



OPEN Influence of race and gender on variability in clinical decision-making in neurology

Randi A Paynter¹, Nina Bozinov², Saadiya Hawa³ & Julia F. Simard^{1,4}✉

Disparities in outcomes have been reported in numerous neurologic diseases. Examining clinical decision-making patterns and behaviors can inform points of intervention to improve disease management and outcomes. This experimental vignette study sought to measure the effects of race (Black, White) and gender (female, male) on variation in clinical decision-making for remitting-relapsing multiple sclerosis (RRMS), primary progressive MS (PPMS), stroke, Parkinson disease, and epilepsy. A factorial survey study using five clinical vignettes was distributed to neurologists in the United States. Race and gender were randomly modified across all vignettes. Correct diagnoses for each vignette were the main outcome and additional outcomes included response time and planned next steps (as a proxy of certainty). 621 neurologists completed the study. Participants consistently diagnosed stroke correctly, independent of race/gender variant. However, we noted some variability for the other conditions. RRMS was correctly diagnosed among White female vignettes 89% of the time, compared to 81–82% for the other vignette types ($p = 0.13$). For PPMS vignettes, females were more likely to be correctly diagnosed than males, and White vignettes more likely to be correctly diagnosed than Black vignettes. We observed some variability in treatment initiation as the specified next step across vignettes, as well as response time. We found that diagnosis and certainty varied by race and gender. The extent to which frequency of correct diagnosis varied suggests uncertainty and reliance upon the epidemiologic knowledge base, particularly as the distribution of correct diagnoses corresponds with the evidence base.

Keywords Diagnosis, Multiple sclerosis, Cognitive bias, Clinical vignettes

Across a range of conditions – e.g., osteoporosis, left-sided colon cancer, and systemic lupus erythematosus (SLE) – patients who represent the less commonly impacted demographic have poorer outcomes and more severe manifestations^{1,2,3,4}. For example multiple sclerosis (MS) is generally considered a female-predominant disease but male patients tend to have worse prognoses. Some have proposed that these types of disparities may be due, in part, to diagnostic delays or misdiagnosis. Interrogating variation in clinical decision-making can identify actionable points of intervention that can improve disease management and outcomes and inform the evidence base for the practice of evidence-based medicine.

Experimental vignette studies offer a methodological approach to understanding variation in clinical decision-making^{5,6}. These factorial studies use identical clinical case vignettes that vary only by the experimental factors under study. This approach has been used to study variation in the delivery of care for common yet complex diseases and conditions in areas such as cancer, cardiovascular disease, pain management, and mental health⁵. Using this approach, we previously showed that case vignettes presented as male were less likely to be correctly diagnosed with SLE by physicians, particularly White males compared with Black females⁷.

To date, there have been little to no data about variation in clinical decision-making among neurologists about common yet complex neurologic diseases and conditions. This study aimed to use an experimental vignette design to measure the effects of race (Black, White) and gender (female, male) on variation in clinical decision-making for remitting-relapsing MS (RRMS), primary progressive MS (PPMS), stroke, Parkinson disease (PD), and epilepsy.

¹Department of Epidemiology and Population Health, Stanford School of Medicine, Stanford, California, US.

²Kootenai Clinic, Neurology, Kootenai Health, Coeur d'Alene, ID, USA. ³Louis A. Weiss Memorial Hospital, Chicago, Illinois, US. ⁴Division of Immunology and Rheumatology, Department of Medicine, Stanford School of Medicine, Stanford, CA94305-5405, US. ✉email: jsimard@stanford.edu

Methods

Standard protocol approvals, registrations, and patient consents

This study was approved by the Institutional Review Board of Stanford University, Stanford, California, and followed AAPOR guidelines⁸. Participants provided informed consent to participate in the study survey before proceeding to provide any data or responses.

Survey design

Five vignettes were drafted by a neurologist (NB), then reviewed and pilot-tested by 30 Stanford neurologists who received a \$20 gift card through Tango Rewards. The pilot and final survey were distributed via Qualtrics (Qualtrics, Provo, UT) and used block randomization to randomize which race-gender version was shown. Following consent, background and practice-related questions were asked before participants were presented the case vignettes. The case vignettes were given in the following order: stroke, MS, PD, epilepsy, MS, where the sequence of the two MS subtypes (RRMS and PPMS) was randomized to account for potential ordering effects (Fig. 1). Each vignette and its associated questions were included on the same survey page. Participants were asked, “What is the most likely diagnosis?” (free text response), “Is your next step diagnostic or therapeutic?” (dichotomous), and “Please briefly tell us what this next step would be:” (free text response; see eMethods 3 in **Supplement 1**).

