Evaluating diversity in neurosurgery through the use of a multidimensional statistical model: a pilot study

Jonathan H. Borden, MD,1 Uma V. Mahajan, BS,2 Lud Eyasu, BA,3 William Holden, BS,3 Brian Shaw, BS,3 Peter Callas, PhD,4 and Deborah L. Benzil, MD5

1Department of Psychiatry, University of Vermont, Burlington, Vermont; 2School of Medicine, Case Western Reserve University, Cleveland, Ohio; 3Larner College of Medicine, University of Vermont; 4Department of Biostatistics, University of Vermont, Burlington, Vermont; and 5Department of Neurosurgery, Cleveland Clinic, Cleveland, Ohio

OBJECTIVE There is a growing body of evidence demonstrating the benefits of diversity across many domains. However, neurosurgery consistently lags most of medicine in many aspects of diversity. Any inability to make progress in this arena is likely due to the multifactorial and complex nature of the issue, which makes it difficult to meaningfully measure and track diversity within the workforce. The goal of this pilot study was to assess the utilization of a multidimensional statistical model to quantify and assess diversity within neurosurgery. The authors sought to 1) assess the diversity of neurosurgery residents using Simpson's Diversity Index and Sullivan's Composite Diversity Index (CDI) and 2) determine if a medical school's intrinsic academic opportunities and resources, indicated by US News & World Report's (USNWR's) best research medical schools ranking, are related to the number of neurosurgery residents produced per medical school.

METHODS A cross-sectional study of all neurosurgery residents (projected graduation years 2020–2026) and 1st-year medical students (matriculating years 2016–2019) was undertaken. Biographical diversity data (gender and matriculation data) were collected from institutional websites between December 2019 and June 2020. The CDI expresses the diversity of a given population by representing the effective proportion of categories present across all diversity attributes and was calculated for neurosurgery residents and medical students. Statistical results are reported as the median and interquartile range.

RESULTS Neurosurgery residency program CDI (0.21, IQR 0.16–0.25) was significantly less (p < 0.001) than medical school CDI (0.42, 0.37–0.48). There was no significant difference in CDI between top-40 and non–top 40 Doximity ranked research output neurosurgery residency programs (p = 0.35) or between top-40 and non–top 40 USNWR ranked research medical schools (p = 0.11). Over a 7-year period, top-40 ranked research medical schools produced significantly more (p < 0.001) neurosurgery residents (11.9, IQR 7.1–18.9) than the non–top 40 ranked research medical schools (5.6, IQR 2.6–8.5).

CONCLUSIONS The authors demonstrated the feasibility of using a multidimensional statistical model as a measure to understand the complex issues of diversity. Their preliminary data suggested that neurosurgery's challenge in achieving the desired diversity relates to uneven attraction and/or recruitment across an increasingly diverse medical student body. In recent years, neurosurgery has made great progress in the arena of diversity and has shown a strong desire to do more. Utilization of these diversity measures will help the neurosurgery field to monitor progress along this valuable journey.

https://thejns.org/doi/abs/10.3171/2021.10.JNS211006

KEYWORDS neurosurgery; residency; diversity; education; equality

Increasingly, the benefits of diversity across many domains are being demonstrated. Diversity is a reality created by individuals from a broad spectrum of demographics, experiences, and beliefs.1 On a population level, biodiversity improves the overall performance of a system by enhancing productivity, stability, and resilience.2–6 Within medical schools, a diverse faculty and student body promote a culture that enhances learning, skill development, and intellectual engagement.7 Within healthcare, a diverse physician workforce leads to im-
proved care for underserved groups, enhanced culturally sensitive care, and a more robust quality of research. In recognition of this, most healthcare systems strive to have the diversity of the nation and their communities represented within their healthcare workforce. Yet, it remains a challenge to meaningfully measure diversity and, in doing so, to understand barriers to achieving increased diversity. Current approaches used to measure diversity focus on quantifying the representation of inherent diversity attributes (those we are born with) for a given population. This method is limited because it does not present a methodology to incorporate acquired diversity attributes (those gained through life experiences) or to meaningfully compare diversity across multiple attributes simultaneously. The use of ecological diversity indices to create a single numeric measure of diversity, for a given population, that embodies multiple diversity attributes, both inherent and acquired, has the potential to mitigate these challenges.

