Information Collected During the Residency Match Process Does Not Predict Clinical Performance

Stephen M. Borowitz, MD; Frank T. Saulsbury, MD; William G. Wilson, MD

Objective: To determine whether information collected during the National Resident Matching Program (NRMP) predicts clinical performance during residency.

Methods: Ten faculty members rated the overall quality of 69 pediatric house officers as clinicians. After rating by the faculty, folders were reviewed for absolute rank on the NRMP match list; relative ranking (where they ranked in their postgraduate year 1 [PGY-1] group); scores on part I of the National Board of Medical Examiners (NBME) examination; grades during medical school pediatrics and internal medicine rotations; membership in the Alpha Omega Alpha Medical Honor Society; scores of faculty interviews during intern application; scores on the pediatric in-service examination during PGY-1; and scores on the American Board of Pediatrics certification examination.

Results: There was substantial agreement among faculty raters as to the overall quality of the residents (agreement rate, 0.60; $\kappa = 0.50$; $P = .001$). There was little correlation between faculty ratings and absolute ($r = 0.19$; $P = .11$) or relative ($r = 0.20$; $P = .09$) ranking on the NRMP match list. Individuals ranked in the top 10 of the match list had higher faculty ratings than did their peers (mean ± SD, 3.66 ± 1.22 vs 3.0 ± 1.27; $P = .03$), as did individuals ranked highest in their PGY-1 group (mean ± SD, 3.88 ± 1.45 vs 3.04 ± 1.24; $P = .03$). There was no correlation between faculty ratings and scores on part I of the NBME examination ($r = 0.10$; $P = .49$) or scores on the American Board of Pediatrics certification examination ($r = 0.22$; $P = .11$). There were weak correlations between faculty ratings and scores of faculty interviews during the intern application process ($r = 0.27$; $P = .02$) and scores on the pediatric in-service examination during PGY-1 ($r = 0.28$; $P = .02$). There was no difference in faculty ratings of residents who were elected to Alpha Omega Alpha during medical school (mean ± SD, 3.32 ± 1.21) as compared with those who were not (mean ± SD, 3.08 ± 1.34) ($P = .25$).

Conclusions: There is significant agreement among faculty raters about the clinical competence of pediatric residents. Medical school grades, performance on standardized examinations, interviews during the intern application process, and match-list ranking are not predictors of clinical performance during residency.

Arch Pediatr Adolesc Med. 2000;154:256-260

Editor's Note: This is definitely a "we see that" article. So why do we waste so much time on this process?

Catherine D. DeAngelis, MD

FROM THE perspective of a residency training program, the principal goal of the National Resident Matching Program (NRMP) is to identify medical students who will perform well during residency. The faculty of residency training programs expend considerable time and effort evaluating potential residents.[1,2] During the evaluation process, intern selection committees typically gather data on medical school performance through dean's letters, letters of recommendation, medical school transcripts, standardized test scores, formal interviews, and personal statements.[2]

It is not clear whether an applicant's match rank or the data collected to determine the rank correlate with performance during residency training.[2,3] Studies examining the relationship between medical school performance and performance during residency have yielded inconsistent results.[4] Some studies have demonstrated a positive relationship between medical school performance and performance during residency,[5-11] whereas others have found that no objective or subjective factors seem to predict performance during residency.[12-15]

The present study was designed to examine whether NRMP match ranking

From the Department of Pediatrics, University of Virginia Health Sciences Center, Charlottesville.

SUBJECTS AND METHODS

POPULATION STUDIED

Sixty-nine pediatric residents who completed all 3 years of pediatric residency training at the University of Virginia, Charlottesville, during a 7-year period were studied. All of the study subjects were pediatric residents and all had completed training at the time of the study. Five of the 69 residents completed fellowship training at the University of Virginia and 3 of these house officers joined the pediatric faculty.

EVALUATION OF CLINICAL PERFORMANCE

Ten faculty members from a wide variety of pediatric disciplines were asked to retrospectively rate the overall quality of these 69 house officers as clinicians. Faculty raters included both general pediatricians and pediatric subspecialists who had frequent and in-depth contact with house staff in a variety of clinical settings (inpatient, outpatient, emergency department, intensive care units, and rehabilitation). Faculty rated the house officers using a 5-point scale and were asked to consider the house officer's knowledge, technical skills, maturity, and individual judgment. A score of 1 indicated that the resident was in the bottom 20% of this group and a score of 5 indicated that the resident was in the top 20% of this group. All 10 raters were full-time faculty during the entire residency of all 69 house officers and every faculty member had frequent contact with all of the residents. Faculty raters were blinded as to the other faculty members' ratings as well as to the data contained in the house officers' folders.

