Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Anna Bratcher Date

Cumulative Infectious Disease Burden's Association with Stroke and Mood Disorders in NHANES III

By

Anna Bratcher

Master of Science in Public Health

Epidemiology

Signature:

_____

Brad Pearce

Committee Chair

Cumulative Infectious Disease Burden's Association with Stroke and Mood Disorders in NHANES III

By

Anna Bratcher

Bachelor of Science

University of Georgia

2012

Thesis Committee Chair: Brad Pearce, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Epidemiology

2016

Abstract

Cumulative Infectious Disease Burden's Association with Stroke and Mood Disorders in NHANES III

By Anna Bratcher

Neuropsychiatric conditions such as stroke and mood disorders affect millions of Americans each year. The etiology of these conditions is complex, with the body of knowledge on their causes constantly growing. Given evidence of chronic infection affecting the central nervous system, we hypothesized that higher cumulative infectious disease burdens could lead to more neuropsychiatric conditions, including stroke and mood disorders. In this thesis, cumulative infectious disease burden was measured as cumulative seroprevelance to hepatitis A, hepatitis B, cytomegalovirus, *Toxoplasma gondii* and *Toxocara spp.*. Using the third national health and nutrition examination survey (NHANES III) data, logistic regressions of history of stroke, major depression, severe major depression, dysthymia, dysthymia with major depression, bipolar disorder and atypical bipolar disorder were done to assess associations with cumulative infectious disease burden. Stroke showed a significant relationship with an odds ratio of 1.37(CI: 1.09, 1.73), with evidence of interaction with age. Among 20 to 59 year olds, the odds ratio was 1.96 (CI: 1.24, 3.11). In comparison, the older age groups, 60 to 69 year olds and above 70 year olds, failed to show significant relationships (OR=1.06 CI: 0.69, 1.62; OR=1.12 CI:0.84,1.51 respectively). There were no significant relationships found between cumulative infectious disease burden and any of the mood disorder outcomes (P-Value>0.05 in all cases). These findings suggest that cumulative infectious disease burden my lead to a higher risk of stroke, particularly in younger ages, but not mood disorders.

Cumulative Infectious Disease Burden's Association with Stroke and Mood Disorders in NHANES III

By

Anna Bratcher

Bachelor of Science

University of Georgia

2012

Thesis Committee Chair: Brad Pearce, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Epidemiology

2016

Acknowledgements

This thesis would not have been possible without the guidance and support of many people. Firstly, I would like to express my gratitude to my thesis advisor, Associate Professor Brad Pearce, for his dedication to this thesis. Secondly, I would like to thank Drs. Deanna Kruszon-Moran and Jeffrey Jones for guidance and direction with the work shown here and related research. It was a pleasure to closely work with these epidemiologist in the analyses presented and in a forthcoming paper. I owe much of my epidemiology skill to these wonderful teachers.

I would also like to thank the rest of the Rollins faculty and staff who have contributed to my education, particularly Jena Black and all her ADAP wisdom. It has been a blessing to be included in such a compassionate, enthusiastic, and excellent department, school and university.

Lastly, I would like to thoroughly thank my family and friends for supporting me on my way to and at Rollins. This thesis is dedicated to future endeavors that we will take together.

Table of Contents

Introduction

The brain is vulnerable to a number of adverse exposures that can affect motor function, cognition, and mood. Pathologies induced by cerebrovascular events include stroke, both ischemic and hemorrhagic. Ischemic strokes are characterized by a blockage occurring in a brain vessel causing damage to the surrounding area, limiting blood flow and therefore tissue oxygenation. This leads to varying degrees of injury, inflammation, and ultimately neuronal death in the affected brain regions. Hemorrhagic strokes occur when a blood vessel ruptures, and the consequential bleeding into the parenchyma of brain causes some of the same symptoms as seen in ischemic stroke.

Mental disorders exemplify brain disorders with complex causation and intertwined risk factors with some other brain disorders. Examples include major depression, dysthymia, bipolar disorder, and schizophrenia.

The terminology used to explain mental disorders and stroke often very based on historical origins and areas of specialization such as psychology, neurology, or neuropathology. For the purpose of this thesis, I use the term neuropsychiatric disorder to refer to either a mental disorder or cerebral vascular disorders such as stroke.

Approximately 20% of Americans experience a mental disorder each year.[1] Additionally, around 795,000 Americans experienced a stroke in the year 2009.[2] In addition to affecting large populations, these two types of neuropsychiatric disorders form a significant healthcare burden in the United States. In 1996, direct costs of mental disorders came out to 69 billion dollars.[1] Stroke costs totaled 72 billion dollars in 2013 with projected 2030 costs estimated around 183 billion dollars.[3] Part of these costs are

determined by the large number of people affected each year. However this is not the only reason for such significant financial burdens.

One reason for high annual costs is that these diseases are complex disorders that require advanced treatments. Each mental disorder and stroke type has many mechanisms and possible causes.[4] Multiple causes complicate the curing and treatment of these diseases and raise the cost by increasing research needs and making effective treatments hard to identify. Mental disorders and stroke both have much research already performed that attempts to tease out major causes. Yet, in both cases, the exact list of causes has yet to been identified. Therefore, further work identifying as many possible sources for these problems may be difficult, but necessary for more efficient treatment.

In this thesis, infectious disease as a potential cause behind the neuropsychiatric disorders of stroke and mental disorders will be examined. The idea that infectious disease may be involved with the alteration of brain function is not a new concept. The majority of individuals who have contracted influenza or mononucleosis have experienced disease induced mental changes such as mood and vegetation symptoms.[5] In extreme cases, an infectious disease can cause brain changes that can be seen by MRI. In the case of rabies, "nonenhancing, ill-defined, mild hyperintensity changes in the brain stem, hippocampi, hypothalami, deep and subcortical white matter, and deep and cortical gray matter" were observed in human cases causing both furious and dumb psychotic episodes.[6]

In the case of stroke, there are many infections that have been shown to be associated with both ischemic and hemorrhagic strokes.[7] There is a good link established between acute infection and stroke in the literature. Infections, typically respiratory and

bacterial, particularly those that occur in the week prior are a significant risk factor for stoke, even after other traditional stroke risk factors are taken into account. Additionally, chronic infections may be linked to stroke, though this has not been as well established.[8]

Not only have infectious diseases been associated with stroke, but they have also been associated with mental disorders.[5,9] Mental diseases, including mood disorders, have been linked to many infectious diseases, either through a specific disease with a specific disorder, or through non-specific combinations of disease. There are several mechanisms proposed for this relationship that are described below.

Not only is infectious disease associated with both mental health and stroke, but mental health and stroke are also associated. It has been shown in observational studies that depressive disorders can be quite common following stroke.[10,11] This indicates an extra layer of complexity to the infectious disease and neuropsychiatric disorder relationship. This thesis will attempt to unravel the associations at play here to describe infectious disease burden's relationship with the neuropsychiatric disorders of stroke and mental disorders.

Background

Infectious Agents

There are many infectious diseases that have established links to altered brain function, either through stroke or mental disorders. This paper focuses on a small subset of these infections which can have subclinical or silent manifestations. In these diseases, when subclinical infections occur there are not enough symptoms to cause the affected individual to seek out care. As a result, these diseases are often not treated and may become persistent. This unresolved nature of these infections contributes to the cumulative disease burden of the individual. The infectious agents behind subclinical diseases examined here are hepatitis A, hepatitis B, cytomegalovirus, *toxoplasma gondii*, and *toxocara*.

Hepatitis A

Hepatitis A is the leading cause of acute hepatitis in the United States. The infectious agent is a spherical virus found in, and transmitted by stool.[12] In clinical cases, infected individuals develop non-specific constitutional symptoms followed by gastrointestinal problems. Less common symptoms include chills, pain, cough, upper respiratory symptoms, constipation and diarrhea among others. Once the icteric phase is entered, the urine darkens and jaundice becomes apparent. After the icteric stage, clinical symptoms resolve at a rate proportional to the duration of the jaundice. [13]

Though clinical symptoms can be expressed, they are not always manifested in Hepatitis A infected individuals. In one outbreak 30% of those infected did not show jaundice, and 14% were completely asymptomatic. These proportions can be attributed to

the age of those affected by this outbreak. While symptoms can manifest at any age, cases in young children tend to be more asymptomatic.[13]

While hepatitis A is generally an acute infection, there is evidence for prolonged infections that result in relapse. The duration of jaundice is variable, with some cases lasting up to 120 days. Some individuals showed symptoms up to 30 months after hospitalization. Most patients completely recover in 6 months.[13]

These long lasting symptoms include fatigue. Fatigue is the major psychological symptom that comes from hepatitis A infection. This symptom is present in roughly 80% of cases. In some cases, patients will experience mental symptoms after the infection is resolved, such as fatigue and loss of appetite.[14]

