

Identifying Medical Students Likely to Exhibit Poor Professionalism and Knowledge During Internship

David L. Greenburg, MD, Steven J. Durning, MD, Daniel L. Cohen, MD, David Cruess, PhD, and Jeffrey L. Jackson, MD, MPH

Uniformed Services University of the Health Sciences, Bethesda, MD, USA.

CONTEXT: Identifying medical students who will perform poorly during residency is difficult.

OBJECTIVE: Determine whether commonly available data predicts low performance ratings during internship by residency program directors.

DESIGN: Prospective cohort involving medical school data from graduates of the Uniformed Services University (USU), surveys about experiences at USU, and ratings of their performance during internship by their program directors.

SETTING: Uniformed Services University.

PARTICIPANTS: One thousand sixty-nine graduates between 1993 and 2002.

MAIN OUTCOME MEASURE(S): Residency program directors completed an 18-item survey assessing intern performance. Factor analysis of these items collapsed to 2 domains: knowledge and professionalism. These domains were scored and performance dichotomized at the 10th percentile.

RESULTS: Many variables showed a univariate relationship with ratings in the bottom 10% of both domains. Multivariable logistic regression modeling revealed that grades earned during the third year predicted low ratings in both knowledge (odds ratio [OR]=4.9; 95%CI=2.7–9.2) and professionalism (OR=7.3; 95%CI=4.1–13.0). USMLE step 1 scores (OR=1.03; 95%CI=1.01–1.05) predicted knowledge but not professionalism. The remaining variables were not independently predictive of performance ratings. The predictive ability for the knowledge and professionalism models was modest (respective area under ROC curves=0.735 and 0.725).

CONCLUSIONS: A strong association exists between the third year GPA and internship ratings by program directors in professionalism and knowledge. In combination with third year grades, either the USMLE step 1 or step 2 scores predict poor knowledge ratings. Despite a wealth of available markers and a large data set, predicting poor performance during internship remains difficult.

KEY WORDS: predicting; intern; professionalism; knowledge; medical education.

J Med Intern Med 22(12):1711–7

DOI: 10.1007/s11606-007-0405-z

© Society of General Internal Medicine 2007

BACKGROUND

Medical educators have a societal obligation to ensure that graduates are competent to practice. However, competence is complex and cannot be measured with a single metric. The American Council for Graduate Medical Education requires that programs certify learners in 6 domains: patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice.¹ Achievement in these domains is not mutually exclusive and acquiring competence in any realm requires proficiency in the others.

Despite decades of medical education research, there is limited data on how well assessment tools used in 1 level of medical education predict performance at subsequent levels of education and future practice. Factors contributing to the paucity of information in this area include difficulty assessing clinical competency, the lack of standard assessment tools between different levels of training, and suboptimal communication between different levels of training and across institutions. The limited literature suggests that measures in 1 domain tend to predict subsequent performance in similar domains.² For example, standardized tests tend to predict future standardized test scores and grades earned at 1 educational level tend to predict grades earned at subsequent training levels.^{2–6} However, the relationship between commonly available measures of academic performance and future competency is largely unknown.² Andriole et al. found that medical school grades and USMLE exam scores predicted program director ratings of surgical interns.⁷

Knowledge of which markers predict poor performance during internship would be useful to medical school admission committees, medical educators, and intern program directors.

This work was presented at the Mid-Atlantic regional SGIM conference at USU on March 9, 2007. It was also presented at the annual SGIM conference as a Lipkin finalist in Toronto, CA on April 27, 2007. It has been accepted for presentation at the USU Research Week conference May 15, 2007 and has been submitted for presentation at the annual Research in Medical Education conference. This was an unfunded research.

The opinions expressed in this paper are those of the authors and should not be construed, in any way, to represent those of the US Army or the Department of Defense.

