

Replicating the 2015 BRFSS Study: Assessing the Impact of Race/Ethnicity on Stroke Prevalence in U.S. Adults.

Dr. Ami Soni.

Type of Article: Replicating article.

Abstract:

This replication study reexamines the 2015 Behavioral Risk Factor Surveillance System (BRFSS) data to assess the impact of race/ethnicity on stroke prevalence among U.S. adults, following the original work by Aldayel et al. (2017). By replicating the study design, including the questionnaire, data collection, and analysis methods, this research aims to validate and confirm the original findings regarding stroke disparities across racial and ethnic groups. This study utilized the same dataset (n=441,456) and statistical methods, including logistic regression and chi-square tests, to evaluate associations between race/ethnicity and stroke prevalence, adjusting for confounding variables.

The study findings generally align with the original study, confirming higher stroke prevalence among Black individuals compared to other racial groups. However, discrepancies were noted in the odds ratios for Hispanic and Multiracial groups, suggesting potential issues with data labelling and coding in the original study. Notably, variations in software used for analysis (R vs. STATA) may also contribute to these differences. Despite these discrepancies, the replication underscores the need for accurate data handling and methodological consistency in research. The study highlights ongoing racial disparities in stroke prevalence and emphasizes the importance of continued research and targeted public health strategies to address these disparities. Future research should focus on standardized data analysis tools and precise data labelling to enhance the reliability of findings across studies.

Keywords:

Stroke, race/ethnicity, BRFSS, replication study, odds ratios, health disparities.

Introduction:

This research project focuses on the article titled “The Association Between Race/Ethnicity and Prevalence of Stroke Among United States Adults in 2015: A Secondary Analysis Study Using Behavioural Risk Factor Surveillance System (BRFSS) Data.” The aim of this investigation is to clarify the impact of race and ethnicity on stroke prevalence among U.S. adults by analysing 2015 BRFSS data. The study provides significant insights into disparities in stroke prevalence across different racial and ethnic groups, thereby enhancing our understanding of health inequalities. (Yousef Aldayel et al., 2017)

The principal goal of this study is to confirm the findings, assess the findings' replicability, and ascertain whether the conclusions can be reliably validated. An acceptable standard for validating scientific claims and reproducing scientific results is established by reproducible research. However, the replication process may face difficulties due to time and resource limitations. In accordance with these specifications, computer code and datasets must be available to others to ensure they can conduct independent analyses and validate published results. (Peng, 2009) which in turn will help to enhance the credibility of research by creating resources which are easy to maintain and reuse, which benefits audiences as well as researchers. Furthermore, when significant findings are consistently repeated by numerous independent researchers using diverse data sources, analytic methods, labs, and equipment, scientific evidence is strengthened. (Peng et al., 2006).

Stroke is one of the five leading causes of mortality and disability in the USA and the second leading cause worldwide, accounting for approximately 160.4 million cases (148.0-171.7) (Ferrari et al., 2024; Guzik & Bushnell, 2017; Roth et al., 2020) it is crucial to investigate the factors influencing its prevalence. Along racial, cultural, and regional boundaries, there are still significant disparities in the epidemiology and outcomes of stroke. (Prust et al., 2024) By reassessing the association between race/ethnicity and stroke prevalence and validating the findings of the initial study through investigations into any discrepancies or changes in the results, if present, and evaluating their potential implications, this project the intent to establish the consistency and credibility of the published results. Further, this study intends to emphasise the crucial role replication plays in maintaining the accuracy and rigour of science in research.

Methods:

The methodology outlined in this article which includes the study design, questionnaire, data collection, processing, and analysis will precisely abide by the procedures followed in the original work.

Study design:

This observational study was conducted by the Behavioral Risk Factor Surveillance System (BRFSS) using both cell phones and landlines, encompassing all 50 states, the District of Columbia, and three U.S. territories. (*Centers for Disease Control and Prevention (CDC), Behavioral Risk Factor Surveillance System (BRFSS).*, n.d.)