Neurosurgery consistently lags most of medicine in many aspects of diversity. Any inability to make progress in this arena is likely due to the multifactorial and complex nature of the issue, which makes it difficult to meaningfully measure and track diversity within the workforce. One factor that may present a barrier to diversity within neurosurgery, and has yet to be explored, is the residency selection process. Neurosurgery is consistently one of the most, if not the most, competitive specialties to match into. In the 2020 application cycle, there were roughly 1.65 applicants per resident position. For comparison, the applicant to resident position ratios for other specialties within the same match cycle are as follows: plastic surgery, 1.57; general surgery, 1.42; internal medicine, 1.20; and family medicine, 1.05. Competitive neurosurgery applicants have, on average, a United States Medical Licensing Examination (USMLE) Step 1 score of 246, a USMLE Step 2 score of 251, 4.3 first/last author publications, 2.1 neurosurgical first/last author publications, and 18.3 total research projects. Neurosurgery residency program directors have indicated that they utilize these aforementioned factors as initial screening tools of applicants. This process may favor students at institutions with the intrinsic opportunities and resources necessary to achieve these academic benchmarks.

The purpose of this pilot study was twofold: 1) to determine the feasibility of using Simpson’s Diversity Index and Sullivan’s Composite Diversity Index (CDI) to quantify and assess diversity among neurosurgery residency programs and medical schools, and 2) to determine if a medical school’s intrinsic academic opportunities and resources, indicated by US News & World Report’s (USNWR) best research medical schools ranking, are related to the number of neurosurgery residents produced per medical school.

**Methods**

This study was determined to be not human subjects research by the University of Vermont IRB. Data were accessed between December 2019 and June 2020 via publicly available websites.

**Data Collection**

**Resident Program Data**

One hundred seventeen neurosurgery residency programs were identified via Doximity. Seventeen did not provide complete residency data at the time of collection, and 1 was excluded as it is part of the military residency match process. Residency program websites (n = 99) were reviewed for the most updated list of current residents (n = 1510, projected graduation years 2020–2026) and biographical diversity data. Doximity neurosurgery residency program research output rankings were recorded to aid in statistical analysis. The Doximity research output rank measures the collective h-index of program alumni and is thus interpreted as a representation of the quality of and resources received in training regarding research support.

**Resident Diversity Data**

Using individual residency program websites, we recorded the following diversity data per resident: medical school matriculation data (allopathic, osteopathic, or international) and perceived gender identity (male or female). Table 1 provides additional details on these diversity attributes. Gender information was identified for 95.4% (n = 1440) of neurosurgery residents using available resident pictures and names. Gender information for 4.6% (n = 70) of the neurosurgery residents was not recorded because their perceived gender identity was unclear. Each resident’s gender was categorized by a single observer, and in cases in which gender identity was not clear, we did not record it. This is consistent with how past gender analyses have been done.

**Resident Medical School Data**

Using individual residency program websites, we recorded the medical school attended per neurosurgery resident. The number of neurosurgery residents produced per medical school was normalized to a class size of 152 to allow direct comparison of residents produced per medical school. USNWR’s best research medical schools rankings were recorded to aid in statistical analysis (n = 123). USNWR’s best research medical schools ranking is calculated utilizing a mathematical model that includes factors such as the number and quality of research projects for students, NIH research funding, and research activity per faculty and can thus be interpreted as a representation of the quality of academic opportunities and resources available during medical school training.