Hepatitis B

Chronic hepatitis B infection affects 360 million people world-wide. As the 10th leading cause of death in the world, each year around 1 million deaths are attributed to hepatitis B. The hepatitis B virus (HBV) is a double-stranded DNA virus with perinatal and horizontal transmission. Given that a small amount of body fluids has the ability to carry a high viral load, HBV is highly infectious through intimate contact. Sexual transmission has been well documented, as well as through intravenous drug use with infected needles. [15]

Hepatitis B commonly causes acute infections both with and without symptoms. However, it is also possible for hepatitis B to cause a chronic infection.[15] In the acute infections, around one-third of adults show symptoms, while the remaining two-thirds do not exhibit any clinical disease. Serious consequences mainly included cirrhosis and liver

cancer in a fraction of those with cirrhosis. 5% of hepatitis B infections in adults result in chronic cases, though rates differ in younger ages. 90% of infants and 30% of children infected with HBV develop chronic infection. Chronic infections seem to be a result of poor immune clearance of the virus infected cells and is indicated by persistence of high levels of HBV DNA and other markers in the infected individual's serum.[16]

Beyond symptoms involving the liver, hepatitis B chronic infection has shown to produce symptoms involving the central nervous system. These symptoms include depression and cognitive dysfunctions. Notably, these neuropsychiatric symptoms increase in intensity as cirrhosis becomes more pronounced in infected individuals. However, patients with hepatitis B chronic infection have less severe depression and cognitive dysfunction than patients suffering from other chronic liver diseases.[17]

Cytomegalovirus

Human cytomegalovirus is a DNA virus in the herpesviridae family. This virus is extremely common, with seroprevalance in 45-100% adults of reproductive age with large variations by geography. In the United States, prevalence can vary by up to 30% between states.[18] Transmission occurs through person-to-person contact. In childhood transmission generally occurs through breast-feeding or contact with other children in nurseries or day care centers. Cytomegalovirus infection in adolescents and adults generally spreads through salivary and sexual transmission.[19]

Generally, cases of cytomegalovirus that present with symptoms are among immunocompromised individuals, particularly newborns, transplant recipients and HIV infected persons. These clinical cases can arise from either novel infection or reactivation

of a latent virus. The symptoms that manifest vary by the source of a patient's immunocompromised state. HIV infected individuals generally develop retinitis. Organ transplant recipients show symptoms that vary by the organ received. For example, bone marrow transplant recipients tend to develop cytomegalovirus pneumonia while liver transplant patients more commonly experience cytomegalovirus induced hepatitis. Congenital cytomegalovirus presents with petechiae, jaundice, and neurologic abnormalities including encephalopathy, seizures and hearing defects, among other symptoms. In some immunocompetent persons, mononucleosis-like illness can develop.[19]

Though there are clinical cases of cytomegalovirus infection, many infections are subclinical and even silent. Like other herpes viruses, primary infection by cytomegalovirus is followed by a latent infection. The sites of latent infection is thought to be the bone marrow and peripheral blood monocytes.[19]

Though the majority of cytomegalovirus cases are subclinical, there are established links between this virus and the central nervous system. Congenital cytomegalovirus infection can result in permanent disabilities such mental retardation and milder cognitive impairment which tend to manifest in infants who are symptomatic at birth.[20] In immunocompromised adults, cytomegalovirus has been shown to create inclusions in the central nervous system, creating variable degrees of inflammation. These inclusions were found predominantly near small blood vessels.[21] Additionally, cytomegalovirus is associated with mononucleosis-like fatigue in less symptomatic cases among immunocompetent patients.[19]

Toxoplasma gondii

Toxoplasma gondii is a cyst-forming parasite that produces toxoplasmosis in humans. *T. gondii* is prevalent in most of the world, with some studies showing that up to one third of the global population has been exposed to the parasite. *T. gondii* can be transmitted by several different routes. If first contracted during pregnancy, it can be vertically transmitted to the fetus via the placenta. In the United States, horizontal transmission generally occurs from ingesting cyst from cat feces. However, over recent decades the prevalence of *T. gondii* in commercially available meat has decreased significantly. Transmission by blood products and tissue transplants can also take place.[22]

While T. gondii affects a large proportion of the population, overt clinical disease is generally only seen in risk groups including fetuses and immunocompromised patients. In the case of congenital toxoplasmosis, about 10% of prenatal infections result in abortion or death. Around 20% of congenital toxoplasmosis cases show clinical signs of toxoplasmosis at birth. In immunocompromised patients, particularly organ transplant recipients and AIDS patients, *T. gondii* is an important opportunistic pathogen that often produces symptoms. The classical symptoms of toxoplasmosis are retinochoroditis, intracranial calcifications and hydrocephalus.[22]

Outside of high risk groups, toxoplasmosis cases are most often asymptomatic. In some cases mild symptoms, including lymphadenopathy can be observed. However, serious symptoms are very rarely seen in healthy individuals.[22]

There is much evidence that links *T. gondii* to the central nervous system and psychological symptoms. *T. gondii* has been shown to produce behavior changes in infected individuals, even in asymptomatic cases. Some personality traits shown to change in strength over a change in a toxoplasmosis seroconversion were guilt proneness,

group dependency and superego strength or willingness to accept group moral standards.[23] In one study toxoplasmosis was linked to bipolar disorder in the American population.[24] While these links are present, the most established link between toxoplasma and the brain is its relationship with schizophrenia. Studies have shown that in utero infection by *T. gondii* shows a significant increase in risk of developing schizophrenia later in life. This link is hypothesized to involve immune system components' interaction with the developing fetal central nervous system.[25]

Toxocara canis and Toxocara cati

Toxocara canis and *Toxocara cati* are the infectious agents behind toxocarosis. This is the clinical disease caused by the zoonotic roundworm infection of dogs and cats. Toxocarosis affects humans world-wide and if frequently underreported. The estimated prevalence of toxocarosis varies greatly from study to study ranging from 4.6-7.3% in children in the United States.[26] The primary route of transmission is through oral ingestion of *Toxocara* eggs. Generally exposure includes contaminated soil and vegetables combined with lack of hygiene, rather than contact with an infected animal.[27]

Clinical disease caused by *Toxocara* species varies widely and is dependent upon the prevalence of larvae among the host tissue. There are three subtypes of clinical disease cause by this parasite. The first is Visceral Larva Migrans which is characterized by a marked, inflammatory immune response. This is most often found in patients 1-7 years of age. The second subtype of Toxocara infection is Ocular Larva Migrans which is characterized by retinal lesions. The mean age of Ocular Larva Migrans is 8 years. Lastly, covert toxocarosis is characterized by non-specific symptoms that do not resemble

either Visceral Larva Migrans or Ocular Larva Migrans. These non-specific symptoms include couch, sleep disturbances, and headaches.[27] While many cases of toxocarosis will have these clinical presentations, most infections remain subclinical or self-limiting. As toxocarosis is often underreported, it is hard to tell what proportion of infections go undiagnosed.[26]

The major psychological symptom generated by toxocarosis is behavioral change exhibited by patients with covert toxocarosis. These behavior changes link *Toxocara* infection to the central nervous system. Furthermore, in rare cases toxocarosis can interact with the central nervous system to become neurological toxocarosis. In some reports, computed tomography scans have shown subtle density changes in the cerebellar region.[26]

Infectious Diseases and the Central Nervous System

Though there is clear evidence that infectious diseases can cause neuropsychiatric disorder, there is very little known about the mechanisms behind how they cause altered brain states. Furthermore, while the link between certain infectious agents, such as rabies, and dramatic symptoms they produce is established, there are many infectious diseases that may cause a mental disorder or stroke but the relationship has not been proven. In addition to the proven and proposed relationships, there is the possibility of relationships that have yet to be identified. Many infectious diseases, including hepatitis, cytomegalovirus, toxoplasma and toxocara strains are subclinical and can even be silent in a portion of infections.[13,15,19,22,27] Given the lack of proper diagnosis on such

subclinical infections, it is hard to clearly link or separate them from neuropsychiatric disorders.

Though there is varying evidence for links between specific infectious agents and mental disorders and stroke, there is a select number of proposed mechanisms through which these links may function. Though some proposed mechanisms apply to all infectious diseases, a few of these mechanisms are specific to viruses. Additionally, some of these proposed mechanisms only apply to mental disorders and not stroke.

