HHS Public Access

Author manuscript

Int J Stroke. Author manuscript; available in PMC 2020 December 29.

Published in final edited form as:

Int J Stroke. 2018 July ; 13(5): 481–495. doi:10.1177/1747493017729269.

Viral and parasitic pathogen burden and the association with stroke in a population-based cohort

Brad D Pearce[1], **Anna Bracher**[1], **Jeffrey L Jones**[2], **Deanna Kruszon-Moran**[3]

[1]Department of Epidemiology, Emory University, Atlanta, GA, USA

[2]Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA

[3]Division of Health and Nutrition Examination Surveys, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD, USA

Abstract

Background: Higher cumulative burden of viral and bacterial pathogens may increase the risk of stroke, but the contribution of parasitic infections in relation to cumulative pathogen burden and risk of stroke has rarely been examined.

Aim: To estimate the association of multiple persistent viral and parasitic infections with stroke in a representative sample of adults in the United States.

Methods: Serological evidence of prior infection was categorized as positive for 0–1, 2, 3, or 4–5 infections based on immunoglobulin G seropositivity to cytomegalovirus, hepatitis A virus, hepatitis B virus, *Toxoplasma gondii*, and *Toxocara* spp. in 13,904 respondents from the National Health and Nutrition Examination Survey III. Regression analysis was used to estimate the cross-sectional association between serological evidence of prior infection and history of stroke adjusting for demographic risk factors, and potential mediators of stroke.

Results: Age-adjusted models that included serological evidence of prior infection to cytomegalovirus, hepatitis A virus, hepatitis B virus, *Toxoplasma gondii*, and *Toxocara* spp. showed that adults in the highest serological evidence of prior infection category (4–5 infections) had a higher prevalence of stroke (5.50%, 95% confidence interval 2.44–10.46%) than those in the lowest serological evidence of prior infection categories (1.49%, 95% confidence interval 1.01–2.11%), and a trend test suggested a graded association between serological evidence of prior infection and stroke ($p = 0.02$). In multivariable logistic regression models, the positive association of serological evidence of prior infection with stroke prevalence remained significant after adjustment for other significant risk factors (odds ratio = 1.4, $p = 0.01$) but was only significant

Corresponding author: Brad D Pearce, Department of Epidemiology, Rollins School of Public Health, Emory University, 1518-002-3BB, Atlanta, GA 30322, USA. bpearce@emory.edu.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CDC Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention.

among those aged 20–59 (odds ratio = 2.0, $p = 0.005$) and not among those aged 60–69 ($p = 0.78$) or 70 and older ($p = 0.43$).

Conclusion: We found support for a connection between serological evidence of prior infection to cytomegalovirus, hepatitis A virus, hepatitis B virus, *Toxoplasma gondii*, and *Toxocara* spp. and stroke among those aged 20–59. There may be a need to consider common parasitic infections in addition to viral and bacterial pathogens when calculating serological evidence of prior infection in relation to cerebrovascular disease.

Keywords

Epidemiology; infection; *Toxoplasma gondii*; *Toxocara*; hepatitis virus; cytomegalovirus

Introduction

There is considerable interest in the role of chronic infections and inflammation in the etiopathogenesis of stroke. The causal chains leading to stroke are multifactorial, which may explain the difficulty in corroborating any single infectious agent as a stroke risk factor.[1,2] Conversely, exposure to multiple pathogens over the life course may provide a more meaningful indicator of an infectious etiology for atherothrombotic stroke.[3,4] Aggregate pathogen exposure in atherosclerosis has been examined in numerous studies, and analogous mechanisms for cerebrovascular ischemic pathophysiology have been considered in recent studies.[2] Studies of pathogen burden (PB) and stroke in humans have mainly been focused on correlations between serological evidence of viral and bacterial infection and stroke.[3,5,6] While it has been reported that PB was associated with increased risk of stroke,[3,4,6] this association was null in the Framingham Heart Study.[7]

The reason for the discrepancies between these findings is unclear. Luna and Elkind argued that the homogeneity of the Framingham sample (e.g. in age and socioeconomic class) may not have captured the wide spectrum of predisposing conditions that participate in the connection between infections and stroke.[2] Indeed, conventional risk factors for stroke (e.g. family history, race/ethnicity, hypertension, and diet) may contribute to a relationship between PB and stroke by influencing the exposure or susceptibility to these infections, or by amplifying downstream inflammatory mediators leading to endothelial dysfunction and thrombosis.[2,4,5] Mechanisms of stroke may differ by age, and infection-related mechanisms may not have been captured in the Framingham study with a mean age at the baseline examination of 69 years.

Persistent infections, which are not completely cleared by the immune system, are particularly strong candidates as mediators of chronic inflammation in pathophysiology of stroke.[2] However, few studies have considered the role of parasites in PB studies and the effect on stroke outcome, even though infections with parasites such as *Toxoplasma gondii* and *Toxocara* spp. are common in the United States,[8,9] and these organisms are mediators of inflammation.[10,11] Notably, Sealy-Jefferson et al. included *T. gondii* in their study of PB and stroke, but the association between *T. gondii* seropositivity and incident stroke was not statistically significant.[12] However, since their study was focused on Mexican Americans living in the Sacramento Valley of California, there is still a need for a nationally

representative of the United States. Moreover, given that conventional risk factors for cerebrovascular disease may interact with PB to impact the risk of stroke, a large sample is desirable because of the number of potential confounder and mediator variables that must be considered.[2,13] In the current study, we tested the association between prevalence of stroke and serological evidence of prior infection (SEPI) from a variety of viral and parasitic infections using data from 13,904 adults in the third National Health and Nutrition Examination Survey (NHANES III).

Methods

Study participants

To examine the relationship between burden of infection with prevalence of stroke, we analyzed data from the third National Health and Nutrition Examination Survey conducted from 1988 to 1994 by the National Center for Health Statistics, Centers for Disease Control and Prevention. The NHANES was based on a stratified, multistage, probability cluster design to obtain nationally representative statistics on the civilian, noninstitutionalized US population for health measures and conditions through household interviews, standardized physical examinations, and collection of biological specimens in mobile examination centers.[14] Detailed descriptions of the design of the survey and the sample have been described elsewhere.[14]

Stroke outcome and measures of infection

Outcome data on previously experienced strokes were collected by self-report ("Has a doctor ever told you that you had a stroke?") during the household interview. Ischemic versus hemorrhagic stroke was not delineated. Based on prior literature and the availability of serological data in NHANES III, the following 11 organisms were considered in this study: *T. gondii*, *Toxocara* spp., *Helicobacter pylori*, varicella, cytomegalovirus, herpes simplex virus 1 and 2 (HSV1 and HSV 2), herpes virus type 8, and hepatitis A, B and C viruses (HAV, HBV, and HCV). Two infections, HSV1 and HSV 2, were measured on those aged 6+ in NHANES III phase I but only among those aged 6–49 in phase II and therefore were not included because analysis of only phase I NHANES III would decrease our sample size by approximately 50%. Similarly, H. pylori was only measured in phase I and was not included. There were too few HCV-positive (2.6% of the sample) and too few varicella-negative (1.4% of the sample) cases to include those variables. Serologic data on the remaining six infectious organisms measured in NHANES III on adult participants ages 20 and older for all six years of data collection were examined. Herpes virus type 8 did not have a univariate association with stroke outcome and was also excluded from the index. The remaining five organisms showing a univariate association with stroke outcome were ultimately included in the (SEPI) index. They included: *T. gondii*, *Toxocara* spp., cytomegalovirus and HAV and HBV. Details on serologic testing for antibody for past or current infection to the five organisms included in our study (*T. gondii*, *Toxocara* spp., cytomegalovirus, and hepatitis A and B) have been previously described.[9,15–18]

A measure of burden of infection was created for each sample person based on the number of infections for which they tested positive. This index was termed "Serological Evidence of

Prior Infection" (SEPI). Because very few sample persons were positive for either zero or for all five infectious agents, the SEPI variable was coded with values of 1 through 4 designating whether the sample person was positive for 0–1, 2, 3, or 4–5 infectious agents.