FOLDER REVIEW

The following information was gathered from residents' files: (1) absolute rank on the NRMP matching list; (2) relative ranking on the NRMP list (where the house officers ranked in their individual intern group, ranging from 1st to 12th); (3) score on part I of the National Board of Medical Examiners (NBME) examination; (4) grade during third-year medical school pediatrics and internal medicine rotations; (5) membership in the Alpha Omega Alpha (AOA) Medical Honor Society; (6) scores of faculty interviews during the intern application procedure; (7) score on the pediatric in-service examination during the first year of residency; and (8) score on the American Board of Pediatrics certification examination.

During the intern application process, all applicants were interviewed by 2 members of the pediatric faculty (all pediatric faculty members participated in the interview process). Interviews were scored on a 6-point scale, with 1 being unacceptable and 6 being superior. Because of the diversity of grading systems at different medical schools, all medical school grades were converted to a 3-point scale, with 1 being equivalent to "pass" or "C" and 3 being equivalent to "honors" or "A."

STATISTICAL METHODS

Agreement among faculty raters was assessed with the multirater κ statistic for categorical variables.[16] Linear regression analysis was used to compare faculty ratings with NRMP rankings, standardized examination scores, and interview scores. Additional comparisons were performed with either unpaired t tests or analysis of variance. Differences were considered significant if $P<.05$.

---

or information collected during the application process to our pediatric residency training program at the University of Virginia was predictive of overall clinical performance during the 3 years of residency. We compared NRMP match ranking; medical school achievements; performance on standardized examinations prior to, during, and after residency training; and interviews during the intern application process with aggregate performance during the 3 years of residency training as assessed by clinical faculty.

RESULTS

There was substantial agreement among the 10 faculty raters as to the overall quality of the 69 residents (agreement rate, 0.60; κ = 0.50; P = .001). There was no significant difference in faculty ratings from year to year, suggesting that faculty were no more likely to highly rate residents who had completed the program in the distant past than residents who had just completed their training (F = 0.64; P = .70).

There was little correlation between faculty ratings and the absolute ranking on the NRMP match list (**Figure 1**) (r = 0.19; P = .11). However, those house officers who were ranked in the first 10 places of the original NRMP rank list for their respective years had higher faculty ratings than did their peers (mean ± SD, 3.66 ± 1.22

vs 2.99 ± 1.27) (t = 1.93; P = .03). Similarly, there was little correlation between faculty ratings and relative ranking on the NRMP match list (**Figure 2**) (r = 0.20; P = .09), although those residents who were ranked the highest in their postgraduate year 1 (PGY-1) group had higher faculty ratings than did their peers (mean ± SD, 3.88 ± 1.45 vs 3.04 ± 1.24) (t = 1.86; P = .03). Those residents who were ranked the lowest in their PGY-1 group had no lower faculty ratings than did their peers (mean ± SD, 2.89 ± 1.18 vs 3.15 ± 1.16) (t = 0.57; P = .57).

Residents who were elected to AOA during medical school had no higher faculty ratings than did those who were not members (mean ± SD, 3.32 ± 1.21 vs 3.08 ± 1.34) (t = 0.68; P = .25). Residents who received an A or equivalent grade during their third-year medical school rotation in pediatrics had no higher faculty ratings than did those residents with lower grades (mean ± SD, 2.99 ± 1.34 vs 3.19 ± 1.31) (t = −0.68; P = .25). Similarly, residents who received an A or equivalent grade during their third-year medical school rotation in internal medicine had no higher faculty ratings than did those residents with lower grades (mean ± SD, 3.28 ± 1.38 vs 3.09 ± 1.29) (t = 0.59; P = .28).

