Characteristics of Highly Ranked Applicants to General Surgery Residency Programs

Steven C. Stain, MD; Jonathan R. Hiatt, MD; Ashar Ata, MBBS, MPH; Stanley W. Ashley, MD; Kevin K. Roggin, MD; John R. Potts, MD; Richard A. Moore, MD; Joseph M. Galante, MD; L. D. Britt, MD, MPH; Karen E. Deveney, MD; E. Christopher Ellison, MD

Importance: With duty hour debates, specialization, and sex distribution changes in the applicant pool, the relative competitiveness for general surgery residency (GSR) is undefined.

Objective: To determine the modern attributes of top-ranked applicants to GSR.

Design: Validation cohort, survey.

Setting: National sample of university and community-based GSR programs.

Participants: Data were abstracted from Electronic Residency Application Service files of the top 20–ranked applicants to 22 GSR programs. We ranked program competitiveness and blinded review of personal statements.

Main Outcomes and Measures: Characteristics associated with applicant ranking by the GSR program (top 5 vs 6-20) and ranking by highly competitive programs were identified using t and χ2 tests and modified Poisson regression.

Results: There were 333 unique applicants among the 440 Electronic Residency Application Service files. Most applicants had research experience (93.0%) and publications (76.8%), and 28.4% had Alpha Omega Alpha membership. Nearly half were women (45.2%), with wide variation by program (20.0%-75.0%) and a trend toward fewer women at programs in the South and West (38.0% and 37.3%, respectively). Men had higher United States Medical Licensing Examination Step 1 scores (238.0 vs 230.1; P=.001) but similar Step 2 scores (245.3 vs 244.5; P=.54). Using bivariate analysis, highly competitive programs were more likely to rank applicants with publications, research experience, Alpha Omega Alpha membership, higher Step 1 scores, and excellent personal statements and those who were male or Asian. However, the only significant predictors were Step 1 scores (relative risk [RR], 1.36 for every 10-U increase), publications (RR, 2.20), personal statements (RR, 1.62), and Asian race (RR, 1.70 vs white). Alpha Omega Alpha membership (RR, 1.62) and Step 1 scores (RR, 1.01) were the only variables predictive of ranking in the top 5.

Conclusions and Relevance: This national sample shows GSR is a highly competitive, sex-neutral discipline in which academic performance is the most important factor for ranking, especially in the most competitive programs. This study will inform applicants and program directors about applicants to the GSR program.


SIGNIFICANT CHANGES HAVE occurred in surgical education during the past decade, including the implementation of duty hour limitations, the initiation of new integrated subspecialty training paradigms, and a potentially changing applicant pool with more women entering medical school. It is possible that these developments have altered the characteristics of applicants for general surgery residency (GSR) training.

At the time of residency application, medical students and their schools submit subjective and objective personal data as part of the Electronic Residency Application Service (ERAS) program; these data are used by surgical faculty and program directors to choose applicants for interviews and subsequent ranking for the residency match. The purpose of this study was to analyze ERAS files to identify the attributes that resulted in a high ranking for surgical residency. We chose to focus on the top 20 applicants ranked by individual programs, believing that this cohort would be likely to match at those programs and would therefore stand as a measure of the attributes of the applicant pool to surgical residency programs in general.

See Invited Critique at end of article
Approval by the Albany Medical Center Institutional Review Board was obtained with the understanding that specific applicant identifiers from the ERAS file would be removed, and names would be redacted from the personal statements. Following an e-mail solicitation to program directors or chairs at 34 GSR programs approved by the Accreditation Council for Graduate Medical Education, 22 programs (64.7%) agreed to provide the ERAS files of their 20 highest-ranked applicants to the first postgraduate year class in 2011. Variables of interest were then extracted from the ERAS file for analysis. When applicants applied to multiple programs in the study, each application (not applicant) was considered individually. For the purpose of analysis, programs were divided into geographic regions (East, Northeast, South, Midwest, and West) and ranked by each author for competitiveness (1 signifies highly competitive; 2, very competitive; and 3, competitive). The program ranks were then averaged across authors and classified into the above 3 categories based on tertiles: top 7 as highly competitive, middle 7 as very competitive, and bottom 8 as competitive. The blinded personal statements were then ranked on a 1-to-5 scale by all authors, with each applicant being reviewed by at least 3 authors. The ranks were then averaged, rounded to the nearest whole number, and reclassified into 3 categories: best (very high or high),1,2 good (average),3 and below average (fair or poor).4,5 The individual applicant characteristics were described by geographic region and program competitiveness. Last, we grouped the applicants by their ranking within the program, comparing the top 5 with those ranked 6 to 20. Applicant characteristics were compared by competitiveness and region of the program and by ranking of the programs. Both χ² and t tests were used for statistical comparison of characteristics across categories. Modified Poisson regression was used to assess the characteristics independently associated with applicant ranking by GSR program (top 5 vs 6-20) and by ranking by highly competitive programs. Statistical significance was set at .05 for all analyses. STATA 11.1 statistical software (StataCorp LP) was used for analysis.