The first of these proposed mechanisms evolves from the viral entry process. In some cases, viruses enter the body through binding to neurotransmitter receptors. This binding may interfere with the normal function of neurotransmitters and subsequently interfere with normal brain function. This is a mechanism that applies mostly to viruses, and only to mental disorders.[5]

The second of the proposed mechanisms is developed from how viruses take up residence inside the host's cells. It has been suggested that this residence may interfere with the differentiated or accessory functions of the neuron, resulting in changes in neuron function that are not easily observed. This mechanism is also specific mostly to viruses, and is more often used to explain mental disorders than stroke.[5]

The third proposed mechanism involves infection during pregnancy. One variation of this mechanism suggests that if an infectious agent were to cross over to the fetus, it may interfere with brain development, leading to impaired brain function in the child. Another variation suggests that even if the fetus is not infected, its brain development may be affected by the mother's activated immune system, specifically the

inflammation created by the infection. This mechanism is not limited to viruses, but is still limited to mental disorders.[5]

The last proposed mechanism applies to all infectious diseases, and to both stroke and mental disorders. It is because of its wide applicability we focus on this mechanism in this paper. This mechanism proposes that the inflammation created by infectious diseases adversely affects brain tissue that results in an alteration of brain function characterized by either stroke or a mental disorder. In the case of stroke, some inflammation products such as C reactive protein, plasma fibrinogen, and interleukin-6 have been shown to affect the stability of plaques in vessels. In some cases decreased stability can cause a plaque to dislodge and become the obstruction in a brain blood vessel behind an ischemic stroke.[28] In the case of mental disorders, these cytokines have been shown to interact with neurotransmitters in animal models. In one of these interactions, inflammatory cytokines such as interferon α can cause decreases in tryptophan, a precursor to serotonin. Additionally, inflammatory cytokines have been shown to interact with the hypothalamic-pituitary-adrenal axis in ways that may affect anxiety in the infected individual.[5]

This thesis attempts to gather evidence for the link between infectious disease and the neuropsychiatric disorders of stroke and mental disorders. In response to the proposed mechanisms, specifically inflammation, this thesis will examine the associations between chronic infectious disease burden with stroke and mental disorders.

Inflammation and the Central Nervous System

Though the infectious agents addressed in this thesis do not always produce serious, or even clinical disease, all of these infections have been shown to interact with the human immune system. [13,16,19,22,27] Immune responses to infection activate effector cells and cytokines which damage cells. The purpose of this cell damage is to hopefully injure infected cells so that the infection will be neutralized. Unfortunately, it is common for effector cells and cytokines to damage surrounding tissue.[29] Additionally, the central nervous system can be affected by an immune response taking place in the periphery. Moreover, an infection in a location other than the central nervous system can incite damage in the brain mediated through the immune response to the infectious agent.[5]

It is this evidence for a link between infectious disease and the central nervous system through immune response consequences that has led to the research presented here. This thesis examines the association between infectious disease burden and neuropsychiatric disorders such as mental disorders and stroke. The hypothesis explored here states that if one infectious agent can incite a detrimental immune response in the brain, multiple infectious agents could compound the negative effects of the activated immune system, leading to synergistic damage in central nervous system tissue.

This hypothesis relies on how the immune system responds to multiple concurrent infections. In reality, concurrent infection can affect immune response in various ways. It has been shown to increase immune response to both infectious agents, decrease the immune response for both, and even increase the response for one while lowering the response for another.[30] There is little to no research on the immune system effects with co-infection of any combination of our diseases. Therefore it is possible that the

cumulative burden of these disease could either instigate or alleviate inflammation. However, it is expected that inflammation will be increased as this is typical of co-infection.[29]

The level of inflammation produced by our infections is relevant as our hypothesis proposes that this is the mechanism through which our infectious diseases cause dysfunction of the central nervous system. In the case of stroke, inflammation caused by chronic infections may contribute to the progression of atherosclerotic plaques. These plaques may be susceptible to inflammatory activity generated by the infection, becoming more unstable as the activity increases. It is possible that this instability could cause a portion of the plaque to dislodge and move to brain tissue where it could cause an ischemic stroke.[8]

Furthermore, infections including both acute and chronic can be associated with systemic inflammation producing molecules such as plasma fibrinogen, C-reactive protein, and interleukin-6. These molecules have the ability to injure brain tissue, possibly leading to higher risk of intracranial hemorrhage. In one case-control study, there was a borderline statistically significant association between subarachnoid hemorrhage and an infection within the past month (OR 4.5 [CI 1.0-21.0]). [7]

Not only is there an association between inflammation and both ischemic and hemorrhagic stroke, there is evidence that the level of inflammation could play a role in determining the level of risk in an individual. There is evidence that neuroprotection can be developed as a response to previous insults. These insults confer neuroprotection if they occur below the subthreshold, a level that separates deleterious ischemic events from more minor ones.[8] In this way a low level of inflammation that results in minor ischemic

events could in fact be protective against major stroke events. Concurrently, higher inflammation, perhaps produced by multiple infections at once, could cross the subthreshold. If these two conditions hold true, cumulative infectious disease burden could better predict stroke risk as opposed to infection by any singular disease.

Not only does inflammation cause the central nervous system to become at risk for stroke, it can also cause the brain to malfunction in ways that can lead to mental illness. As previously described, inflammatory molecules can damage tissues other than the infected target cells. Sometimes this damage is done to brain tissue, creating dysfunction.[5] For some infectious disease associated mental disorders, this mechanism has been described further.

In one proposed mechanism, inflammatory molecules reduce the level of serum tryptophan. Since tryptophan in a crucial precursor to serotonin, this reduction would result in serotonin dysregulation associated mental illnesses such as depression and dysthymia. In fact, one study found that inflammatory markers such as interlerukin-6 and interleukin-RA were negatively correlated with tryptophan levels. The same study revealed that persons with sleep disorders or major depression also had higher levels of inflammatory markers, interleukin-6 and interleukin-8.[31]

Beyond interference with tryptophan levels, mental disorders can be induced through other mechanisms. For example, corticotrophin-releasing hormone (CRH) has been associated with depression. This hormone is a regulator of stress response characterized by activation of the hypothalamic-pituitary-adrenal axis and the sympathetic nervous system. It has been shown that cytokines, a certain set of inflammatory molecules, can have profound effects on CRH function.[32] Therefore, if an

infection or group of infections causes the activation of cytokines in the central nervous system, it is foreseeable that these cytokines could disrupt CRH function, resulting in a mental disorder.

Both these molecule-specific mechanisms are only a few examples of how inflammation could possibly lead to mental disease. In reality, there are many more mechanisms that could contribute to the observed association.[5,33]

Unlike with the ischemic subthreshold that offered neuroprotection against ischemic stroke, there is no evidence that a lower level of inflammation may be beneficial with infections and mental disease. However, this does not remove the possibility that multiple infections could produce more inflammation than a singular infection, leading to elevated or even synergistically higher risk of disease. If this is the case, cumulative infectious disease burden could better predict mental disease risk as opposed to infection by any singular disease.

Research Objective

In this thesis, the association between cumulative infectious disease burden and neuropsychiatric disorders is explored. This research is based off of the evidence that inflammation caused by infectious agents interacts and can adversely affect normal brain function. The association will be investigated through two different analysis:

1. An analysis describing the association of cumulative disease burden with the risk of stroke, with an emphasis in interaction between infections and traditional stroke risk factors.
2. An analysis describing the association of cumulative infectious disease burden with the risk of specific mood disorders.

Through these analysis, this research will attempt to describe in detail the association between cumulative infectious disease burden and their role in stroke and mental disorders.

Methods

Data

The dataset used is the open-source National Health and Nutrition Examination Survey III (NHANES III) provided by the Centers for Disease Control (CDC). This survey was designed by the National Center for Health Statistics to assess the health and nutrition of the United States population. The NHANES III includes demographic, socioeconomic, dietary, health-related, medical, dental, physiological and laboratory data gathered in an interview, medical examination and laboratory tests.[34]

Dataset Components

The data source used for this project will be the laboratory, examination, and survey data sets collected as components of the complete open-source NHANES III data set. All data sets were cleaned and had implausible values verified wherever possible prior to release to the public.[35]

Laboratory

Laboratory data was provided in the NHANES III Laboratory Data File. The physiological data included were collected through urine and venipuncture tests. Also included is fasting times, time of day and conditions before the venipuncture. For values below the limit of detection, a value of the limit of detection divided by the square root of two were entered.[36]

Examination

Examination Data was acquired through the NHANES III Examination Data File. This data was collected in the mobile examination center and in some cases a home examination. The method of collection was not recorded, as the mobile examination center and home examinations were assumed to be comparable. This file included data from questionnaires that gathered nutrient and diet information and data from the physical examination.[37]

Interview/Survey

Interview data was obtained from the NHANES III Household Adult Questionnaire Data File. This data set includes demographic, sampling weight, and health variables. Health variable questions included topics such as tobacco use, alcohol use, and conditions. Blood pressure was also measured during the interview as a supplement to the physical examination.[35]

Data Disclaimer

The data described above was provided by the National Center for Health Statistics (NCHS). Only the initial data was provided by NCHS. Any analyses, interpretations, or conclusions included in this thesis should be credited to the author and not to NCHS.

Study Population

The resulting data is representative of the civilian, non-institutionalized United States population in the 50 states and District of Columbia that includes interview, examination and serologic data points. Data was collected on individuals two months and older, with no upper age limit, though only those 20 and older were used in the analyses contained in this thesis.[35]

Sample Design

This was a stratified, multistage, probability cluster data collection project that was conducted from 1988-1994. The survey was conducted so that the first three years (1988-1991) and the second three years (1991-1994) each contained a national probability sample.[35] This is notable because the variables collected across these two phases were not consistent. While analysis in this thesis attempts to use variables collected across both phases whenever possible, we included supplemental analyses performed on a half sample if an important variable was only collected during one phase.

