Medical Student Performance Evaluations in 2005: An Improvement Over the Former Dean's Letter?

Judy A. Shea, PhD, Elizabeth O'Grady, Gail Morrison, MD, Barbara R. Wagner, MS, and Jon B. Morris, MD

Abstract

Purpose
To collect information regarding preparation, content, and format of Medical Student Performance Evaluations (MSPEs) and evaluate a sample of 2005 MSPEs to assess compliance with the 2002 Association of American Medical Colleges-issued MSPE guidelines.

Method
Cross-sectional survey with all 126 U.S. allopathic medical schools. Associate deans of students affairs were sent an eight-item questionnaire in June 2006 and asked to submit a sample of redacted MSPEs for 2005 graduates, choosing one from each tertile of the class. Survey data are summarized. MSPEs were abstracted, and results are presented descriptively.

Results
The survey response rate was 84%. Most associate deans (71%) reported having primary responsibility for composing MSPEs; 78% adhered to the format and content guidelines three fourths of the time. The abstraction of 293 MSPEs (78%) showed that more than 80% adhered to format recommendations. However, only 70% to 80% stated grades clearly, avoided the word *recommendation*, and stated whether the student had completed remediation. Fewer than 70% indicated whether the student had had any adverse actions or provided adequate comparative data. Strikingly, only 17% provide comparative data in the summary paragraph. Overall, 75% of the MSPEs were judged to be "adequate."

Conclusions
MSPEs are somewhat variable in terms of which specific items are included. There has been steady quality improvement since prior surveys, primarily in formatting and labeling. However, a sizable minority of writers are still using the MSPE as a recommendation, and too few are providing helpful comparative data.

Acad Med. 2008; 83:284–291.

The Medical Student Performance Evaluation (MSPE, formerly called the dean's letter) is a summary evaluation of a graduating medical student and has the potential to provide valuable information to residency directors about the comparative medical school performance of applicants.[1] Graduating students send these evaluations along with their transcript and letters of recommendation when they apply for a medical residency program. In most cases, the MSPE is the only place to find some types of information about the student's performance, such as data formally comparing the student with his or her peers, and comments about the student's performance in the clinical setting. Because of the importance of this evaluation in the residency application process, the Association of American Medical Colleges (AAMC) set out specific guidelines for writing the dean's letter in 1989, stating explicitly that it is "not a letter of recommendation; it is a letter of evaluation," and asked that comparative data be included.[2] Furthermore, the AAMC, in subsequent guidelines publication in 2002, changed the name of the dean's letter to Medical Student Performance Evaluation to reflect its purpose as an evaluation, not a recommendation or predictor of future performance.[3]

Previous studies of the former dean's letter have been of two types. In the first, which involved surveys of residency program directors (who are the ultimate consumers of the product), a fairly consistent message emerged that the letters were not especially useful.[4,5] The accuracy of the information included in the letters was called into question,[6] and, in a commentary, the letters were described as a "wonderland of positive adjectives" and a "fantasyland."[7] In fairness, others have found dean's letters to be quite accurate.[8]

The second type of study has been designed as abstracted reviews of dean's letters, sometimes accompanied by a survey of letter writers. Yager and colleagues[9] rated 212 dean's letters from 103 medical schools in the United States for 1981 graduates applying to psychiatry residencies. Overall, 27% of the letters were judged to be "excellent"; 49% "average"; and 24% "poor." In another study conducted in September 1984, the deans of student affairs at all 124 U.S. medical schools that existed at the time were sent a survey to determine what information was included in dean's letters, how they were composed and edited, and the nature of the final summary statements, especially if comparative data were included.[10] Results showed that letters often concluded with a final summary statement evaluating the student (85%). However, only 24% of the letters gave the percentages of students at the same institution falling into comparative categories, such as *outstanding*, *excellent*, *very good*, and *good*. More than half of the respondents indicated that class rank was never included.

In early 1992, a survey was administered to the letter writers at all of the then 125 U.S. medical schools; 550 letters for the

Please see the end of this article for information about the authors.

Correspondence should be addressed to Dr. Shea, University of Pennsylvania, 1223 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021; telephone: (215) 573-5111; fax: (215) 573-8778; e-mail: (sheaja@mail.med.upenn.edu).