Received May 12, 2007

Revised August 30, 2007

Accepted September 24, 2007

Published online October 20, 2007

The Uniformed Services University (USU) is unique because our medical school graduates typically enter our residency training programs providing an opportunity to assess student outcomes at the next level of training. Furthermore, our program directors have a vested interest in providing the school with outcome measurements, as about 25% of our graduate medical education (GME) positions are filled with USU graduates.

Using a previously validated program director rating form, we assessed performance at the end of the postgraduate year (PGY)-1 (intern) year.⁸ In this analysis, we found that the 18 survey questions collapsed into 2 domains using factor analysis, which we have labeled professionalism and knowledge. This 2-domain factor analysis has been seen in other studies of physician competence.³ In this study, our goal was to assess whether data available to medical schools predicts poor ratings of knowledge or professionalism by internship program directors at the end of the PGY-1 year.

METHODS

Sample Subjects

The Uniformed Services University offers a traditional 4-year curriculum. Upon matriculation, students become officers in the United States uniformed services (Army, Navy, Air Force, and Public Health Service), and after graduation, receive their GME training from USU-affiliated programs that offer the full range of GME training opportunities at hospitals dispersed across the country. We drew our sample from graduates who received their MD degree from USU between 1993 and 2002. Students who matriculated at USU but did not graduate or enter an internship were excluded.

Study Measurements

We used 3 sources of data. The USU Admissions, Promotions, Registrar's office provided demographic (age, sex, race) and academic information (prior degrees, cumulative grades earned in college and during graduate classes, MCAT scores, USU admission interview scores, USU grades and class rank, USMLE exam scores and attempts, and whether the student was identified as having an academic difficulty, decelerated, or repeated any years while at USU). Use of grades earned at USU was limited to annualized and cumulative annual grade point averages as we did not have course-specific grades. At USU, grades earned during the first and second years largely reflect performance on course-specific tests, performance in small group discussions, and scores on course-specific NBME "shelf" exams. Like most medical schools, grades during third and fourth year clerkship rotations vary from service to service. However, USU grades are largely based on clinical performance. Many clerkships incorporate in-house written exams, national standardized written examinations, and OSCEs with specific weightings of these components varying between different clerkships.

An additional data source was a survey of all graduates from USU between 1980 and 1999. Whereas this survey assessed the student's perspective about the quality of their medical education in a broad range of domains, we restricted our analysis to: experiencing a major life crisis and the student's marital and parental status at both matriculation and gradu-

ation. Graduates were asked if they had experienced a "major life crisis" while enrolled at USU (yes/no) and if so what effect did it have on their academic performance at USU (major, moderate, minimal, or none).

The final data source was the program director rating form completed at the end of the intern year (PGY-1 Survey, Appendix). Survey items were developed by interdepartmental academicians with expertise spanning undergraduate and graduate medical education programs, and is a modification of the American Board of Internal Medicine resident assessment tool. The survey consisted of 18 items scored on a 6-point Likert scale (5=outstanding, 4=superior, 3=average, 2=needs improvement, 1=not satisfactory, 0=unable to judge). In a previous study, we found this form to be both feasible (annual response rate of 72–90%) and reliable (Cronbach's alpha=.96).⁸ In that study, we found that the 18 survey items collapse into 2 domains, professionalism and knowledge.