Questionnaire:

The questionnaire for the survey was collaboratively designed by the BRFSS and the Centers for Disease Control and Prevention (CDC). All the core questions remaining unchanged throughout the survey period. (*Centers for Disease Control and Prevention (CDC), 2015. BRFSS 2015 Annual Data.*, n.d.) Initiated in 1984, the BRFSS is designed to collect state-level data on residents' health-related risk behaviors, chronic health conditions, and the utilization of preventive services. The system operates on a comprehensive scale, conducting approximately 400,000 interviews annually across these regions. (*Centers for Disease Control and Prevention (CDC), Behavioral Risk Factor Surveillance System (BRFSS).*, n.d.) This research addresses the core questions about the relationship between prevalence of stroke, race/ethnicity, age, Gender, health coverage, hypertension, education level, alcoholism (heavy alcohol consumption), cigarette smoking, diabetes, obesity, and myocardial infarction.

Data Collection and Processing:

Trained interviewers conducted BRFSS using a computer-assisted telephone interviewing system known as CATI. Participants' responses to the question, "Has a doctor, nurse, or other health professional ever told you that you have had a stroke?" were recorded as either "yes," "no," or "don't know/not sure." This response served as the dependent variable or outcome of interest. The response to the "five-level race/ethnicity" variable was used as the independent variable or exposure. For this analysis, participants were categorized into the following groups: "White only, Non-Hispanic," "Black only, Non-Hispanic," "Other race only, Non-Hispanic," "Multiracial, Non-Hispanic," "Hispanic," and "Don't know/Not sure/Refused." In addition to these two variables, other pertinent variables included: age groups (18-44, 45-54, 55-64, ≥65 years old), sex (male and female), educational level (did not graduate from high school, graduated from high school, attended college or technical school, or graduated college or technical school), health coverage (yes or no), smoking (yes or no), heavy alcohol consumption (yes or no), obesity (obese or not obese), hypertension (yes or no), diabetes (yes or no), and history of myocardial infarction (yes or no). (*Centers for Disease Control and Prevention (CDC), 2015. BRFSS 2015 Annual Data.*, n.d.)

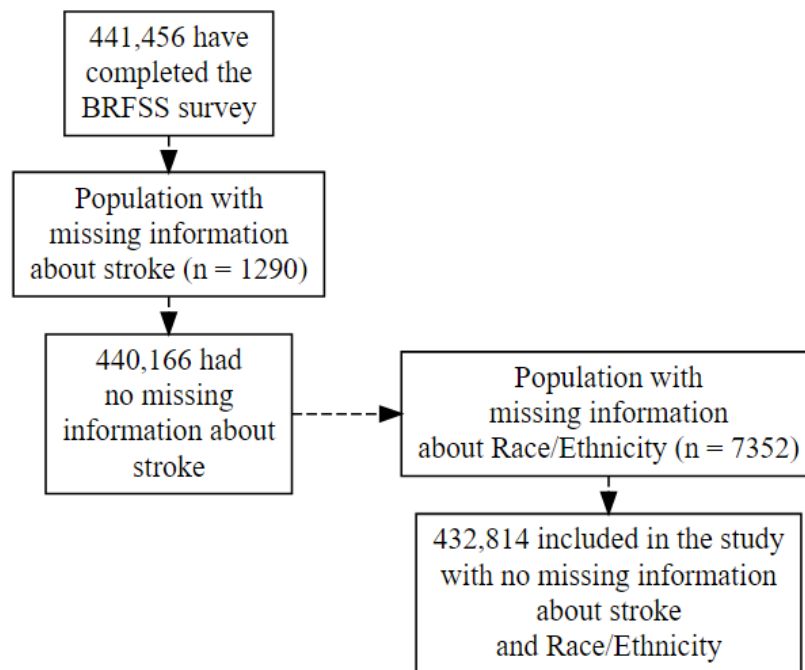


Figure 1: The above figure shows the Exclusion Criteria of the Stroke and Race Variables.

Data Analysis:

The BRFSS survey conducted interviews with approximately 441,456 U.S. adults. Inclusion criteria required complete information on both stroke and race/ethnicity; individuals with missing data on these variables were excluded. Participants with incomplete responses, refusals, or lack of knowledge regarding other relevant variables were also excluded.