1992 graduating class were collected and rated according to the 1989 guidelines published by the AAMC.[11] Only 38% of schools introduced their letters as "letters of evaluation." Fifteen percent of the schools failed to use the AAMC guidelines for format. When the schools were rated for overall quality of format and the presence of comparative data elements, 55% of the schools passed and 45% failed. A follow-up study involved a survey administered to the writers of dean's letters incorporating items from two previous surveys (1981 and 1992) and provided a separate analysis of the content of the letters for 1998 graduates.[1] Many improvements were noted. Adequate letters were produced by 65% of schools on the basis of the 1989 AAMC guidelines, compared with 55% of schools in 1992. Notably, schools were organizing the letters into more readable formats, and more schools were providing comparative data to allow the comparison of medical students with their peers. However, more than one third of U.S. medical schools still produced unacceptable letters in 1998, nearly 10 years after the original AAMC guidelines were published.[2]

The updated guidelines, published in 2002, seek to ensure consistency in the redesigned and renamed MSPEs, specifically in evaluating the student and providing comparative data of the student relative to his or her peers, preferably via easy-to-interpret graphs. The 2002 guidelines also reaffirm the MSPE's purposes, which include improving the collaboration and communication between senders and recipients of the MSPE, and establishing an ongoing quality-improvement process. The guidelines suggest that the MSPE should be a two- to three-page, single-spaced document with six distinct, labeled sections: Identifying Information, Unique Characteristics, Academic History, Academic Progress, a Summary Statement, and Appendices. Of note, the summary section of the assessment should contain the student's performance relative to his or her peers, including information about any comparative categories used by the school (e.g., outstanding, excellent, very good, good). Lastly, the Appendices section should include five appendices and should contain graphical representations of the student's performance relative to his or her peers, as well as specific information about, and policies in place at the particular medical school, to provide a context in which to help interpret the MSPE.

The objective of our study is to assess compliance with the 2002 AAMC MSPE guidelines, using methods similar to those reported in earlier studies of the dean's letter: surveys administered to the associate deans of student affairs of all 126 U.S. medical schools and simultaneous review of current MSPEs. One difference between our study and previous ones is that we asked student affairs deans for a sample of three MSPEs, one each from the upper, middle, and lower tiers of students. Earlier studies primarily selected letters from files of multiple residency directors. Our specific objectives were to collect information regarding the preparation, content, and format of MSPEs and to evaluate a sample of 2005 MSPEs to assess compliance with guidelines put forth in the 2002 AAMC *Guide to the Preparation of the Medical Student Performance Evaluation*.[3] Our goal was to determine the quality of more recent MSPEs as compared with the results of previous studies, and to identify areas for continued improvement.

Method

This national study was originally requested by members of the AAMC Medical Student Performance Evaluation Advisory Committee, an advisory appointed by the AAMC president and initiated by the Group on Student Affairs (GSA) at their February 2006 meeting. AAMC staff provided names and contact information for the associate deans of student affairs at all 126 AAMC-member U.S. medical schools. In June 2006, the associate dean of student affairs at each school was mailed a packet that included a letter of invitation and a study description, an eight-item questionnaire that could be completed electronically or on paper, and a request for a sample of redacted MSPEs, one from each of three tiers of students (upper, middle, and lower). Participants could return their responses by regular mail in a prestamped envelope or electronically to an e-mail account established for the study. An endorsement memo from the AAMC GSA Steering Committee was also included in the packet. We sent a second mailing and e-mail communication to nonresponders in July 2006, and we sent a final e-mail and made targeted phone calls to nonresponders in September/October 2006. The institutional review boards at the University of Pennsylvania and the AAMC approved the study.

Instruments

In the questionnaire, we asked respondents to explain how MSPEs were composed and by whom, the student's role in the MSPE process, and how professionalism issues were addressed in the MSPE (the latter topic is not included in this manuscript). We developed the abstraction form for MSPEs to parallel the 2002 guidelines published by the AAMC. Following those guidelines, the major sections of the abstraction form focused on label and appearance, content, and comparative performance data. Raters indicated whether the targeted piece of information was present or absent. For a few items, they assigned a code of *partial*. The abstraction form in this study was quite different from those in earlier studies. In earlier studies, raters made more global assessments for sections of the dean's letter, and dean's letters were assigned grades of *honors*, *pass*, and *fail*.