There was no correlation between faculty ratings and scores on part I of the NBME examination (**Figure 3**) (r = 0.10; P = .49) or the certifying examination of the
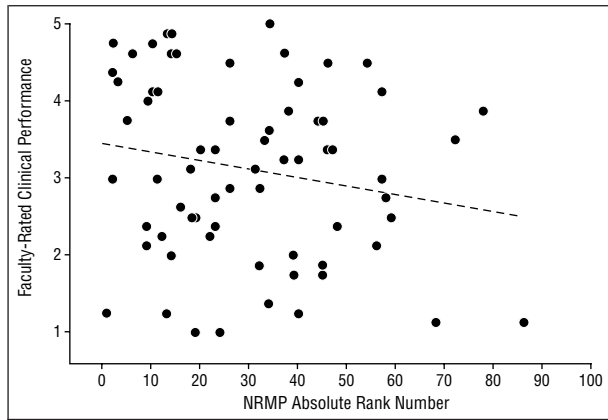
Figure 1. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and absolute rank on the National Resident Matching Program (NRMP) list (r = 0.19; P = .11).*
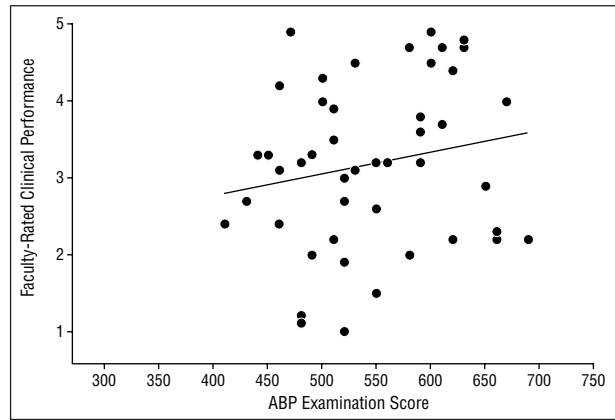

Figure 4. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and absolute scores on the American Board of Pediatrics (ABP) certification examination (r = 0.22; P = .11).*
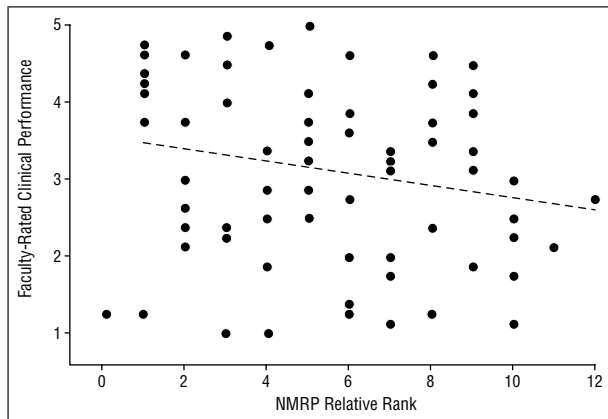

Figure 2. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and relative ranking on the National Resident Matching Program (NRMP) list (where the house officers ranked in their individual intern group, ranging from 1st to 12th) (r = 0.20; P = .09).*
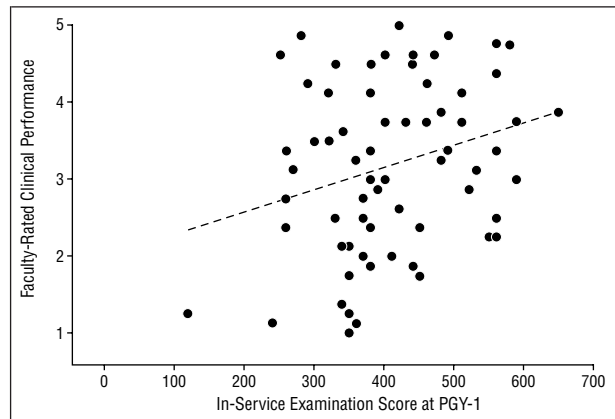

Figure 5. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and absolute scores on the pediatric in-service examination administered during postgraduate year 1 (PGY-1) (r = 0.28; P = .02).*
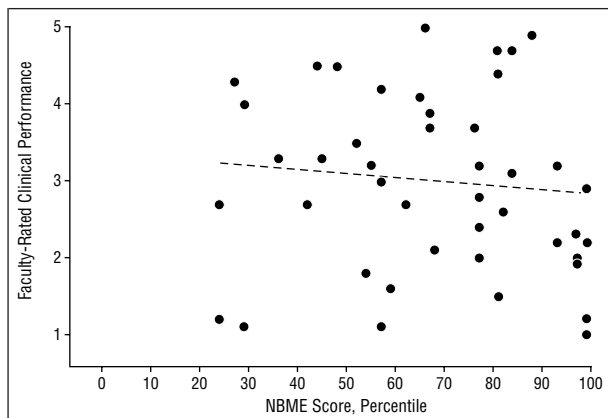