### RESULTS

The 440 applications from the 22 GSR programs included 333 unique applicants. The mean age was 27.4 years (range, 23.2-44.5 years; median, 26.8 years; and 75th percentile, 28.1 years), and 45.2% were women, with wide variation by program (20%-75%). There was a trend toward fewer women ranked in the top 20 in programs in the South and West (Table 1).

The mean United States Medical Licensing Examination (USMLE) I score for all applications was 234.4, with significantly higher scores for men than for women (238.0 vs 230.1; *P < .001*). The mean USMLE II score was 245.1 and was similar for both sexes. When USMLE scores were analyzed by competitiveness of the program as ranked by the authors (Table 2), there was a linear and significant correlation for USMLE I scores. The USMLE II scores were significantly lower for competitive programs but were similar for highly competitive and very competitive programs. Highly competitive programs were more likely to receive applications from students who had been elected to the Alpha Omega Alpha (AOA) Honor Medical Society and those with more research experience and publications. The only statistically significant finding from analysis of racial status was that highly competitive programs were more likely to have applicants who were self-described as Asian. On multivariate adjustment, the likelihood of being ranked by highly competitive programs increased 1.36 times (95% CI, 1.23-1.50, *P < .001*) for every 10-unit increase in the USMLE I scores, 2.20 times for students with publications (1.34-2.46, *P = .001*), 1.62 times for students with better personal statements (1.02-2.60, *P = .04*), and 1.70 times for Asian students compared with white students (1.25-2.31, *P = .001*).

### DISCUSSION

The process of selecting medical students for surgical residency is complex and endeavors to reconcile the attri-
Table 2. Comparison of Applicants by Program\(^a\)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Program</th>
<th>(P) Value</th>
<th>(P) Value</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N) ((95% CI))</td>
<td>(HC^b)</td>
<td>(VC^c)</td>
<td>(Cd^d)</td>
<td>All Programs</td>
</tr>
<tr>
<td>USMLE I</td>
<td>242.7 (239.8-245.5)</td>
<td>237.1 (234.0-240.2)</td>
<td>225.1 (222.8-227.4)</td>
<td>234.4 (232.6-236.1)</td>
</tr>
<tr>
<td>USMLE II</td>
<td>250.6 (247.4-253.8)</td>
<td>249.8 (247.3-252.3)</td>
<td>238.1 (235.5-240.7)</td>
<td>245.1 (243.4-246.9)</td>
</tr>
<tr>
<td>Women</td>
<td>37.9</td>
<td>47.9</td>
<td>49.4</td>
<td>45.2</td>
</tr>
<tr>
<td>Personal statement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best</td>
<td>27.3</td>
<td>25.2</td>
<td>16.4</td>
<td>22.8</td>
</tr>
<tr>
<td>Good</td>
<td>60.4</td>
<td>59.0</td>
<td>54.6</td>
<td>57.9</td>
</tr>
<tr>
<td>Fair to poor</td>
<td>12.2</td>
<td>15.8</td>
<td>29.0</td>
<td>19.3</td>
</tr>
<tr>
<td>AOA</td>
<td>47.1</td>
<td>33.6</td>
<td>7.5</td>
<td>28.4</td>
</tr>
<tr>
<td>Research</td>
<td>97.1</td>
<td>96.4</td>
<td>86.3</td>
<td>93.0</td>
</tr>
<tr>
<td>Publications</td>
<td>90.7</td>
<td>77.9</td>
<td>63.8</td>
<td>76.8</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>5.0</td>
<td>4.3</td>
<td>11.9</td>
<td>7.3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2.1</td>
<td>6.4</td>
<td>17.5</td>
<td>9.1</td>
</tr>
<tr>
<td>Not provided</td>
<td>10.7</td>
<td>5.7</td>
<td>8.8</td>
<td>8.4</td>
</tr>
<tr>
<td>Other or Asian</td>
<td>25.7</td>
<td>13.6</td>
<td>13.1</td>
<td>17.3</td>
</tr>
<tr>
<td>White</td>
<td>56.4</td>
<td>70.0</td>
<td>48.8</td>
<td>58.0</td>
</tr>
</tbody>
</table>

