ORIGINAL INVESTIGATION
Does it Match? Analyzing Internal Medicine Residency Match Data on Student Doctor Network and Charting Outcomes

Asad Loya, BS1, Khaled M. Abdelrahman, BS2, Vignesh Ramachandran, MD3
1Office of Undergraduate Medical Education, Baylor College of Medicine, Houston, Texas
2Office of Undergraduate Medical Education, Dell Medical School-University of Texas, Austin, Texas

ABSTRACT

Introduction: Applying to Internal Medicine residency has become increasingly competitive. The National Resident Matching Program (NRMP) bi-annual report (Charting Outcomes in the Match) details match-specific data for each specialty, including Internal Medicine. However, recently, applicants use online forums. The forum Student Doctor Network (SDN) includes anonymized and individual Internal Medicine data beyond NRMP data. The data validity and its concordance with NRMP have yet to be verified.

Methods: SDN Internal Medicine data (2016-17 and 2017-18) was analyzed. Factors influencing interview rate were delineated. Only cases with entirely complete data 55.8% (149/267) were included. NRMP 2018 applicant characteristics (aggregated matched and unmatched groups) were compared to SDN data.

Results: Mean Step 1 and Step 2 Clinical Knowledge (CK) scores from SDN were significantly higher than those reported on NRMP (P < 0.001). Significantly more SDN applicants reported AOA membership (P < 0.001). Also, 74.6% (106/142) reported Honors on the Internal Medicine clerkship. In unadjusted SDN data analysis, higher Step 1, higher Step 2 CK, AOA membership, and Internal Medicine Honors significantly increased interview offer rates (P < 0.001).

Discussion: NRMP data lacks complete individual applicant data, which aids applicants in identifying programs they are competitive for, saving application costs, and personalizing the application process. Although SDN data is skewed towards higher scores/ AOA status, it spans a wide range of data, making it widely applicable. Furthermore, it may be useful to applicants considering top-ranked programs.

INTRODUCTION

The number of applicants for U.S. medical residency programs through the National Resident Matching Program reached an all-time high in 2018 with 37,103 applicants submitting program choices for 33,167 total positions. Similarly, Internal Medicine (categorical) has increased the number of positions offered yearly since 2005, culminating in an all-time high of 7,542 spots in 2018. In order to navigate the residency application process, the National Resident Matching Program (NRMP) produces a bi-annual report (Charting Outcomes in the Match) detailing match-specific data for each specialty, including Internal Medicine. The NRMP report describes many mean statistics of matched and unmatched residency program candidates such as United States Medical Licensing Exam (USMLE) Step 1 and Step 2 Clinical Knowledge (CK) scores, number of research experiences and publications, number of volunteer and work experiences, and number of contiguous ranks by applicants. This dataset has served as a barometer of specialty competitiveness, including Internal Medicine,
and allows applicants to assess their portfolio against the means outlined. However, applicants may also seek information regarding the Internal Medicine residency application process beyond what is described in Charting Outcomes in the Match. The rise of online medical forums has drastically changed the residency application process in recent years. Many prospective Internal Medicine applicants utilize Student Doctor Network (SDN) to learn about residency programs and to assess their competitiveness in the match, a phenomenon described in the literature of other specialties. For instance, since its advent in during the 2015-2016 application cycle, the “Official Eternal Internal Medicine What Are My Chances Thread” (in which applicants anonymously post USMLE scores, grades, and other pertinent application information to receive feedback from other users) has yielded an ever-growing 298,387 views and 1,561 posts.

In particular, the members of SDN collaborate annually to produce an Internal Medicine spreadsheet outlining forum members’ application information. This serves as a resource for current and future applicants. This “IM Interview Offer” spreadsheet found on SDN includes self-reported and anonymous data uploaded by users of the online community, including data not included in Charting Outcomes in the Match such as clerkship grades, number of interview offers, names of programs where interviews were received by applicants, and more. The SDN spreadsheet also includes the many statistics listed in Charting Outcomes in the Match.

SDN continues to grow as an important resource for Internal Medicine, with over 181,000 posts in over 13,000 discussion topics. While the SDN “IM Interview Offer” spreadsheet includes such data of interest from hundreds individualized self-reported applicants, the validity of this data and its concordance with NRMP data has not been verified. The authors aim to a) summarize SDN Internal Medicine applicant data; b) compare this data to Charting Outcomes in the Match data; and c) assess the utility of SDN Internal Medicine data in the residency application process.