In the first stage of the design 81 primary sampling units (PSUs) that were mostly counties were selected with probability proportional to size. Thirteen of these counties were selected with certainty. The other 68 PSUs were chosen by taking two PSU from each of 34 strata. Next these locations were randomly divided into two groups (44 and 45 locations) to be surveyed in each phase.[35]

In the second stage, census blocks or other area segments were randomly chosen. The third stage randomly sampled households and group quarters, such as dormitories.

The fourth stage randomly selected individuals that lived in the selected residencies. After the sampling was completed the survey included observations on around 40,000 individuals.[35]

Further Information

For more details on data collection, study population and sampling design of NHANES III consult the Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94.[38]

Analysis

The data described above will be used to determine the association of multiple infectious disease burden with stroke using logistic modeling techniques. Interaction in the resulting models will be thoroughly described. The second step will be another set of logistic models looking at the association of multiple infectious disease burden with mood disorders recorded in NHANES. Several facets of this work will be carried-out in conjunction with and advised by Drs. Deanna Kruszon-Moran (CDC/OPHSS/NCHS) and Jeffrey L. Jones (CDC/CGH/DPDM).

Building the Data Set

The data set used is a combination of the data sets described above (laboratory, examination, and interview/survey from the NHANES III). Each of these data sources were imported from the open-source files. After importation, the laboratory, examination, and interview/survey data were merged by the unique subject identifier consistent for study individuals across all NHANES III data sets.

The data sets described above were imported, merged and cleaned using SAS 9.4 (Research Triangle Institute, Research Triangle Park, NC, USA) programming. For all analyses described below, the full merged and cleaned dataset and SAS-callable SUDAAN 11.0 (Research Triangle Institute, Research Triangle Park, NC, USA) programming was used.

Variables

Outcome and Exposure

The outcome of stroke was obtained from the Interview/Survey data set and was therefore self-reported. ("Has a doctor ever told you that you have had a stroke?") The stroke outcome was coded as binary with the categories defined as participants who have had a stroke versus participants with no history of stroke.

Mood disorder outcomes come from the Examination data set, and therefore were physician diagnosed using DSM-III criteria. Major depression was coded as a binary outcome with the categories defined as participants who are experiencing a non-severe or severe major depressive episode versus those who are not experiencing this. A subset of depression, referred to as major depression with severity, was created to examine only those with severe cases of depression. Major depression with severity was coded as a binary outcome with the categories defined as participants who are experiencing a severe major depressive episode versus those who are not experiencing this including those with non-severe depression. Both major depression and major depression with severity excluded those individuals in a period of bereavement. Dysthymic disorder was coded as binary with the categories defined as those with dysthymic disorder versus those without the disorder. Dysthymic disorder with major depression was coded as binary with the categories defined as those with both dysthymia and a depressive episode compared to those with either one or neither of these. Bipolar Type I was coded as a binary variable with the categories defined as those who experience both manic and depressive episodes who meet DSM-III severity and exclusion criteria, those who only experience mania meeting DSM-III severity and exclusion criteria, those who do not experience mania or

depression but meet DSM-III severity and exclusion criteria versus all others. Atypical bipolar was coded as a binary variable with the categories defined as those who are experiencing a non-severe or severe depressive episode, no manic episodes, and some symptoms of mania including euphoria or irritability.

The cumulative infectious disease burden exposure was formulated using data from the laboratory data set and therefore came from serologic tests. These serologic tests collected on the full sample included hepatitis A, B, and C viruses, Cytomegalovirus, *Toxoplama gondii*, and *Toxocara Spp*., and Varicella. Among those who had stroke, varicella and hepatitis C produced small numbers for presence or absence of infection. There were 0 Varicella seronegative and 11 Hepatitis C seropositive participants who experienced a stroke. Similar patterns were found in all mood disorder outcomes. In order to avoid unstable estimates we excluded these diseases from our analysis.

Using the five remaining infections, Hepatitis A and B, *Toxoplama gondii*, *Toxocara Spp*., and Cytomegalovirus, a cumulative burden measure was calculated for each participant. This measure was categorized as an ordinal categorical variable with the categories defined as 0-1, 2, 3, or 4-5 positive serological tests to our five infectious agents. The outer ends of the distribution were combined to avoid small numbers among those with either 0 or 5 positive serological tests.

Further information on serologic testing for our infectious agents can be found elsewhere.[39-43]

Covariates

Sociodemographic variables included age in years (categorized into 20-59, 60-69, and 70+ years age groups), gender, race/ethnicity (self-reported as non-Hispanic white, non-Hispanic black, Mexican American or other), metropolitan residence (categorized as yes/no), foreign birth (categorized as yes/no), poverty (categorized as at or above poverty line and below poverty line using family income divided by the U.S. Department of Health and Human Services' guidelines poverty threshold specific for family size) and education level (categorized as less than high school and high school completed and beyond).

Risk behavior variables included were smoking (categorized as ever versus never smoked) and alcohol (categorized as past, current or never used). Health status variables included were history of diabetes diagnosis (categorized as yes/no), hypertension and hypertension medication usage (categorized as no hypertension or medication use, hypertension and no medication use, or hypertension and medication use, where a diagnosis refers to a previous diagnosis with hypertension or having a mean systolic blood pressure greater than 150 or a mean diastolic pressure greater than or equal to 90), total cholesterol (categorized as above or equal to 240 versus below 240), HDL (categorized as above 40 versus 40 and below), C-Reactive protein (categorized as less than 0.21 mg/dL, between 0.22 and 0.99 mg/dL, and greater than 1 mg/dL), bilirubin (categorized into tertiles of as less than 0.5 mg/dL, between 0.5 and 0.6 mg/dL, and greater than 0.7 mg/dL) and triglycerides. Triglycerides were only measured on a random half-sample who fasted prior to the blood draw. Because of this, we conducted any analysis that includes triglycerides separately.

Descriptive statistics

Demographics and cofactor distributions were produced for each outcome sample. These are represented in Table 1-7. This was repeated for each outcome because the different data requirements for each analysis produced varying sample sizes. These tables show the size of each subgroup from the sample that belongs to each demographic/covariate category along with the prevalence of the outcome of that subgroup by percentages. Also included are the significance levels of prevalence comparisons between the subgroups, indicating whether the subgroups are significantly different from each other in regards to the outcome.

All prevalence percentages are weighted to represent the total uninstitutionalized US population. Tables 1-7 present age-adjusted estimates. For all tables, the Taylor Series Linearization method was used to calculate standard errors and the exact binomial method was used to calculate 95 percent confidence intervals. T-statistic P-values presented were calculated from linear orthogonal contrast procedures in SUDAAN.

Stroke

Modeling of Stroke Incidence

The first analysis of this thesis will examine a DAG and interaction effects that were found in a previously done analysis. This work was done by Deanna Kruszon-Moran, Brad Pearce and Jeffrey Jones with involvement by this thesis' author, Anna Bratcher. This previous analysis examined the relationship of cumulative infectious disease burden and stroke in NHANES III. A manuscript is currently in preparation for publication.

In this previously done analysis, a univariate relationship between each of the cofactors, including each individual infectious disease and the cumulative infectious disease burden indicator, and stroke was examined. Then a univariate relationship between each of the cofactors and the exposure, cumulative infectious disease burden was examined. The combination of these two univariate analyses helped identify possible important cofactors to be included in logistic regression.

After important cofactors were identified the group built a final model in two different ways. In the first method, the researchers added in cofactors stepwise and deleted the insignificant variables. The groups of cofactors added were demographic cofactors (age, gender, race and Hispanic origin, foreign birth, metropolitan residence, poverty index, education, smoking, and alcohol use), metabolic conditions (diabetes and obesity), cardiovascular disease indicators (high total cholesterol, low HDL, and hypertension/high blood pressure diagnosis and medication use) and metabolic activity indicators (high CRP and bilirubin). This stepwise approach was used to observe the cumulative disease burden's effect at each level, given that increased level provided an increased probability that the cofactor would be in the pathway.

The second method of modeling used was a simplified backwards elimination technique. Also for this model, we stratified the model by age, due to results in the stepwise modelling which indicated that different cofactors were more important among various age groups.

Note: All analyses in the section above, labeled "Previous Research", were performed by Deanna Kruszon-Moran, Brad Pearce, Jeffrey Jones, and Anna Bratcher.

DAG

To inform the analysis performed above, a directed acyclic graph (DAG) was produced to visually represent possible underlying causal relationships that might be at play. This DAG is presented as Figure 1. All relationships shown are based on previous literature.