Survey distribution

A mailing list of neurologists was provided by SPAN Global Services/LakeMedia Group, a medical marketing company. 22,085 email addresses for neurologists were received on 6/1/2022. 332 duplicate addresses were removed. 21,753 survey invitations to neurologists were sent on 8/16/2022, with randomized incentives of \$0, \$10, \$20, \$50, or \$75 gift cards (J. Simard, ScD, unpublished data, February 2024). Four follow-up emails were sent, the final on 9/9/2022. On 10/12/2022, those who had neither opted out nor completed the survey were offered \$75 gift cards to participate. Two follow-up emails were sent, and the survey was closed on 11/15/2022.

Study population

We received 949 survey responses but initially excluded 257 who responded “No” to the first question (“Are you a Neurologist?”). An additional nine were excluded because the question was not answered. Of the 949 survey responses, Qualtrics flagged 8 as possible spam; 7 of which were excluded based on responses to the first question. We further excluded 1 survey because the participant reported living outside the US, and an additional 13 surveys missing location. One of the surveys excluded for no location was flagged as spam. Of the remaining, we excluded surveys with no ($n = 27$) or only some ($n = 21$) vignettes completed, yielding 621 US neurologists participating in the present study (Fig. 2).

Outcomes

First, we identified correct diagnoses for each vignette blinded to the participant’s data and which race-gender version they saw (details provided in eMethods 1 in **Supplement 1**). As outlined in the supplementary material, we used string searches to parse write-in text responses. These were categorized and then compared to the correct diagnosis. Planned next steps (“therapeutic” or “diagnostic”) among the correct diagnoses was a proxy of providers’ certainty in their diagnoses, assuming that a choice of “therapeutic” indicates more certainty than selecting “diagnostic”. We assumed that individuals who opted to treat a condition (therapeutic next step) would be more certain in their diagnosis than those who opted to seek more information through laboratory and other instrumental findings (diagnostic next step). For each vignette a secondary outcome of response time (i.e., estimated time spent on the vignette) was calculated as the difference between the date/time stamps for submitting the final answer (finishing the vignette) and first opening the page.

Survey demographics

Demographic data included self-reported gender (Female or Male or Other) and year of birth. Age was calculated by subtracting year of birth from 2022. Participants could specify “Hispanic” or “Not Hispanic” ethnicity and separately asked about race: “What is your race? Please select all that apply.” Participants were provided with multiple choice check boxes, along with an “Other” option with an accompanying text box to fill in. Classification of “Other” responses, along with US geographic information, are provided (eMethods 2 in **Supplement 1**). The complete survey is included in eMethods 3 in **Supplement 1**.

Analysis

Data were exported from Qualtrics and cleaned and analyzed in SAS version 9.4 (SAS Institute Inc., Cary, NC). After all vignette responses for diagnoses were classified, we merged the classified correct/incorrect codes with the participants’ demographic data, timestamps, and indicator for which race-gender vignette versions were shown for each participant. Descriptive statistics were calculated to characterize the study population and presented as frequencies and proportions. Our primary comparisons were between two subtypes of MS and comparing both MS subtypes with the common neurologic conditions of stroke, PD, and epilepsy. The proportion of correct responses by race-gender version for each vignette was calculated and heterogeneity by race-gender version was assessed using a Chi-square test on an RxC table to generate P values. eFigures 1–3 in **Supplement 1** were produced using R version 4.3.1⁹. The distribution of participants’ next steps (diagnostic vs. therapeutic) was then estimated by vignette race-gender version, separately for correct and incorrect responses. To calculate the theoretical proportion of patients successfully diagnosed and treated without delay, the fraction of correct diagnoses that were also recommended “therapeutic” as next step was calculated from all vignettes randomized to each race-gender version (eTable 3 in **Supplement 1**). As a secondary analysis, we estimated