Statistical Analysis

We performed confirmatory principal components factor analysis with varimax rotation, keeping factors with eigenvalues >1. The 18 items again collapsed into 2 domains that we labeled knowledge and professionalism based on the pattern of factor loads. These domains were scored using a least-squared method so that each variable had a mean of 0 and a standard deviation of 1. Performance in these domains was dichotomized at the 10th percentile. We explored the univariate association between variables in our data set and the domains of professionalism and knowledge using either Student's *t* tests or Pearson's chi-square. We then performed logistic regression analyses on the 2 domains of professionalism and knowledge with the goal of developing models that would predict which medical school graduates would be rated as poor performers in professionalism and knowledge during their intern year. In model building, we a priori selected potential variables that had either prior literature showing a relationship with outcomes or that had face validity for a possible association. These variables included: age, sex, race, marital status, children, undergraduate GPA, MCAT subtest and total scores, attainment of advanced degrees before USU matriculation, USU admission interview scores, grades earned at USU, USU class rankings, USMLE step 1 and step 2 scores and attempts, self-report of a major life crisis while at USU, academic deceleration, repetition of any year, and having an academic difficulty noted in the student's academic record. We tested for potential collinearity (Pearson's and Spearman's correlation coefficients) and avoided simultaneously modeling variables with correlation coefficients >0.6. Model fit was assessed using the Hosmer–Lemeshow test.⁹ All statistical analyses were performed using SPSS (version 14.0).

RESULTS

Demographic information

Our sample was 76% male with an average matriculation age of 25. Eighty-one percent were white, 10% were Asian, and 2% were black. During the students' time at USU, the fraction married and with children increased from 59% to 71% and from 18% to 40%, respectively. Over half of our sample had

Table 1. Ratings of Intern's Knowledge by Intern Characteristics

	Lowest 10% (n=107)	Upper 90% (n=962)	p value
Age at USU matriculation	25.3	24.7	.1
Female	30.0%	23.0%	.14
Race			.022
White	71.0%	81.7%	
Black	4.7%	1.9%	
Hispanic	3.7%	4.4%	
Asian	17.8%	9.5%	
Other	2.8%	2.6%	
Financial significance of attending USU			.15
Very important	75.5%	59.0%	
Somewhat important	22.4%	36.1%	
Not very important	2.0%	3.4%	
Not important at all	0.0%	1.6%	
Marital status at USU matriculation			.06
Single	68.8%	58.0%	
Married	27.1%	40.9%	
Divorced	4.2%	1.1%	
Marital status at USU graduation			.77
Single	30.6%	24.7%	
Married	67.3%	71.5%	
Divorced	2.0%	3.6%	
Widowed	0.0%	0.2%	
Children at USU matriculation	14.0%	19.0%	.4
Children at USU graduation	33.0%	41.0%	.26
Prior military affiliation	45.0%	54.0%	.15
Highest degree before USU			.035
Bachelors	85.0%	90.7%	
Masters	11.2%	8.2%	.03
Doctorate	3.7%	1.0%	.039
MCAT			
Total	29	30	.009
Biological sciences	9.8	10.1	.08
Physical sciences	9.7	10	.12
Verbal reasoning	9.5	9.9	.02
Undergraduate grades			
Cumulative	3.4	3.46	.03
Science	3.36	3.44	.025
Other	3.47	3.51	.27
Average USU interview score	1.9	1.9	.99
Annualized USU grades			
MS1	2.76	3.03	<.001
MS2	2.74	3.1	<.001
MS3	3.03	3.33	<.001
MS4	3.31	3.52	<.001
Cumulative USU grades			
MS1	2.76	3.03	<.001
MS2	2.76	3.06	<.001
MS3	2.86	3.17	<.001
MS4	2.97	3.23	<.001
Clinical science (MS3 and MS4)	3.19	3.43	<.001
Class rank			
After MS1	47.9	36.9	<.001
After MS2	67	49.1	<.001
After MS3	65.2	46	<.001
After MS4	62.8	44.3	<.001
USMLE scores			
Step 1	198.5	210.1	<.001
Step 2	194.2	206.8	<.001
USMLE attempts			
Step 1	1.16	1.03	<.001
Step 2	1.12	1.02	<.001
Academic deceleration for any reason	13.1%	4.8%	<.001
Repeated any year	2.8%	0.8%	.09

Table 1. (continued)