**MATERIALS AND METHODS**

**Data collection**

Resident application data was acquired through aggregation of two publicly available datasets on Student Doctor Network (SDN) and compared to data reported in NRMP’s 2018 Charting Outcomes in the Match. The two analyzed SDN datasets contained self-reported data in the spreadsheets entitled “2016-17 IM Interview Offer” and “Official 2017-2018 Internal Medicine Official Application Sheet”, both of which encompassed the time span between the 2016 and 2018 NRMP Charting Outcomes in the Match reports. Cases with missing data (n = 118) for any of the variables of interest were excluded from examination with a total of 55.8% (149/267) utilized (Figure 1). NRMP matched (n = 3070) and unmatched (n = 59) cases were aggregated with recalculation of means in order to compare the group to the SDN dataset, which may have included unmatched applicants. Institutional Review Board review was not required for this study.

**Data analysis**

Statistical analysis was conducted using Microsoft Excel (Microsoft, Redmond, WA, USA) and SPSS v25 (IBM Corp, Armonk, NY, USA). Normality tests were conducted using the Shapiro-Wilk test. Mann-Whitney U test was performed to assess associations when the input variable was nominal and simple regression was performed when the input variable was continuous. One-sample t-test and one-sample binomial test were used to compare the SDN combined dataset to NRMP Charting Outcomes in the Match reference values. Cases in which a factor was reported as not applicable were excluded from analysis for that specific factor. A
A P-value of < 0.05 was used to determine statistical significance.

<table>
<thead>
<tr>
<th></th>
<th>SDN Data (n=149)</th>
<th>NRMP Combined (n=3129)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1 Score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>244</td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>243.28 (14.19)</td>
<td>232.51</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Step 2 Score</strong></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Median</td>
<td>258</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>256.19 (11.90)</td>
<td>245.57</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AOA status</strong></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes</td>
<td>45 (30.2%)</td>
<td>~ 513 (16.4%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>80 (53.7%)</td>
<td>~ 2616 (83.6%)</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>24 (16.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IM Honors</strong></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Yes</td>
<td>106 (71.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>36 (24.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>7 (4.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interview Invite %</strong></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Median</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>46.90 (25.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong># Programs Applied</strong></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Median</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>50.43 (42.69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>288</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong># Interview Offers</strong></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Median</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>17 (9.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 1 and 2 score p-values reflect one sample T-test using NRMP means as a reference. AOA Status p-value reflects one sample binomial test using NRMP probability as a reference.

Table 1. Comparison of Student Doctor Network applicant data with Charting Outcomes in the Match data.
RESULTS

SDN Data Descriptive Analysis

The aggregated SDN dataset including complete cases from 2016-18 consisted of 149 applicants’ data. The applicant characteristics for the cohort are provided in Table 1. The mean Step 1 and Step 2 CK scores of applicants are 243.28 (+/-14.19) and 256.19 (+/- 11.90) respectively. 125 (83.9%) students attended medical schools with Alpha Omega Alpha Honors Society (AOA) denoted on applications; 45 (36.0%) of these students reported induction into AOA. Similarly, 142 (95.3%) attended medical schools denoting Honors as a possible grade received on the Internal Medicine core clinical clerkship; 106 (74.6%) of these individuals reported receiving Honors. The mean number of programs applied to, number of interviews received, and interview offer rates were 50.43 (+/-42.69), 17 (+/-9.20), and 46.90% (+/-25.60%), respectively. The Shapiro-Wilk test determined that all the analyzed variables were non-normal at the $P < 0.05$ level of significance.

Comparison to NRMP

Charting Outcomes in the Match applicant characteristics of the combined matched and unmatched groups (n = 3129) are reported (Table 1). The mean Step 1 and Step 2 CK scores were calculated as 232.51 and 245.57 respectively, both of which were significantly decreased compared to SDN means of 243.28 and 256.19, respectively ($P < 0.001$). The calculated number of applicants with AOA membership was approximately 513 students (16.4%), which was also significantly decreased compared to the SDN mean of 36.0% ($P < 0.001$). A comparison of other variables was not possible due to lack of data collection by NRMP for those fields (Table 1).

Factors influencing Interview Invite Rate

In unadjusted models, higher Step 1 score, higher Step 2 CK score, AOA membership, and Honors in the Internal Medicine clerkship were all significantly associated with increased interview rates ($P < 0.001$) (Table 2).

<table>
<thead>
<tr>
<th>Simple Linear Regression</th>
<th>Beta Stand. Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>0.377</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Step 2</td>
<td>0.439</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mann - Whitney U test</th>
<th>Mean Rank</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOA status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>80.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>53.3</td>
<td></td>
</tr>
<tr>
<td>IM Honors (MS3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>80.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>45.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Univariate analysis of continuous and categorical variables association with increased interview rates.