MSPE letter abstraction and analysis

We assigned each school a study number and coded all materials to protect school and student confidentiality. Most MSPEs we received from participating institutions were received as redacted documents. When we received the documents, we reviewed and redacted them further to remove letterheads and logos, identifying regional programs, and recognizable hospital and physician information. A team of four raters was trained to use the abstraction form; each rater received approximately three hours of training. We sorted MSPEs into sets that contained at most one MSPE per school per set. Two individuals independently rated each set of letters. The two raters met to review ratings, confer when there was disagreement on a particular item, and come to consensus. The number of disagreements among pairs of abstracted forms with 57 unique elements ranged from 0 to 16, with a median of 4, or 7.2% overall. Raters discussed elements about which there were large numbers of disagreements with one of the investigators (J.A.S.), and, in most cases, the element definition was refined to add clarity, then abstractions for that element

were redone using the revised coding scheme. After an initial review of the data with the Medical Student Performance Evaluation Advisory Committee, we created a brief revised abstraction form, and all MSPEs were reabstracted to provide greater detail on some data elements. All data were analyzed using SAS version 9.1 statistical software (SAS Institute, Inc., Cary, NC). We present overall percentages and other descriptive statistics below.

Results

We received questionnaire responses from 106 (85%) of the 126 schools. Ninety-eight (78%) schools sent samples of redacted letters, though one school provided only two MSPEs instead of the requested three, for a total of 293 MSPEs. Ninety-two schools (74%) returned all requested materials. There were no differences in response rates between public (59 of 78, or 76%) and private schools (37 of 47, or 79%) or by any other discernable school characteristic. One school (Uniformed Services University of the Health Sciences) was not classified as public or private, but it returned the survey and the sample MSPEs.

Survey

According to the 106 questionnaire responses, primary responsibility for creating the MSPE rests with student affairs (71%) and academic affairs (7%) deans. At another 5% of the schools, primary responsibility rested with a faculty team. Two percent of respondents gave responsibility to the student's advisor. Most respondents (60%) indicated that the actual composition of MSPEs at their medical school was completed primarily by a faculty member or administrator. Only 15% of MSPEs were composed by a team of faculty/administrators. Fourteen percent of MSPEs were composed primarily by an administrative staff person and then reviewed/approved by a faculty member or administrator. Others who authored MSPEs were student advisors (3%) and advisory deans. Two thirds (66%) reported that students review the MSPE as recommended by the AAMC guidelines (i.e., only for the accuracy of the factual data presented). However, 32% reported that students review the MSPE for accuracy of data as well as other content.

Seventy-eight percent of the MSPE writers indicated that three quarters or more of their MSPEs adhered to guidelines regarding format, content, and comparative data. Nearly all reported implementing the sections recommended by the AAMC: Unique Characteristics (88%), Academic History (91%), Academic Progress–Preclinical (94%) and –Clinical (92%), and a Summary Statement (92%). Graphical data as appendices were reportedly used by 63% for basic science performance, 83% for clinical clerkship performance, and 42% for overall comparative data. Lastly, 86% of respondents claimed to include a medical school information page, detailing the evaluation policies and characteristics of the individual medical school.

Most (69%) respondents indicated that narrative comments about students' clinical rotations were reproduced exactly as they were written, with corrections only for spelling and/or punctuation, although 25% reported that they provided a summary of these comments. Six percent reproduce comments exactly as written, without any corrections. Comments about professionalism were typically (77%) embedded within comments about other activities, although 13% addressed them in a distinct Professionalism section.

Abstracted MSPE letters

Table 1 provides the percentage of the 293 MSPEs we received that met each element of the 2002 AAMC recommendations. These results are detailed here.

Format. Compliance with format recommendations was generally high. Almost all (92%) were labeled as an MSPE or "evaluation." The median length of the MSPE was four pages (range: 2–12), and the median length of the included appendices was three pages (range: one to nine). In most MSPEs, the five suggested sections plus appendices were present and labeled. Moreover, the prevalence of specific sections observed in the abstraction was consistently a little lower than those reported above by the survey respondents, perhaps because abstractors were looking for a distinct, labeled section rather than reporting only whether certain content was present. Only 78% of the MSPEs included the location of the medical school.