Although the use of USNWR’s best research medical school ranking introduces potential confounds due to differences in student populations between medical schools, we believe that it remains a valid measure of academic opportunities and resources. The average Medical College Admission Test (MCAT) score and grade point average of matriculating students differ between schools, with higher ranked schools having higher averages, so this can be interpreted as higher ranked schools selecting the best and brightest individuals. This interpretation assumes an individual’s performance and effort are static and meaningfully represented using the traditional measures used for acceptance into medical school. However, performance
and effort are fluid, changing in response to personal life events and career aspirations, and are not so easily represented utilizing traditional measures. Moreover, this interpretation does not take into consideration the complex decision-making in choosing a medical school, which may often result in choosing not to attend the highest ranked school into which an individual is accepted.

Medical School Diversity Data

Publicly available data through the American Association of Medical Colleges (AAMC) Medical Student Admissions Requirements (MSAR) include only allopathic medical schools, so osteopathic and international medical school diversity data were excluded. Using USNWR’s 2020 best research medical schools rankings, we identified 108 allopathic medical schools. The following self-reported medical school diversity data for 1st-year medical students (n = 17,954), years 2016–2019, were recorded from the AAMC MSAR (Table 1): gender, undergraduate matriculation data, race and ethnicity, age, medically underserved background, and proficiency of additional languages spoken. Matriculation data refer to the type of academic institution attended prior to medical school (out of state, in state, or international). Medically underserved areas are designated by the US Health Resources and Services Administration as having too few primary care providers, high infant mortality, high poverty, or a high elderly population. Table 1 provides additional details on these diversity attributes.

Diversity Measurements

Simpson’s Diversity Index

\[ D = 1 - \sum p_i^2 \]

where \( p_i \) represents the proportion of individuals in the \( i \)th category, and \( D \) is interpreted as the probability that any two individuals from a given population and drawn at random will be from a different category within a specific diversity attribute. \( D \), within a given diversity attribute, is maximized with increasingly equal representation among many categories within a given diversity attribute. The maximum possible \( D \) for a given diversity attribute increasingly approaches 1 as the number of categories within that diversity attribute increases. Thus, to meaningfully compare different indices with each other, their diversity attributes must have a similar number of categories. Simpson’s Diversity Index was calculated for each diversity attribute identified within neurosurgery residency programs and medical schools.

Sullivan’s CDI

Derived from Simpson’s Diversity Index, the CDI (\( A_w \)) expresses the diversity of a given population by representing the relative proportion of categories present across all diversity attributes, using the following equation:

\[ A_w = 1 - \frac{1}{V} \sum_{i=1}^{k} \left( (p_i) \right)^2 \]

where there are \( V \) attributes, \( k \) categories, and \( p_i \) proportions in each category. That is, \( A_w \) represents the CDI of an individual or cohort of interest, which encompasses many
different diversity attributes of interest. \( A_n \) is interpreted as the probability that any two individuals from a given population and drawn at random will be from different categories across all diversity attributes.\(^{10}\) \( A_n \), within a given population, is maximized with increasingly equal representation among many categories across multiple attributes. The maximum possible \( A_n \) for a given population increasingly approaches 1 as the number of categories within any diversity attribute increases. Thus, to meaningfully compare different indices with each other, the diversity attributes included in the CDI must have a similar number of categories. Additionally, since CDI represents a composite, its use requires at least two diversity attributes. A CDI was calculated for neurosurgery residency programs and medical schools. To facilitate as direct a comparison as possible between neurosurgery residency programs and medical schools, the diversity attributes included in the CDI calculation were limited to gender and matriculation data. Gender diversity, an inherent attribute, represents the gender distribution within medical schools and residency programs. Matriculation diversity, an acquired attribute, represents the unique experiences and training of students from different educational institutions. Regarding the utilization of matriculation diversity, we acknowledge that there are limitations to this approach. Perhaps our country’s medical education system has a duty to optimize the residency matches of our country’s own medical students, so we may never aspire for maximum diversity in this regard. However, it is an important variable to consider within this pilot study, as it clearly represents an acquired diversity attribute.