Selection of covariates

Sociodemographic factors related to stroke that we examined in our analysis included age in years, sex, race and Hispanic origin, metropolitan residence, country of birth, poverty–index ratio, and education level. Health status and risk behavior variables associated with stroke outcome also examined were smoking, alcohol use, obesity, history of diabetes, history of hypertension or high blood pressure and hypertension medication use, high total cholesterol, low levels of high-density lipoprotein (HDL), C-reactive protein (CRP) level, bilirubin levels, and high triglyceride level. Details on how each cofactor was coded can be found in the Appendix 1. Details on all wording of all questions and tests for laboratory results can be found in the survey documentation.[19]

Response rates

There were 18,825 individuals (80.9% of those sampled) ages 20 and older interviewed in NHANES III. Of those interviewed, 91% (17,030) were examined, 17,025 (99.97% of those examined) had self-reported data on stroke and 13,904 (81.6% of those examined) had data on all five infectious agents as well as stroke history. The percent with complete data on all five infectious agents as well as self-reported stroke varied by >5% for only one variable, age. Complete serologic data were available on 74% of those aged 70 years and older compared with 83% among those aged 20–49 and 86% among those aged 60–69.

Statistical analysis

Estimates of prevalence were weighted to represent the total resident US population and to account for oversampling and nonresponse to the household interview and physical examination. [20,21] Examination weights were used since serologic data were collected during the Mobile Examination Center or home examination. Analyses involving triglyceride level utilized the fasting subsample weights. Statistical analyses were conducted with SUDAAN, statistical software for analysis of complex survey data (version 10.0.1) to account for the complex survey design (Research Triangle Institute, Research Triangle Park, North Carolina). Standard error estimates were calculated using the Taylor series linearization method. Ninety-five percent confidence limits for prevalence estimates were estimated using the exact binomial method.[22] Estimates with relative standard errors >40 or with fewer than 10 persons reporting a history of stroke were not reported, because they are considered highly unstable.

Univariable analyses examining the association of each cofactor with stroke, each individual infection with stroke, and SEPI index with stroke were conducted using a t statistic from a linear orthogonal contrast procedure. A linear regression model, adjusted for age-group, was used to evaluate the association of each cofactor with the summary variable measuring SEPI. All analyses were conducted in SUDAAN with a p value <0.05 considered significant. No corrections for multiple comparisons were made, but exact p values are provided.

To examine whether SEPI was associated with stroke after adjustment for all relevant predictors of stroke, we used logistic regression modeling procedures. Individual logistic regression models that included age, the variable for SEPI as a continuous term, and each individual cofactor were examined. In addition, the interaction between each cofactor and SEPI was evaluated. To further determine which cofactors were significant predictors of stroke when considered simultaneously, we built more complex models with various levels of adjustment. At each step, variables that were not significantly associated with stroke outcome were deleted from the model in a backward stepwise fashion before proceeding to the next step. Steps included adjusting for—(1) all demographic, socioeconomic, and risk behavior variables (age, sex, race and Hispanic origin, foreign birth, metropolitan residence, poverty index, education, smoking, and alcohol use); (2) all variables significant in the previous step as well as diabetes diagnosis and obesity; (3) all significant variables from the first two steps but now including cardiovascular disease indicators (high total cholesterol, low HDL, and hypertension/high blood pressure medication use or diagnosis); and (4) all previous significant variables plus those variables potentially on the pathway between infectious disease burden and stroke (bilirubin and high CRP). At each step, we deleted those not significant at $p > 0.05$ in a backwards stepwise process, reported the significance of our measure of SEPI, and evaluated any interactions with SEPI among each of the remaining predictors. Significant interactions were further evaluated in stratified models.

To further validate our findings, we conducted an additional simplified backward stepwise analysis considering all cofactors simultaneously and evaluating interactions at each step. An interaction between our measure of SEPI and age remained for all models, so we ran the final model from the simplified procedure on each age strata. The simplified model was also run on the fasting half sample to evaluate the effects of high triglyceride level on the association of SEPI and stroke and again repeated for each age strata. In all multivariable analyses, variables with a $p < 0.05$ were considered significant, and no adjustments for multiple comparisons were made.

Results

Univariable analysis

Results from the univariable age-adjusted analysis examining prevalence of stroke by positivity status for each individual infection and by level of SEPI are shown in Table 1. All the age-adjusted analyses were adjusted for a three-level age-group variable. The number of strokes for each age-group was 73 in the 20–59 years of age-group, 94 in the 60–69 years of age-group, and 226 in the 70+ years of age-group. For each of the five infections, prevalence of stroke was between 40 and 108% higher among those seropositive for the infection compared to those seronegative ($p < 0.05$) (Table 1). The prevalence of stroke also increased from 1.49% (95% confidence interval (CI) 1.01–2.11) among those seropositive to 0–1 infections to 5.5% (95% CI 2.44–10.46) among those seropositive to 4–5 infections. Results from the univariable age-adjusted analysis examining prevalence of stroke by individual level of each cofactor are also shown in Table 1. In addition, the relationships between each cofactor and SEPI from a linear model are shown in Table 2. SEPI increased with older age (Table 2). The Pearson correlation coefficient for SEPI and age was 0.30 (95 % CI 0.28–

0.33). Cofactors that had a significant association with both SEPI and stroke were age, poverty index, education, alcohol use, diabetes, high total cholesterol, low HDL, high triglycerides, hypertension medication and/or diagnosis, high CRP levels, and bilirubin levels. Cofactors that were only associated with SEPI were race/ethnicity, foreign birth, and obesity. Smoking was associated only with stroke. Sex and metropolitan residence did not have a significant relationship with either SEPI or stroke.

