

# Educational Research: Why Medical Students Choose Neurology

## A Computational Linguistics Analysis of Personal Statements

Sarah Grzebinski, BA, Helen Cheung, MD, Charles Sanky, MPH, Jessica Ouyang, PhD, and Stephen Krieger, MD

*Neurology*® 2021;97:e103-e108. doi:10.1212/WNL.00000000000011753

### Correspondence

Dr. Krieger  
stephen.krieger@mssm.edu

## Abstract

### Objective

To understand medical students' motivations for choosing neurology and how applicants conceptualize the field, as this information can be used to enhance interest in neurology and develop educational programs to help identify, support, and recruit future neurologists.

### Background

Applicants to neurology residencies submit personal statements describing themselves and their motivations. Textual analysis of personal statements has been performed in internal medicine and general surgery, but never before in neurology. We hypothesized that specific words and themes would be mentioned in residency personal statements with high frequencies indicating students' motivations.

### Methods

We used computational linguistics software to assess key words and thereby study motivations, expectations, and themes present among neurology applicants. A total of 2,405 personal statements submitted over 5 years to our institution were de-identified and compiled into a database for evaluation through 3 computational linguistics software programs. We performed calculations of term frequencies (TF) and TF-inverse document frequencies and performed K-means clustering to identify unique words and common themes.

### Results

Specific disease states were discussed. For example, stroke (TF 2,178), epilepsy (TF 970), and dementia (TF 944) were referenced more often than amyotrophic lateral sclerosis (TF 220) and carpal tunnel (TF 10). The most common proper names cited were Oliver Sacks (TF 94) and Sherlock Holmes (TF 41). Common themes included fascination with the brain, interest in research, desire to help patients, early interests in neurology, continued pursuit of learning, appreciation for time with patients, family history with neurologic illness, and intellectual curiosity.

### Conclusions

This first computational linguistic analysis of neurology personal statements provides understanding into medical students' motivations and interests. Ongoing subgroup and thematic analyses may inform educational strategies and enhance recruitment to our field.

---

From Columbia University (S.G., J.O.); and Icahn School of Medicine at Mount Sinai (H.C., C.S., S.K.), New York, NY.

Go to [Neurology.org/N](https://www.neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

## Glossary

AAMC = Association of American Medical Colleges; ERAS = electronic residency application service; SIGN = Student Interest Group in Neurology; TF = term frequency; TF-IDF = term frequency–inverse document frequency.

Each year, thousands of medical students apply to residency programs through the Association of American Medical Colleges (AAMC) electronic residency application service (ERAS).<sup>1</sup> The application requires a personal statement as well as transcripts, letters of recommendation, a “dean’s letter” or medical student performance evaluation, extracurricular activity descriptions, and a curriculum vitae.<sup>2</sup> Without a specific prompt, personal statements allow students to personalize their application by describing themselves, their motivations, and their aspirations for entering a specific field of medicine. Consequently, personal statements offer insights into what drew applicants to specific specialties and may yield important insights for medical education and specialty recruitment. Previous studies have analyzed personal statements in pediatrics, internal medicine, and general surgery resident applications using computational linguistics.<sup>3–5</sup> These studies identified shared narrative themes in personal statements and found significant differences between the personal statements of male and female applicants. The findings of these studies informed our hypothesis and helped us identify themes and differences in the field of neurology. To date, there have been no such analyses of neurology applicants’ personal statements.

Using computational linguistic analysis of personal statements, we sought to investigate common themes, motivations, and expectations of medical students applying to neurology residency.

## Methods

### Study Population

We collected 2,405 personal statements submitted through ERAS to the Icahn School of Medicine at Mount Sinai residency program in neurology between 2013 and 2017.<sup>1</sup> Each personal statement was de-identified and compiled into a large secure database on an institutional encrypted server, termed a “corpus.” Demographics and applicant characteristics were extracted from the database using R-programming. We then grouped essays by self-reported gender, medical school location (United States or international medical graduate), race, and age. As the qualifiers of self-reported gender and medical school location provided the largest grouping of essays and therefore an increased chance of significance in data analysis, these qualifiers were chosen to then divide the personal statement corpus into secondary databases “corpora” for further analysis. The project was reviewed and deemed exempt by the Mount Sinai institutional review board, and permission to use the ERAS data was received from the AAMC.