A secondary analysis was performed to assess the association between race/ethnicity and stroke prevalence. Statistical significance was determined using a p-value threshold of less than 0.05, after adjusting for confounding variables. The Chi-square test was employed to examine associations between variables. Pearson's product-moment correlation was used to measure collinearity. Logistic regression analysis was conducted to calculate odds ratios for odds ratio while a multivariable logistic regression was carried out while adjusting the confounding variables. All statistical analyses were carried out using R open-source software.

Results:

A total of 441,456 individuals completed the BRFSS survey. Of these, 1,290 had missing information regarding stroke status, leaving 440,166 participants with complete stroke data. Additionally, 7,352 participants had missing information on race/ethnicity, resulting in 432,814 individuals with complete data on both stroke and race/ethnicity.

Age Distribution:

In the original study, 43% of participants were aged 44 years or younger, with a balanced gender distribution. In contrast, this analysis found that approximately 46% of participants were aged 44 years or younger, while gender distribution remained evenly balanced.

Educational Attainment:

The original paper reported that 86% of participants had at least graduated from high school, with the exception of the “Other, Non-Hispanic” group, where 37.8% had not completed high school. In this study the findings similarly showed that 86% had at least graduated from high school. Furthermore, around 37% of Hispanics did not finish high school, compared to 8% of Whites. Additionally, 15% of Whites were undergraduates, whereas the highest graduation rates were observed among the “Other race, non-Hispanic” group (45%), with Blacks having the lowest at 19%.

Healthcare Coverage & Myocardial Infarction:

Both the original study and the results in this study indicated that over 70% of each racial group had healthcare coverage.

The findings related to myocardial infarction are consistent with those reported in the original study.

Table 1: The Association between the Race/Ethnicity and some characteristics among the adult population in the U.S, BRFSS 2015

Characteristics		Race/Ethnicity					
Variables		White only,Non-Hispanic	Black only, Non-Hispanic	Other race only, Non-Hispanic	Multiracial, Non-Hispanic	Hispanic	P-value
Age							
'18-44'		75826 (39.8%)	10571 (50.8%)	8662 (59.6%)	3383 (61.4%)	17475 (62.4%)	0.00e+00
'45-54'		53106 (17.8%)	6103 (17.9%)	3382 (15.2%)	1299 (13.8%)	6535 (16.4%)	0.00e+00
'55-64'		77924 (18.5%)	7944 (16.2%)	3558 (13.7%)	1588 (13.5%)	5501 (11.6%)	0.00e+00
'>=65'		128288 (24%)	9649 (15.2%)	3993 (11.4%)	1850 (11.3%)	6177 (9.6%)	0.00e+00
Sex							
SEX_Male		142189 (48.3%)	12360 (45.4%)	9335 (49.6%)	3696 (50.4%)	15098 (48.6%)	7.18e-07
SEX_Female		192955 (51.7%)	21907 (54.6%)	10260 (50.4%)	4424 (49.6%)	20590 (51.4%)	7.18e-07
Health Coverage							
HealthPlan_Yes		316586 (92.7%)	30602 (85.8%)	17695 (90.7%)	7376 (88%)	27568 (71.7%)	0.00e+00
HealthPlan_No		17431 (7.3%)	3515 (14.2%)	1777 (9.3%)	708 (12%)	7903 (28.3%)	0.00e+00
Education							
Under High School		16995 (8.3%)	4226 (15.2%)	1542 (7.7%)	648 (10.8%)	9907 (37%)	0.00e+00
High School		92552 (28.4%)	11255 (32.1%)	4844 (19.4%)	2410 (29.1%)	10049 (26%)	0.00e+00
College Candidate		93191 (33.1%)	9473 (33.3%)	4882 (27.6%)	2681 (36.8%)	8131 (24.3%)	0.00e+00
College		131510 (30.1%)	9171 (19.4%)	8226 (45.2%)	2364 (23.2%)	7443 (12.7%)	0.00e+00
Hypertension							
Hypertension_No		198591 (66.6%)	16000 (59%)	13140 (76.7%)	5006 (70.5%)	24351 (75.3%)	2.07e-178
Hypertension_Yes		135674 (33.4%)	18192 (41%)	6387 (23.3%)	3091 (29.5%)	11221 (24.7%)	2.07e-178
Obesity							
Not_Obese		222111 (73.5%)	18320 (63.9%)	13603 (85.1%)	5101 (72.3%)	21215 (71.3%)	1.76e-166
Obese		88345 (26.5%)	12851 (36.1%)	4219 (14.9%)	2490 (27.7%)	10170 (28.7%)	1.76e-166
Alcoholism							
Alcoholism_No		300161 (93.4%)	30343 (95.7%)	17239 (96.3%)	7184 (92.6%)	31664 (95.8%)	1.76e-37
Alcoholism_Yes		18279 (6.6%)	978 (4.3%)	735 (3.7%)	513 (7.4%)	1241 (4.2%)	1.76e-37
Smoking							
Smoking_No		278183 (82.7%)	26469 (81.1%)	15173 (88%)	6038 (76.8%)	29405 (87.6%)	1.45e-71
Smoking_Yes		45447 (17.3%)	5751 (18.9%)	3207 (12%)	1787 (23.2%)	4243 (12.4%)	1.45e-71
Diabetes							
Diabetes_No		294524 (90.1%)	27217 (85.5%)	16985 (89.8%)	7020 (91.3%)	30497 (89.1%)	1.28e-32
Diabetes_Yes		40231 (9.9%)	6987 (14.5%)	2563 (10.2%)	1083 (8.7%)	5096 (10.9%)	1.28e-32
Myocardial Infarction							
MI_Yes		20065 (4.8%)	1863 (4%)	922 (2.7%)	536 (4.9%)	1356 (2.8%)	7.07e-28
MI_No		313784 (95.2%)	32179 (96%)	18531 (97.3%)	7540 (95.1%)	34158 (97.2%)	7.07e-28