Individual logistic regression models

The relationship between age with SEPI on stroke outcome as well as the association with stroke for each individual cofactor after adjusting for age and SEPI was further examined in individual logistic regression models. Besides age, other cofactors that showed significant associations with stroke after adjustment for age and SEPI included living below poverty, lower education, any history of smoking, past alcohol use and no alcohol use, diagnosis of diabetes, high total cholesterol, low HDL, high triglycerides, high CRP levels, low bilirubin levels, and history of hypertension/high blood pressure with or without medication compared to those without any history of hypertension or high blood pressure. SEPI remained significantly associated with stroke in all of these individual models. These results are presented in Table 3. Possible interactions between the levels of each cofactor and the effect of SEPI on the prevalence of stroke were further examined in individual logistic regression models. The *p* values for the interaction term for each model are presented in the last column of Table 3. From these individual models, three cofactors showed possible evidence of modifying the effect of SEPI on stroke prevalence: age, alcohol use, and hypertension/high blood pressure diagnosis with medication use. However, after adjusting for the interaction between age and SEPI, no other interactions with SEPI remained statistically significant. The odds of stroke for each level of increasing SEPI was highest among those in the youngest age-group and only reached statistical significance in that age-group (age 20–59 years, odds ratio (OR) = 2.3 (95% CI 1.6–3.2), *p* value <0.01; age 60–69 years, OR = 1.1 (95% CI 0.7–1.7), *p* value = 0.67; age 70 years and above, OR = 1.2 (95% CI 0.9–1.6), *p* value = 0.12) (contrasts for age 20–59 versus age 60– 69, *p* value = 0.03; age 20–59 versus age 70 years and above, *p* value < 0.01).

Complex models

In the first step of complex modeling, all demographic, socioeconomic, and risk behavior variables were included in the model with SEPI. These variables included age, sex, race and Hispanic origin, foreign birth, metropolitan residence, poverty index, education, smoking, and alcohol use. Variables no longer significant (*p* value >0.05) were eliminated from the model. The remaining significant cofactors included age, race and Hispanic origin, education, smoking status, and alcohol use along with SEPI. Odds of stroke increased by 30% (OR = 1.3, 95% CI 1.0–1.6, *p* value = 0.03) with increasing level of SEPI after adjusting for these initial cofactors (Table 4).

In the second step of complex modeling, the metabolic conditions of diabetes and obesity were considered by adding them to the reduced model produced by step 1. Obesity was not significantly associated with stroke and was therefore eliminated from the model. The effect

of SEPI on stroke remained marginally significant (OR = 1.3, 95% CI 1.0–1.6, *p* value = 0.05).

In the third step of complex modeling, cardiovascular disease indicators, (not including high triglycerides), were considered. In this step, we added variables for high total cholesterol, low HDL, and hypertension/high blood pressure diagnosis and medication use to the reduced model produced by step 2. High total cholesterol, alcohol use, and race and Hispanic origin were no longer significantly associated with stroke outcome and were eliminated from the model. SEPI continued to be significantly associated with stroke (OR = 1.4, 95% CI 1.1–1.7, *p* value = 0.01).

In the fourth step of complex modeling, variables that measure metabolic activity that may be in the pathway between SEPI and stroke outcome were considered. High CRP and bilirubin were added to the reduced model produced by step 3. In this final model, older age, lower education, having ever smoked, diagnosis of diabetes, low HDL, hypertension/high blood pressure diagnosis with medication use, high CRP levels, and low bilirubin levels were significantly associated with stroke. SEPI remained significantly associated with stroke (OR = 1.4, 95% CI 1.1– 1.7, *p* value = 0.01). Results from this complex stepwise modeling are summarized in Table 4.

After step 4, the interaction between SEPI and age remained statistically significant when comparing those aged 20–59 (OR for SEPI 2.0 (95% CI 1.3–3.1)) with those ages 60–69 years (OR for SEPI 1.1 (95% CI 0.7– 1.6), *p* value for interaction = 0.05) and comparing those aged 20–59 with those aged 70+ years (OR for SEPI 1.1 (95% CI 0.8–1.5), *p* value for interaction 0.01). As with the individual regression models, no other interactions with SEPI were statistically significant once the interaction with age and SEPI was included in the model.

Simplified backwards stepwise procedure

To further validate our findings from the four-step modeling above, we conducted a more simplified backwards stepwise analysis allowing all cofactors including SEPI to be considered simultaneously in the initial model except for high triglycerides because it was measured on a fasting half sample. Using a simplified backwards stepwise approach, all cofactors whose *p* value >0.05 were removed. In the final model, the prevalence of stroke was associated with older age, male sex, lower educational level, ever smoked, diabetes diagnosis, higher total cholesterol, lower levels of HDL, higher CRP, lower bilirubin, and history of hypertension/ high blood pressure with medication compared to those without any history of hypertension or high blood pressure. The ORs and p values for each cofactor from the final simplified model are given in Table 5. As before, SEPI remained significantly associated with stroke with no difference in effect as compared to the final model from the more complex modeling procedure (OR = 1.4 (95% CI 1.1–1.7), *p* value = 0.01).

Because the interaction with age remained significant, the final model from the simplified backwards stepwise procedure was run separately for each age strata. Results are presented in Table 5. Similar to the previous findings, we saw that SEPI was significantly associated

with stroke outcome among those aged 20–59 (p value = 0.01), but not among those aged 60–69 (p value = 0.78) or those aged 70 and above (p value = 0.43).

Half sample model

To evaluate both the main effect of high triglyceride levels (≥150 vs. <150mg/dL) on stroke as well as the possible confounding effect of high triglyceride on SEPI, we utilized the fasting half sample and included a binary variable for high versus normal triglyceride level to the final model from the simplified backward stepwise procedure described above. This model was run on the combined ages (20 and over) as well as stratified by age-group (20–59, 60–69, and 70+ years). High triglycerides were significantly associated with stroke in the combined model and the model for those aged 20–59 years. As in the previous models, SEPI was significantly associated with stroke in both the combined model (OR = 1.8 (95% CI 1.3–2.7), p value <0.01) and the 20–59-year-old age strata (OR = 4.3 (95% CI 1.7–10.4), $p < 0.01$) but not among those aged 60–69 or 70 and older. As a result of the reduced power due to the reduced sample size, many of the associations with the other possible risk factors of stroke were no longer statistically significant (model details not shown).

Discussion

We found a strong cross-sectional association between our measure of SEPI to cytomegalovirus, HAV, HBV, *T. gondii*, and *Toxocara* spp. and stroke in a large representative sample of the US population. In contrast to most previous studies of infectious disease burden and stroke,[2] we included parasitic infections (*T. gondii*, *Toxocara* species) and performed extensive statistical adjustment for covariates that could explain the relationship between infection burden and stroke.

Among respondents who were seropositive (showed evidence of past infection) for 4–5 of the infections tested, the age-adjusted prevalence of stroke was 5.50%, compared to 1.49% among respondents who were seropositive for only 0–1 of these infections. There was a graded dose–response linear trend for SEPI on the prevalence of stroke. Moreover, the association between SEPI and stroke remained statistically significant after adjustment for demographic factors associated with stroke including race/ethnicity, smoking, and education. It should be noted that using prevalent, nonfatal strokes would bias our analysis to inclusion of more minor strokes. Therefore, our results are more applicable to minor strokes.

We also considered medical conditions that increase risk of stroke.[13] Diabetes is a well-established risk factor for stroke,[13] and our univariable analysis confirmed this association and also showed a positive correlation between diabetes prevalence and SEPI. This finding is congruent with a prior study in a different cohort consisting exclusively of middle-aged men that found a positive correlation between insulin resistance and infectious disease burden.[23] This prior study did not include parasitic infections, and the association of infectious disease burden with insulin resistance was mainly driven by enteroviruses and *C. pneumonia*, though herpes simplex also likely contributed.[23] The authors suggested that simultaneous persistent infections may drive inflammation, which in turn increases the risk for diabetes. However, it is also possible that diabetes is increasing the susceptibility or consequences of these infections.[23] In our study, the increasing prevalence of stroke outcome with increasing levels

of SEPI was not substantially affected by adding diabetes prevalence to our multivariable models, which suggests that SEPI is not acting through diabetes-related mechanisms to increase stroke prevalence.