### Data Analyses

Linguistic analysis was performed using term frequency (TF), TF–inverse document frequency (TF-IDF), and K-means clustering. TF identified patterns and frequently repeated phrases and words in a given corpus through Antconc, a text analysis toolkit.

TF outputs the most common words in the 2,405 personal statements and their frequencies. TFs were calculated for general terms and disease states. TF-IDF analyses indicated how characteristic a word was of the entire collection of personal statements and whether the use of that word distinguished the document from the rest of the collection. TF-IDF analyses and calculations of the largest absolute difference between male and female cohorts were performed on male and female corpora to determine what words were characteristic of each cohort. Larger TF-IDF values indicated that a specific word was more unique to its cohort than words with lower scores.

Batched K-means clustering is a method of finding similar topics on a sentence level using the Sci-kit Learn Python Library for machine learning. The program takes each sentence in the corpus and sorts it into clusters of similar sentences to identify common topics. The algorithm was run 3 different times to ensure coherent and concise clusters with no extreme outliers. K-means clustering was also performed after filtering for sentences that included the words “looking forward” to assess applicant aspirations. Authors independently and blindly reviewed clusters and scores to identify common themes.

## Results

Of the 2,405 applicants whose personal statements were analyzed, 53% self-identified as male, 45% as female, and 2% did not indicate a preferred gender. The applications were received from 97 different countries and 641 medical schools worldwide: 56% of applicants applied from international medical schools and 44% of applicants applied from US medical schools.

Of the most referenced disease states identified by Antconc, the 3 most frequently mentioned were stroke (TF 2,178), epilepsy (TF 970), and dementia (TF 944). TF-IDF analyses by gender determined words that were most characteristic of male cohorts and female cohorts and identified the largest differences between male and female cohorts (table 1). The presence of these words indicated a higher likelihood of having been written by a female or male respondent.

**Table 1** Term Frequencies for Female and Male Applicants: Words Characteristic of Female Corpora, Male Corpora, Both Female and Male Corpora, as well as the Largest Differences Between Male and Female Corpora

<b>Greatest differences:</b> F > M (TF-IDF)	<b>Female</b> (TF-IDF)	<b>Greatest overlap</b> (female and male: TF-IDF)	<b>Male</b> (TF-IDF)	<b>Greatest differences:</b> M > F (TF-IDF)
Grandmother (0.000755)	Health (0.002563)	Stroke (female: 0.002561) (male: 0.002371)	Human (0.002368)	Computer (0.000814)
Puzzle (0.00067)	Love (0.002327)	Brain (female: 0.00247) (male: 0.002303)	Train (0.002366)	Technology (0.000581)
Language (0.000652)	Student (0.002239)		Teach (0.002291)	Highlight (0.000444)
Dance (0.00063)	Family (0.00223)			EEG (0.00044)
Culture (0.000621)	Neuroscience (0.002214)			Neurophysiology (0.00043)
Mother (0.000594)				Benefit (0.000418)
Support (0.000547)				Suffer (0.000412)
Women (0.000531)				

Abbreviation: TF-IDF = term frequency–inverse document frequency.

K-means analysis identified consistent themes found in personal statements (salient examples in table 2). Themes included (1) fascination with the brain and interest in neuroscience and the mysteries encountered while studying the brain, (2) references to past research and interest in future research experience, (3) the desire to help those in need to heal, (4) early interest in and exposure to neurology, (5) continued pursuit of learning and potential for discovery, (6) appreciation for time and relationships with patients, (7) shared family connection or history with neurologic illness, and (8) intellectual curiosity/interest in puzzles and problem-solving.

After sorting for “looking forward” statements, the following desires were found: (1) contribution to research, (2) expansion of knowledge and continuity of learning, (3) interaction with a diverse patient population, (4) enhancement of patient quality of life, (5) motivation to become a better physician and improve personal character, and (6) contribution and connection to a team.