Table_1: The table shows the association between Race/Ethnicity and some of the characteristics of US adults' population, BRFSS 2015.

Hypertension and Diabetes:

In the original study, Black population had a higher prevalence of hypertension (40.5%) while in this analysis the prevalence of hypertension in Blacks population is slightly higher about 41%, the lowest prevalence was found in the “Other race, non-Hispanic” group at 23.3%.

Diabetes prevalence was around 14.5% in this study results, closely aligning with the 14.3% reported in the original study. Multiracial individuals had the lowest diabetes rate at approximately 8.6%.

Heavy Drinking & Smoking:

The findings from this analysis and the original paper are generally consistent, with minor variations in smoking rates. Specifically, the proportion of individuals who smoke was 17.3% in this study, slightly lower compared to 17.5% in the original paper.

Regarding racial profile for smoking, the highest percentage was observed in the Multiracial group, at 23.2%. For drinking habits, the highest proportion was also found in the Multiracial group, at 7.3%.

Table_2: The Unadjusted and Adjusted (Odds Ratio) association between the prevalence of Stroke and some characteristics such as Race/Ethnicity, sex, and other confounding variables among the US adults population, BRFSS 2015.

Characteristics	Unadjusted_OR (95% CI)	Adjusted_OR (95% CI)
Race/Ethnicity		
White only, Non-Hispanic	Ref	Ref
Black only, Non-Hispanic	1.3(1.18-1.44)	1.29(1.15-1.43)
Other race only, Non-Hispanic	0.61(0.5-0.74)	0.96(0.78-1.19)
Multiracial, Non-Hispanic	1.25(1.03-1.52)	1.55(1.28-1.89)
Hispanic	0.56(0.49-0.65)	0.73(0.62-0.84)
Age		
18-44	Ref	Ref
45-54	3.51(3.03-4.06)	2.35(2.02-2.73)
55-64	5.96(5.26-6.75)	3.16(2.77-3.61)
>=65	10.7(9.54-12.02)	4.62(4.05-5.28)
Sex		
Sex_Male	Ref	Ref
SEX_Female	1.05(0.99-1.12)	1.16(1.09-1.25)
Health Coverage		
Health_coverage_No	Ref	Ref
Health_coverage_Yes	1.75(1.49-2.04)	1.03(0.87-1.21)
Education		
College	Ref	Ref
College Candidate	1.73(1.59-1.88)	1.4(1.29-1.53)
High School	2.14(1.98-2.32)	1.51(1.39-1.64)
Under High School	3.37(3.06-3.71)	2.14(1.91-2.39)
Hypertension		
Hypertension_No	Ref	Ref
Hypertension_Yes	5.75(5.35-6.18)	2.53(2.32-2.75)
Obesity		
Not_Obese	Ref	Ref
Obese	1.38(1.29-1.47)	0.95(0.88-1.02)
ALcoholism		
Alcoholism_No	Ref	Ref
Alcoholism_Yes	0.65(0.56-0.75)	0.78(0.67-0.92)
Smoking		
Smoking_No	Ref	Ref
Smoking_Yes	1.53(1.42-1.65)	1.65(1.52-1.8)
Diabetes		
Diabetes_No	Ref	Ref
Diabetes_Yes	3.89(3.63-4.17)	1.46(1.34-1.58)
Myocardial Infarction		
Myocardial_Infarction_No	Ref	Ref
Myocardial_Infarction_Yes	10.67(9.91-11.48)	4.52(4.16-4.91)

Table_2: The table displays the Unadjusted and Adjusted association between Stroke and some characteristics among US adult's population, BRFSS 2015.

Logistic Regression Analysis:

Logistic regression analysis carried out in this paper revealed statistically significant results for all variables.