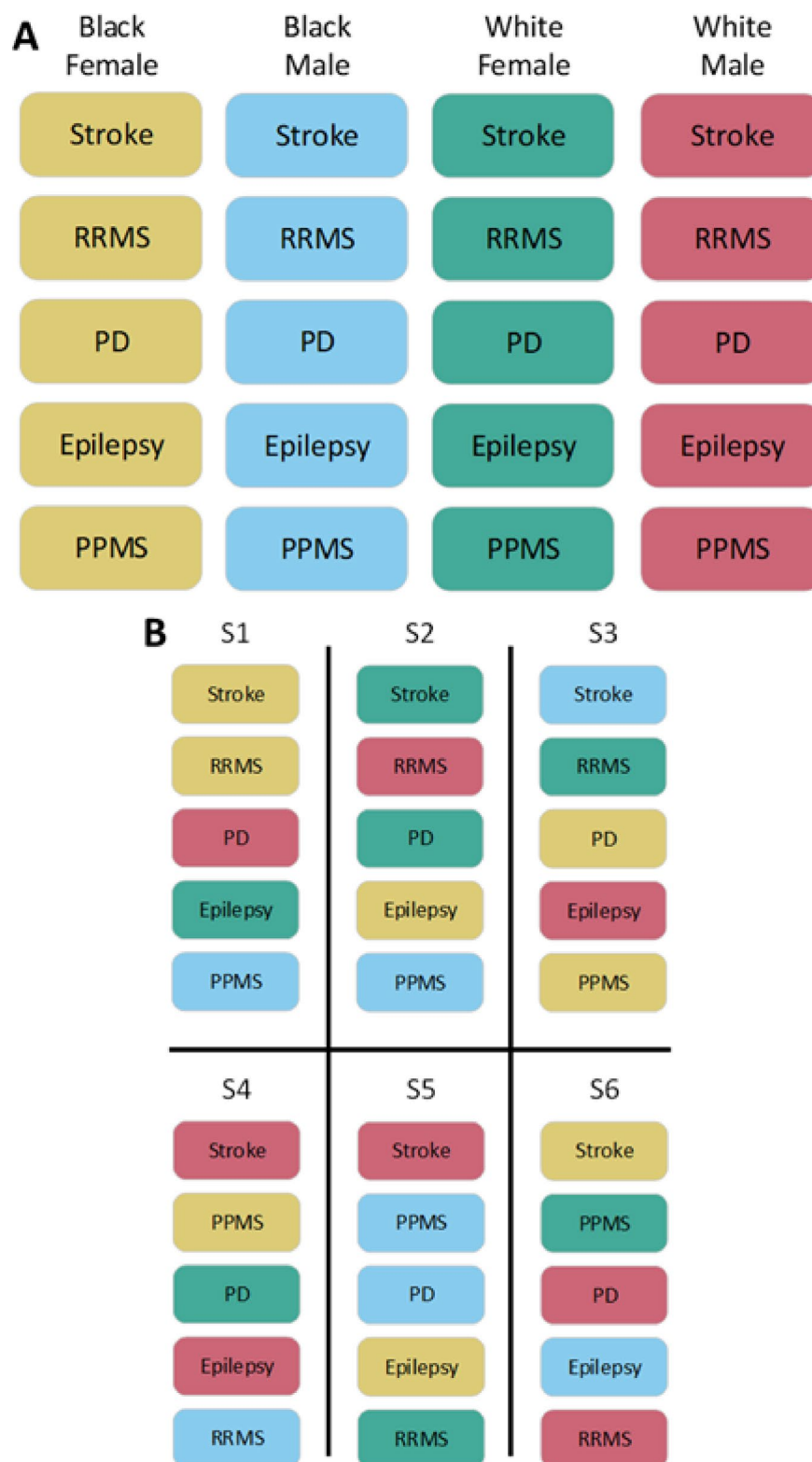


Fig. 1. Schematic for pool of case vignettes and randomization mechanism used in a survey-based study of how patient race and gender influence clinical decision-making among neurologists, United States, August—November 2022. Abbreviations: RRMS, relapsing-remitting multiple sclerosis; PD, Parkinson disease; PPMS, primary progressive multiple sclerosis. **(A)** Pool of case vignettes by race and gender; **(B)** an example of 6 possible sets of randomized surveys. Each participant received the same 5 clinical cases, randomized by race and gender. The scenarios were presented in the same order, with the exception of RRMS and PPMS, as shown in S1—S3 compared with S4—S6.