Interaction Assessment

Given the results to the simplified backwards elimination technique an interaction assessment was performed to further examine the role of age among the cumulative infectious disease burden and stroke relationship. For this assessment, the ORs for the cumulative infectious disease variable for the simplified backwards elimination model and each of the stepwise models when stratified by age group were compared. Figure 2 visually illustrates the observed effects.

Mood Disorder Outcomes

As done in the previous research, univariate analyses were conducted to examine the association of each covariate with the exposure and the outcome. The mood disorder outcomes of interest here were major depression, severe major depression, dysthymia, dysthymia and major depression, bipolar disorder and atypical bipolar disorder. The covariates used in this analysis are a subset of those looked at in the stroke analysis, each of which were previously identified to be informative of the infectious disease and mood disorder relationship[24]. This abbreviated list consisted of sociodemographic and behavioral variables along with C reactive protein levels. The association with the

outcome is presented along with the descriptive statistics in Tables 2-7. Associations between the exposure and the covariates are summarized in Table 12. For these univariate analyses, t-statistics from a linear orthogonal contrast procedure were used to determine significance, where $\alpha=0.05$, of the associations. There was no correction made for multiple comparisons.

Once univariate relationships with covariates of interest were described, multivariate relationships were examined using logistic regression models. For each mood disorder outcome of interest a simplified backwards stepwise procedure was used to identify models with only significant covariates. As with the univariate analysis, we used a significance level of $\alpha=0.05$ for retention of covariates. Cumulative infectious disease burden was forced into each model to obtain an estimate effect in spite of significance.

Results

Descriptive statistics

Stroke

For the stroke analysis we included individuals over 19 with stroke and infectious disease data. These criteria brought our complete sample size to 13,904 participants. This sample was largely in the 20-59 years age group (67.5%), and less so in the 60-69 year and 70+ year age groups (14.6% and 17.8% respectively). They were also predominantly white (41.9%), with substantial proportions of black (27.2%) and Mexican Americans (27.0%). This sample was 51.9% female. Other distributions of covariates can be seen in Table 1.

Mood disorders

For the mood disorder analyses we included individuals over 19 with mood disorder and infectious disease data. These criteria brought our complete sample size to 6,515 participants. This sample size was limited by the collection of mood disorders which was only done on those under 40 years of age. This sample had a nearly even split of age groups with 49.6% in the 20-29 years age group and 50.4% in the 30-39 years age group. They were also evenly spread across race/ethnicity categories with 30.9% white, 31.7% black and 33.2% Mexican Americans. This sample was 53.4% female. Other distributions of covariates can be seen in Table 2-7. These distributions are prior to being weighted. After using the provided weights, the sample should be representative of the entire US non-institutionalized population.

Stroke Analysis

Previous Research

In the previous analysis Deanna Kruszon-Moran, Brad Pearce, Jeffrey Jones and Anna Bratcher (this thesis' author) ultimately found that the relationship between cumulative infectious disease burden and stroke was significant. These results can be seen in Tables 1 and 8-11. This conclusion was arrived at given the results detailed below.

The univariate relationships between each cofactor and stroke, which are displayed in Table 1, the only cofactors to not have a significant relationship with stroke were gender, foreign birth, metropolitan residence and obesity. In Table 8, which shows the relationship between cumulative infectious disease burden and the cofactors, the only non-significant cofactors were gender, metropolitan residence and smoking. Given these results we would expect that age, race and Hispanic origin, poverty, education, alcohol use, diabetes, high total cholesterol, low HDL, triglycerides, hypertension/hypertension medication, high CRP and bilirubin would all act as important cofactors in the model.

In the stepwise modeling, cumulative infectious disease burden was retained in every step as the exposure of interest. A summary of the stepwise procedure can be seen in Table 9. The significant cofactors in the demographic step were age, race and Hispanic origin, education, smoking and alcohol use. In the metabolic conditions step, only diabetes was significant. Low HDL, Hypertension and high blood pressure medication use were kept in the model after looking at cardiovascular disease indicators. Lastly, the significant metabolic activity indicators that remained significant were High CRP and

Bilirubin level. Cumulative infectious disease burden remained significant at each step (P-Values at each step: 0.034, 0.049, 0.009, and 0.008 respectively).

In the simple backwards stepwise procedure, the final model consisted of age, sex, education, smoking, diabetes, total cholesterol, HDL, hypertension and hypertension medication use, CRP, bilirubin and cumulative infectious disease burden. Across all ages, cumulative infectious disease burden was significant (OR: 1.37, CI: 1.09, 1.73). Furthermore, when stratified by age group, the effect of cumulative infectious disease burden was stronger (OR: 1.96, CI: 1.24, 3.11) among the youngest age group. Given this result a more in depth interaction was performed as part of this thesis. A summary of the models from the simple backwards stepwise procedure can be seen in Table 10.

Note: All analyses in the section above, labeled "Previous Research", were performed by Deanna Kruszon-Moran, Brad Pearce, Jeffrey Jones, and Anna Bratcher. Any misrepresentations of their analysis are entirely my fault.

DAG

In the DAG shown, possible causal relationships between cumulative infectious disease burden (CIDB), stroke and the cofactors used in the previous research described above were visually graphed. If we break down this DAG by the stepwise modeling used to model this relationship between CIDB and stroke, we see that the groups of covariates have increasing complexity in their relationships. This DAG illustrates how the social and demographic cofactors are expected to act as classical confounders as indicated by the backdoor path between CIDB and stoke through them. From this step on, the cofactors are no longer classical confounders since they have relationships to other covariates beyond the confounding triangle seen with the demographic cofactors. In the

second step, diabetes and obesity are both indicated to be confounders as covariates along biasing paths, and therefore should be controlled for. In the third step, all variables are the same except, blood pressure and blood pressure medication, which were combined into one variable in the quantitative analysis. In the fourth step, both bilirubin and CRP are indicated to be intermediate covariates in the CIDB and stroke relationship.

Given the complexity of the relationships shown, these arrows should not be seen as definitive. Each of the arrows shown in the "Complex" portion of the DAG, though based on the literature, may not be accurate. This uncertainty is reflected in the choice to use steps in the modelling technique, with each added step including a higher level of uncertainty as to the appropriateness of the variables included. The DAG made and presented by this thesis' author, though not included in the original paper, helped inform this decision to use stepwise modelling.

Interaction

Given the wide range of effect across age levels seen in the backwards stepwise model, an interaction assessment was performed on each step of the stepwise and simple backwards models. Figure 2 shows the ORs for cumulative infectious disease burden at each age group level for each step of our modeling procedure.

These figures show a general trend of cumulative infectious disease burden having the greatest effect in the lowest age group (20-59 years). All of the ORs produced for this age group were significant (P-Values: 0.02, 0.03, <0.01, <0.01, and <0.01 respective to stepwise order followed by simple backwards stepwise model). Though the other two age groups both show non-significant increases in OR for every increase in

cumulative infectious disease burden (P-value >0.05 for all steps with both age groups), the middle age group (60-69 years) shows an elevated OR compared to the highest age group (70+ years). This produces an interaction trend where cumulative infectious disease burden losses its effect as age increases.

Mental Health Analysis

Univariate relationships between each outcome and the cofactors are summarized in Tables 2-7. Age, race/ethnicity, gender, foreign birth, and cumulative infectious disease burden were all significantly associated with major depression. The same cofactors were significantly associated with severe major depression. For dysthymia, age, gender, poverty, education, and smoking all have significant associations with the outcome. Dysthymia with major depression was significantly associated with age, gender, poverty, and smoking. Race/ethnicity, foreign birth, and poverty are significantly associated with bipolar disorder. Only foreign birth and smoking are associated with atypical bipolar disease.

The most notable of these significant univariate associations is the inverse relation of cumulative infectious disease burden and major depression. When only adjusted for age, it appears that the more diseases a participant is seropositive for, the less likely that participant is to experience major depression (P-value= 0.03). This direction of association is opposite of our hypothesis that a higher CIDB will raise your risk for mood disorders, including major depression.

Univariate relationships between each cofactor and the exposure are displayed in Table 12. Age, race/ethnicity, foreign birth, poverty, and education all had significant associations with cumulative infectious disease burden.

Multivariate modeling results are displayed in Table 13. The models described in this table are the result of the simplified backwards stepwise procedure performed on the whole set of cofactors in relation to each mood disorder outcome with the cumulative infectious disease burden variable forced into the model at each step. CIDB was not significant in any of the simplified models.

CIDB was forced into all of the simplified model to observe the main effect of interest. In addition to CIDB, varying sets of cofactors were retained in each model. For the model of major depression, age, race, sex, poverty and smoking were retained. For this reduced model, those with depression were significantly older and richer, and were more often white, female and smokers. The same set of cofactors identifying the same risk factors was retained for severe major depression. Age, sex, poverty, education and smoking were retained in the model of dysthymia. Those with dysthymia were significantly older, richer and more educated, and were more often female and smokers. Age, sex, poverty and smoking were retained in the model of dysthymia with major depression. Those who had dysthymia with major depression were significantly older and richer, and were more often female and smokers. Only poverty was retained in the model for bipolar disorder, showing that those with bipolar disorder were less poor. Only foreign birth and smoking were retained in the model of atypical bipolar disorder, indicating those with atypical bipolar disorder were more often US born and smokers.