The original paper's unadjusted odds ratios (OR) for stroke indicated 1.33 for Hispanics (95% CI: 1.10-1.61) and 1.30 for Blacks (95% CI: 1.17-1.41) compared to the reference group. In this analysis, the unadjusted OR for Blacks was approximately 1.30 (95% CI: 1.18-1.44), consistent with the original study, although minor differences in the last decimal place were noted.

For Hispanics, the unadjusted OR was around 0.56 (0.49-0.65, 95% CI). The Multiracial group had an OR of 1.25 (1.03-1.52, 95% CI), which closely resembles the Hispanic results reported in the original study.

In the adjusted analysis, the OR for Black individuals was 1.29 (1.15-1.43), aligning closely with the original study's findings. For Hispanics, the adjusted OR was 0.73 (0.62-0.84), whereas the original study reported an OR of approximately 1.57 (95% CI: 1.28-1.91). This discrepancy is similar to the findings for the Multiracial group in this study, with an OR of 1.55 (1.28-1.89). The differences in the Hispanic, Multiracial and Other race only, categories suggest potential labeling errors in the original paper. This study followed the BRFSS 2015 codebook for accurate labeling, which highlights inconsistencies in the original study's categorization.

Discussion:

The replication of the 2015 BRFSS study results demonstrates both similarities and discrepancies when compared to the original findings.

Similarities: -

The age distribution and educational attainment data were broadly consistent between the original and replicated studies, with both showing a high percentage of participants having graduated from high school and a similar proportion having healthcare coverage. -

Additionally, the prevalence rates of Hypertension and diabetes among Black population were closely aligned, reinforcing the finding of elevated health risks within this group.

Discrepancies: -

Data Labeling and Coding Issues:

Significant discrepancies were noted, one particularly concerning the labelling of the race variable. Specifically, the reference numbers for the categories "Other race, non-Hispanic", "Multiracial, Non-Hispanic", and "Hispanic" were inaccurately assigned in the original study, deviating from the specifications outlined in the BRFSS codebook.

Inaccurate labelling and coding between the original study and this replication may have led to differences in reported results. Accurate and consistent labelling is vital for replicating findings and making valid comparisons.

Furthermore, there was an error in the reported number of individuals with missing information on race/ethnicity; the correct figure is 7,352, rather than 7,452 as previously stated in the original paper. This discrepancy affects the alignment of the population count, as using the incorrect figure of 7,452 does not reconcile with the total of 432,814 participants with complete data on both stroke and race/ethnicity. Such issues underscore the importance of meticulous data handling and accurate labeling to ensure the reliability of the research findings.