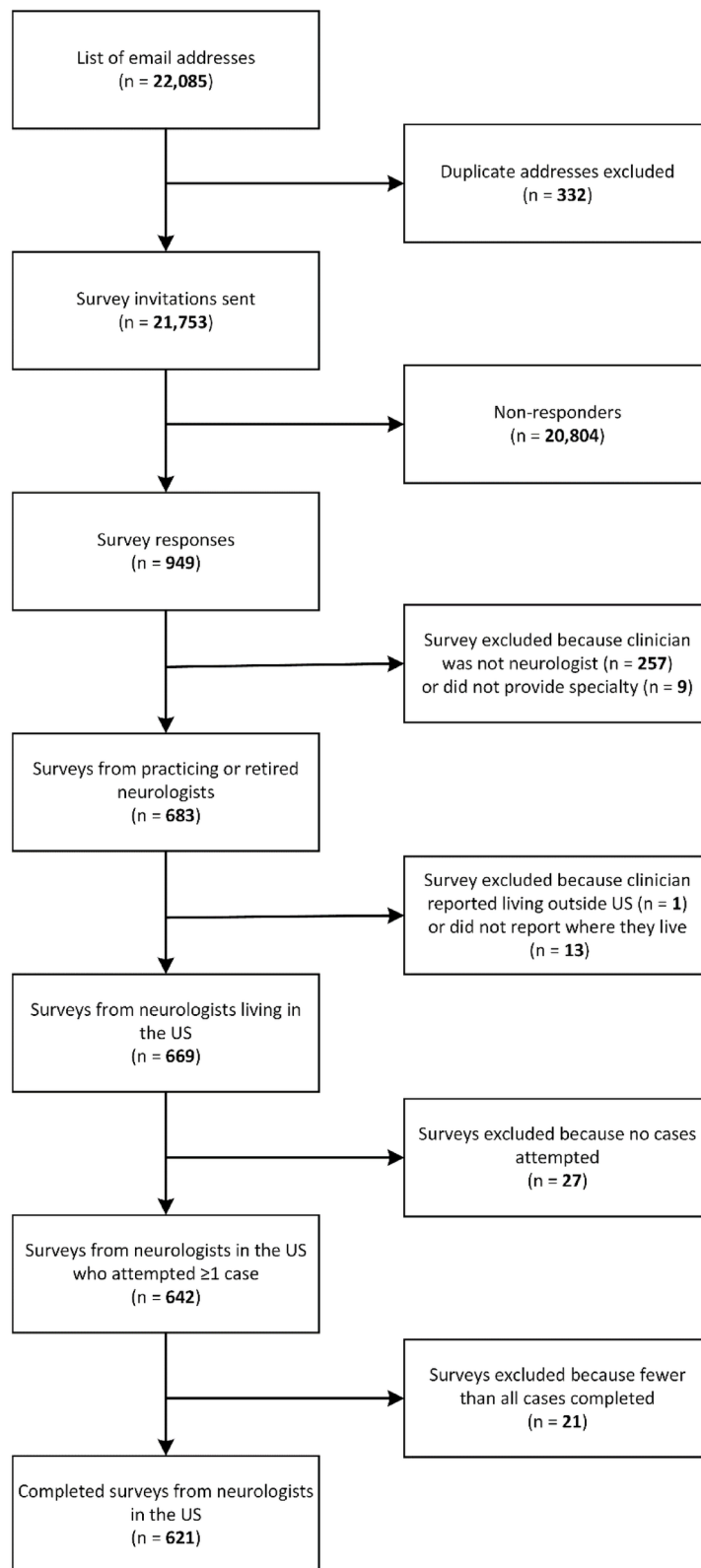


Fig. 2. Study Population Flowchart.

median time taken to complete vignettes by “next step”, by race-gender version and vignette, among correct responses only.

Results

The demographics of the 621 neurologist participants are described in eTable 2 in **Supplement 1**. Our study population is representative of the US neurologist population reported in the American Medical Association (AMA) Physician Masterfile with respect to distribution of gender, ethnicity, and race, and the focus on patient care¹⁰. However, our sample is younger (age <55 years, sample 62% vs. AMA 40.7%), has fewer international medical graduates (sample 25% vs. AMA 31.5%), and reports engaging in more research as a type of work (sample 13% vs. AMA 4.5%). Compared with the US distribution of neurologists in 2012¹¹, our sample has a similar distribution by US geographic division.

Comparing the proportion of correct diagnoses across race-gender versions, stroke was correctly diagnosed consistently across the four race-gender versions (minimum 91%, maximum 93%, $P = 0.96$). (Table 1) The other clinical vignettes were less uniform. For PD, vignettes describing Black persons (Black females, 92%; Black males, 94%) were diagnosed correctly more frequently than White persons (White females, 87%; White males, 91%), and males were diagnosed correctly more frequently than females, but not statistically significant ($P = 0.18$). Correct diagnoses were similar for epilepsy among vignettes describing Black males (91%), White females (93%), and White males (89%), but somewhat lower among Black females (85%) ($P = 0.13$).

In contrast, the two MS case results were distinct and differed from each other. For the RRMS subtype, correct diagnoses were similar among vignettes describing Black females (82%), Black males (81%), and White males (81%), but higher among White females (89%) ($P = 0.13$). PPMS overall was less frequently correctly diagnosed than was RRMS. For PPMS vignettes, females (Black females, 76%; White females, 81%) were more likely to be correctly diagnosed than males (Black males, 70%, White males, 75%), and White vignettes more likely to be diagnosed than Black vignettes ($P = 0.17$).

When examining whether race and/or gender influenced next step among the correct responses, we found some variability across vignettes and the race-gender versions (Table 2, and eFigure 1 in **Supplement 1**). For the PD case vignette, the planned next step was “therapeutic” consistently across all race-gender versions (minimum 80%, maximum 84%). For the stroke vignette, the planned next step was “therapeutic” at comparable levels for Black females (18%), White females (18%), and White males (18%), but higher for Black males (25%). In the epilepsy case vignette, the planned next step as “therapeutic” was lower for Black females (41%) compared with 48% the other three versions.