	Lowest 10% (n=107)	Upper 90% (n=962)	p value
Major life crisis during medical school	39.0%	24.0%	.02
Reported effects of major life crisis			0.47
Major effect	31.6%	23.8%	
No effect-moderate effect	68.4%	76.2%	
Academic difficulty noted	24.3%	11.7%	<.001

some type of military affiliation before matriculating at USU and almost 10% of students obtained a masters or doctorate degree before starting medical school. The average cumulative MCAT scores was 29.6. During enrollment, 13% of the students were identified as having an academic difficulty resulting in a mandatory appearance before the Student Promotion Committee. The fraction of students repeating the first, second, and third years of medical school were 1.0%, 0.1%, and 0.3%, respectively. No student in our study repeated the fourth year. USMLE step 1 and step 2 scores averaged 211 and 210 with a mean of 1.04 and 1.03 attempts, respectively. The mean GPA increased from year to year (first year GPA=3.00, fourth year GPA=3.50), whereas the variance decreased (first year SD=0.51, fourth year SD=0.34). Strong correlations were noted between first and second year grades ($r=.73, p<.001$) and USMLE step 1 and step 2 scores ($r=.72, p<.001$). There was a moderate correlation between third and fourth year grades ($r=.49, p<.001$).

Program Director Rating Form

From the classes of 1993–2002, there was a total of 1,559 graduates from USU. We collected intern program director surveys from 1,247 (80%). Of the collected surveys, 1,069 (representing 69% of all graduates) were scored on all items with an answer other than “unable to judge” allowing for inclusion in our factor analysis. Mean scores across items on the program director survey ranged from 3.80 to 4.32 for all items. Median scores on all items were 4 except for “professional demeanor, maturity, and ethical conduct” for which the median score was a 5. Standard deviations for all items ranged from 0.76 to 0.86 (Appendix).

Long-term Career Outcome Survey

Data concerning major life crises and academic deceleration comes from graduates of the 1993–1999 classes. Of the 860 graduates of these classes, 700 (81%) responded to our survey. About 25% of the respondents reported experiencing a major life crisis while enrolled at USU. Of these, 25% felt the crisis had a major effect on their performance. The fraction reporting academic deceleration for any reason was 5.6%.

KNOWLEDGE

Variables associated with increased risk of poor knowledge ratings included (Table 1): lower undergraduate cumulative

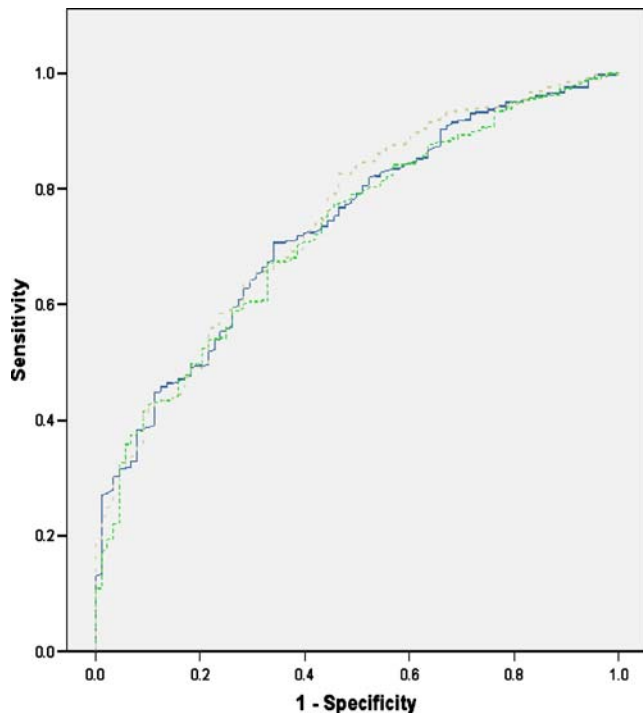


Figure 1. ROC curves for multivariable logistic regression models predicting intern knowledge ratings. *Solid blue line* represents the third year GPA and step 1; area under ROC curve=.735. *Dashed green line* represents the third year GPA and step 2; area under ROC curve=.726. *Dashed gray line* represents the third and fourth year GPA and step 1; area under ROC curve=.746.