Figure 3. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and percentile scores on part I of the National Board of Medical Examiners (NBME) examination (r = 0.10; P = .49).*
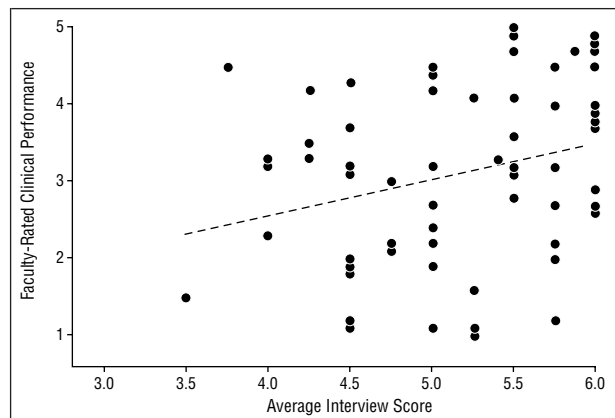

Figure 6. *Association of average faculty performance rating (from 1, bottom 20%, to 5, top 20%) and scores of faculty interviews during the intern application process. Interviews were scored using a 6-point scale, with 1 being unacceptable and 6 being superior (r = 0.27; P = .02).*

American Board of Pediatrics completed after residency (**Figure 4**) ($r = 0.22$; $P = .11$). There was a weak but significant correlation between faculty ratings and the in-service examination administered during PGY-1 (**Figure 5**) ($r = 0.28$; $P = .02$).

There was a weak but significant correlation between faculty ratings and scores of faculty interviews during the intern application process (**Figure 6**) ($r = 0.27$; $P = .02$). Those residents who were awarded all "superiors" during intern application interviews had signifi-

cantly higher faculty ratings than did their peers (mean ± SD, 3.84 ± 0.67 vs 2.94 ± 1.32) ($t = 3.34$; $P = .001$).

COMMENT

Residency training programs hope to select candidates who will succeed and achieve at their highest potential while in the program. As postgraduate positions become increasingly competitive and residency programs have many more applicants than positions, some preliminary screening of applications must be performed so that interviews can be scheduled.[1] During the evaluation process, most residency programs typically gather data on medical school performance through dean's letters, letters of recommendation, medical school transcripts, standardized test scores, formal interviews, and personal statements.[2] Students most likely to be ranked highest are those who have a high academic standing in medical school, perform well during interviews, and are perceived by program directors to be well-rounded individuals.[17]

The utility of these data as a means of identifying those medical students who will be successful residents is based on the unproven assumption that performance during medical school is a good predictor of performance during residency. While dean's and faculty letters, transcripts of grades, assessments during interviews, and applicants' autobiographies are predictive of high match ranking,[3,18] performance during medical school does not reliably differentiate applicants who will perform well during residency from those who will perform poorly.[3]

Evaluating the residency selection process is difficult because there is no uniformly accepted or objective means of measuring performance during residency other than scores on certifying examinations.[2] While high scores on standardized tests are objective and quantifiable, they have not been shown to be associated with strong performance during residency.[12] The concept of general performance is not well defined for residents in training. Previous studies have suggested that valued resident characteristics vary depending on the clinical setting[19]; because of this, obtaining faculty consensus of overall resident performance is often difficult.[11,19] In this study, we attempted to overcome this problem by choosing a group of faculty raters with a wide variety of backgrounds to obtain a global measure of resident quality.[2] Despite the diversity of our 10 faculty raters, there was remarkable agreement among them as to the overall quality of the 69 residents.

While the faculty agreed about the overall quality of residency performance, none of the traditional measures of medical school performance predicted performance during residency. In agreement with other studies, we found no correlation between NBME scores and performance during residency, nor was there any correlation between grades in required clerkships and subsequent performance during residency.[14] Perhaps more surprising, there was little correlation between NRMP ranking and subsequent clinical performance. While the small number of residents with the highest absolute ranking as well as the highest relative ranking on the NMRP rank list tended to perform somewhat better than their peers, for all other residents there was no association between NRMP ranking and subsequent performance.

