicine Residency, University of California, Irvine Medical Center, Orange, California, 2010
Medical Education Fellowship, Stanford University, Palo Alto, California, 2011
American College of Emergency Physicians (ACEP) Teaching Fellowship, Dallas, Texas, 2011

TEACHING
Residency Program Director, University of California, Irvine Emergency Medicine Residency Program, 2015-Present
Course Director, Clinical Foundations IV, University of California, Irvine, 2011-Present
Associate Residency Director, University of California, Irvine Emergency Medicine Residency Program, 2011-2015
Assistant Professor of Clinical Emergency Medicine, University of California, Irvine Medical Center, 2011-Present
Clinical Instructor, Stanford University, 2010-2011
Chief Resident, University of California, Irvine Emergency Medicine Residency Program, 2009-2010

ABSTRACTS
Boysen M, Arcila M, Luloff L, Fox JC, Scruggs W. Optic Nerve Sheath Diameter in Patients Undergoing Head CT in the Emergency Department. SAEM Western Regional Annual Meeting, March 2006, Redondo Beach, CA
Boysen M, Gharahbaghian L, Fox JC. The Sensitivity and Specificity of Focused Assessment with Sonography for Trauma in Detecting Clinically Significant Injuries Following Blunt Abdominal Trauma in the Pediatric Population. SAEM Western Regional Annual Meeting, March 2006, Redondo Beach, CA
Trauma in the Pediatric Population. SAEM Annual Meeting, May 2006, San Francisco, CA


Boysen MM, Henderson JL, Langdorf MI. Positive Cerebrospinal Fluid Cultures in Immunocompetent Emergency Department Patients with Normal Cell Counts and Gram Stains are All Contaminants. Cal/ACEP Annual Meeting, June 2006, San Diego, CA

Boysen MM, Henderson JL, Langdorf MI. Positive Cerebrospinal Fluid Cultures in Immunocompetent Emergency Department Patients with Normal Cell Counts and Gram Stains are All Contaminants. ACEP Scientific Assembly, October 2006, New Orleans, LA.


Mattson J*, Chakravarthy B, Boysen-Osborn M (senior author). Medical Student Performance Evaluations: Adherence to AAMC guidelines 10 years later. Western Group on Educational Affairs (WGEA) Regional Meeting, Mach 2014, Honolulu, HI. *Won Medical Education Scholarship Research and Evaluation (MESRE) award for Excellence for Outstanding Oral Presentation of an Educational Research in Medical Education


PUBLICATIONS


Coppe JP, Boysen M, Sun CH, Wong BJF, Kang MK, Park NH, Desprez PY, Campisi J, Krtolica A. A Role for Fibroblasts in Mediating the Effects of Tobacco-Induced Epithelial Cell Growth and Invasion. Mol Cancer Res 2008;6(7): 1085-98. PMID: 18644973