Statistical Analyses

Statistical analyses were completed by one of the authors (P.C.). Data were analyzed using Wilcoxon rank-sum tests. “Top-40 NIH-Ranked Medical School” is a commonly cited grouping by the National Resident Matching Program (NRMP) and was used to dichotomize data per USNWR’s best research medical schools ranking and Doximity’s residency research output ranking. Within USNWR’s research medical school rankings, because of ties, 44 schools were included in the top-40 rankings. Statistical significance was set at \( p < 0.05 \). Statistical results are reported as the median and interquartile range. All analyses were conducted using SAS version 9.4 (SAS Institute Inc.).

Results

Descriptive Data

Descriptive statistics for Simpson’s Diversity Indices for neurosurgery residents and medical students are shown in Table 1.

Diversity Attributes Analysis

Diversity Attributes Between Neurosurgery Residents and Medical Schools

Descriptive statistics for comparisons of diversity attributes between neurosurgery residents and medical students are shown in Table 2. Compared to neurosurgery residents, medical students had significantly greater gender diversity (\( p < 0.001 \)), matriculation diversity (\( p < 0.001 \)), and CDI (\( p < 0.001 \)). Abundance plots visually depicting the differences in diversity between neurosurgery residency programs and medical schools are shown in Fig. 1.

Diversity Attributes Within Neurosurgery Residency Programs and Medical Schools

Descriptive statistics for comparisons of diversity attributes within neurosurgery residency programs and medical schools are shown in Table 3. Top-40 research output residency programs had significantly more gender diversity than non–top 40 research output residency programs (\( p = 0.04 \)). Between top-40 and non–top 40 research output residency programs, there was no significant difference in matriculation data diversity (\( p = 0.49 \)) or CDI (\( p = 0.35 \)).

Between top-40 and non–top 40 research medical schools, there was no significant difference in gender diversity (\( p = 0.49 \)), matriculation data diversity (\( p = 0.12 \)), or CDI (\( p = 0.11 \)). Other notable findings indicated that top-40 research medical schools had significantly more racial/ethnic diversity than non–top 40 research medical schools (Simpson’s indices of 0.65 [IQR 0.61–0.68] vs 0.61 [0.53–0.65], respectively, \( p = 0.008 \)). Top-40 research medical schools had significantly more language diversity (e.g., proficiency of additional languages spoken) than non–top 40 research medical schools (Simpson’s indices of 0.51 [IQR 0.48–0.52] vs 0.47 [0.43–0.49], respectively, \( p < 0.001 \)).

Medical School Representation Within Neurosurgery Residents

Comparison of the normalized number of neurosurgery residents produced per medical school over a 7-year period showed that significantly more neurosurgery residents were produced from top-40 research medical schools (\( n = 44, 11.9 \) [IQR 7.1–18.9]) than from non–top 40 research medical schools (\( n = 79, 5.6 \) [2.6–8.5]; \( p < 0.001 \)). These findings are depicted in Fig. 2.

Discussion

Establishing a method to measure and track diversity is the first step in achieving a medical workforce that reflects the diversity of the community and country. This alone is
a challenging task. Our study demonstrates the feasibility of a multidimensional statistical model to meaningfully quantify diversity from a mathematical perspective, using both inherent and acquired attributes. Not surprisingly, our data confirm that neurosurgery resident cohorts are less diverse than medical student cohorts. While our data show no significant difference in CDI within neurosurgery residency programs or medical schools, top-40 USNWR ranked research medical schools were significantly overrepresented within neurosurgery residents. Post hoc analyses using the NRMP 2020 charting outcomes data support this notion. As compared to applicants from top-40 research medical schools, there is a greater percentage of applicants from non–top 40 research medical schools who do not match into neurosurgery (72.4%) versus applicants from non–top 40 who do match (61%), although this difference is not statistically significant (p = 0.11). Taken as a whole, this suggests that the diversity challenges within neurosurgery may be related to uneven recruitment and/or attraction across a diverse medical student body.