Unique Characteristics. The Unique Characteristics section contained a median of 16 lines per letter, with a range of 0 to 123 lines. Information from the student's medical school years was about twice as long (median 10 lines) as premedical school information (five lines).

Academic History. The 2002 AAMC guidelines provided a list of topics that should be included in a complete discussion of a student's history, and it also provided a template for summarizing the information. The relatively low percentages we observed for mention of topics from the list, such as gaps/leaves of absence (58%), adverse actions (63%), and dates in and out of other degree programs (32%), presumably occurred because these events were not mentioned when they did not occur. Writers constructed the Academic History section fully by template in 41% of the MSPEs and partially by template in 23%.

Academic Progress. The Academic Progress section of the MSPEs consistently included a narrative about the student's basic science years (97%) and avoided a course-by-course description (95%), in keeping with the 2002 recommendations. The median length of the basic science performance description was six lines (range: 0–52 lines). Only 70% of MSPEs explicitly stated or made clear that the information regarding the student's clinical clerkships was presented in chronologic order. Almost all MSPEs followed recommendations regarding provision of a brief narrative describing each of the student's core clerkships.

Summary paragraph. In nearly all (98%) of the summary paragraphs, a summative evaluation of the student was presented. However, the MSPEs were quite diverse in what was said about a student. In 41% of the MSPEs there were adjectives and descriptors, but no sense of summary categories (e.g., "he is a very motivated student, works hard and is a fast learner"). (In 5%, however, there was actually a very detailed account of comparative categories in the appendices that was not mentioned in the summary paragraph.) Twelve percent presented bolded or underlined summary descriptors but did not provide the full range of descriptors or the percentage of students from that institution in each category. Overall, only 17% presented

Table 1

Percentage of Abstracted Sample Medical Student Performance Evaluations (MSPEs) from 2005 Containing Abstracted Elements*

Element	No. (%)
Format†	
Labeled as MSPE	271 (92)
Single spaced	293 (100)
Times New Roman font	250 (85)
At recommended length of two to three pages	105 (36)
Identifying Information section	215 (73)
Unique Characteristics section	216 (74)
Academic History section	242 (83)
Academic Progress section	238 (81)
Summary section	254 (87)
Appendices (includes 70 that were present but not labeled)	251 (86)
Identifying information§	
City and state of medical school	228 (78)
Academic History section	
Expected graduation data	274 (94)
Matriculation date	283 (97)
Gaps/leaves/absences	170 (58)
Repeats/remediations	209 (71)
Adverse actions	185 (63)
Dates in and out of other degree programs	95 (32)
Academic Progress–Preclinical	
Narrative on basic science years	284 (97)
Avoid course-by-course description	278 (95)
Academic Progress–Clinical	
Clerkship information in chronologic order	204 (70)
Narrative on each of six core clerkships	271 (92)
Narrative about student's enthusiasm and motivation	287 (98)
Narrative about compatibility with faculty, team members, peers, and patients	287 (98)
Summary	
Comparative data	51 (17)
NOT a recommendation	179 (61)
Appendices‡	
Overall comparative data available somewhere in appendices	217 (74)

* Associate deans for student affairs at all 126 Association of American Medical Colleges member schools in the United States submitted three MSPEs (one each from the upper, middle, and lower class tiers) for analysis. One dean submitted only two MSPEs, for a total of 293 from all schools.
† Font size (12 point) and margins (1 inch) were not formally analyzed, because the letters were received in multiple formats.
§ Student's legal name and name of school were redacted.
‡ Provided with 251 (86%) MSPEs.

some type of comparative data in the summary paragraph, such as quartiles (11%) or clearly stated categories such as outstanding, excellent, very good, and good (6%). In an additional 30%, comparative data were present, but the information required to fully interpret the category was located in the appendices. Despite the focus of the guidelines on *evaluation*, 39% of the MSPE summary statements officially contained a *recommendation*.

Appendices. We evaluated the 85% of MSPEs that included appendices to determine what data were included. Results are shown in Table 2. Many MSPEs had appendices with the institution's total group performance provided as a graph or similar distribution, but the individual student's data were not shown. The reader needed to return to the main body of the MSPE to find the student level performance data. Overall, the topic which most often was included in appendices was information about the student's clinical performance (78%). Most MSPEs (79%) also included a medical school information page.