and science GPA, lower MCAT total and verbal reasoning scores, lower performance in any academic marker during medical school, more attempts at USMLE exams, and non-white race (RR=1.6, 95%CI=1.1–2.2). Self-report of experiencing a major life crisis (RR=1.6; 95%CI=1.1–2.4), having an academic difficulty noted in the student's USU record (RR=2.1; 95%CI=1.4–3.0), and decelerating (RR=2.7; 95%CI=1.6–4.8) were also associated with elevated risk of being rated in the lowest 10%. Having an advanced degree was associated with a trend toward poor knowledge ratings but did not reach statistical significance (RR=1.6; 95%CI=0.99–2.64).

Despite strong univariate associations between several potential predictor variables and knowledge ratings on multivariate modeling, only the GPA during the clinical years and USMLE exam scores were independent, statistically significant predictors of ratings by program directors. There was significant colinearity between the USMLE step 1 and step 2 scores ($r=.723$, $p<.001$) and moderate colinearity between the third and fourth year grades ($r=.490$, $p<.001$). Incorporating fourth year grades did not significantly improve the area under the ROC curves of our models (Fig. 1). Also, as Program Directors may not have the fourth year grades at the time they select medical students for postgraduate training, we report odds ratios for models using the third year GPA (odds ratio (OR)=4.9; 95%CI=2.7–9.1). Because of the high degree of colinearity, we did not simultaneously use both USMLE exam scores in models. Either predicted knowledge ratings during internship or as step 2 scores may not be available at the time of intern selection, we report models using the USMLE step 1 score (OR=1.03; 95%CI=1.02–1.05). The resulting model was well

fitting (Hosmer–Lemeshow test, $p=.39$), but only modestly predictive (area under ROC curve=.735, $p<.001$).

PROFESSIONALISM

Variables associated with poor professionalism ratings included (Table 2): poor performance in any academic marker during medical school (annual grades, cumulative grades, annual class rank, USMLE exam scores), nonwhite race (RR=1.5, 95%CI=1.04–2.2), earning a masters degree before USU matriculation (RR=1.9, 95%CI=1.1–3.2), and having an academic difficulty noted in the student's academic record (RR=1.8, 95%CI=1.2–2.8).

Again, despite numerous strong univariate associations between potential predictors of poor professionalism in multivariable models, only the third year GPA was an independent predictor (OR=7.29, 95%CI=4.1–13.0). This model (Fig. 2) was also well fitting (Hosmer–Lemeshow test, $p=.72$) and modestly predictive (area under ROC curve=.725, $p<.001$).

DISCUSSION

Despite a wealth of possible predictor variables and a large prospectively assembled cohort, the ability to predict which students will be rated poorly during internship remains difficult. Although many of the candidate variables seemed intuitively appealing and were significant predictors on univariate screening, when we explored ratings using multivariable techniques, no variables contributed significantly to predictive models after controlling for third year GPA and USMLE scores. For the domain of knowledge, the GPA during the third year and the USMLE step 1 score were modestly predictive of poor ratings by program directors. For professionalism, only the third year GPA was predictive.

Based on our findings, we feel that the most important markers for program directors to consider when approaching applicants for the match are the clerkship year grades and USMLE step 1 or step 2 scores. Our data suggests that regardless of performance during college or the basic science years, students who perform competently during the clinical years are likely to continue exhibiting competent performance during internship. It is important to note that demographic factors such as age, race, gender, marital status, and presence of children played no role in intern ratings received by program directors in our study.

Not surprisingly, there was a strong correlation between USMLE step 1 and step 2 scores as well as between the third and fourth year GPAs. Program directors have to make decisions regarding accepting trainees for residency before the fourth year GPA will be available. Fortunately, the third year GPA, among our cohort, was a better marker. The third year grades exhibited a stronger effect size and had better test characteristics as it had a broader range and suffered less grade inflation.