## Discussion

In a recent large survey study on characteristics of medical students pursuing a career in neurology, Gutmann et al.<sup>6</sup> note that recruiting future neurologists requires a deeper understanding of the reasons why medical students choose to enter this field. Using a large-scale computational linguistics approach to neurology residency application personal statements, we used applicants’ own words to identify several shared themes for choosing neurology, and found differences in the way female and male applicants conceptualized the field. These data may offer a

better understanding of the factors that motivate students to pursue training in neurology, which can in turn inform educational approaches. Incorporating these themes into medical school neuroscience courses, neurology clerkships, and Student Interest Group in Neurology (SIGN) events may offer neurology educators an opportunity to enhance recruitment.

Many applicants referenced an early interest and exposure to neurology, often describing their first courses in neurology sparking their desire to learn more about how the brain works. Similar findings were identified in the Gutmann et al.<sup>6</sup> survey study, in which high ratings of basic neuroscience courses and clerkships were associated with a greater likelihood of entering the field of neurology. However, as they point out, one limitation of their study is that the AAMC surveys do not capture students’ perceptions of neurology as a specialty. Negative student preconceptions about neurology—that it is impenetrably complicated and offers little by way of treatments for patients with neurologic disease—are often grouped together under the concept of “neurophobia.”<sup>7</sup> While neurology educators have endeavored to dispel those negative stereotypes about neurology, our study identifies that students develop connections with neurology drawing from their personal and family health experiences, the desire to forge relationships with patients, and the potential for future research and discovery. Our study suggests that incorporating these specific themes into the preclinical neuroscience curriculum could help destigmatize the field early in students’ medical training. Surveying medical students at the end of their first-year introductory neurology course regarding the themes we have identified could allow educators to target students who respond positively with increased early exposure to neurology.

**Table 2** K-Means Clustering: General Themes and Emblematic Sample Sentences

**Theme 1: Fascination with the brain**

"After further study of the brain, I became even more and more inquisitive, which led to so many questions on how the brain worked: why people act the way they do, how the body functions as one, and how thoughts and emotions emerged."

"As my studies progressed, I became fascinated, not just by how the brain worked, but also how the mind worked; memory, learning and cognition."

"Such cases challenge me and make me want to learn more about 'the way the brain works.'"

"Over time and through my life, research, and clinical experiences, I have realized that nothing fascinates me more than the complexity and mystery of the human brain."

**Theme 2: Research (inspired by; desire to contribute)**

"I hope to continue my medical training in a neurology residency program with a strong combination of both clinical education and research."

"As my extensive research experience shows, I have a great deal of interest in clinical research and hope to apply my research skills to expand the current knowledge of neurologic disease."

"During the residency program, I want to continue to refine my clinical and interpersonal skills, pursue clinical research, and share teaching experiences with other residents and medical students."

"I am specifically interested in the different research opportunities and the research residency track of your program since it offers the combination of clinical work and simultaneous scientific research I am looking for in my medical career."

**Theme 3: Helping patients**

"My goal is to help serve my community as a helpful, compassionate, and knowledgeable physician."

"In doing so, patients are better served as we help them heal and as we help them die."

"I was captivated and knew from that day I wanted to be able to help people the way the neurologist helped me."

"I have a passion to help those in need and I believe that as a neurologist I will play a vital role in helping both patients and their families."

**Theme 4: Early interest in neurology**

"Neurology has been my passion from the very first year of medical school."

"I first became interested in the nervous system in my first-year neurosciences course."

"My passion for neurology began at the end of my first year of medical school, with my first neuroanatomy lectures."

"My interest in neurology began in my first year of medical school."

**Theme 5: Continued pursuit of learning**

"What fascinated me the most about the field was not how much is known about it, but how much is yet to unfold."

"There is still much to discover."

"There is so much left to learn and so much new research going on in the field that I can think of no better time to begin training in neurology."

"There is still much to be learned about the human brain and great room for improvement in the neurology field."

**Theme 6: Time and relationship with patients**

"I appreciate that neurology provides a wide variety of opportunities from direct patient care to clinical research, and that it reaches patients in almost every care setting from acute emergencies through long-term and intensive care."