Summary student assessment in MSPEs. After abstracting the individual elements of the MSPE, abstractors made a summary classification: on the basis of all available information, would the reader know how the student performed in relation to peers? Mutually exclusive categories are provided in Table 3. Overall, 32% of the MSPEs provide only (usually positive) descriptors about a student's performance. Often, words were used to suggest comparative categories (e.g., "this is a superior student"), but the full range of categories or distribution of categories was not given. Smaller numbers of MSPEs provided (with boldface type or italics) a set of labels or categories, but either the distribution was not provided (14%) or the writer noted the number of students in only the top one or two categories (7%). Almost half provided clear and complete summative, comparative data through rank/quartiles (20%), or a combination of a labeling system with a noted distribution (25%).

In general, we observed two types of labeling systems for comparative categories. One system used words such as *outstanding*, *excellent*, *very good*, and *good*. However, the lowest category might be called *satisfactory*, *marginal*, *solid*, or *qualified*. *Superior* was a particularly tricky adjective, appearing anywhere in the top three categories. A second system used some variation of the word *recommend*. For example, one might recommend "enthusiastically, strongly, without reservation, and with confidence," or "most highly, very highly, and highly." The distributions aligned with the categories were also variable. Though about one in five used quartiles, other schools used a normal distribution, and still others assigned almost all students to the top two categories.

Overall view of MSPEs. We abstracted a total of 21 elements, which became part of criterion scores. The median number of elements found in an MSPE was 17

Table 2

Data Contained in the Appendices of Abstracted Sample Medical Student Performance Evaluations (MSPEs)* from 2005

Performance component	Graphs showing institutional and student data, no. (%)	Graphs showing institutional data only, no. (%)	No graphs, no. (%)	School policy precludes comparing students, no. (%)	Student data provided elsewhere, no. (%)
Preclinical	99 (41)	56 (23)	37 (15)	27 (11)	20 (8)
Clinical	137 (55)	89 (36)	13 (5)	3 (1)	6 (2)
Professional attributes	23 (9)	4 (2)	193 (77)	3 (1)	27 (11)
Overall performance	80 (32)	46 (19)	92 (37)	18 (7)	14 (6)

* Associate deans for student affairs at all 126 Association of American Medical Colleges member schools in the United States submitted three MSPEs (one each from the upper, middle, and lower class tiers) for analysis. One dean submitted only two MSPEs, for a total of 293 from all schools. Abstracts were provided with 251 of the 293 (86%) submitted MSPEs.

(range: 8–21). Overall, five MSPEs (1.7%) had all elements, and 14 (4.8%) had all but one. As shown in Table 4, in no areas except Identifying Information and Preclinical Curriculum did the majority of MSPEs meet all recommendations. However, when we relaxed standards to be either all elements or all elements except one when there were multiple elements, most MSPEs met recommendations. The Academic History and Summary Paragraph sections, at 68% and 67%, remained the lowest.

Our final analyses looked jointly at overall format and provision of comparative data, consistent with the earlier studies. We assigned individual MSPEs to a category based on format and provision of comparative data. For overall format, we classified MSPEs as *fail* if three or more categories in Table 4 were *few* (7%); we classified MSPEs as *honors* if *all or most* was assigned to each category in Table 4 (8%); we classified everything else as *pass* (85%). For provision of comparative data, we classified as *adequate* MSPEs in which, between the body of the document and the appendices, it was possible to compare the student's performance with his or her peers' in the basic sciences/preclinical curricula and in the clinical curricula, *or* in which there were comparative data with student performance indicated for an overall classification (80%). Table 5 shows cross-tabulation of the format and comparative data indicators, in comparison with earlier studies. Overall, 75% of the 2005 MSPEs were found to be *adequate*.

Table 3

Characteristics of Sample Abstracted Medical School Performance Evaluations (MSPEs) from 2005* That Affected Abstractors' Ability to Compare a Student with His or Her Peers

Characteristic	No. (%)
Only nonspecific adjectives, descriptors provided without comparative data	95 (32)
Comparative category labels† provided, but without percentages of students in each category	43 (15)
Comparative category labels† provided, but only with partial percentages of students in each category	20 (7)
Class rank or quartile provided	60 (20)
Comparative category labels† provided, along with percentages of students in each category	72 (25)
School policy excludes comparative data	3 (1)

* Associate deans for student affairs at all 126 Association of American Medical Colleges member schools in the United States submitted three MSPEs (one each from the upper, middle, and lower class tiers) for analysis. One dean submitted only two MSPEs, for a total of 293 from all schools.
† Examples of category labels are Outstanding, Excellent, Very Good, and Good.