LIMITATIONS

There are several important limitations of our study. Our study includes graduates from only 1 institution. Whereas USU

Table 2. Ratings of Intern's Professionalism by Intern Characteristics

	Lowest 10% (n=107)	Upper 90% (n=962)	p value
Age at USU matriculation	24.8	24.8	.9
Female	19.8%	24.5%	.32
Race			.015
White	72.1%	81.4%	
Black	4.7%	1.9%	
Hispanic	2.3%	4.5%	
Asian	14.0%	10.0%	
Other	7.0%	2.2%	
Financial significance of attending USU			.61
Very important	71.0%	59.9%	
Somewhat important	25.8%	35.3%	
Not very important	3.2%	3.2%	
Not important at all	0.0%	1.5%	
Marital status at USU matriculation			0.58
Single	66.7%	58.5%	
Married	33.3%	40.0%	
Divorced	0.0%	1.5%	
Marital status at USU graduation			.26
Single	25.8%	25.3%	
Married	64.5%	71.5%	
Divorced	9.7%	3.0%	
Widowed	0.0%	0.2%	
Children at USU matriculation	29.0%	18.0%	.16
Children at USU graduation	32.3%	40.7%	.35
Prior military affiliation	51.8%	53.3%	.83
Highest degree before USU			.043
Bachelors	84.9%	90.6%	
Masters	15.1%	7.9%	.026
Doctorate	0.0%	1.4%	.34
MCAT			
Total	30.3	29.8	.22
Biological sciences	10.1	10	.83
Physical sciences	10.5	9.9	0.012
Verbal reasoning	9.8	9.8	.72
Undergraduate grades			
Cumulative	3.47	3.45	.59
Science	3.45	3.43	.45
Other	3.5	3.5	.96
Average USU admission interview score	2	1.9	.12
Annualized USU grades			
MS1	2.82	3.02	.001
MS2	2.86	3.08	<.001
MS3	3.01	3.33	<.001
MS4	3.32	3.52	<.001
Cumulative USU grades			
MS1	2.82	3.02	.001
MS2	2.84	3.05	<.001
MS3	2.9	3.15	<.001
MS4	3	3.23	<.001
Clinical science (MS3 and MS4)	3.18	3.43	<.001
Class rank			
After MS1	45.6	37.3	.001
After MS2	62.9	49.8	<.001
After MS3	63.3	46.5	<.001
After MS4	62.3	44.6	<.001
USMLE scores			
Step 1	204.6	209.3	.018
Step 2	200.7	205.9	.016
USMLE attempts			
Step 1	1.01	1.05	.24
Step 2	1.03	1.03	.71
Academic deceleration for any reason	8.1%	5.4%	.29
Repeated any year	0.0%	1.1%	.32

Table 2. (continued)

	Lowest 10% (n=107)	Upper 90% (n=962)	p value
Major life crisis during medical school	29.0%	24.8%	.6
Reported effects of major life crisis			0.028
Major effect	55.6%	22.6%	
No effect-moderate effect	44.4%	77.4%	
Academic difficulty noted	22.1%	12.2%	.009

draws students from all over the United States, students interested in attending a military medical school may be inherently different from those who choose to attend a civilian medical school, for example, the preponderance of white, male, and married students with children. Conversely, whereas all graduates came from a single medical school, they all participated in fully accredited training programs scattered all over the United States and were distributed across the entire range of residency training possibilities.