Abbreviations: AOA, Alpha Omega Alpha; C, competitive; HC, highly competitive; USMLE, United States Medical Licensing Examination; VC, very competitive.
\(^a\)Data are given as percentages unless otherwise noted.
\(^b\)Institutions included Brigham and Women’s Hospital; Cornell University; Emory University; Oregon Health Sciences University; University of Chicago; University of California, Los Angeles; and Washington University in St Louis.
\(^c\)Institutions included Eastern Virginia Medical School; Medical University of South Carolina; Ohio State University; University of California, Davis; University of California, San Diego; University of Maryland; and University of Texas at Houston.
\(^d\)Institutions included Albany Medical Center, Abington Memorial, Atlanta Medical Center, University at Buffalo, Midwest Community Program (hospital chose not to be named), Howard University, University of Connecticut, and University of Tennessee at Chattanooga.
\(^e\)P values are for \(\chi^2\) tests comparing proportions.

Table 3. Comparison of Applicants by Ranking\(^a\)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ranking</th>
<th>(P) Value</th>
<th>(P) Value</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N) ((95% CI))</td>
<td>(1-5)</td>
<td>(16-20)</td>
<td>All Programs</td>
<td>(HC) vs (VC)</td>
</tr>
<tr>
<td>USMLE I</td>
<td>237.7 (234.4-240.9)</td>
<td>233.2 (231.1-235.3)</td>
<td>234.4 (232.6-236.1)</td>
<td>.03</td>
</tr>
<tr>
<td>USMLE II</td>
<td>246.7 (243.4-250.0)</td>
<td>244.6 (242.5-246.6)</td>
<td>245.1 (243.4-246.9)</td>
<td>.29</td>
</tr>
<tr>
<td>Female sex</td>
<td></td>
<td>47.6</td>
<td>45.2</td>
<td>.09</td>
</tr>
<tr>
<td>Personal statement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best</td>
<td>21.3</td>
<td>23.3</td>
<td>22.8</td>
<td>.85</td>
</tr>
<tr>
<td>Good</td>
<td>60.2</td>
<td>57.1</td>
<td>57.9</td>
<td>19.3</td>
</tr>
<tr>
<td>Fair to poor</td>
<td>18.5</td>
<td>19.6</td>
<td>19.3</td>
<td>.005</td>
</tr>
<tr>
<td>AOA</td>
<td>39.1</td>
<td>24.9</td>
<td>28.4</td>
<td>.045</td>
</tr>
<tr>
<td>Research</td>
<td>94.6</td>
<td>92.4</td>
<td>93.0</td>
<td>.70</td>
</tr>
<tr>
<td>Publications</td>
<td>78.2</td>
<td>76.4</td>
<td>76.8</td>
<td>.12</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td>7.6</td>
<td>6.4</td>
<td>.91</td>
</tr>
<tr>
<td>Black</td>
<td>7.3</td>
<td>14.6</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>7.3</td>
<td>6.4</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>Not provided</td>
<td>9.1</td>
<td>12.7</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>Other or Asian</td>
<td>18.8</td>
<td>60.0</td>
<td>58.0</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: AOA, Alpha Omega Alpha.
\(^a\)Data are given as percentages unless otherwise noted.
depends in large part on the information contained in
the ERAS file submitted by students, augmented by face-
to-face interviews (the results of which were not part of
this study).