Discussion

Results Summary

Given that both stroke and mood disorders affect a large number of American each year, research into the various causes of these conditions is beneficial. The research presented in this thesis adds to the body of knowledge on these conditions, particularly how they are related to infectious disease. This research builds off of the literature that suggests inflammation from infectious diseases effects the brain and its function.[7,33] In particular, work by Brad Pearce, Deanna Kruszon-Moran, Jeffrey Jones and Anna Bratcher has shown that stroke is significantly associated with cumulative infectious disease burden in NHANES III.

While previous research shows that cumulative infectious disease burden is associated with stroke, this thesis further examines this relationship's interaction with age. The results presented above show that this association between CIDB and stroke is more pronounced in younger ages. This suggests that the mechanism through which infections cause neurological insults, suspected here to be inflammation, may be more prominent in stroke victims of younger ages.

In contrast to these results, no significant association of CIDB with mood disorders was found in NHANES III. Additionally, no interaction effects were found. This suggests that CIDB may not be associated with mood disorders. Therefore, unlike the analysis with stroke, we did not find any evidence of mood disorders being caused by our suspected mechanism of inflammation from infectious diseases.

However, more research must be done to fill remaining gaps in knowledge. The research here suggest that CIDB may be causing stroke, particularly in younger age

groups. Future research is needed to assess the causality behind the observed associations. Additionally, though we did not find evidence of association in CIDB and mood disorders, future research may be able to verify this null association, especially if the limitations of our study are addressed.

Implications

The results alone of CIDB being associated with stroke and not mood disorders are informative and important in the future treatment of neuropsychiatric disease. However, if we were to look further into this analysis, there are certain implications made by the final models that may speak to the mechanisms behind these two relationships, significant and non-significant. Specifically, the final models obtained for stroke and each of the mood disorder outcomes varied in the confounders that were retained in the model. The presence and absence of each confounder may speak to the causality underlying the observed relationships.

Though the NHANES III was a cross-sectional study and causal conclusions cannot be drawn here, the retaining of different confounders tells us that different factors are important, either causally or associative. Perhaps most informative is how CRP was retained in the various models. Previous research has analyzed CRP as a possible cause in stroke and depression. In one meta-analysis[44], CRP was found to be associated with stroke. This meta-analysis also attempted to describe the association between CRP and depression, but was unable to draw conclusions because of inconsistency in the literature.

As an inflammation marker, CRP is important to our analysis in that it addresses our suspected mechanism, which is that these neuropsychiatric disorders could be due to

inflammation. Examining CRP behaves in our analysis allows us to discern how inflammation might be associated with CIDB and both stroke and mood disorders. Also, unlike other biochemical markers included only in the stroke analysis, we included CRP in the initial full models for all outcomes which allows us to compare its effects across neuropsychiatric disorders.

Given that CRP was retained in both final models for stroke (stepwise and simple backwards stepwise) and none of the mood disorder outcomes, this analysis may inform our mechanism hypothesis. The absence of a significant relationship involving CRP in the mood disorder analyses tells us that inflammation may not be important here. On the other hand, the inclusion of CRP as significant in the stroke analysis tells us that inflammation may be involved in the relationship of CIDB with stroke. However, CIDB's retention along with CRP indicates that CRP is not in the pathway between CIDB and stroke. If CRP were in the pathway, controlling for it would eliminate CIDB's effect in the model.

These implications, along with the fact that we found an association between CIDB and stroke and not CIDB and mood disorders, give evidence that inflammation from multiple infections is important in etiology of stroke and not mental illness. Since these are merely implications, future research is needed to explore the mechanisms and causal relationships that are occurring here before definite conclusions can be drawn.

Stroke and Depression

As mentioned in the introduction, there is a level of interplay between the two neuropsychiatric disorders being examined here. In particular, there is a large literature

regarding post-stroke depression. Given the previous research[10], we suspect that the above analysis could be bolstered by an analysis including stroke and depression together. Unfortunately, we were not able to perform this analysis due to sample collection in NHANES III and low prevalence of stroke in younger age groups. Since mood disorder data was only collected on those younger than 39 years of age and strokes were a rare outcome in these ages, any analysis looking at stroke and depression in this data set would most likely lack the power needed to draw conclusions. Future research should make an effort to address this remaining question.

Strengths and Limitations

As with any epidemiological work, there are strengths and weaknesses to the work presented here. These strengths and weaknesses both stem from the study design of the third National Health and Nutrition Examination Survey.

The first weakness, as mentioned above, is that this survey is cross-sectional. Since cross-sectional data gives us a screenshot of what is occurring in the population, we cannot assume any observed exposure was incident before an observed disease. This limits us from making causal conclusions. Therefore, all significant association we observe are just that, associations.

The second weakness comes from the NHANES III's method of obtaining variables for the diseases used to build the cumulative infectious disease variable. These variables (Hepatitis A, hepatitis B, cytomegalovirus, *toxoplasma gondii*, and *toxocara*) were determined by the presence antibodies in the participants' serum. The presence of these antibodies indicate a past infection with the disease in question. However, the time

that the person was infected was not assessed. Therefore a participant could have been infected recently or even decades in the past. This is particularly important to keep in mind when looking at the age interactions, given that these variables are more likely to represent recent infections in the younger age groups.

In addition to weaknesses, there are particular strengths to the study design. The first of these strengths is the representative nature of the study design. When analyzed correctly-as done here- this sample represents the entire non-institutionalized population of the United States that is above two months of age. This representation gives us the ability to apply our conclusions to a wide population and makes us confident that the associations we have found are widespread.

The second strength is the sample size of the analysis. In our two analysis groups, stroke and mood disorder outcomes, the sample sizes were all in the thousands (13,904 and 6,515 individuals, respectively). This sample size allowed us to run thorough analysis that were able to contain many covariates that we hypothesized were important without over-parameterizing our models.

Impact on Public Health

This thesis impacts public health by adding to the current knowledge on infectious causes of stroke and mood disorders. Each analysis presented here builds of past research and draws conclusions that can inform healthcare practices.

For the stroke-related analyses, this thesis complements the work done by Brad Pearce, Deanna Kruszon-Moran, Jeffrey Jones and Anna Bratcher (this thesis' author) on cumulative infectious disease burden and stroke incidence. The interaction analysis

further informs this relationship in NHANES III, indicating that the association found between cumulative infectious disease burden and stroke might be more important among younger patients. The DAG also supports the previous research by illustrating the complex relationships at play. The mood disorder analyses shift focus to other ways cumulative infectious disease burden can impact health. Though no significant relationship was found, this work builds off of previous suggestions that inflammation may contribute to mood disorders[32,33]. By failing to show any significant associations, this thesis added to the knowledge of which sources of inflammation may or may not contribute to mood disorders.

All of these analyses assess the answer to original public health questions that have an impact of public health. Both the treatment and prevention of stroke and mood disorders benefit from the knowledge presented here, as we have identified the association, and possible risks, of cumulative infectious disease burden in relation to these diseases.

Conclusions

In conclusion, this thesis has presented analyses that speak to the two objectives detailed in the introduction. For the first objective, an analysis detailing cumulative infectious disease burdens role in stroke was carried out. In this analysis we found that our CIDB variable, looking at viral and parasitic infectious disease, was a significant predictor of stroke when assessed in various ways. Furthermore, CIDB showed a significant interaction with age, suggesting that cumulative infectious disease burden is especially important to stroke incidence in those between 20 and 59 years of age. These

findings have produced strong evidence for a relationship between CIDB and stroke that previous literature has failed to demonstrate.

In addition to showing that CIDB is meaningful, this thesis has also demonstrated that this variable is specific. In the analysis of the second objective, we found that CIDB was not related to any of the mood disorders assessed. This helps strengthen our conclusions drawn in the first objective by showing that our CIDB variable is specific in nature, and will not always produce significant results when assessed.

Given the evidence presented by this thesis, we can conclude that CIDB, as an epidemiologically specific representation of cumulative viral and parasitic disease burden, is important in stroke incidence, particularly those in between 20 and 59 years of age.