It is well established that neurosurgery lags medical schools in diversity. Creating a truly diverse neurological workforce is an iterative process in which each new effort builds on past accomplishments to develop new and expanded potential. Initial efforts in neurosurgery relied on specific individuals stepping up or stepping out, including the formation of Women in Neurosurgery (WINS) in 1990 by a pioneering group of female neurosurgeons. Nearly 2 decades later (2008), WINS published a white paper in which they challenged the neurosurgery community with a call to action: attain a 20% female complement within residency programs and a 20% female faculty complement by 2020. As of 2020, 19.5% of current neurosurgery residents were female and a growing proportion of top leadership positions are held by women.

Over the last decade, progress has been made; however, significant opportunity for diversification exists for both women and other underrepresented minorities. To optimize patient care, we must strive to achieve a diverse workforce that recognizes the value of inherent attributes (those we are born with) as well as acquired attributes (those we gain through life experience). By conceptualizing individuals as unique, when grouped in diverse combinations, we are able to utilize each individual’s experiences, skills, and training to promote innovation and creativity. The proposed statistical model in this study represents a methodology to mathematically quantify diversity using both inherent and acquired attributes, which can help achieve this goal.

Understanding barriers to diversity is also valuable in order to design and implement appropriate and meaning-

### Table 3. Summary comparison of diversity indices within medical schools and neurosurgery residency programs by top-40 vs non–top 40 research rank

<table>
<thead>
<tr>
<th>Index</th>
<th>Medical Schools (n = 108)</th>
<th>Neurosurgery Residency Programs (n = 99)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Top 40</td>
<td>Non–Top 40</td>
</tr>
<tr>
<td>Simpson’s Diversity Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.50 (0.50–0.50)</td>
<td>0.50 (0.49–0.50)</td>
</tr>
<tr>
<td>Matriculation data</td>
<td>0.37 (0.28–0.47)</td>
<td>0.34 (0.19–0.43)</td>
</tr>
<tr>
<td>CDI</td>
<td>0.43 (0.39–0.48)</td>
<td>0.42 (0.35–0.46)</td>
</tr>
</tbody>
</table>

Values are expressed as the median (interquartile range), unless indicated otherwise. Boldface type indicates statistical significance.
ful systemic change. This study suggests that there may be intrinsic factors specific to top-ranked research medical schools that contribute to the overrepresentation of their students entering neurosurgery residencies. Indeed, there is a growing body of literature supporting this notion. Higher ranked medical schools offer more exposure to neurosurgery faculty and research environments, which leads to more opportunities for mentorship, publications, and a higher graduating h-index.

Given the increasing interest in and competition for neurosurgery residency positions, some method is required to reliably assess and screen applicants. Program directors indicate that they rely on traditional academic factors as initial screening tools of applicants. To date, there is limited available peer-reviewed evidence supporting the use of academic benchmarks as predictors of future ability or success in medicine as a whole. This represents, in part, second-generation bias, wherein program directors utilize markers such as USMLE scores and number of publications to represent an applicant’s neurosurgical potential. The use of these factors as an initial applicant screening tool is rooted in a meritocratic ideology—successes or failures are defined primarily by individual effort, agency, and resilience. Within a meritocracy, it is assumed that upward societal mobility can be achieved through hard work. Using this ideology may be self-fulfilling but can only be valid for complex societies if individuals have equal opportunities and resources. Socially, this is not probable, and there is also a growing body of evidence demonstrating that this is not the case academically.

Rather than recognizing the intrinsic differences in the available opportunities and resources between medical schools and individuals’ unique lives, a student’s inability to meet the required academic benchmarks to pass the initial screening process is viewed as a representation of their lack of neurosurgical potential. Students often readily accept this as being a reflection of their individual effort rather than a reflection of the larger system of which they are a member. The impact of this is that applicants who have the potential to be a successful neurosurgeon may decide not to apply to neurosurgery or may never have the opportunity to interview at programs because they are unable to pass the initial screening process. The use of a numbers-driven screening process unintentionally perpetuates academic inequalities, which decreases the real and potential applicant pool and thus decreases potential diversity in the field.