A number of interrelated pathogenic mechanisms have been considered for the putative role of infections in stroke.[2] Specifically, infections may exacerbate atherosclerosis, induce vasculitis, or cause plaque instability. Various cellular and molecular processes have been tentatively implicated, including low-grade inflammation, pathogen–host epitope cross reactivity, aberrant adhesion molecule expression, and abnormal lipid metabolism.[2] However, these mechanisms are mainly relevant to persistent infections, and some of the infections we examined are not typically persistent (e.g. hepatitis A), and hence SEPI is not an indicator of simultaneous persistent infections but rather of the history of several infections, only some of which may have remained persistent. Specifically, HAV is not considered a persistent infection, though it produces a long-lived immunoglobulin G response.[24] HBV infection develops into a long-term chronic infection in less than 5% of immunocompetent adults, but the rate of chronic infection is much higher in neonates and young children.[25]

More broadly, we considered the relationship of SEPI and other components of metabolic syndrome (hypertension or dyslipidemia) which have both been tentatively linked to infectious disease burden (reviewed in Ebrahimi et al.[26]) and are established risk factors for stroke.[13] However, the addition of these exposure variables to our multivariable models had little effect on the statistically significant relationship between SEPI and stroke.

A prominent theory for the mechanistic link between infectious disease burden and ischemic stroke is that multiple infections lead to chronic inflammation—either systemically or locally within arteriosclerotic plaques—which increase the risk for an ischemic stroke.[2] While we cannot use our cross-sectional study to directly address this mechanism, the inclusion of CRP (an inflammation marker) in our models did not attenuate the increased odds of stroke among respondents ith high SEPI. Therefore, we found no evidence that CRP is in the path between SEPI and stroke. However, the CRP assay used for NHANES III samples had low sensitivity compared to the high-sensitivity CRP assays used in current clinical practice. We applied the cut points for CRP from the older NHANES III literature, and the imprecision of CRP values from NHANES III may have limited our ability to detect a role for CRP in the pathway.

We also found no evidence that bilirubin was a mediator or effect modifier of the relationship between SEPI and stroke. We did, however, replicate results from a prior study showing the protective effects of higher bilirubin on stroke.[27]

The relationship between SEPI and stroke was most prominent among the younger age strata (20–59 years) of respondents. Many of the biological mechanisms for stroke in older patients (e.g. atherosclerosis) are applicable to younger individuals as well, but there are also some pathophysiological pathways that may be more common among young adult stroke patients.[28] On the other hand, it is possible that the association between SEPI and stroke in older age-groups could have been attenuated by a higher presence of other major risk

factors. We did not have specific data on some known risk factors for stroke such as atrial fibrillation. Regardless, our data suggest that pathways involving chronic infections may be particularly important in these young stroke patients.

Limitations

In addition to some of the limitations described above (our association is only with nonfatal strokes, the nonpersistent nature of hepatitis A and some hepatitis B infections, and the low sensitivity of the NHANES III CRP assay), our conclusions should be viewed cautiously because of the cross-sectional nature of our study and the lack of data on particular etiological subtypes of stroke. Stroke was self-reported and not verified by an independent clinical evaluation, which is a limitation of this study and prior studies of stroke in the NHANES III dataset. We studied only a few parasitic infections and did not have serological data on some infections known to be linked to stroke (e.g. Trypanosoma cruzi[29]). Because we only considered infections with serology data already present in the NHANES dataset, our study was constrained to only a subset of infections with data collected for other purposes. Our study is also limited by the use of serological indices of infection, and hence we cannot know when the infection occurred. We cannot know, for example, if exposure to one or more of the infectious agents occurred before or after the reported stroke. Likewise, the temporal relationship between other stroke risk factors in our models (e.g. hypertension) and the stroke outcome cannot be discerned in this cross-sectional sample. Moreover, our approach of considering each infection as equivalent to each other infection may be overly simplistic and does not capture the possibility that some infections (e.g. hepatitis B) may contribute more to the overall burden than other infections.

NHANES does not include individuals in some sociodemographic categories with potentially higher risk for severe stroke or infection (i.e. those institutionalized, incarcerated, those in nonhospital based long-term care facilities, injection drug using, homeless populations, etc.). Given that individuals in these high-risk groups for stroke and infection would not have been included in our analysis, our estimates of the association between infection and stroke outcome may be conservative. On the other hand, those who were infected and presented to a physician for care may have been more likely to have been diagnosed with stroke than would have otherwise have gone undiagnosed. This and other residual confounders may have inflated our estimate of SEPI's effect on stroke.

There were only 73 cases of stroke in the youngest age-group with only 9 cases among those with the highest level of SEPI; therefore, the results for this age-group are based on smaller numbers than the other age strata. Interestingly, we found the protective effect of bilirubin was most evident in the younger age strata.

There is a bidirectional relationship between host immune responses and stroke. As discussed above, inflammatory responses to infection may trigger stroke, but stroke itself has a complex effect on immune responses, including B cells and antibody production.[30] Thus, in our cohort, a prior stroke could influence the antibody titer to some of infections we tested.

Conclusion

The connection between exposure to multiple pathogens and stroke risk remains controversial.[2] Most prior studies of the possible link between infectious disease burden and stroke have been case-control analyses or have focused on specific high-risk populations or residents of specific cities or geographic regions. We examined the relationship between stroke and SEPI to cytomegalovirus, HAV, HBV, *T. gondii*, and *Toxocara* spp. in a nationally representative population with extensive control for health-related covariates. In support of prior studies connecting infectious disease burden with stroke,[2] we found a significant association between stroke and SEPI to cytomegalovirus, HAV, HBV, *T. gondii*, and *Toxocara* spp. overall and specifically among those aged 20–59 upon age stratification. Notably, this association was independent of various demographic and health status covariates. Our inclusion of chronic parasitic infections, specifically *T. gondii* and *Toxocara* spp., in the SEPI index suggests the need for other studies to consider parasites as contributors to the multifaceted pathophysiological mechanisms underlying stroke. Our study differs from most prior studies in that it has broad regional representation in the United States and included the aforementioned parasitic pathogens in addition to viruses. Luna and Elkind have suggested that geographical variables may be important in the putative connection between infection burden and stroke risk because different pathogen combinations may circulate in a population based on spatiotemporal factors.

Future studies of infection burden and risk of stroke should better define the dynamics of pathogen exposure and stroke using longitudinal designs and consideration of geographical factors as well as genetic vulnerabilities, life course exposures, and immunosenescence. Such studies are needed to determine whether there is any causal connection between infection burden and stroke, which would add to the broad interest in preventing infections.

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: National Institute of Mental Health (Award Number: 1R01MH092512-01).