References

1. U. S. Department of Health and Human Services. *Mental Health: A Report of the Surgeon General—Executive Summary.* Rockville, MD: National Insitute of Mental Health;1999.
2. Go AS, Mozaffarian D, Roger VL, et al. Executive summary: heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation.* 2013;127(1):143-152.
3. Colona K. The Financial Burden of Stoke. 2014. Accessed September 30, 2014.
4. Susser E, Schwartz S, Morabia A, Bromet EJ. *Psychiatric epidemiology: searching for the causes of mental disorders.* Oxford University Press; 2006.
5. Pearce BP. Viruses and Psychiatric Disorders. *Neuroimmunological Basis of Behavior and Mental Disease*. London: Springer New York; 2088:383-401.
6. Laothamatas J, Hemachudha T, Mitrabhakdi E, Wannakrairot P, Tulayadaechanont S. MR imaging in human rabies. *American journal of neuroradiology.* 2003;24(6):1102-1109.
7. Fugate JE, Lyons JL, Thakur KT, Smith BR, Hedley-Whyte ET, Mateen FJ. Infectious causes of stroke. *The Lancet Infectious Diseases.* 2014;14(9):869-880.
8. Emsley HCA, Hopkins SJ. Acute ischaemic stroke and infection: recent and emerging concepts. *The Lancet Neurology.* 2008;7(4):341-353.
9. Bilbo SD, Yirmiya R, Amat J, Paul ED, Watkins LR, Maier SF. Bacterial infection early in life protects against stressor-induced depressive-like symptoms in adult rats. *Psychoneuroendocrinology.* 2008;33(3):261-269.
10. Robinson RG, Kubos KL, Starr L, Rao K, Price T. Mood disorders in stroke patients. *Brain.* 1984;107(Pt 1):81-93.
11. Robinson RG, Starr LB, Kubos KL, Price T. A two-year longitudinal study of post-stroke mood disorders: findings during the initial evaluation. *Stroke.* 1983;14(5):736-741.
12. Feinstone SM, Kapikian AZ, Purcell RH. Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science.* 1973;182(4116):1026-1028.
13. Cuthbert JA. Hepatitis A: old and new. *Clinical Microbiology Reviews.* 2001;14(1):38-58.
14. Tong MJ, El-Farra NS, Grew MI. Clinical manifestations of hepatitis A: recent experience in a community teaching hospital. *Journal of Infectious Diseases.* 1995;171(Supplement 1):S15-S18.
15. Chen DS. Toward elimination and eradication of hepatitis B. *Journal of gastroenterology and hepatology.* 2010;25(1):19-25.
16. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok ASF. Management of hepatitis B: summary of a clinical research workshop. *Hepatology.* 2007;45(4):1056-1075.
17. Pavić S, Švirtlih N, Delić D, Radovanović-Špurnić A. Depressive symptoms and cognitive dysfunctions in patients with chronic hepatitis B. *Srpski arhiv za celokupno lekarstvo.* 2015;143(1-2):6-11.
18. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Reviews in medical virology.* 2010;20(4):202-213.
19. de Jong MD, Galasso GJ, Gazzard B, et al. Summary of the II international symposium on cytomegalovirus. *Antiviral research.* 1998;39(3):141-162.
20. Dollard SC, Grosse SD, Ross DS. New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection. *Reviews in medical virology.* 2007;17(5):355-363.

21. Said G, Lacroix C, Chemouilli P, et al. Cytomegalovirus neuropathy in acquired immunodeficiency syndrome: a clinical and pathological study. *Annals of neurology.* 1991;29(2):139-146.
22. Tenter AM, Heckeroth AR, Weiss LM. Toxoplasma gondii: from animals to humans. *International journal for parasitology.* 2000;30(12):1217-1258.
23. Flegr J, Zitková Š, Kodym P, Frynta D. Induction of changes in human behaviour by the parasitic protozoan Toxoplasma gondii. *Parasitology.* 1996;113(01):49-54.
24. Pearce BD, Kruszon-Moran D, Jones JL. The relationship between Toxoplasma gondii infection and mood disorders in the third National Health and Nutrition Survey. *Biological psychiatry.* 2012;72(4):290-295.
25. Mortensen PB, Nørgaard-Pedersen B, Waltoft BL, Sørensen TL, Hougaard D, Yolken RH. Early infections of Toxoplasma gondii and the later development of schizophrenia. *Schizophrenia bulletin.* 2007;33(3):741-744.
26. de Almeida Carvalho EA, Rocha RL. Visceral Larva Migrans Syndromes Associated with Toxocariasis: Epidemiology, Clinical and Laboratory Aspects of Human Toxocariasis. *Current Tropical Medicine Reports.* 2014;1(1):74-79.
27. Overgaauw PAM, Nederland V. Aspects of Toxocara epidemiology: human toxocarosis. *Critical reviews in microbiology.* 1997;23(3):215-231.
28. Emsley HC, Hopkins SJ. Acute ischaemic stroke and infection: recent and emerging concepts. *The Lancet. Neurology.* 2008;7(4):341-353.
29. Chaplin DD. 1. Overview of the human immune response. *Journal of Allergy and Clinical Immunology.* 2006;117(2):S430-S435.
30. Cox FEG. Concomitant infections, parasites and immune responses. *Parasitology.* 2001;122(S1):S23-S38.
31. Song C, Lin A, Bonaccorso S, et al. The inflammatory response system and the availability of plasma tryptophan in patients with primary sleep disorders and major depression. *Journal of affective disorders.* 1998;49(3):211-219.
32. Raison CL, Capuron L, Miller AH. Cytokines sing the blues: inflammation and the pathogenesis of depression. *Trends in immunology.* 2006;27(1):24-31.
33. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nature reviews neuroscience.* 2008;9(1):46-56.
34. Statistics NCfH. National Health and Nutrition Examination Survey, 2013–2014 Overview. 2013.
35. Statistics NCfH. NHANES III HOUSEHOLD ADULT DATA FILE DOCUMENTATION In: Services USDoHaH, ed1996.
36. Statistics NCfH. NHANES III LABORATORY DATA FILE DOCUMENTATION. In: Services USDoHaH, ed1996.
37. Statistics NCfH. NHANES III EXAMINATION DATA FILE DOCUMENTATION. In: Services USDoHaH, ed. Vol Catalog Number 762001996.
38. Statistics NCfH. *Plan and operation of the third National Health and Nutrition Examination Survey, 1988-94.* Natl Ctr for Health Statistics; 1994.
39. Jones JL, Kruszon-Moran D, Wilson M, McQuillan G, Navin T, McAuley JB. Toxoplasma gondii infection in the United States: seroprevalence and risk factors. *American journal of epidemiology.* 2001;154(4):357-365.
40. Won KY, Kruszon-Moran D, Schantz PM, Jones JL. National seroprevalence and risk factors for zoonotic Toxocara spp. infection. *The American journal of tropical medicine and hygiene.* 2008;79(4):552-557.

41. Staras SA, Dollard SC, Radford KW, Flanders WD, Pass RF, Cannon MJ. Seroprevalence of cytomegalovirus infection in the United States, 1988–1994. *Clinical Infectious Diseases.* 2006;43(9):1143-1151.
42. Bell BP, Kruszon-Moran D, Shapiro CN, Lambert SB, McQuillan GM, Margolis HS. Hepatitis A virus infection in the United States: serologic results from the Third National Health and Nutrition Examination Survey. *Vaccine.* 2005;23(50):5798-5806.
43. McQuillan GM, Coleman PJ, Kruszon-Moran D, Moyer LA, Lambert SB, Margolis HS. Prevalence of hepatitis B virus infection in the United States: the National Health and Nutrition Examination Surveys, 1976 through 1994. *American Journal of Public Health.* 1999;89(1):14-18.
44. Kuo H-K, Yen C-J, Chang C-H, Kuo C-K, Chen J-H, Sorond F. Relation of C-reactive protein to stroke, cognitive disorders, and depression in the general population: systematic review and meta-analysis. *The Lancet Neurology.* 2005;4(6):371-380.

Tables

Table 1. Age Adjusted[a] and Age Specific Prevalence (Weighted Percent) of Stroke by each individual infection, cumulative infectious disease burden (CIDB), and Specific Cofactors, among adults age 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	5,944	1.63	1.25	2.08	
	Seropositive	7,960	2.41	1.85	3.09	0.02
Hepatitis B	Seronegative	12,734	1.81	1.54	2.12	
	Seropositive	1,170	3.69	2.05	6.09	0.04
Cytomegalovirus	Seronegative	10,995	1.05	0.77	1.39	
	Seropositive	2,909	2.18	1.76	2.66	<0.01
Toxoplasma gondii	Seronegative	10,006	1.76	1.41	2.16	
	Seropositive	3,898	2.47	1.96	3.07	0.02
Toxocara	Seronegative	11,793	1.79	1.54	2.06	
	Seropositive	2,111	2.87	1.85	4.23	0.05
Cumulative Infectious Disease Burden	0-1 Diseases	4,906	1.49	1.01	2.11	
	2 Diseases	5,166	1.71	1.30	2.22	
	3 Diseases	2,905	2.75	2.11	3.51	
	4-5 Diseases	927	5.50	2.44	10.46	0.02[c]
Age	*Reference* 20-59	9,390	0.70	0.48	0.98	
	60-69	2,033	3.99	2.78	5.53	<0.01
	70+	2,481	8.35	7.07	9.79	<0.01
Race and Hispanic Origin	*Reference* non-Hispanic White	5,823	1.82	1.52	2.16	
	Non-Hispanic Black	3,788	2.59	1.90	3.43	0.05
	Mexican American	3,757	1.73	1.29	2.27	0.75
Gender	Female	7,211	1.79	1.39	2.27	
	Male	6,693	2.08	1.64	2.60	0.38
Foreign Birth	US Born	11,006	1.89	1.62	2.18	
	Born outside the US	2,861	2.27	1.16	3.99	0.55
Metropolitan Residence	No	7,196	2.09	1.61	2.67	
	Yes	6,708	1.73	1.44	2.06	0.22
Poverty Index	At or above	9,674	1.76	1.47	2.10	
	Below	2,956	3.38	2.39	4.63	<0.01
Education	Completed High School or Greater	8,194	1.51	1.16	1.92	
	Less than High School	5,624	2.99	2.35	3.75	<0.01
Smoking	Never	6,223	1.37	1.07	1.73	
	Any	7,623	2.30	1.93	2.72	<0.01
Alcohol Use	*Reference* Current	6,220	1.29	0.96	1.70	