One potential method that has more recently been deployed to overcome the noted limitations of a meritocratic screening process is a competency-based approach. A competency is defined as a set of skills and attributes necessary to perform a role. Success or failure for any given skill or attribute can be demonstrated within a multitude of scenarios. Thus, a competency-based approach facilitates the evaluation of individuals within the context of their lives, which includes factors such as their available
opportunities and resources. Moreover, this approach allows for the conceptualization that each medical specialty requires different competencies to be successful. Such a holistic evaluation system is likely to achieve a truer understanding of an applicant’s future potential within a field of interest.\textsuperscript{38} Utilization of this approach has allowed parallel systems, such as medical schools, to broaden the diversity of their applicant interview pool, as well as subsequent classes, with no sacrifice to future performance.\textsuperscript{39,40}

The institution of medicine does not currently have a contemporary data-driven understanding of what predicts a “successful” physician or specialist. The change of USMLE Step 1 to pass/fail will necessitate the development of a new approach toward evaluating residency applicants. We recognize that recruitment is a complicated topic for every specialty, particularly understanding diversity in the context of recruitment. However, we view this complexity as an opportunity to explore this topic using the scientific method. Thus, our team’s future research will focus on elucidating a clearer understanding of recruitment and how this impacts diversity across all medical specialties in the hope of identifying a data-driven approach toward the residency selection process. The proposed multidimensional statistical model in this pilot study represents a starting point in this effort.\textsuperscript{10} Future research will utilize a machine learning computational model founded on the principles of a natural language processing system and mathematical modeling in order to clarify which attributes (including skills, knowledge, past achievements including academic markers such as test scores and publications, and life experience) of medical students predict “success” as a physician or specialist. To do this, we aim to obtain formal data on residency applicants as well as current residents and their clinical performance. Access to these data will allow a more robust model to be created.

**Study Limitations**

First, our analyses, interpretations, and conclusions presented in this paper are inherently limited by publicly available data. Acquisition of additional diversity data on medical students and applicants to residency programs is necessary for a comprehensive and consistent assessment of diversity. Second, we acknowledge that our use of the USNWR research rank as a measure of academic opportunities and resources intrinsic to a medical school may not be a perfect proxy. Additionally, we acknowledge a variety of factors besides medical school rank and research resources may influence an institution’s ability to produce a high volume of students that match into neurosurgery residency.

**Conclusions**

We demonstrate the feasibility of a multidimensional statistical model as a measure to understand the complex issues of diversity. In recent decades and years, neurosurgery has made great progress in the arena of diversity and has shown a strong desire to do more. Our present study allows the neurosurgery community to better understand how to optimize systemic efforts to turn this desire into a reality. Utilization of our proposed model will help the neurosurgical field, and specifically residencies, to identify and utilize attributes of interest that reflect their definition of diversity to monitor their progress along this valuable journey.

**References**


Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Borden, Callas, Benzil. Acquisition of data: Borden, Eysau, Holden, Shaw. Analysis and interpretation of data: Borden, Mahajan, Callas. Drafting the article: Borden, Mahajan, Callas, Benzil. Critical revising the article: all authors. Reviewed submitted version of manuscript: Borden. Approved the final version of the manuscript on behalf of all authors: Borden. Statistical analysis: Borden, Callas. Administrative/technical/material support: Borden. Study supervision: Benzil.

Supplemental Information
Previous Presentations
An abstract of this work was submitted for presentation at the 2021 AANS Annual Scientific Meeting held in Orlando, Florida, on August 21–25, 2021.

Correspondence
Jonathan H. Borden: University of Vermont Medical Center, Burlington, VT. jonathan.borden@uvmhealth.org.