In Table_1 of the original paper, a discrepancy was observed regarding the count of White, non-Hispanic high school graduates, which should be 92,552, rather than the erroneously reported 922,552.

Variations in smoking and drinking habits among different racial groups were consistent with the original study, though minor discrepancies were observed in certain areas.

The Odds Ratio for stroke among the Hispanic, Multiracial and Other race only groups notably differ in this paper compared to the original findings. This difference likely stems from mislabeling in the original study; adherence to the BRFSS 2015 codebook in this analysis corrected these inconsistencies.

The discrepancy in statistical results between this study and the original paper can be attributed to one of the reasons being the differences in the software used for analysis—specifically, the use of R in this study versus STATA in the original analysis. Different software packages may employ distinct algorithms for stratification and weighting, which can influence the precision of statistical estimates and significance. (Harris et al., 2019)

Contextual Findings: -

Previous studies, such as those by Jones et al., and G. Howard et al., support the observation of higher stroke risk among younger Black populations, consistent with the results of this study, reinforcing the need for targeted interventions. - Jacobs et al. highlighted disparities among Black women, particularly in lower socioeconomic groups, while Fang et al. noted a decrease in stroke rates over time. These studies complement the findings of this replication by providing a broader context of stroke disparities. ((Fang et al., 2014; G. Howard et al., 1994; Jacobs & Ellis, 2023; Jones et al., 2020). as demonstrated by multiple research such as REGARDS study (REasons for Geographic and Racial Differences in Stroke Study) (V. J. Howard et al., 2005).

One such study being carried out by Howard et al. (1994), where their findings indicated that the increased risk of stroke mortality is more pronounced at younger ages among the Black population in the United States than among their White counterparts, corroborating earlier findings. (G. Howard et al., 1994)

Conclusions:

This replication study successfully affirmed and extends the original research findings, demonstrating that the observed patterns—such as the high prevalence of stroke among various racial groups, particularly Black populations—persist. While the results align closely with the original study, minor discrepancies were noted.

Several limitations should be acknowledged. Variations in statistical software (R vs. STATA) and potential issues with data labeling. Different software packages can yield varying results due to differences in algorithms and weighting methods. Therefore, standardizing the weighting process and ensuring consistent methodological practices are essential for enhancing reliability.

Despite these limitations, this study provides valuable insights into the relationship between race, health behaviors, and socioeconomic factors. It highlights the importance of precise data reporting and methodological consistency in research replication. Future research should address these issues by employing uniform data analysis tools and maintaining accurate data labeling to improve the reliability and comparability of findings.

Overall, the research highlights the need for continued investigation into health disparities and supports the development of targeted public health strategies to address these issues effectively. The replication process reaffirms the robustness of the original study's findings while stressing the critical importance of methodological transparency and consistency. To further understand stroke disparities, additional studies using diverse datasets are recommended.

Data Availability:

The datasets analysed during the current study are publicly available at CDC's website, under the 2015 BRFSS optional module. (*Centers for Disease Control and Prevention (CDC), 2015. BRFSS 2015 Annual Data.*, n.d.)

Note:

The source code for this project is available at <https://github.com/Dr-A-Soni>.

References:

1. *Centers for Disease Control and Prevention (CDC), 2015. BRFSS 2015 Annual Data.* (n.d.). Retrieved September 5, 2024, from https://www.cdc.gov/brfss/annual_data/annual_2015.html
2. *Centers for Disease Control and Prevention (CDC), Behavioral Risk Factor Surveillance System (BRFSS).* (n.d.). Retrieved September 5, 2024, from <https://www.cdc.gov/brfss/>
3. Fang, M. C., Coca Perrailon, M., Ghosh, K., Cutler, D. M., & Rosen, A. B. (2014). Trends in Stroke Rates, Risk, and Outcomes in the United States, 1988 to 2008. *The American Journal of Medicine*, 127(7), 608–615. <https://doi.org/10.1016/j.amjmed.2014.03.017>