Second, we excluded students who failed to pass the USMLE step exams and graduate from medical school. Poor performance during the basic science years is the most common reason for disenrollment from USU. Hence, we excluded the most poorly performing medical students. This may explain why basic science grades were not associated with our multivariable logistic regression outcomes. However, as one cannot enter residency training without graduating medical school, these excluded students who would not be eligible

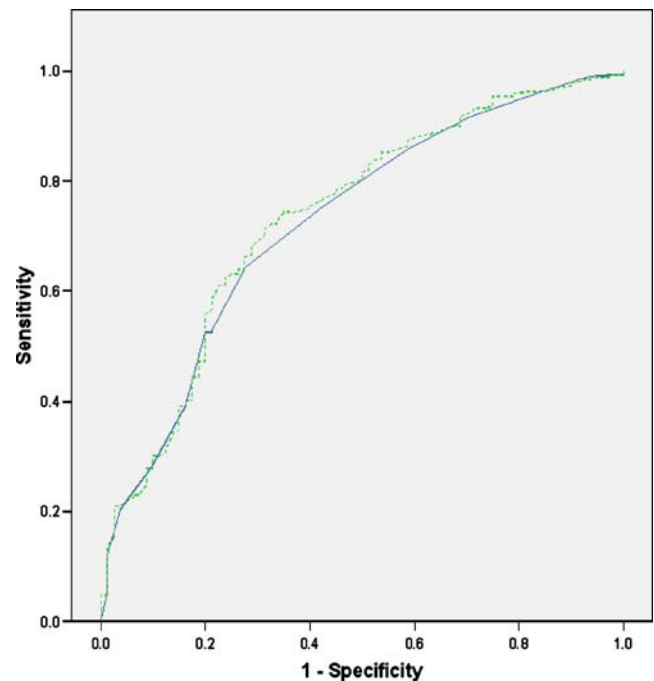


Figure 2. ROC curves for multivariable logistic regression models predicting intern professionalism ratings. Solid blue line represents the third year GPA; area under ROC curve=.725. Dashed green line represents the third and fourth year GPA; area under ROC curve=.735.

for GME selection and do not affect the implications of our study for program directors.

A third limitation involved the decision to exclude interns for whom the program director did not rate the intern on all items. We chose to exclude these students as we felt that these items were not likely to have been scored "unable to rate" randomly and that attempts at imputing the data could result in biased results. This decision reduced our sample by 14% from 1,247 to 1,069. Retrospectively, we evaluated for the effects of excluding these students by imputing responses for students who were scored "unable to rate" on an item by substituting the average score for this item from among the population of students who had been excluded for scores of "unable to rate" on other survey items. We performed multivariable logistic regression analyses using the imputed data. No new predictive markers were found and the odds ratios associated with the imputed responses did not change significantly. Furthermore, grades and USMLE scores for students excluded for receiving a score of "unable to rate" were not significantly different from those students included in our analysis. Therefore, we feel that the decision to exclude these students did not affect our results.

A fourth limitation is the surprisingly small number of markers we found which were independently predictive of performance during internship. We were concerned that our decision to dichotomize performance at the 10th percentile may have limited the ability of our analysis to detect markers that may truly have predictive abilities. To assess this, we did several sensitivity analyses. We varied our threshold from the 5th to 25th percentile and found similar results. We also compared the lowest 10th percent against the top 40th percent. Despite these adjustments, no new predictive markers were found.

A fifth limitation is that the last class evaluated in our study graduated in 2002. As the USMLE added the successful completion of an objective structured clinical exam (step 2 clinical skills examination) for medical licensure in 2004, we cannot assess whether the OSCE would have additional predictive power.

A sixth limitation is that we examined only the quantitative aspects of the intern program director survey. Qualitative information may be more sensitive for detecting individuals likely to exhibit poor professionalism.¹⁰⁻¹² Future works will examine the qualitative information from the PGY-1 Program Director surveys.

A seventh limitation is that we only had grades by year. The lack of clerkship-specific grades makes it impossible to look at a potential relationship between performance on a specific clerkship and subsequent performance in a specific internship.

Finally, our results pertain to predicting performance of physicians at the end of their intern year. Our results may not apply to predicting performance beyond internship.