More variation was observed for planned next steps across race-gender versions for the two MS subtypes (Table 2, and eFigure 1 in **Supplement 1**). Overall, there was a higher recommendation for a “therapeutic” next step for RRMS (range 58–69%) than for PPMS (range 25–36%). For the RRMS correct diagnoses, White female vignettes were the least likely to be recommended a “therapeutic” next step (58%) despite being most likely to be correctly diagnosed. White male vignettes were the most likely to have recommended a “therapeutic” next step

Disease Case	Race and Gender			
Measure	Black Female	Black Male	White Female	White Male
Stroke				
Randomized, No.	128	163	161	169
Correct diagnosis, No. (%)	118 (92)	150 (92)	147 (91)	157 (93)
<i>P</i> for heterogeneity	0.96	–	–	–
RRMS				
Randomized, No.	159	147	150	165
Correct diagnosis, No. (%)	130 (82)	119 (81)	134 (89)	133 (81)
<i>P</i> for heterogeneity	0.13	–	–	–
Parkinson Disease				
Randomized, No.	150	155	166	150
Correct diagnosis, No. (%)	138 (92)	146 (94)	145 (87)	137 (91)
<i>P</i> for heterogeneity	0.18	–	–	–
Epilepsy				
Randomized, No.	150	155	162	154
Correct diagnosis, No. (%)	127 (85)	141 (91)	150 (93)	137 (89)
<i>P</i> for heterogeneity	0.13			
PPMS				
Randomized, No.	139	153	162	167
Correct diagnosis, No. (%)	105 (76)	107 (70)	131 (81)	125 (75)
<i>P</i> for heterogeneity	0.17	–	–	–

Table 1. Proportion of correct diagnoses identified among disease case vignettes with race and gender jointly randomized. Abbreviations: RRMS, relapsing-remitting multiple sclerosis; PPMS, primary progressive multiple sclerosis.

Disease Case	Correct Diagnosis				Incorrect Diagnosis			
	Race and Gender				Race and Gender			
Next Step ^a	Black Female	Black Male	White Female	White Male	Black Female	Black Male	White Female	White Male
Stroke								
Diagnostic, No. (%)	97 (82)	113 (75)	120 (82)	128 (82)	9 (90)	13 (100)	10 (71)	10 (83)
Therapeutic, No. (%)	21 (18)	37 (25)	27 (18)	29 (18)	1 (10)	0 (0)	4 (29)	2 (17)
RRMS								
Diagnostic, No. (%)	45 (35)	46 (39)	56 (42)	41 (31)	21 (72)	24 (86)	13 (81)	25 (78)
Therapeutic, No. (%)	85 (65)	73 (61)	78 (58)	92 (69)	8 (28)	4 (14)	3 (19)	7 (22)
Parkinson Disease								
Diagnostic, No. (%)	27 (20)	24 (16)	24 (17)	27 (20)	8 (67)	7 (78)	12 (57)	8 (62)
Therapeutic, No. (%)	111 (80)	122 (84)	121 (83)	110 (80)	4 (33)	2 (22)	9 (43)	5 (38)
Epilepsy								
Diagnostic, No. (%)	75 (59)	74 (52)	78 (52)	71 (52)	10 (43)	8 (57)	7 (58)	9 (53)
Therapeutic, No. (%)	52 (41)	67 (48)	72 (48)	66 (48)	13 (57)	6 (43)	5 (42)	8 (47)
PPMS								
Diagnostic, No. (%)	71 (68)	80 (75)	84 (64)	91 (73)	27 (79)	39 (85)	25 (81)	33 (79)
Therapeutic, No. (%)	34 (32)	27 (25)	47 (36)	34 (27)	7 (21)	7 (15)	6 (19)	9 (21)

Table 2. Distribution of choices for “next step” among correctly and incorrectly identified diagnoses for case vignettes with race and gender jointly randomized. Abbreviations: RRMS, relapsing-remitting multiple sclerosis; PPMS, primary progressive multiple sclerosis. ^aThe percentage indicated for “next step” is the percent choosing either “diagnostic” or “therapeutic” among correct diagnoses or incorrect diagnoses within each race/gender group.

(69%), followed by Black females (65%) and Black males (61%). For PPMS, on the other hand, we observed the following proportions of recommendation for “therapeutic” next step (in decreasing order): White female, 36%; Black female 32%; White male, 27%; Black male, 25%.

The proportion of diagnoses that were correct and also recommended “therapeutic” as the next step, with the denominator as the total randomized, represents the theoretical proportion that will be successfully diagnosed and promptly treated within each race-gender version (eTable 3 in **Supplement 1**). Although the White female case vignettes were the most likely to receive a correct RRMS diagnosis, they were the least likely to be recommended a therapeutic next step compared with male and Black female case vignettes. In absolute terms, the White female RRMS vignettes yielded a similar probability of treatment as the other race-gender versions. When correct diagnoses were evaluated for next steps for the epilepsy case vignette, Black females were recommended for therapy at a lower rate than the other three race-gender versions. The result is that 35% of the Black female vignettes would be recommended for therapy, compared with 43–44% of the other race-gender versions. Among the stroke case vignettes, 23% of the Black male vignettes were diagnosed correctly and recommended therapy compared with 16–17% of each of the other race-gender versions. Among PD case vignettes, 79% of the Black male vignettes were diagnosed correctly and recommended therapy compared with 73–74% of each of the other race-gender versions. Finally, the proportions for race-gender versions among the PPMS vignette followed the same relative pattern as for correct diagnoses; Black female 34%, Black male 18%, White female 29%, and White male 20%.