4. Ferrari, A. J., Santomauro, D. F., Aali, A., Abate, Y. H., Abbafati, C., Abbastabar, H., Abd ElHafeez, S., Abdelmasseh, M., Abd-Elsalam, S., Abdollahi, A., Abdullahi, A., Abegaz, K. H., Abeldaño Zuñiga, R. A., Aboagye, R. G., Abolhassani, H., Abreu, L. G., Abualruz, H., Abu-Gharbieh, E., Abu-Rmeileh, N. M., ... Murray, C. J. L. (2024). Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, 403(10440), 2133–2161. [https://doi.org/10.1016/S0140-6736\(24\)00757-8](https://doi.org/10.1016/S0140-6736(24)00757-8)
5. Guzik, A., & Bushnell, C. (2017). Stroke Epidemiology and Risk Factor Management. *CONTINUUM: Lifelong Learning in Neurology*, 23(1), 15–39. <https://doi.org/10.1212/CON.0000000000000416>
6. Harris, J. K., Wondmeneh, S. B., Zhao, Y., & Leider, J. P. (2019). Examining the Reproducibility of 6 Published Studies in Public Health Services and Systems Research. *Journal of Public Health Management and Practice*, 25(2), 128–136. <https://doi.org/10.1097/PHH.0000000000000694>
7. Howard, G., Anderson, R., Sorlie, P., Andrews, V., Backlund, E., & Burke, G. L. (1994). Ethnic differences in stroke mortality between non-Hispanic whites, Hispanic whites, and blacks. The National Longitudinal Mortality Study. *Stroke*, 25(11), 2120–2125. <https://doi.org/10.1161/01.STR.25.11.2120>
8. Howard, V. J., Cushman, M., Pulley, L., Gomez, C. R., Go, R. C., Prineas, R. J., Graham, A., Moy, C. S., & Howard, G. (2005). The Reasons for Geographic and Racial Differences in Stroke Study: Objectives and Design. *Neuroepidemiology*, 25(3), 135–143. <https://doi.org/10.1159/000086678>
9. Jacobs, M. M., & Ellis, C. (2023). Stroke in women between 2006 and 2018: Demographic, socioeconomic, and age disparities. *Women's Health*, 19. <https://doi.org/10.1177/17455057231199061>
10. Jones, E. M., Okpala, M., Zhang, X., Parsha, K., Keser, Z., Kim, C. Y., Wang, A., Okpala, N., Jagolino, A., Savitz, S. I., & Sharrief, A. Z. (2020). Racial disparities in post-stroke functional outcomes in young patients with ischemic stroke. *Journal of Stroke and Cerebrovascular Diseases*, 29(8), 104987. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.104987>
11. Peng, R. D. (2009). Reproducible research and Biostatistics. *Biostatistics*, 10(3), 405–408. <https://doi.org/10.1093/biostatistics/kxp014>
12. Peng, R. D., Dominici, F., & Zeger, S. L. (2006). Reproducible Epidemiologic Research. *American Journal of Epidemiology*, 163(9), 783–789. <https://doi.org/10.1093/aje/kwj093>

13. Prust, M. L., Forman, R., & Ovbiagele, B. (2024). Addressing disparities in the global epidemiology of stroke. *Nature Reviews Neurology*, 20(4), 207–221.
<https://doi.org/10.1038/s41582-023-00921-z>
14. Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T. J., Carapetis, J., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., ... Fuster, V. (2020). Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019. *Journal of the American College of Cardiology*, 76(25), 2982–3021. <https://doi.org/10.1016/j.jacc.2020.11.010>
15. Yousef Aldayel, A., Mousa Alharbi, M., Mustafa Shadid, A., & Carlos Zevallos, J. (2017). The association between race/ethnicity and the prevalence of stroke among United States adults in 2015: a secondary analysis study using Behavioural Risk Factor Surveillance System (BRFSS). *Electronic Physician*, 9(12), 5871–5876.
<https://doi.org/10.19082/5871>