"I also realized how crucial it is to educate patients and spend time needed to communicate to patients and their families about their illness."

"I envision a career as a neurologist who will be actively involved in patient care, research, and education dedicated to improving the lives of patients and populations of patients."

"Patient-focused care is the most important aspect of medicine to me, so I look forward to being part of a multidisciplinary team offering the best available care for patients and their families."

**Theme 7: Family history of neurologic illness**

"My grandmother suffered from Alzheimer's disease and my father is living with Parkinson's disease."

"I come from a family whose members suffer from neurologic diseases like Parkinson and Alzheimer diseases."

"I remember being a young child watching my grandfather struggle with both Alzheimer's disease and Parkinson's disease."

"Having grappled with neurologic diseases in my own family I know all too well that the opportunity to preserve a life, or even improve a person's quality of life is a reward without monetary value."

Continued

**Table 2** K-Means Clustering: General Themes and Emblematic Sample Sentences (continued)

**Theme 8: Interest in problem solving and deductive reasoning**

"I love problem-solving and puzzles, in general, and always have."

"I will bring this commitment and love of problem-solving to my neurology residency."

"So often, patients with neurologic disease come to clinic with problems that cannot easily be solved."

"Neurology fulfilled my intellectual curiosity and passion for problem solving."

Applicants also referenced a fascination with the brain and its mysteries. The most frequently named proper nouns were Oliver Sacks and Sherlock Holmes, often mentioned as sources of inspiration. Sacks was known for his stories of finding solutions to diagnostic mysteries.<sup>8</sup> The problem-solving distinguished by Sacks and Holmes is a significant factor that leads medical students to neurology. Additional emphasis on neurology's unique physical examination and localization methodologies, particularly in clinical venues such as neurology clerkships, may distinguish our field on the basis of the intellectual approach essential to neurology, and may more overtly appeal to students who find gratification in this sort of clinical problem-solving.

According to applicants, the ability to research and make discoveries in neurology strongly drew them to the field. Creation of more research opportunities and highlighting new therapeutic advancements in neurology, for instance in neuroscience curricula and SIGN events, may facilitate medical student recruitment. Similarly, applicants were motivated by the ability to serve patients and develop relationships with them. Medical student clerkships can foster an interest in neurology with increased first-hand experiences in history-taking, patient care, and goals of care discussions, towards the goal of yielding the meaningful student-patient relationships often discussed in applicants' personal statements.

There are several differences between the male and female personal statements, specifically in their most frequently mentioned terms. Male respondents were more likely to highlight technology and training; female respondents were more likely to mention love, culture, family members, and support. Men seemed to highlight experiences and accomplishments; women were more likely to emphasize teamwork and mentorship. However, our analysis does not necessarily parse out whether these gender differences stem from distinct reasons for interest in neurology or from societal, gender-normative expectations as to how men and women are expected to communicate in a professional setting. Specifically, group achievement over individual achievement and emphasis on nurturing are well-documented societal gender norms, also consistent with the studies performed in internal medicine and general surgery.<sup>4,5</sup> One study demonstrated that when societal pressures to please are removed, such as through an explicit writing prompt for a legal memo, gender differences disappeared.<sup>9</sup> As such, these differences may be in part attributed to gender-normative traditions of communication, bias, the types of mentor and role models available, and

professional development opportunities that may differ for men and women. The historical and current demographic composition of neurology physicians may also shape how underrepresented social groups approach and ultimately decide upon the field. Neurology training programs should recognize diverse motivations and nurture an inclusive environment of these nuances when fostering professional development and mentorship.

Furthermore, a more detailed appreciation of the most common motivations expressed by those joining the field of neurology can potentially inform neurology education beyond recruitment alone. Neurology residency programs are broadly responsible for creating a fulfilling educational environment for their trainees. A question raised by our study is how often the topics and themes that drew residents to neurology comprise their residency experience. In this respect, it was notable that the frequency that specific disease states were mentioned by applicants does not reflect the true prevalence of those conditions. Headaches were mentioned by 8.6% of applicants, and carpal tunnel syndrome by only 0.4%, though the prevalence of these conditions—particularly in neurology resident clinic—is considerably higher.<sup>10,11</sup> Noting that burnout among neurologists and neurology residents is a concern in our field,<sup>12</sup> we hypothesize that a misalignment between trainee expectations and career motivations (why someone seeks to become a neurologist) and the realities of clinical neurology could be one component of professional dissatisfaction. To explore this potential contributor to burnout, a future study could examine the extent to which both the diseases and the most common themes we have identified in neurology application personal statements are congruent with the clinical experiences in neurology residency training and clinical practice.