Discussion

Several earlier studies using similar methods of a survey to the associate deans of student affairs in tandem with a review of actual MSPEs suggested that gradual improvements have been made over time and continue to be made. Between 1992 and 1998, the percentage of dean's letters assessed as adequate increased from 55% to 65%.[1] Results of our study suggest a continuing improvement—75% of MSPEs were adequate. Of particular importance is the observation that about half of the MSPEs contained a fully detailed set of descriptive, comparative data. We suspect this is a conservative estimate and that more schools used such data in the actual MSPEs but did not convey it in the sample materials. However, it is also likely that MSPE composers are, at least for lower-performing students, reluctant to provide information that directly compares these students with their peers, because it could disadvantage students seeking spots in the more competitive fields or programs.

Our study has some limitations. First, because the response rates to the survey and request for redacted MSPEs were not 100%, the possibility of bias exists. However, response rates of 85% for the survey and 78% for the MSPEs are really very high for medical literature.[13] Second, associate deans were asked to select their own samples of MSPEs; in some of the earlier studies, samples had been drawn from existing residency applicant pools.[1,11] Thus, it is possible that the student affairs deans may have selectively submitted their "best" letters. Third, we made the assumption that the appendices were the same for all students from a school unless otherwise

Table 4

Percent of Sample Abstracted Medical Student Performance Evaluations (MSPEs) from 2005* Containing Elements Recommended by the Association of American Medical Colleges in 2002

MSPE section and recommended elements	Has all elements, no. (%)	Has all or most elements,† no. (%)	Has some elements,‡ no. (%)	Has few elements,§ no. (%)
Identifying information (1 element): city/state¶	228 (78)	228 (78)		65 (22)
General format (5 elements): at least five or six labeled sections, length of two or three pages, labeled as MSPE, Times New Roman font, single spaced	65 (22)	213 (73)	78 (27)	2 (1)
Academic History (5 elements): expected graduation data, matriculation data, noted if there were leaves/absences/gaps, repeat/remediations, or adverse actions	132 (45)	196 (67)	78 (27)	19 (6)
Academic Progress–Preclinical (2 elements): narrative regarding overall performance; avoid course-by-course description	278 (95)	285 (97)		9 (3)
Academic Progress–Clinical (5 elements): narrative for all core clerkships plus electives; chronologic order, narrative regarding student initiative, enthusiasm; narrative regarding compatibility with faculty, peers, students; grades clearly stated	174 (59)	272 (93)	18 (6)	3 (1)
Summary paragraph (2 elements): does not mention "recommendation," and comparative data are provided within paragraph with labels and/or percentages	32 (11)	198 (68)		95 (32)
Appendices—comparative data (1 element): provided in graphs for basic science/preclinical work *and* each core clerkship, *or* overall performance provided	217 (74)	217 (74)		76 (26)

* Associate deans for student affairs at all 126 Association of American Medical Colleges member schools in the United States submitted three MSPEs (one each from the upper, middle, and lower class tiers) for analysis. One dean submitted only two MSPEs, for a total of 293 from all schools.
† Meets total possible number of criteria or total possible minus 1.
‡ Meets two or three criteria of five possible.
§ Meets zero criteria of one or two possible; meet zero or one criteria of five possible.
¶ Redacted student legal name and school name.

noted. We feel confident in doing so because in the instances when schools attached appendices to all three MSPEs, there were rarely differences among the appendices. In these infrequent cases, schools provided somewhat less-detailed comparative data in the MSPE for the student from the lowest tier.