Several authors have examined the relationship of
applicant characteristics to performance as residents.
Bell et al1 found that an online survey tool (TriMetrix
Personal Talent Report; Target Training International,
Ltd) was not predictive of applicant ranking by the
program. Evaluating only residents matched at their
program, Alterman et al2 reported that USMLE I perfor-
mance, high performance outside of medicine, and
interview data had predictive value for residency perfor-
maance as measured by graduation rates, American
Board of Surgery In-Training Examination scores, and
Accreditation Council for Graduate Medical Education
core competency evaluations. In a study of 77 residents
from 1 university and 1 community-based university-
affiliated program, Tolan et al3 found that USMLE scores
were predictive only of medical knowledge and that other factors, including female sex, AOA
membership, and number of honors received during medical school, were predictive of higher overall
competency.

The current study focused on the match process rather
than subsequent residency performance. The 333 appli-
cants in the study are a select group and likely are a rep-
resentative sample of the best applicants to GSR pro-
grams nationwide. If each of them matched during the
first postgraduate year in 2011, they would occupy nearly
one-third of the available 1108 categorical general sur-
gery spots in the 2011 match.4 While it is a convenience
sample of program directors willing to participate, we
compiled applications from a geographically and aca-
demically diverse group of programs from which to ana-
lyze the characteristics of top applicants for general sur-
gery training. It is a fair assumption that ranking the top
20 candidates for a program would almost guarantee
matching at the program if the student ranked that pro-
gram highly enough. Stated another way, most pro-
grams go below the 20th spot on their rank list to fill their
resident complement.

In a survey of 262 surgical program directors and
chairs, Makdisi et al5 found that USMLE I was the
single most important factor in screening, although
final selection was relatively subjective and based on a
combination of interviews, USMLE scores, research
experience, and personal judgment. The mean USMLE
I and II scores among our applicants (234 and 245)
were higher than the nationwide sample of US senior
medical students from the 2011 match (227 and 238).4
The mean USMLE scores of applicants to GSR are
lower than some other surgical specialties (plastic sur-
gery, orthopedic surgery, otolaryngology, and neuro-
surgery) but still higher than most other specialties
(anesthesia, emergency medicine, family medicine,
neurology, OB/GYN, pathology, pediatrics, physical
medicine and rehabilitation, and psychiatry).2 While
the USMLE score is not the sole criterion for ranking
applicants for residency training, our categorization of
the competitiveness of the program had a strong corre-
lation with the USMLE scores (Table 2). The most
highly competitive training programs were more likely
to receive applications from and assign a high rank to
students with higher USMLE scores and AOA mem-
bership. Many of us who advise students in applying
for surgical residency spend significant time counsel-
ing students about their personal statements. Surpris-
ingly, there was very little correlation with our grading
of the personal statements, the competitiveness of the
programs to which the students applied, or top 5 rank-
ning by the individual program. After reviewing the
personal statements for this study, one of the authors
(S.W.A.) stated, “I have to admit that I seldom read the
personal statements, and now I remember why.”
Similarly, White et al6 evaluated personal statements
from applicants to the Scott and White surgical resi-
dency and found little interrater reliability and a lack
of objective criteria for evaluation.

During the past 20 years, there has been an increase
in the percentage of women entering medical school,
from 38.6% in 1990 to 47% in 2010.7 A recent analysis
by Davis et al8 showed that the proportion of graduates
of US medical schools entering surgical residency in-
creased from 32% of accounted-for positions in 2000 to
40% in 2005. Our data, showing that 45.2% of highly
ranked applicants were women, suggest that GSR train-
ing has kept pace with the proportion of medical school
classes that are female. This bodes well for the future of
general surgery, which depends on the quality of our
trainees.