DISCUSSION

The Charting Outcomes in the Match provides applicants with an extensive dataset to assess competitiveness for residency programs. However, it is lacking in its ability to provide complete individual applicant data, which has a different utility to prospective applicants beyond mean data. Such a dataset not only allows applicants to assess their competitiveness for Internal Medicine residency in general, but also candidacy for programs. As the number of applications per applicant to Internal Medicine and costs associated with the application and interview process continue to increase, such insights are invaluable.8 This year, the NRMP released additional data for applicants to utilize, including the At-A-Glance Program Director Survey and the Interactive Charting Outcomes in the Match.9,10 While these applications provide insight into components of the application that Program Directors in Internal Medicine most value and how scores may influence interview chances, they are again aggregated data with modest response rate
which do not provide personalized insights or program-specific chances of receiving an interview or matching. These are major factors limiting utility.

Our analysis showed a significantly higher Step 1, Step 2 CK, and AOA status of SDN applicants compared to those reported on NRMP. Despite this difference in distributions however, SDN data spans a wide range of Step 1 (range = 65 points) and Step 2 CK (range = 57 points) scores, making it accessible to most applicants; NRMP data encompasses applicants with Step 1 (range > 70 points) and Step 2 CK (range >70 points) of a similar range. Furthermore, due in part to having the highest number of residency spots, Internal Medicine has remained one of the lesser competitive specialties with a 96.7% match rate for U.S. senior medical students. Nevertheless, the competition for top tier programs can be fierce. This is highlighted by discussion threads of “top tier” programs on SDN and other forums, partially driven by annual rankings published by Doximity and U.S. News of Internal Medicine programs. In qualitatively assessing the interview invitations of applicants on the SDN spreadsheet, we noted a disproportionate number of applicants receiving interviews at programs within the Doximity and U.S. News top 20 ranked programs. While this highlights a reporting bias in the SDN data, it may also point to the SDN serving as a unique and valuable tool for applicants with aspirations of such programs. Charting Outcomes in the Match is unable to provide such insights due to its reporting of solely mean data for Internal Medicine. Despite an overrepresentation of cases with higher statistics, SDN still provides potentially valuable case data with program-specific insights for applicants with middle-to-lower end objective data, which Charting Outcomes in the Match is unable to provide. Although less extensive, SDN’s individualized data may still serve as a valuable tool to applicants of all backgrounds.

Step 1, Step 2 CK, AOA, and Honors in the Internal Medicine clerkship each were independently associated with increased interview rate in univariate analysis (Table 2). A common misconception among students is Step 1 may be the most integral part of their application to Internal Medicine residency, causing them to possibly undervalue other aspects of their application. This misconception may stem from the notion that Internal Medicine is a less competitive specialty in terms of mean NRMP scores and percentage of AOA applications. However, our analysis shows that other aspects of the application (Step 2 CK, AOA, Honors in IM clerkship) are all associated with increased interview rate and match success. Furthermore, Step 2 CK had a higher impact on interview rate per point than Step 1 (Standardized Beta coefficients of 0.439 versus 0.377, respectively). Step 2 CK is designed as an overall test of clinical knowledge and IM residency training is built upon developing and utilizing much of this foundation; this may be why IM residency programs place an emphasis on Step 2 CK in the application process. Such a univariate analysis could not be performed for Charting Outcomes in the Match as it provides summary statistics rather than case data; nonetheless NRMP offers many other forms of statistics and interactive graphs that may be instrumental to applicants.

Part of the appeal of such online forums that offer self-reported data is the notion of longitudinal application cycle insights. As residency admissions become more competitive, some applicants can select Internal Medicine specialization earlier on in medical school. The information derived from such forums can then be used to enhance competitiveness. SDN and other venues offer a “before, during, and after” role for applicants. For instance, the “Official 2017-2018 Help Me Rank Megathread [Internal Medicine]” thread (797 posts and 109,305 views) dissects a portion of the application cycle beyond the scope of any
Match-related dataset; this thread allows applicants to request insights into how to best construct their Rank Order List by aggregating feedback from forum members. This highlights the role of the forum beyond the mere pre-application period. Furthermore, the immense number of thread views disproportionate to the number of posts highlights that online viewers can extract information from others’ posts without having to create an account and submit posts themselves. In this manner, the growth of such online anonymous self-reporting data for the Internal Medicine match can be expected to inevitably grow as the forums hook applicants before, during, and after the application cycle. In the future, the sample size of SDN self-reported data may be large enough to serve an even more important resource for applicants, perhaps eventually displacing Charting Outcomes in the Match as the gold standard resource utilized by students and faculty.