Overall, the quality and consistency of MSPEs has improved continuously, albeit slowly. A "glass half full" view would suggest there are some areas where MSPEs are particularly strong, such as identifying the MSPE as an evaluation, providing narrative for each core clerkship, using standard formatting, and including appendices. However, a "glass half empty" perspective points to several areas that remain weak. One particularly weak area mentioned earlier is that only 17% provided "stand alone" comparative data in the summary paragraph. Herein lies the Achilles heel of the MSPE: the dean of students must advocate, and the program directors must discriminate. Nomograms with grade distributions provide useful information for the preclinical and clinical courses. However, given that no national objective standards exist for these course grades, with the exceptions of NBME shelf examinations, a meaningful bottom line is best determined by members of the teaching faculty who best understand the intricacies of a given medical school curriculum. We would argue that insofar as the dean of students must advocate on behalf the students, for the dean of students to make a bottom-line determination is an inherent conflict of interest. Ideally, the evaluation process should separate the MSPE writer from the bottom-line determination. One option that would be most helpful to program directors would be an unbiased committee composed of diverse teaching faculty that compares the MSPEs side-by-side with an objective ranking of the students relative to their peers, for example, outstanding (25%), excellent (25%), very good (25%), and good (25%).

Other opportunities for improvement include the excessive length of some MSPEs, and continued use of the term *recommendation*. Arguably, provision of comparative data and clarity that the document is an evaluation, not a recommendation, are among the more important elements in terms of producing a useful document for program directors. In retrospect, it may have been useful to review the abstraction criteria a priori and gather expert opinions about the relative importance of the various components. Certainly it is easy to argue that Program Directors probably prefer to see, for

Table 5

Classification of the Quality of Dean's Letters (1992 and 1998)* and Medical Student Performance Evaluations (2005)† as Reported in Three Studies‡

Classification	1992 (%)	1998 (%)	2005 (%)
Adequate			
"Honors" format/"adequate" comparative data	8	25	7
"Pass" format/"adequate" comparative data	47	40	68
Total adequate	55	65	75
Inadequate			
"Fail" format/"adequate" comparative data	8	4	4
"Pass" format/"inadequate" comparative data	28	30	17
"Fail" format/"inadequate" comparative data	9	1	4
Total inadequate	45	35	25

* *Source*: Hunt DD, MacLaren CF, Scott CS, Marshall SG, Braddock CH, Sarfaty S. A follow-up study of the characteristics of dean's letters. Acad Med. 2001;76:727–733.

† Associate deans for student affairs at all 126 Association of American Medical Colleges member schools in the United States submitted three MSPEs (one each from the upper, middle, and lower class tiers) for analysis. One dean submitted only two MSPEs, for a total of 293 from all schools.

‡ The abstraction form and analytic details differed among studies.

example, performance data for each clerkship and comparative performance data, than to read MSPEs that conform to recommendations for margins, spacing, font, and section headings (though arguably standardizing these features allows the reader to focus on content). Future studies might start with views of Program Directors and query their preferences.

From four previously inexperienced MSPE readers, one of the enduring conclusions is that each school produced an MSPE with a unique look and style. For the most part, the readers informally felt that they "got to know" the students. But there were a small set of MSPEs that were so dense, unstructured, and long that they proved to be very frustrating to abstract. One might think such MSPEs could hurt the residency candidate. Interestingly, there were no quantifiable differences among MSPEs within a school. Occasionally, students in the upper tiers would have a clear "bottom line" label attesting to their strong performance. Differences between the middle- and lower-tier students were usually less apparent and sometimes only detectable when all three letters were reviewed together (after the main abstraction).

This survey and review of the 2005 MSPE process shows that associate deans of student affairs continue to invest many resources annually in producing MSPEs. What can be done to assist the process, that simultaneously leads to higher-quality letters without increasing the burden and/or removing the individuality of the school? One possibility is to publish samples of redacted "good" MSPEs that show a variety of useful styles and formats. Many MSPE writers are new to the task and would welcome this type of modeling of high-quality MSPEs. Another possibility is to create a process to give MSPE writers feedback about the quality of their MSPEs. Ultimately, it will be time to return to program directors and ask what they want—it would at least be interesting to know if the improvements that have been observed from analyses of MSPEs are noted by the end consumers.