We observed some variability in the secondary outcome of median response time among correct responses by race-gender version (eTable 4 and eFigure 2 in **Supplement 1**) as well as when stratified by next step (eTable 5 and eFigure 3 in **Supplement 1**).

Discussion

As physicians experience uncertainty they may rely on prior probabilities from population-based epidemiologic studies to inform the diagnosis of a particular case^{12,13,14,15}. Five clinical vignettes with race and gender randomly assigned were evaluated by 621 participating neurologists. Because of the randomization and the sizable sample drawn from across the US, we would expect participants to perform equally well across each race-gender version within the same clinical case vignette if there was no variation in clinical decision-making due to race or gender.

We observed that stroke, PD, and epilepsy had the highest overall proportions of correct diagnoses. This is consistent with each of the conditions having classic symptoms associated with affected brain anatomy. The two MS vignettes, RRMS and PPMS, both were correctly “diagnosed” less frequently. This is consistent with the heterogeneity in presentation as well as the multiple criteria that must be met for an MS diagnosis. RRMS was diagnosed correctly more frequently than PPMS, consistent with PPMS being more difficult to diagnose due to its insidious and progressive course.

We observed the most variation in clinical decision-making for the RRMS vignette compared with all other conditions. Specifically, the proportion of RRMS vignettes for White females diagnosed correctly was higher than the three other vignettes, but these differences were not statistically significant. Until recently the prevailing belief about MS has been that it is rare among Black persons and more common among females

than males^{16,17,18,19}. Recent US studies also described an overall female predominance, but reported comparable prevalence among White and Black females, especially at the median age at onset^{20,21}. The results from our study suggest reliance on a well-accepted but less well-documented prior, i.e., that MS occurs more frequently among White females^{16,17,18,19}. This is consistent with our prior findings studying rheumatologists and SLE⁷. On the other hand, PPMS, which reportedly occurs with approximately equal frequency between males and females, did not follow the same pattern as RRMS^{22,23,24}. Instead, PPMS results followed the same pattern of overall US MS incidence recently reported for corresponding race-gender versions²⁰.

As a post hoc comparison we considered results by whether conditions onset acutely or insidiously. Both PD and PPMS are progressive and neurodegenerative with insidious onset. The proportion correct among race-gender versions for these two conditions followed a pattern consistent with the currently accepted relationship between gender and risk of each condition. Specifically, PD is diagnosed more frequently in men than in women^{25,26,27}, and MS generally is diagnosed more frequently in women than in men^{20,23}. It should be noted that PPMS comprises only about 15% of all MS cases, and the ratio of women: men is estimated to be more even^{22,23,24}. In contrast, three clinical case vignettes represented conditions with acute onset. Two of these (RRMS and epilepsy) had one race-gender version with a somewhat different proportion of correct diagnoses from the other three race-gender versions. For RRMS, the White female case vignettes had a higher proportion of correct diagnoses than the other versions; for epilepsy, the Black female case vignettes had a lower proportion of correct diagnoses than the other race-gender versions.

We also sought to understand the potential impact of certainty or uncertainty on the management of the neurologic conditions. Although the White female case vignettes were the most likely to receive a correct RRMS diagnosis, they were the least likely to be recommended a therapeutic next step. All versions of the RRMS vignette yielded a similar absolute probability of treatment ranging from 50 to 56%. It may be that RRMS is easier to recognize in “classical” patient, i.e., one who is White and female, but does not translate to higher likelihood of treatment. In contrast we observed differences in diagnosis and initiation of treatment in the epilepsy case vignettes. The Black female epilepsy vignettes were less likely to be diagnosed correctly and less likely to be recommended for therapy than the other race-gender versions. The result in absolute terms is that 35% of the Black female vignettes would be recommended for therapy, compared with 43% of the Black male, 44% of the White female, and 43% of the White male epilepsy vignettes. These results at both diagnosis and next step suggest persistent uncertainty around Black women and epilepsy that warrants further investigation.

Next steps for the stroke vignettes had a high proportion of diagnostic testing, which may be expected as the vignette stated that MRI results were pending. Black male vignettes were recommended therapy more often than others, perhaps reflecting more certainty in the diagnosis consistent with overlap in the risk profile for cardiovascular disease among the US Black male population²⁸. The next step recommendation of treatment for PD was uniformly high, reflecting high certainty in the diagnosis and standard treatment. Finally, next steps for PPMS were generally low for recommending treatment, reflecting low certainty in the diagnosis. Both female versions of the PPMS vignettes had higher proportions of treatment recommended, reflecting higher certainty.