Forums such as SDN are not without their limitations, however. Other studies have also demonstrated findings suggestive of reporting bias on student doctor network, with an overrepresentation of students with higher statistics. Despite this bias, the range of data is extensive allowing a broad spectrum of applicants to identify similar students and gain advice and insight into the application process. Another important limitation of the SDN dataset was that of cases with missing data. To address this, we excluded any cases which did not have inputs for all variables assessed. By doing so, we aimed to reduce bias, as applicants may feel more comfortable reporting their strengths and omitting their weaknesses on such online forums. Another proposed limitation of online forums, which remains unstudied, is the validity of data available. Unfortunately, due to the anonymous nature of the forum it is difficult to authenticate its reliability. Nevertheless, such a study is cutting-edge and has never been performed before. It is of value to students applying to IM residency and educators as is reflected by similar studies performed in other specialties. It also expands upon the growing desire and need for students to have access to individualized data to navigate the complex residency process. Lastly, compared to the NRMP dataset, SDN was limited by a much smaller case count and a non-parametric distribution. As a result, our multiple regression analysis poorly explained the variance (R^2<25%) and was not displayed in this study.

Medical school administrators and medical school education leaders should take steps to provide students with tools to better navigate the residency application process. While the NRMP data is of great utility, it may better serve as an overall baseline assessment of competitiveness for Internal Medicine residency. What may be of greater utility to students is data similar to the SDN spreadsheet. Some programs and institutions have taken steps to work towards this goal. For instance, the Texas STAR (Seeking Transparency in Application to Residency) database was created in collaboration with dozens of medical schools. Using it, applicants are able to filter and search anonymous applicant data across all specialties, including Internal Medicine, amongst participating schools. Data points in Texas STAR include Step 1, Step 2 CK, AOA, number of clerkships Honored, Research publications, and more. Furthermore, this dataset, like SDN data, shows where interviews were received and where the applicants matched. Such data is immensely valuable to applicants. In this era of personalized medicine, the residency application process for Internal Medicine ought to be personalized.
REFERENCES

1. Charting Outcomes in the Match: U.S
Allopathic Seniors. National Resident
www.nrmp.org/wp-content/
uploads/2018/06/ChartingOutcomes-in-
the-Match-2018-Seniors.pdf. Published

2. Sura K, Wilson LD, Grills IS. Comparison
of Self-Reported Data on Student Doctor
Network to Objective Data of the National
Resident Matching Program. J Am Coll
Radiol. 2017 Dec;14(12):1594-1597. doi:
10.1016/j.jacr.2017.08.011.

How Does the Current Generation of
Medical Students View the Radiology
Match?: An Analysis of the AuntMinnie
and Student Doctor Network Online
Forums. Acad Radiol. 2018 Jun;25

4. Official Eternal Internal Medicine What
Are My Chances Thread. Student Doctor
forums.studentdoctor.net/threads/official-
eternal-im-wamcthread.1132592/.
Published 2015. Accessed November 1,
2018.

5. 2016-2017 IM Interview Offer. Google
Sheet public Web site. https://
docs.google.com/spreadsheets/
d/1RicH7w5X4itti389QdFC1WimTGDZI
WNlZYcLTPwtWx8/edit#gid=505432853.
P Published 2016. Accessed November 1,
2018.

6. Official 2017-2018 Internal Medicine
Official Application. Google Sheet Web
site. https://docs.google.com/spreadsheets/
/u/1/d/
e/2PACX1vQDyXpBO3FvSQEtdeTPQeH
qO4dDNP2fOGXt3MNKfpaOpObBuw1
EvjrhLNROppUJtmvypOIMPTI2yY/
pubhtml?

fbclid=1wAR2soV0xL_A0xk_P-91sgZKcs
R_T52DjX8p_2QmuEulifyjTABfAhS2A
Wmo. Published 2017. Accessed
November 1, 2018.

7. Internal Medicine and IM Subspecialties.
Student Doctor Network Web site.
https://forums.studentdoctor.net/forums/
internal-medicine-and-
imspecialties.41/. Published November
1, 2018.

8. Pereira AG, Chelminski PR, Chheda SG, et
al. Application Inflation for Internal
Medicine Applicants in the Match:
Drivers, Consequences, and Potential
(8):885-891. doi:10.1016/
j.amjmed.2016.04.001.

National Resident Matching Program Web
site. http://www.nrmp.org/at-a-glance-
program-director-survey/. Published 2018.

10. Interactive Charting Outcomes in the
Match. National Resident Matching
interactive-charting-outcomes-in-the-
match/. Published 2018. Accessed
November 1, 2018.

Program for You. New England Journal of
Medicine Knowledge+ Web site. https://
knowledgeplus.nejm.org/blog/match-day-
residency-program/. Published 2017.

Megathread [Internal Medicine]. Student
forums.studentdoctor.net/threads/
oficial-2017-2018-
help-me-rank-megathread-internal-
medicine.1289993/. Published 2017.

13. Texas Star (Seeking Transparency in
Application to Residency). UT
Southwestern Medical Center Web site.