Appendix 1

Coding of sociodemographic factors

- *Age*: Measured in years and grouped as those aged 20–59, 60–69, and 70+ years.
- *Race and Hispanic origin*: Based on the respondents' self-reported information and categorized as non-Hispanic white, non-Hispanic black and Mexican American; respondents who did not self-identify into these three groups were classified as "other," which also included all non-Mexican American Hispanics and individuals reporting multiple races.
- *Metropolitan residence*: coded as metropolitan or non-metropolitan.

- *Country of birth*: Categorized as being born in the 50 U.S. states or D. C. versus not being born in the U.S. states and D.C.
- *Poverty–index ratio*: Calculated by dividing family income by a poverty threshold specific for family size produced by the US census and categorized as below (<1.0) or at or above the poverty line (1.0).
- *Education*: Measured as last year of school completed using individual education and grouped as less than a high school graduate and high school completed or beyond.

Health status and risk behavior variables

- *Smoking*: Coded as ever a smoker versus never.
- *Alcohol use*: Coded as current user, past user, and never used.
- *Obesity*: Body mass index calculated from measured height and weight and coded as ≥30 kg/meters^2 versus <30 kg/meters^2
- *History of diabetes*: Coded as received a previous diagnosis of diabetes versus no diagnosis
- *Hypertension/hypertension medication*: Three-level variable that included history of hypertension or high blood pressure and hypertension medication use coded as (1) nonhypertensives—those who had no history of diagnosed hypertension/high pressure and whose blood pressure was normal at exam; (2) unmedicated hypertensives—those with a history of diagnosed hypertension or high blood pressure or who measured ≥140 systolic or ≥90 diastolic blood pressure at the NHANES exam but who are NOT currently taking hypertension medication; or (3) medicated hypertensives—those with history of diagnosed hypertension or high blood pressure and are currently taking medication for it.
- *Total cholesterol*: Coded as high levels of total cholesterol (≥240 mg/dL) versus normal (<240 mg/dL).
- *HDL*: Coded as low levels of HDL versus normal (<40 vs. ≥40 mg/dL for males and <50 vs. ≥50 mg/ dL for females).
- *CRP*: C-reactive protein (CRP) coded as low (≤0.21), middle (0.22–0.99), or high (≥1.0).
- *Bilirubin levels*: Based on tertiles and coded as low (<0.5 mg/dL), middle (0.5–0.6 mg/dL), and high (≥0.7 mg/dL).
- *Triglyceride levels*: Coded as high (≥150 mg/dL) versus normal (<150 mg/dL). Triglyceride level was measured on only a random half sample of individuals who fasted prior to blood draw.

Details on all questions and laboratory results can be found in the survey documentation.[19]

References

1. Yu M, Zhang Y, Yang Z, Ding J, Xie C and Lu N. Association between Helicobacter pylori infection and stroke: a meta-analysis of prospective observational studies. J Stroke Cerebrovasc Dis 2014; 23: 2233–2239. [PubMed: 25263434]
2. Luna JM and Elkind MS. Infectious burden and risk of stroke In: Chen J, Hu X, Stenzel-Poore M and Zhang JH (eds) Immunological mechanisms and therapies in brain injuries and stroke. Heidelberg: Springer, 2014, pp. 177–210.
3. Elkind MS, Ramakrishnan P, Moon YP, et al. Infectious burden and risk of stroke: the northern Manhattan study. Arch Neurol 2010; 67: 33–38. [PubMed: 19901154]
4. Corrado E, Rizzo M, Tantillo R, et al. Markers of inflammation and infection influence the outcome of patients with baseline asymptomatic carotid lesions: a 5-year follow-up study. Stroke 2006; 37: 482–486. [PubMed: 16373649]
5. Kis Z, Sas K, Gyulai Z, et al. Chronic infections and genetic factors in the development of ischemic stroke. Microbiol Q J Microbiol Sci 2007; 30: 213–220.
6. Smieja M, Gnarpe J, Lonn E, et al. Heart Outcomes Prevention Evaluation (HOPE) Study Investigators. Multiple infections and subsequent cardiovascular events in the Heart Outcomes Prevention Evaluation (HOPE) Study. Circulation 2003; 107: 251–257. [PubMed: 12538424]
7. Haider AW, Wilson PW, Larson MG, et al. The association of seropositivity to Helicobacter pylori, Chlamydia pneumoniae, and cytomegalovirus with risk of cardiovascular disease: a prospective study. J Am Coll Cardiol 2002; 40: 1408–1413. [PubMed: 12392829]
8. Jones JL, Kruszon-Moran D, Sanders-Lewis K and Wilson M. Toxoplasma gondii infection in the United States, 1999 2004, decline from the prior decade. Am J Trop Med Hyg 2007; 77: 405–410.
9. Won KY, Kruszon-Moran D, Schantz PM and Jones JL. National seroprevalence and risk factors for zoonotic Toxocara spp. infection. Am J Trop Med Hyg 2008; 79: 552–557. [PubMed: 18840743]
10. Dupont CD, Christian DA and Hunter CA. Immune response and immunopathology during toxoplasmosis. In: Seminars in immunopathology 2012; 34: 793–813. [PubMed: 22955326]
11. Fan CK, Liao CW and Cheng YC. Factors affecting disease manifestation of toxocarosis in humans: genetics and environment. Vet Parasitol 2013; 193: 342–352. [PubMed: 23290279]
12. Sealy-Jefferson S, Gillespie BW, Aiello AE, Haan MN, Morgenstern LB and Lisabeth LD. Antibody levels to persistent pathogens and incident stroke in Mexican Americans. PLoS ONE 2013; 8: e65959. [PubMed: 23799066]
13. Goldstein LB, Bushnell CD, Adams RJ, et al. Guidelines for the primary prevention of stroke a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011; 42: 517–584. [PubMed: 21127304]
14. NCHS. Sample design: Third National Health and Nutrition Examination Survey. In: U.S. Department of Health and Human Services, ed. Vol Publication No. (PHS) 92–1387. Washington, DC: US GPO, 1992.
15. Staras SA, Dollard SC, Radford KW, Flanders WD, Pass RF and Cannon MJ. Seroprevalence of cytomegalovirus infection in the United States, 1988–1994. Clin Infect Dis 2006; 43: 1143–1151. [PubMed: 17029132]
16. McQuillan GM, Coleman PJ, Kruszon-Moran D, Moyer LA, Lambert SB and Margolis HS. Prevalence of hepatitis B virus infection in the United States: the National Health and Nutrition Examination Surveys, 1976 through 1994. Am J Publ Health 1999; 89: 14–18.
17. Jones JL, Kruszon-Moran D, Wilson M, McQuillan G, Navin T and McAuley JB. Toxoplasma gondii infection in the United States: seroprevalence and risk factors. Am J Epidemiol 2001; 154: 357–365. [PubMed: 11495859]
18. Bell BP, Kruszon-Moran D, Shapiro CN, Lambert SB, McQuillan GM and Margolis HS. Hepatitis A virus infection in the United States: serologic results from the third National Health and Nutrition Examination Survey. Vaccine 2005; 23: 5798–5806. [PubMed: 16307834]
19. Web resource: http://cdc.gov/nchs/nhanes/nh3data.htm (accessed 4 March 2015).
20. Mohadjer LWJ. National Health and Nutrition Examination Survey III: weighting and estimation methodology In: National Center for Health Statistics, Hyattsville, MD, https://cdc.gov/nchs/Data/Nhanes3/Manuals/wgt_exec.pdf (1996, accessed 4 March 2015).