The new resident duty hour requirements, which set
a global limit of 80 hours per week, were instituted in
2003 and subsequently amended in 2011.9,10 In 2008, the
American College of Surgeons expressed concern that re-
strictions on duty hours would result in poorly trained
surgeons, which would then adversely affect patient safety
and quality of care.11 In a systematic review of articles
regarding effects of Accreditation Council for Graduate
Medical Education duty hour restrictions on surgical resi-
dents and faculty, Jamal et al12 concluded that the limi-
tations had a positive effect on residents but a negative
effect on surgical faculty. In contrast, a survey of Uni-
versity of Wisconsin medical students during their third-
year clerkships found that the 80-hour workweek had
not improved the interest of male or female medical stu-
dents in surgery.13

No comprehensive examination has explored the ef-
fect of integrated training programs in plastic surgery,
vascular surgery, and thoracic surgery on GSR applica-
tions. There is the belief, however, that the integrated
0 + 5 programs will attract better candidates for train-
ing in those disciplines than the traditional 5 + 2 training
programs. Chikwe et al14 compared the applications
to integrated thoracic surgery residency at Mount Sinai
Medical Center. There was no difference in the overall
USMLE I scores of traditional fellowship vs integrated
residency applicants but a significant difference for can-
didates who were shortlisted (score of 252 vs 222,
P = .03). Zayed et al15 examined the characteristics of
applicants to integrated vascular surgery at the Stanford
University training program and found that applicants
to the integrated program had significantly higher
USMLE scores, were more likely to be AOA members,
and were more likely to be female. It is unclear whether integrated surgical training programs will siphon highly qualified medical students from general surgery or if these programs are attracting a different pool of applicants.

Our study has certain limitations. Regarding training programs, our grouping into tiers was subjective and not based on any tangible rating system of the medical center or program, quality of fellowships the residents obtained, pass rates on American Board of Surgery examinations, or other criteria. Regarding applicants, we did not evaluate possibly important characteristics, including interview scores, grades in surgery core clerkships, quality of research publications, completion of an externship at the target program, quality of medical school, and personal qualities. Finally, as some programs do not necessarily rank candidates purely by the quality of the applicant, the top 5 analyses may be inaccurate.

On the basis of our analysis of ERAS files, we conclude that GSR training programs are still attracting high-quality applicants to the specialty. The students who are ranked highly by training programs (at least in the top 20) had high USMLE scores, and 28.4% of the applicants were AOA members. Nearly all have done research, and most have publications. The stratification of programs by competitiveness may allow students applying to surgery to measure their applications against the highest-ranked students at a range of programs. There was a range of USMLE scores of students ranked in the top 20, even at the highly competitive residencies. The personal statement could not differentiate candidates in this highly qualified cohort, suggesting that its inclusion as a required element of applications should be reevaluated. Clearly, additional attributes beyond USMLE scores resulted in high ranking. We hope that our study provides useful information for residency program directors to compare their own programs with representative programs across the country.

Accepted for Publication: October 16, 2012.

Author Affiliations: Departments of Surgery, Albany Medical College, Albany, New York (Drs Stain and Ata); University of California at Los Angeles (Dr Hiatt); Brigham and Women’s Hospital, Boston, Massachusetts (Dr Ashley); University of Chicago, Chicago, Illinois (Dr Roggin); University of Tennessee at Chattanooga (Dr Moore); University of California at Davis, Sacramento (Dr Galante); Eastern Virginia Medical School, Norfolk (Dr Britt); Oregon Health Sciences University, Portland (Dr Deveney); and The Ohio State University, Columbus (Dr Ellison); and Accreditation Council for Graduate Medical Education, Chicago, Illinois (Dr Potts).

Correspondence: Steven C. Stain, MD, Department of Surgery, Albany Medical College, 50 New Scotland Ave, Mail Code 194, Albany, NY 12208-3479 (stains@mail.amc.edu).

Author Contributions: Drs Stain, Hiatt, Ata, Roggin, Potts, Moore, Galante, and Ellison had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Stain, Hiatt, Ata, and Roggin. Acquisition of data: Stain, Ata, Roggin, Potts, Moore, Galante, Deveney, and Ellison. Analysis and interpretation of data: Stain, Hiatt, Ata, Ashley, Roggin, Potts, Galante, and Britt. Drafting of the manuscript: Stain, Ata, and Ellison. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Stain, Ata, and Galante. Obtained funding: Stain. Administrative, technical, and material support: Stain, Ashley, Roggin, Galante, Deveney, and Ellison. Study supervision: Stain.

Conflict of Interest Disclosures: None reported.

Previous Presentation: This paper was presented at the 93rd Annual Meeting of the New England Surgical Society; September 23, 2012; Rockport, Maine; and is published after peer review and revision.

REFERENCES