Several limitations warrant discussion. Our investigation included 2,405 personal statements from 2013 to 2017. Although we only examined personal statements sent to one institution, the large global sample of applicants allows for findings to be generalized to many neurology residency programs across the United States. The software used to perform linguistic analyses provided raw scores and terms without interpretation. Therefore, our study required authors to review the output for interpretation. Authors independently and blindly reviewed findings to overcome subjective biases. Our quantitative linguistic analysis methodology makes it more difficult to capture important themes present in a smaller

cohort of the corpus if they did not reach statistical significance. For this reason, we focused our analyses on the 2 largest subcorpora: that of male and female applicants (comprising 53% and 45%, respectively). These analyses utilized a gender dichotomy due to 2% not identifying as male or female, but we recognize that there are a variety of identities across the spectrum of gender, about which we did not have further information or a large enough subset to draw conclusions. Therefore, our binary approach to gender presents a limitation, possibly due to fear or anticipated bias associated with disclosing nonbinary gender identity or ERAS applications not providing space for further clarification.

In the future, further evaluation of essay groupings using qualitative methods could identify important differences by self-reported race/ethnicity, medical school location, additional advanced degrees, and age. Another limitation is that our current study evaluated only neurology applicants' personal statements. A subsequent study could use this methodology to evaluate personal statements for all neurology and internal medicine applicants to a single institution to look for between-field differences that could further delineate neurology-specific motivations.

We also acknowledge that personal statements may carry an inherent bias as medical students may be writing what they believe an evaluation committee wants to hear rather than what truly draws them to the field of neurology. In addition, the personal statements include a retrospective, ex post facto sense of what inspired their interest in neurology, and real-time assessment through medical school training may further inform future education and recruitment efforts.

Through this first linguistic analysis of personal statements in neurology residency applications, we gained a deeper understanding of how medical students perceive neurology as well as their hopes and motivations for entering the field. We identified that several themes—including personal connections to patients with neurologic disease, the deductive reasoning that distinguishes our field, and the potential for future research and discovery—are powerful motivators for students to apply in neurology, suggesting that embedding these concepts in neurology curricular exposure may foster interest and potentially enhance recruitment to our field. Ongoing efforts include further subgroup thematic analyses, comparative studies with applications to other fields, and extending this study over subsequent years to evaluate changing themes of interest among neurology applicants.

## Acknowledgment

The authors thank Matthew A. Sheridan for his technical expertise in data extraction and compilation.

## Study Funding

No targeted funding reported.

## Disclosure

S. Grzebinski, H. Cheung, C. Sanky, and J. Ouyang have nothing to disclose. S. Krieger reports consulting or advisory work with Biogen, EMD Serono, Genentech, Genzyme, Mallinckrodt, MedDay, Novartis, Teva, and TG Therapeutics and nonpromotional speaking with Biogen, EMD Serono, and Novartis. Go to [Neurology.org/N](http://Neurology.org/N) for full disclosures.

## Appendix Authors

Name	Location	Contribution
<b>Sarah Grzebinski, BA</b>	Columbia University, New York	Interpreted data, drafted the manuscript for intellectual content, revised the manuscript
<b>Helen Cheung, MD</b>	Icahn School of Medicine At Mount Sinai, New York	Interpreted data, revised the manuscript for intellectual content, revised the manuscript
<b>Charles Sanky, MPH</b>	Icahn School of Medicine At Mount Sinai, New York	Compiled database, revised the manuscript for intellectual content, revised the manuscript
<b>Jessica Ouyang, MA, PhD</b>	Columbia University, New York	Performed linguistic analysis
<b>Stephen Krieger, MD</b>	Icahn School of Medicine At Mount Sinai, New York	Designed and conceptualized study, interpreted data, revised the manuscript for intellectual content