21. Ezzati TKM. Nonresponse adjustment in a National Health Survey In: 1992 Proceedings of the section on survey research methods. Alexandria, VA: American Statistical Association, 1993, pp.339–344.
22. Korn EL and Graubard BI. Confidence intervals for proportions with small expected number of positive counts estimated from survey data. Surv Meth 1998; 24: 193–201.
23. Fernández-Real JM, López-Bermejo A, Vendrell J, Ferri MJ, Recasens M and Ricart W. Burden of infection and insulin resistance in healthy middle-aged men. Diabetes Care 2006; 29: 1058–1064. [PubMed: 16644637]
24. Cuthbert JA. Hepatitis A: old and new. Clin Microbiol Rev 2001; 14: 38–58. [PubMed: 11148002]
25. Trepo C, Chan HL and Lok A. Hepatitis B virus infection. Lancet 2014; 384: 2053–2063. [PubMed: 24954675]
26. Ebrahimi A, Nabipour I, Vahdat K, et al. High sensitivity C-reactive protein is associated with the metabolic syndrome independent to viral and bacterial pathogen burden. Diabetes Res Clin Pract 2009; 84: 296–302. [PubMed: 19362745]
27. Perlstein TS, Pande RL, Creager MA, Weuve J and Beckman JA. Serum total bilirubin level, prevalent stroke, and stroke outcomes: NHANES 1999–2004. Am J Med 2008; 121: 781–788. e781. [PubMed: 18724968]
28. Griffiths D and Sturm J. Epidemiology and etiology of young stroke. Stroke Res Treat 2011; 2011: 209370, 9.
29. Carod-Artal FJ and Gascon J. Chagas disease and stroke. Lancet Neurol 2010; 9: 533–542. [PubMed: 20398860]
30. Selvaraj UM, Poinsatte K, Torres V, Ortega SB and Stowe AM. Heterogeneity of B cell functions in stroke-related risk, prevention, injury, and repair. Neurotherapeutics 2016; 13: 729–747. [PubMed: 27492770]

Table 1.

Age-adjusted[a] and age-specific prevalence (weighted percent) of stroke by each individual infection, serological evidence of prior infection (SEPI), and specific cofactors, among adults ages 20 years and older, NHANES III

Cofactor	Level of cofactor	Sample size	Prevalence of stroke	Lower 95% confidence interval	Upper 95% confidence interval	*p* value *t* test[b]
Total		13,904	1.94	1.63	2.29	
Hepatitis A	Seronegative	5944	1.63	1.25	2.08	
	Seropositive	7960	2.41	1.85	3.09	0.02
Hepatitis B	Seronegative	12,734	1.81	1.54	2.12	
	Seropositive	1170	3.69	2.05	6.09	0.04
Cytomegalovirus	Seronegative	2909	1.05	0.77	1.39	
	Seropositive	10,995	2.18	1.76	2.66	<0.01
Toxoplasma gondii	Seronegative	10,006	1.76	1.41	2.16	
	Seropositive	3898	2.47	1.96	3.07	0.02
Toxocara	Seronegative	11,793	1.79	1.54	2.06	
	Seropositive	2111	2.87	1.85	4.23	0.05
Serological evidence of prior infection	0–1 diseases	4906	1.49	1.01	2.11	
	2 diseases	5166	1.71	1.3	2.22	
	3 diseases	2905	2.75	2.11	3.51	
	4–5 diseases	927	5.5	2.44	10.46	0.02[c]
Age	*Reference* 20–59	9390	0.70	0.48	0.98	
	60–69	2033	3.99	2.78	5.53	<0.01
	70+	2481	8.35	7.07	9.79	<0.01
Race and Hispanic origin	*Reference* non-Hispanic white	5823	1.82	1.52	2.16	
	Non-Hispanic black	3788	2.59	1.90	3.43	0.05
	Mexican American	3757	1.73	1.29	2.27	0.75

Cofactor	Level of cofactor	Sample size	Prevalence of stroke	Lower 95% confidence interval	Upper 95% confidence interval	*p* value *t* test[b]
Sex	Female	7211	1.79	1.39	2.27	0.38
	Male	6693	2.08	1.64	2.60	
Foreign birth	US born	11,006	1.89	1.62	2.18	0.55
	Born outside the US	2861	2.27	1.16	3.99	
Metropolitan residence	No	7196	2.09	1.61	2.67	0.22
	Yes	6708	1.73	1.44	2.06	
Poverty index	At or above	9674	1.76	1.47	2.10	<0.01
	Below	2956	3.38	2.39	4.63	
Education	Completed high school or greater	8194	1.51	1.16	1.92	
	Less than high school	5624	2.99	2.35	3.75	<0.01
Smoking	Never	6223	1.37	1.07	1.73	<0.01
	Any	7623	2.30	1.93	2.72	
Alcohol use	*Reference* current	6220	1.29	0.96	1.70	
	Past	4902	2.59	2.00	3.29	<0.01
	Never	2425	2.11	1.47	2.94	0.02
Obesity[d]	Not obese	10,125	1.77	1.39	2.21	
	Obese	3530	2.43	1.76	3.26	0.14
Diabetes	No diagnosis	12,813	1.57	1.32	1.86	
	Diagnosis	1073	6.39	4.24	9.19	<0.01
High total cholesterol[e]	Normal	10,989	1.74	1.41	2.12	0.02
	High	2872	2.89	2.05	3.94	
Low HDL[f]	Normal	8721	1.44	1.14	1.80	
	Low	5043	2.77	2.17	3.47	<0.01
High triglycerides[g]	Normal	4269	1.22	0.88	1.66	

Int J Stroke. Author manuscript; available in PMC 2020 December 29.

Cofactor	Level of cofactor	Sample size	Prevalence of stroke	Lower 95% confidence interval	Upper 95% confidence interval	*p* value *t* test[b]
	High	1874	3.21	2.06	4.74	<0.01
Hypertension/hypertension medication[h]	*Reference* nonhypertensives	8254	1.26	0.90	1.71	
	Nonmedicated hypertensives	2959	1.88	1.25	2.69	0.13
	Medicated hypertensives	2238	4.44	3.06	6.21	<0.01
High CRP	*Reference* <0.22 mg/dL	9060	1.50	1.22	1.83	
	0.22–0.99 mg/dL	3452	2.17	1.54	2.96	0.08
	≥1.0 mg/dL	1350	4.58	2.95	6.74	<0.01
Bilirubin[i]	*Reference* lowest tertile	5051	2.37	1.81	3.03	
	Middle tertile	4684	1.93	1.55	2.39	0.20
	Highest tertile	4068	1.52	1.17	1.93	0.01

CRP: C-reactive protein; HDL: high-density lipoprotein; NHANES: National Health and Nutrition Examination Survey.

[a] Age-adjusted, except for age cofactor.

[b] *p* value from *t* test comparing prevalence of stroke in each individual group to reference group within each cofactor.

[c] *p* values from linear test for trend in stroke prevalence with increasing level of serologic evidence of prior infection.