Despite incentives offered for participation in our study, our response rate was low and therefore our statistical power limited for some analyses, such as exploring gender or race concordance or the influence of international training. These data are consistent with results from our prior rheumatology study. For transparency we include results for time taken to answer the group of three questions for each vignette. However, we refrain from drawing any conclusions because we cannot disentangle hesitation or “thinking time” from the time required to enter a lengthy response. In a prior study of rheumatology clinical case vignettes, we felt more confident using response time because we asked a single question regarding diagnosis. The response time in the present study includes the time required to answer the three questions for each vignette. Any future studies employing this randomized survey strategy may benefit from recording time taken for each question separately. As respondents were simply asked for “the most likely diagnosis” for each vignette, we have no insights into what diagnostic criteria were applied in their thinking. Although response time has been previously used as a proxy for implicit bias, our survey design did not allow for its use in this way. Therefore, we were unable to estimate implicit bias among the neurologist participants.

This novel study evaluated variation in clinical decision-making among neurologists from across the US. We used randomization to assign the race and gender of each case vignette, allowing us to hold everything constant except these factors. We also preemptively randomized order to minimize ordering effects. Our findings are consistent with prior work suggesting that similar biases may influence providers when applying diagnostic criteria to a condition with heterogeneous presentation. This variation is pronounced when there is evidence for a race- or gender-based predominance in prevalence. In summary, using clinical vignettes in a randomized survey experiment, we found that diagnosis and certainty may vary by the gender and race of the hypothetical neurology patient. The extent to which frequency of correct diagnosis varied in the present study suggests uncertainty and reliance upon epidemiologic knowledge base, particularly as the distribution of correct diagnoses corresponds with the evidence base.

Data availability

Anonymized study data used in the preparation of this manuscript are available at <https://doi.org/10.17605/OSF.IO/EF38A>.