## References

- Association of American Medical Colleges. Residency Applicants by Specialty 2013–2017 Accessed March 3, 2021. <https://www.aamc.org/data-reports/interactive-data/eras-statistics-data>.
- Association of American Medical Colleges. About ERAS®. Accessed March 3, 2021. <https://www.aamc.org/services/eras-for-institutions/medical-schools>; Accessed 3.3.2021
- Babal JC, Gower AD, Frohna JG, Moreno MA. Linguistic analysis of pediatric residency personal statements: gender differences. *BMC Med Educ* 2019;19:392.
- Osman NY, Schonhardt-Bailey C, Walling JL, Katz JT, Alexander EK. Textual analysis of internal medicine residency personal statements: themes and gender differences. *Med Educ* 2015;49:93–102.
- Ostapenko L, Schonhardt-Bailey C, Sublette JW, Smink DS, Osman NY. Textual analysis of general surgery residency personal statements: topics and gender differences. *J Surg Educ* 2018;75:573–581.
- Gutmann L, Cahill C, Jordan JT, et al. Characteristics of graduating US allopathic medical students pursuing a career in neurology. *Neurology* 2019;92:e2051–e2063.
- Jozefowicz RF. Neurophobia: the fear of neurology among medical students. *Arch Neurol* 1994;51:328–329.
- Marshall J. In the region of lost minds. *New York Times*; March 2, 1986.
- Larson BN. Gender/genre: the lack of gendered register in texts requiring genre knowledge. *Written Commun* 2016;33:360–384.
- Abu-Arafeh I, Razak S, Sivaraman B, Graham C. Prevalence of headache and migraine in children and adolescents: a systematic review of population-based studies. *Dev Med Child Neurol* 2010;52:1088–1097.
- Atroschi I, Gummeson C, Johnsson R, Ormstein E, Ranstam J, Rosén I. Prevalence of carpal tunnel syndrome in a general population. *JAMA* 1999;282:153–158.
- Busis NA, Shanafelt TD, Keran CM, et al. Burnout, career satisfaction, and well-being among US neurologists in 2016. *Neurology* 2017;88:797–808.

# Neurology<sup>®</sup>

## **Educational Research: Why Medical Students Choose Neurology: A Computational Linguistics Analysis of Personal Statements**

Sarah Grzebinski, Helen Cheung, Charles Sanky, et al.

*Neurology* 2021;97:e103-e108 Published Online before print March 3, 2021

DOI 10.1212/WNL.0000000000011753

**This information is current as of March 3, 2021**

<b>Updated Information &amp; Services</b>	including high resolution figures, can be found at: <a href="http://n.neurology.org/content/97/1/e103.full">http://n.neurology.org/content/97/1/e103.full</a>
<b>References</b>	This article cites 9 articles, 2 of which you can access for free at: <a href="http://n.neurology.org/content/97/1/e103.full#ref-list-1">http://n.neurology.org/content/97/1/e103.full#ref-list-1</a>
<b>Subspecialty Collections</b>	This article, along with others on similar topics, appears in the following collection(s): <b>All Education</b> <a href="http://n.neurology.org/cgi/collection/all_education">http://n.neurology.org/cgi/collection/all_education</a> <b>Professional conduct and ethics</b> <a href="http://n.neurology.org/cgi/collection/professional_conduct_and_ethics">http://n.neurology.org/cgi/collection/professional_conduct_and_ethics</a>
<b>Permissions &amp; Licensing</b>	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: <a href="http://www.neurology.org/about/about_the_journal#permissions">http://www.neurology.org/about/about_the_journal#permissions</a>
<b>Reprints</b>	Information about ordering reprints can be found online: <a href="http://n.neurology.org/subscribers/advertise">http://n.neurology.org/subscribers/advertise</a>

*Neurology*® is the official journal of the American Academy of Neurology. Published continuously since 1951, it is now a weekly with 48 issues per year. Copyright © 2021 American Academy of Neurology. All rights reserved. Print ISSN: 0028-3878. Online ISSN: 1526-632X.