[d] Obesity—body mass index greater than or equal to 30.

[e] High total cholesterol—greater than or equal to 240 mg/dL.

[f] Low HDL—less than 40 mg/dL for males and less than 50 for females.

[g] High triglycerides—greater than or equal to 150 mg/dL, Gathered on a random half sample of participants who fasted.

[h] Hypertension/hypertension medication—(1) *nonhypertensives*: those who had no history of diagnosed hypertension/high pressure and whose blood pressure was normal at exam; (2) *nonmedicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure or who measured ≥140 systolic or ≥90 diastolic blood pressure at the NHANES exam but who are NOT currently taking hypertension medication; or (3) *medicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure and are currently taking medication for it.

[i] Lowest tertile is <0.5 mg/dL, middle tertile 0.5–0.6, and highest tertile ≥0.7 mg/dL.

Table 2.

Summary of regression analyses for variables associated with increasing serological evidence of prior infection,[b] among adults age 20 years and older, NHANES III

Cofactor	Level of cofactor	Beta	*p* value
Age	*Reference* 20–59		
	60–69	0.55	<0.01
	70+	0.75	<0.01
Race and Hispanic origin	*Reference* Non-Hispanic white		
	Non-Hispanic black	0.49	<0.01
	Mexican American	0.67	<0.01
Sex	Female		
	Male	0.00	0.85
Foreign birth	US born		
	Born outside the US	0.89	<0.01
Metropolitan residence	No		
	Yes	0.06	0.23
Poverty index	At or above		
	Below	0.40	<0.01
Education	Completed high school or greater		
	Less than completed high school	0.58	<0.01
Smoking	Never		
	Any	−.01	0.83
Alcohol use	*Reference* current		
	Past	0.17	<0.01
	Never	0.41	<0.01
Obesity[c]	Not obese		

Cofactor	Level of cofactor	Beta	*p* value
	Obese	0.09	<0.01
Diabetes	No diagnosis		
	Diagnosis	0.19	<0.01
Total cholesterol[d]	Normal		
	High	0.08	0.03
Low HDL[e]	Normal		
	Low	0.10	<0.01
High triglycerides[f]	Normal		
	High	0.09	0.01
Hypertension/hypertension medication[g]	*Reference* nonhypertensives		
	Nonmedicated hypertensives	0.07	0.04
	Medicated hypertensives	0.16	<0.01
High CRP	*Reference* <0.22 mg/dL		
	0.22–0.99 mg/dL	0.07	0.02
	≥1.0 mg/dL	0.16	<0.01
Bilirubin[h]	*Reference* lowest tertile		
	Middle tertile	−.07	0.02
	Highest tertile	−.11	<0.01

CRP: C-reactive protein; HDL: high-density lipoprotein; NHANES: National Health and Nutrition Examination Survey.

[a]Age-adjusted, except for age cofactor (please refer to methods section for details).

[b]Serologic evidence of prior infection is a measure of burden of infection based on the number of infections for which a sample person tested positive out of five organisms (*Toxoplasma gondii, Toxocara* spp., cytomegalovirus, hepatitis A and hepatitis B virus).

[c]Obesity—body mass index greater than or equal to 30.

[d]High total cholesterol—greater than or equal to 240 mg/dL.

[e] Low HDL—less than 40 mg/dL for males and less than 50 for females.

[f] High triglycerides—greater than or equal to 150 mg/dL

[g] Hypertension/hypertension medication—(1) *nonhypertensives*: those who had no history of diagnosed hypertension/high pressure and whose blood pressure was normal at exam; (2) *nonmedicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure or who measured ≥140 systolic or ≥90 diastolic blood pressure at the NHANES exam but who are NOT currently taking hypertension medication; or (3) *medicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure and are currently taking medication for it.

[h] Lowest tertile is <0.5 mg/dL, middle tertile 0.5–0.6, and highest tertile ≥0.7 mg/dL.

Table 3.

Odds ratios and *p* values from logistic models adjusting for age, SEPI, and each individual cofactor as well as testing for interaction of SEPI and each cofactor on prevalence of stroke among adults ages 20 and older from NHANES III

Cofactor	Levels of cofactor	OR[a]	95% confidence interval	*p* value	*p* value for interaction of SEPI and cofactor
Age	*Reference* 20–59				
	60–69	4.70	2.86–7.70	<0.01	0.03
	70+	9.66	6.22–15.0	<0.01	<0.01
SEPI		1.49	1.22–1.81	<0.01	
Race and Hispanic origin	*Reference* non-Hispanic white				
	Non-Hispanic black	1.23	0.83–1.83	0.29	0.69
	Mexican American	0.72	0.47–1.10	0.13	0.19
SEPI		1.38	1.09–1.73	0.01	
Sex	Female				
	Male	1.16	0.81–1.67	0.42	0.81
SEPI		1.48	1.22–1.81	<0.01	
Foreign birth	US born				
	Born outside the US	0.90	0.47–1.69	0.73	0.53
SEPI		1.51	1.23–1.84	<0.01	
Metropolitan residence	No				
	Yes	0.82	0.59–1.13	0.22	0.35
SEPI		1.49	1.22–1.82	<0.01	
Poverty index	At or above				
	Below	1.68	1.20–2.36	<0.01	0.47
SEPI		1.46	1.16–1.83	<0.01	
Education	*Reference* completed high school or greater				

Cofactor	Levels of cofactor	OR[a]	95% confidence interval	*p* value	*p* value for interaction of SEPI and cofactor
	Less than high school	1.69	1.21–2.37	<0.01	0.57
SEPI		1.39	1.14–1.69	<0.01	
Smoking	Never				
	Any	1.73	1.32–2.28	<0.01	0.87
SEPI		1.53	1.23–1.90	<0.01	
Alcohol use	*Reference* current				
	Past	1.95	1.27–2.98	<0.01	0.03
	Never	1.62	1.10–2.37	0.02	0.11
SEPI		1.46	1.18–1.81	<0.01	
Obesity[b]	Not obese				
	Obese	1.39	0.91–2.13	0.12	0.16
SEPI		1.48	1.21–1.81	<0.01	
Diabetes	No diagnosis				
	Diagnosis	3.80	2.70–5.35	<0.01	0.05
SEPI		1.45	1.18–1.78	<0.01	
High total cholesterol[c]	Normal				
	High	1.54	1.03–2.28	0.03	0.47
SEPI		1.48	1.22–1.81	<0.01	
Low HDL[d]	Normal				
	Low	1.92	1.37–2.68	<0.01	0.22
SEPI		1.47	1.20–1.81	<0.01	
High triglycerides[e]	Normal				

Int J Stroke. Author manuscript; available in PMC 2020 December 29.