CONCLUSION

Despite our limitations, a number of important conclusions emerged from this large, comprehensive longitudinal database of our graduates. First, competency during internship is hard to predict with commonly used data collected before and during medical school. Despite the broad range of potential predictive variables available, our predictive ability was only modest. Secondly, the USMLE and clinical year GPA are the

best predictors that program directors currently have to identify interns who may perform poorly in professionalism and cognitive domains. Third, our paper has implications for medical educators and program directors. As our multivariable analyses suggest that the many candidate variables evaluated in our study do not predict performance after controlling for the third year grades and USMLE exam scores, we feel that applicants for internship should be evaluated most heavily on their performance during their clinical clerkships and on the USMLE exams.

Acknowledgements: This project was funded by the Uniformed Services University as part of a quality improvement project.

Conflicts of Interest: None disclosed.

Corresponding Author: David L. Greenburg, MD; Medicine-EDP, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814, USA (e-mail: dgreenburg@usuhs.mil).

APPENDIX

PGY-1 Program Director's Survey

5=outstanding, 4=superior, 3=average, 2=needs improvement, 1=not satisfactory, 0=unable to judge

- Initial histories and physicals, and daily patient evaluations
- Technical skills
- Oral communicative skills
- Written communicative skills
- Fund of basic science knowledge
- Fund of clinical science knowledge
- Analysis of clinical data, differential diagnosis, and management plans
- Scholarly approach to patient management
- Clinical judgment
- Quality of medical records
- Efficiency in patient management
- Relationships with patients and families
- Relationships with peers, staff, and other health care professionals
- Effectiveness and potential as a teacher
- Initiative, motivation, conscientiousness, and attitude
- Professional demeanor, maturity, and ethical conduct
- Demonstrated leadership ability
- Overall performance

REFERENCES

1. Principles of good medical education and training. Postgraduate Medical Education and Training Board. General Medical Council; 2005. http://www.gmc-uk.org/education/publications/gui_principles_final_1.0.pdf. Accessed February 25, 2007.
2. Hamdy H, Prasad K, Anderson MB, et al. BEME systematic review: predictive values of measurements obtained in medical schools and future performance in medical practice. *Med Teach*. 2006;28(2):103-16.
3. Durning SJ, Cation LJ, Jackson JL. The reliability and validity of the American Board of Internal Medicine Monthly Evaluation Form. *Acad Med*. 2003;78(11):1175-82.

4. **Huff KL, Koenig JA, Treptau MM, Sireci SG.** Validity of MCAT scores for predicting clerkship performance of medical students grouped by sex and ethnicity. *Acad Med.* 1999;74(10 Suppl):S41-4.
5. **Julian ER.** Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med.* 2005;80(10):910-7.
6. **Silver B, Hodgson CS.** Evaluating GPAs and MCAT scores as predictors of NBME I and clerkship performances based on students' data from one undergraduate institution. *Acad Med.* 1997;72(5):394-6.
7. **Andriole DA, Jeffe DB, Whelan AJ.** What predicts surgical internship performance? *Am J Surg.* 2004;188(2):161-4.
8. **Durning SJ, Pangaro LN, Lawrence LL, Waechter D, McManigle J, Jackson JL.** The feasibility, reliability, and validity of a program director's (supervisor's) evaluation form for medical school graduates. *Acad Med.* 2005;80(10):964-8.
9. **Hosmer D, Lemeshow S.** *Applied Logistic Regression.* New York, NY: Wiley; 2000.
10. **Papadakis MA, Hodgson CS, Teherani A, Kohatsu ND.** Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. *Acad Med.* 2004;79(3):244-9.
11. **Papadakis MA, Loeser H, Healy K.** Early detection and evaluation of professionalism deficiencies in medical students: one school's approach. *Acad Med.* 2001;76(11):1100-6.
12. **Papadakis MA, Teherani A, Banach MA, et al.** Disciplinary action by medical boards and prior behavior in medical school. *N Engl J Med.* 2005;353(25):2673-82.