Received: 25 June 2024; Accepted: 30 April 2025

Published online: 23 May 2025

References

- Rinonapoli, G. et al. Osteoporosis in men: A review of an underestimated bone condition. *Int. J. Mol. Sci.* **22** (4). <https://doi.org/10.3390/ijms22042105> (2021).
- Sun, M., Wang, Y., Sundquist, J., Sundquist, K. & Ji, J. Temporal trends of sex disparity in incidence and survival of colorectal cancer: variations by anatomical site and age at diagnosis. *Clin. Epidemiol.* **12**, 73–81. <https://doi.org/10.2147/CLEPS240006> (2020).
- Mihailovic, J. et al. Worse cardiovascular and renal outcome in male SLE patients. *Sci. Rep.* **13** (1), 18628. <https://doi.org/10.1038/s41598-023-45171-7> (2023).
- Crosslin, K. L. & Wiginton, K. L. Sex differences in disease severity among patients with systemic lupus erythematosus. *Gend. Med.* **8** (6), 365–371. <https://doi.org/10.1016/j.genm.2011.10.003> (2011).
- Sheringham, J., Kuhn, I. & Burt, J. The use of experimental vignette studies to identify drivers of variations in the delivery of health care: a scoping review. *BMC Med. Res. Methodol.* **21** (1), 81. <https://doi.org/10.1186/s12874-021-01247-4> (2021).
- Rossi, P. H. & Nock, S. L. *Measuring Social Judgments: the Factorial Survey Approach* (Sage, 1982).
- Simard, J. F. et al. Are we missing lupus in males?? Evidence of cognitive Bias from a randomized experiment in the united States. *Am. J. Epidemiol.* **191** (1), 230–233. <https://doi.org/10.1093/aje/kwab199> (2022).
- Best Practices for Survey Research. AAPOR: American Association for Public Opinion Research. Accessed January 27. (2024). <https://aapor.org/standards-and-ethics/best-practices/>
- R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (2023). <https://www.R-project.org/>
- Physician Specialty Data Report & American Medical Association (AMA). Physician Masterfile. AAMC (2022). <https://www.aamc.org/data-reports/workforce/report/physician-specialty-data-report>
- Dall, T. M. et al. Supply and demand analysis of the current and future US neurology workforce. *Neurology* **81** (5), 470–478. <https://doi.org/10.1212/WNL.0b013e318294b1cf> (2013).
- Lutfe, K. E., Link, C. L., Grant, R. W., Marceau, L. D. & McKinlay, J. B. Is certainty more important than diagnosis for Understanding race And gender disparities? An experiment using coronary heart disease And depression case vignettes. *Health Policy.* **89** (3), 279–287. <https://doi.org/10.1016/j.healthpol.2008.06.007> (2009).
- Balsa, A. I. & McGuire, T. G. Prejudice, clinical uncertainty and stereotyping as sources of health disparities. *J. Health Econ.* **22** (1), 89–116. [https://doi.org/10.1016/s0167-6296\(02\)00098-x](https://doi.org/10.1016/s0167-6296(02)00098-x) (2003).
- Balsa, A. I., McGuire, T. G. & Meredith, L. S. Testing for statistical discrimination in health care. *Health Serv. Res.* **40** (1), 227–252. <https://doi.org/10.1111/j.1475-6773.2005.00351.x> (2005).
- Howard, J. *Cognitive Errors and Diagnostic Mistakes* (Springer International Publishing, 2018). <https://doi.org/10.1007/978-3-319-93224-8>
- Baum, H. M. & Rothschild, B. B. The incidence and prevalence of reported multiple sclerosis. *Ann. Neurol.* **10** (5), 420–428. <https://doi.org/10.1002/ana.410100504> (1981).
- Langer-Gould, A., Brara, S. M., Beaber, B. E. & Zhang, J. L. Incidence of multiple sclerosis in multiple Racial and ethnic groups. *Neurology* **80** (19), 1734–1739. <https://doi.org/10.1212/WNL.0b013e3182918cc2> (2013).
- Okai, A. F. et al. Advancing care and outcomes for African American patients with multiple sclerosis. *Neurology* **98** (24), 1015–1020. <https://doi.org/10.1212/WNL.0000000000200791> (2022).
- Hamilton, R. & Ciccarelli, O. Multiple sclerosis incidence in black patients: it is time to do away with a Racial medical myth. *Neurology* **98** (18), 739–740. <https://doi.org/10.1212/WNL.0000000000200309> (2022).
- Hittle, M. et al. Population-Based estimates for the prevalence of multiple sclerosis in the united States by race, ethnicity, age, sex, and geographic region. *JAMA Neurol.* **80** (7), 693–701. <https://doi.org/10.1001/jamaneurol.2023.1135> (2023).
- Langer-Gould, A. M., Gonzales, E. G., Smith, J. B., Li, B. H. & Nelson, L. M. Racial and ethnic disparities in multiple sclerosis prevalence. *Neurology* **98** (18), e1818–e1827. <https://doi.org/10.1212/WNL.0000000000200151> (2022).
- McKay, K. A., Kwan, V., Duggan, T. & Tremlett, H. Risk factors associated with the onset of relapsing-remitting and primary progressive multiple sclerosis: a systematic review. *Biomed. Res. Int.* **2015**, 817238. <https://doi.org/10.1155/2015/817238> (2015).
- Miclea, A. et al. Age-dependent variation of female preponderance across different phenotypes of multiple sclerosis: A retrospective cross-sectional study. *CNS Neurosci. Ther.* **25** (4), 527–531. <https://doi.org/10.1111/cns.13083> (2019).
- Ribbons, K. A. et al. Male sex is independently associated with faster disability accumulation in Relapse-Onset MS but not in primary progressive MS. *PLoS One.* **10** (6), e0122686. <https://doi.org/10.1371/journal.pone.0122686> (2015).
- Ben-Shlomo, Y. et al. The epidemiology of Parkinson's disease. *Lancet* **403** (10423), 283–292. [https://doi.org/10.1016/S0140-6736\(23\)01419-8](https://doi.org/10.1016/S0140-6736(23)01419-8) (2024).
- GBD 2016 Parkinson's Disease Collaborators. Global, regional, and National burden of Parkinson's disease, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol.* **17** (11), 939–953. [https://doi.org/10.1016/S1474-4422\(18\)30295-3](https://doi.org/10.1016/S1474-4422(18)30295-3) (2018).
- Cerri, S., Mus, L. & Blandini, F. Parkinson's disease in women and men: what's the difference?? *J. Parkinsons Dis.* **9** (3), 501–515. <https://doi.org/10.3233/JPD-191683> (2019).
- He, J. et al. Trends in Cardiovascular Risk Factors in US Adults by Race and Ethnicity and Socioeconomic Status, 1999–2018. *JAMA.* ;326(13):1286–1298. (2021). 10.1001/jama.2021.15187 Table 1. Proportion of correct diagnoses identified among disease case vignettes with race and gender jointly randomized.

Author contributions

NB, SH, and JFS designed the survey and data collection instrument. SH and RP managed and contributed to data analysis, with RP leading the analysis, preparation of tables and figures. All authors contributed to the writing of the manuscript, which was led by RP and JFS. All authors reviewed and approved the manuscript. This work is supported by National Institutes of Health NIAID R01 AI154533. Dr. Simard is a Scientific Advisory Board member for the Metrodora Foundation.

Declarations

Competing interests

Dr. Simard participated in one Scientific Advisory Board meeting in early 2024 for the Metrodora Foundation.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-00726-8>.

Correspondence and requests for materials should be addressed to J.F.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025