Cofactor	Levels of cofactor	OR[a]	95% confidence interval	*p* value	*p* value for interaction of SEPI and cofactor
	High	2.66	1.81–3.90	<0.01	0.29
SEPI		1.98	1.43–2.74	<0.01	
Hypertension/hypertension medication[f]	Medicated hypertensives				
	Nonmedicated hypertensives	3.49	2.22–5.49	<0.01	0.02
	Reference nonhypertensives	1.55	0.91–2.64	0.10	0.84
SEPI		1.45	1.18–1.77	<0.01	
High CRP	*Reference* <0.22 mg/dL				
	0.22–0.99 mg/dL	1.47	0.99–2.17	0.05	0.30
	≥1.0 mg/dL	3.02	1.95–4.69	<0.01	0.06
SEPI		1.48	1.21–1.81	<0.01	
Bilirubin[g]	*Reference* lowest tertile				
	Middle tertile	0.82	0.60–1.12	0.21	0.75
	Highest tertile	0.62	0.45–0.86	<0.01	0.27
SEPI		1.48	1.21–1.81	<0.01	

CRP: C-reactive protein; HDL: high-density lipoprotein; NHANES: National Health and Nutrition Examination Survey; OR: odds ratio; SEPI: serological evidence of prior infection.

[a] ORs for covariates refer to the increase in odds of stroke with each increase in a level of the cofactor when the interaction term is not in the model; ORs for SEPI refer to the increase in odds of stroke for each increase in level of SEPI.

[b] Obesity—body mass index greater than or equal to 30.

[c] High total cholesterol—greater than or equal to 240 mg/dL.

[d] Low HDL—less than 40 mg/dL for males and less than 50 for females.

[e] High triglycerides—greater than or equal to 150 mg/dL.

[f] Hypertension/hypertension medication—(1) *medicated hypertensives*: those with history of diagnosed hypertension or high blood pressure and are currently taking medication for it; (2) *nonmedicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure or who measured ≥140 systolic or ≥90 diastolic blood pressure at the NHANES exam but who are NOT currently taking hypertension medication; or (3) *nonhypertensives*: those who had no history of diagnosed hypertension/high pressure and whose blood pressure was normal at exam.

[g] Lowest tertile is <0.5 mg/dL, middle tertile 0.5–0.6, and highest tertile ≥0.7 mg/dL.

Table 4.

Association of increasing serological evidence of prior infection (SEPI) on prevalence of stroke from complex stepwise models, among adults ages 20 years and older, NHANES III

	Cofactors[a]	SEPI beta	*p* value	OR for SEPI	OR lower limit, 95% confidence	OR upper limit, 95% confidence
Model 1	SEPI Age, race, and Hispanic origin, education, smoking, alcohol use	0.25	0.03	1.29	1.02	1.63
Model 2	SEPI Age, race and Hispanic origin, education, smoking, alcohol use, **diabetes**	0.24	0.05	1.27	1.00	1.61
Model 3	SEPI Age, education, smoking, diabetes, **low HDL, hypertension/hypertension medication**	0.31	0.01	1.37	1.08	1.73
Model 4	SEPI Age, education, smoking, diabetes low HDL, hypertension/hypertension medication, **high CRP, bilirubin level**	0.32	0.01	1.38	1.09	1.74

CRP: C-reactive protein; HDL: high-density lipoprotein; NHANES: National Health and Nutrition Examination Survey; OR: odds ratio.

[a]Significant cofactors in model—those added are bold and those retained in model are regular font.

Table 5.

Final simplified backwards stepwise model of stroke for total population, among adults ages 20 years and older, NHANES III: All ages and stratified by age

	All ages		Ages 20–59		Ages 60–69		Ages 70+	
Cofactor	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**
Age								
20–59 (Ref)	1.00		N/A		N/A		N/A	
60–69	2.30[a]	1.29–4.10						
≥70	5.55[a]	3.43–8.99						
Sex								
Male	1.55[a]	1.05–2.30	2.16	0.96–4.85	1.30	0.60–2.82	1.16	0.81–1.65
Female (Ref)	1.00		1.00		1.00		1.00	
Education								
Less than high school	1.46[a]	1.04–2.04	1.59	0.71–3.57	1.65	0.97–2.79	1.27	0.94–1.72
High school or greater (Ref)	1.00		1.00		1.00		1.00	
Smoking								
Ever	1.54[a]	1.11–2.14	1.81	0.75–4.38	1.11	0.49–2.49	1.72[a]	1.13–2.61
Never (Ref)	1.00		1.00		1.00		1.00	
Diabetes								
Diagnosis	2.66[a]	1.87–3.78	2.40	0.91–6.31	3.38[a]	1.37–8.32	2.16[a]	1.49–3.14
No diagnosis (Ref)	1.00		1.00		1.00		1.00	
Total cholesterol[b]								
High	1.52[a]	1.07–2.17	2.67[a]	1.19–6.00	0.65	0.33–1.28	1.56[a]	1.07–2.28
Normal (Ref)	1.00		1.00		1.00		1.00	
HDL[c]								

	All ages		Ages 20–59		Ages 60–69		Ages 70+	
Cofactor	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**
Low	1.57[a]	1.05–2.35	1.24	0.47–3.32	1.66	0.75–3.65	1.89[a]	1.29–2.76
Normal (Ref)	1.00		1.00		1.00		1.00	
Hypertension/hypertension medication[d]								
Medicated hypertensives	2.76[a]	1.77–4.29	2.77[a]	1.07–7.15	1.56	0.77–3.16	2.95[a]	1.78–4.90
Nonmedicated hypertensives	1.41	0.84–2.35	1.89	0.72–5.01	0.96	0.40–2.29	1.27	0.81–2.01
Nonhypertensives (Ref)	1.00		1.00		1.00		1.00	
CRP								
≤0.21 (Ref)	1.00		1.00		1.00		1.00	
0.22–099	1.11	0.74–1.69	0.99	0.44–2.25	1.78	0.71–4.49	0.87	0.60,1.26
≥1.0	1.96[a]	1.22–3.16	2.48	0.90–6.84	2.26	0.90–5.63	1.40	0.86–2.29
Bilirubin[e]								
Lowest tertile (Ref)	1.00		1.00		1.00		1.00	
Middle tertile	0.88	0.64–1.22	0.76	0.34–1.71	1.01	0.50–2.07	0.88	0.61–1.27
Highest tertile	0.62[a]	0.44–0.86	0.27[a]	0.09–0.87	0.75	0.33–1.71	0.93	0.59–1.48
Serologic evidence of prior infection	1.37[a]	1.09–1.73	1.96[a]	1.24–3.11	1.06	0.69–1.62	1.12	0.84–1.51

CRP: C-reactive protein; CI: confidence interval; HDL: high-density lipoprotein; NHANES: National Health and Nutrition Examination Survey; OR: odds ratio; Ref: reference group.

[a]Significant at $p < 0.05$.

[b]High total cholesterol—greater than or equal to 240 mg/dL.

[c]Low HDL—less than 40 mg/dL for males and less than 50 for females.

[d]Hypertension/hypertension medication—(1) *medicated hypertensives*: those with history of diagnosed hypertension or high blood pressure and are currently taking medication for it; (2) *nonmedicated hypertensives*: those with a history of diagnosed hypertension or high blood pressure or who measured ≥140 systolic or ≥90 diastolic blood pressure at the NHANES exam but who are NOT currently taking hypertension medication; or (3) *nonhypertensives*: those who had no history of diagnosed hypertension/high pressure and whose blood pressure was normal at exam.

[e]Lowest tertile is <0.5 mg/dL, middle tertile 0.5–0.6, and highest tertile ≥0.7 mg/dL.