The observation that only 6% of the MSPEs included all or all but one of the recommended elements supports several recommendations for future guidelines. First, it is possible that the current recommendation of two to three pages is simply not realistic. It may take more (maybe four or five) to adequately cover all important points in the evaluation while conveying uniqueness data about a student. Second, guidelines can continue to emphasize that some pieces of information, such as the location of the medical school and the fact that clerkships are presented in chronological order, really are important for the reader. Third, the message that the MSPE is an evaluation summary clearly has been heard, but the fact that it is not a recommendation needs continued reinforcement. Fourth, the guidelines should be more explicit in emphasizing the need to state clearly that an item is not applicable or did not occur, rather than leaving it up to the reader to make the assumption, as we discussed with gaps/leaves and absences. Finally, and most importantly, the next version of the guidelines should make it a priority to emphasize the need for comparative data to appear briefly in the summary paragraph as well as appendices.

Acknowledgments

The authors would like to acknowledge the contributions of Austin Morris, Justin Morris, and Amanda Mulholland, as well as the members of the Association of American Medical Colleges Medical Student Performance Evaluation Advisory Committee (in addition to Gail Morrison, MD, chair): David Battinelli, MD, Dennis Boulware, MD, Linda DeCherrie, MD, W. Patrick Duff, MD, Carol MacLaren, PhD, Polly Moss, MEd, Maria Savoia, MD, Henry Schultz, MD, Geoffrey Young, PhD, H. David Wilson, MD, and Robert F. Sabalis, PhD.

Dr. Shea is professor of medicine, Department of Medicine, University of Pennsylvania, and associate dean of medical education research, Academic Programs, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Ms. O'Grady is associate director of evaluation and assessment, Academic Programs, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Dr. Morrison is professor of medicine, Department of Medicine, University of Pennsylvania, and vice dean for education, Academic Programs, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Ms. Wagner is director of student affairs, Academic Programs, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Dr. Morris is associate dean for student affairs, Academic Programs, University of Pennsylvania School of Medicine, and professor of surgery, Department of Surgery, University of Pennsylvania, Philadelphia, Pennsylvania.

References

1 Hunt DD, MacLaren CF, Scott CS, Marshall SG, Braddock CH, Sarfaty S. A follow-up study of the characteristics of dean's letters. Acad Med. 2001;76:727–733.
2 Committee on Dean's Letters, Association of American Medical Colleges. A Guide to the Preparation of the Medical School Dean's Letter. Washington, DC: Association of American Medical Colleges; 1989.
3 Dean's Letter Advisory Committee, Association of American Medical Colleges. A Guide to the Preparation of the Medical Student Performance Evaluation. Washington, DC: Association of American Medical Colleges; 2002.
4 Boyse TD, Patterson SK, Cohan RH, et al. Does medical school performance predict radiology resident performance? Acad Radiol. 2002;9:437–445.

5 Wagoner NE, Suriano JR. Program directors' responses to a survey of variables used to select residents in a time of change. Acad Med. 1999;74:51–58.
6 Edmond M, Roberson M, Hasan N. The dishonest dean's letter: An analysis of 532 dean's letters from 99 U.S. medical schools. Acad Med. 1999;74:1033–1035.
7 Friedman RB. Fantasyland. N Engl J Med. 1983;308:641–653.
8 Vosti KL, Jacobs CD. Outcome measurement in postgraduate year one of graduates from a medical school with a pass/fail grading system. Acad Med. 1999; 74:547–549.
9 Yager J, Strauss DB, Tardiff K. The quality of dean's letters from medical schools. J Med Educ. 1984;59:471–478.
10 Leiden LI, Miller GD. National survey of writers of dean's letters for residency applications. J Med Educ. 1986;61: 943–953.
11 Hunt DD, MacLaren CF, Scott CS, Chu J, Leiden LI. Characteristics of dean's letters in 1981 and 1992. Acad Med. 1993;68: 905–910.
12 Loftus LS, Arnold LK, Willoughby TH, Connolly A. First year residents' performances compared with their medical school class ranks as determined by three ranking systems. Acad Med. 1992;67: 319–323.
13 Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. J Clin Epidemiol. 1997;50: 1129–1136.

Did You Know?

In 2007, with funding from the National Institutes of Health, researchers at the University of Maryland discovered that cocaine exposure in an animal model causes permanent damage to a part of the brain responsible for judgment and learning new behaviors, which may shed light on why drug addicts often relapse after being treated for their addiction.

For other important milestones in medical knowledge and practice credited to academic medical centers, visit the "Discoveries and Innovations in Patient Care and Research Database" at (www.aamc.org/innovations).