Measures of medical school performance are often used as screening tools during the NRMP ranking process. Those students who have a high academic standing in medical school, perform well in an interview, and are perceived by the program directors to be well-rounded individuals are likely to be ranked highest.[17] This was clearly true in our study as well. Those students who were elected to the AOA Medical Honor Society during medical school had much higher NRMP rankings than did their peers who were not AOA members ($t = -5.13$; $P<.001$). Similarly, ranking on the NRMP match list was highly correlated with performance during intern applicant interviews ($r = 0.48$; $P<.001$). However, based on the available data, we conclude that no objective or subjective selection factors can reliably predict the level of residency performance.[13]

It is not surprising that the attempts to predict performance during residency based largely on measures of cognitive ability have been unsuccessful.[15] Searching for medical students who will be successful residents is largely predicated on the assumption that the best predictor of future performance is past performance. While this assumption may be partly correct, professional success during residency is the result of a combination of cognitive abilities, psychomotor skills, experience, interpersonal skills, various motivational and affective attitudes, and quality of character.[4] Some of these skills and attributes are not required to excel during medical school; as a result, some students who excel during medical school do not perform well during residency. Perhaps it is time to develop other tools to predict performance during residency. In one recent study, among all variables of medical school performance, the data-collection score on the clinical skills examination (standardized patient examination) yielded the highest correlation (0.27) with performance as a first-year resident.[14]

REFERENCES

1. Sklar DP, Tandbert DT. The value of self-estimated scholastic standing in residency selection. *J Emerg Med*. 1995;13:683-685.
2. Sklar DP, Tandberg DT. The relationship between national resident match program rank and perceived performance in an emergency medicine residency. *Am J Emerg Med*. 1996;14:170-172.
3. Brown E, Rosinski EF, Altman DF. Comparing medical school graduates who perform poorly in residency with graduates who perform well. *Acad Med*. 1993;68:1041-1043.
4. Papp KK, Polk HC, Richardson JD. The relationship between criteria used to select residents and performance during residency. *Am J Surg*. 1997;173:326-329.

5. Markert RJ. The relationship of academic measures in medical school to performance after graduation. *Acad Med*. 1993;68(suppl 2):S31-S34.
6. Arnold L, Willoughby TL. The empirical association between student and resident physician performances. *Acad Med*. 1993;68(suppl 2):S35-S40.
7. Fincher RM, Lewis LA, Kuske TT. Relationship of interns' performances to their self-assessments of their preparedness for internship and to their academic performances in medical school. *Acad Med*. 1993;68(suppl 2):S47-S50.
8. Case SM, Swanson DB. Validity of the NBME Part I and Part II scores for selection of residents in orthopaedic surgery, dermatology, and preventive medicine. *Acad Med*. 1993:68(suppl 2):S51-S56.
9. Erlandson EE, Calhoun JG, Barrack FM, et al. Resident selection: applicant selection criteria compared with performance. *Surgery*. 1982;92:270-275.
10. Amos DE, Massagli TL. Medical school achievements as predictors of performance in a physical medicine and rehabilitation residency. *Acad Med*. 1996;71:678-680.
11. Kesler RW, Hayden GF, Lohr JA, Saulsbury FT. Intern ranking versus subsequent house officer performance. *South Med J*. 1986;79:1562-1563.
12. Wood PS, Smith AL, Altmaier EM, Tarico VS, Franken EA Jr. A prospective study of cognitive and noncognitive selection criteria as predictors of resident performance. *Invest Radiol*. 1990;25:855-859.
13. Kron IL, Kaiser DL, Nolan SP, Rudolf LE, Muller WH Jr, Jones RS. Can success in the surgical residency be predicted from preresidency evaluation? *Ann Surg*. 1985;202:694-695.
14. Smith SR. Correlations between graduates' performances as first-year residents and their performances as medical students. *Acad Med*. 1993;68:633-634.
15. Quattlebaum TG, Darden PM, Sperry JB. In-training examinations as predictors of resident clinical performance. *Pediatrics*. 1989;84:165-172.
16. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-374.
17. Provan JL, Cuttress L. Preferences of program directors for evaluation of candidates for postgraduate training. *CMAJ*. 1995;153:919-923.
18. Aghababian R, Tandberg D, Iserson L, Martin M, Sklar D. Selection of emergency medicine residents. *Ann Emerg Med*. 1993;22:1753-1761.
19. Kastner L, Gore E, Novack AH. Pediatric residents' attitudes and cognitive knowledge, and faculty ratings. *J Pediatr*. 1984;104:814-818.