	Past	4,902	2.59	2.00	3.29	<0.01
	Never	2,425	2.11	1.47	2.94	0.02
Obesity[1]	Not Obese	10,125	1.77	1.39	2.21	
	Obese	3,530	2.43	1.76	3.26	0.14
Diabetes	No Diagnosis	12,813	1.57	1.32	1.86	
	Diagnosis	1,073	6.39	4.24	9.19	<0.01
High Total Cholesterol[2]	Normal	10,989	1.74	1.41	2.12	
	High	2,872	2.89	2.05	3.94	0.02
Low HDL[3]	Normal	8,721	1.44	1.14	1.80	
	Low	5,043	2.77	2.17	3.47	<0.01
High Triglycerides[4]	Normal	4,269	1.22	0.88	1.66	
	High	1,874	3.21	2.06	4.74	<0.01
Hypertension/ Hypertension Medication[5]	*Reference* Neither diagnosis nor medication	8,254	1.26	0.90	1.71	
	Hypertension diagnosis with no Medication	2,959	1.88	1.25	2.69	0.13
	Hypertension diagnosis and on Hypertension Medication	2,238	4.44	3.06	6.21	<0.01
High CRP	*Reference* <0.22 mg/dL	9,060	1.50	1.22	1.83	
	0.22-0.99 mg/dL	3,452	2.17	1.54	2.96	0.08
	>=1.0 mg/dL	1,350	4.58	2.95	6.74	<0.01
Bilirubin[6]	*Reference* Lowest tertile	5,051	2.37	1.81	3.03	
	Middle tertile	4,684	1.93	1.55	2.39	0.20
	Highest tertile	4,068	1.52	1.17	1.93	0.01

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 cumulative infectious disease burden

1. Obesity – body mass index greater than or equal to 30

2. High total cholesterol – greater than or equal to 240 mg/dL

3. Low HDL – less than 40 mg/dL for Males and less than 50 for Females

4. High triglycerides – greater than or equal to 150 mg/dL

5. Hypertension/ Hypertension Medication - those with a physician diagnosis of hypertension or high blood pressure and are currently taking medication for it compared with those with a physician diagnosis of hypertension or high blood pressure or who measured greater than or equal to 140 systolic or greater than or equal to 90 diastolic blood pressure at the NHANES exam but who are NOT currently taking medication compared to those who had neither been diagnosed with hypertension/high pressure and who's blood pressure was normal at exam.

6. Lowest tertile is <0.5 mg/dL, middle tertile 0.5-0.6 and highest tertile >=0.7 mg/dL

Table 2. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Major Depression by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Major Depression by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	9.14	7.70	10.76	
Age	20-29	3,229	7.76	6.07	9.74	
	30-39	3,286	10.34	8.19	12.83	0.08
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	10.25	8.34	12.44	
	Non-Hispanic Black	2,068	6.47	5.07	8.11	0.01
	Mexican American	2,162	6.79	4.76	9.34	0.02
Gender	Female	3,482	11.88	9.73	14.31	
	Male	3,033	6.44	4.85	8.34	<0.01
Foreign Birth	US Born	4,829	9.89	8.22	11.77	
	Born outside the US	1,674	4.79	3.06	7.10	<0.01
Poverty Index	At or above	1,594	11.01	8.18	14.39	
	Below	4,442	8.97	7.28	10.90	0.28
Education	Completed High School or Greater	1,882	8.97	5.99	12.78	
	Less than High School	4,606	9.19	7.62	10.96	0.90
Smoking	Never	3,363	8.19	6.62	10.01	
	Any	3,107	9.92	7.91	12.23	0.17
High CRP	*Reference* <0.22 mg/dL	4,721	9.08	7.37	11.03	
	0.22-0.99 mg/dL	1,284	9.61	6.67	13.28	0.77
	>=1.0 mg/dL	494	10.10	7.03	13.93	0.61
Hepatitis A	Seronegative	3,946	9.99	8.11	12.13	
	Seropositive	2,569	5.87	4.06	8.17	0.01
Hepatitis B	Seronegative	6,078	9.08	7.58	10.77	
	Seropositive	437	10.15	4.00	20.28	0.79
Cytomegalovirus	Seronegative	2,145	9.79	7.52	12.48	
	Seropositive	4,370	8.43	6.98	10.07	0.33
Toxoplasma gondii	Seronegative	5,379	9.50	7.93	11.27	
	Seropositive	1,136	7.08	4.43	10.63	0.14
Toxocara	Seronegative	5,493	9.57	7.99	11.35	
	Seropositive	1,022	6.53	3.33	11.32	0.15
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	9.94	7.87	12.33	
	2 Diseases	1,992	8.20	5.26	12.06	
	3 Diseases	853	5.44	2.94	9.10	
	4-5 Diseases	232	4.77	1.13	12.64	0.03[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of major depression in each individual group to reference group within each cofactor

c P-values from linear test for trend in major depression prevalence with increasing level of cumulative infectious disease burden

Table 3. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Severe Major Depression by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Severe Major Depression by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	8.17	6.79	9.74	
Age	20-29	3,229	6.73	5.16	8.59	
	30-39	3,286	9.42	7.30	11.92	0.06
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	9.19	7.38	11.27	
	Non-Hispanic Black	2,068	6.03	4.67	7.64	0.01
	Mexican American	2,162	5.77	3.90	8.20	0.02
Gender	Female	3,482	11.03	8.87	13.49	
	Male	3,033	5.35	3.92	7.11	<0.01
Foreign Birth	US Born	4,829	8.96	7.35	10.79	
	Born outside the US	1,674	3.64	1.97	6.09	<0.01
Poverty Index	At or above	1,594	10.21	7.39	13.65	
	Below	4,442	7.94	6.33	9.80	0.22
Education	Completed High School or Greater	1,882	8.50	5.54	12.36	
	Less than High School	4,606	8.12	6.61	9.84	0.83
Smoking	Never	3,363	7.08	5.57	8.85	
	Any	3,107	9.05	7.12	11.30	0.11
High CRP	*Reference* <0.22 mg/dL	4,721	8.00	6.38	9.88	
	0.22-0.99 mg/dL	1,284	8.93	6.03	12.63	0.61
	>=1.0 mg/dL	494	9.61	6.53	13.50	0.42
Hepatitis A	Seronegative	3,946	9.01	7.21	11.08	
	Seropositive	2,569	4.95	3.23	7.22	0.01
Hepatitis B	Seronegative	6,078	8.10	6.65	9.74	
	Seropositive	437	9.45	3.44	19.75	0.73
Cytomegalovirus	Seronegative	2,145	8.62	6.45	11.22	
	Seropositive	4,370	7.69	6.24	9.34	0.50
Toxoplasma gondii	Seronegative	5,379	8.45	6.98	10.12	
	Seropositive	1,136	6.59	4.02	10.10	0.23
Toxocara	Seronegative	5,493	8.48	6.95	10.22	
	Seropositive	1,022	6.33	3.16	11.13	0.31
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	8.93	6.95	11.26	
	2 Diseases	1,992	7.13	4.33	10.95	
	3 Diseases	853	4.86	2.48	8.47	
	4-5 Diseases	232	4.47	1.04	11.93	0.05[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of severe major depression in each individual group to reference group within each cofactor

c P-values from linear test for trend in severe major depression prevalence with increasing level of cumulative infectious disease burden

Table 4. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Dysthymia by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Dysthymia by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	6.39	5.26	7.69	
Age	20-29	3,229	4.48	3.30	5.92	
	30-39	3,286	8.05	5.92	10.65	0.02
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	6.22	4.73	8.00	
	Non-Hispanic Black	2,068	8.05	6.52	9.80	0.14
	Mexican American	2,162	7.20	5.22	9.62	0.45
Gender	Female	3,482	7.95	6.13	10.11	
	Male	3,033	4.86	3.52	6.51	0.02
Foreign Birth	US Born	4,829	6.48	5.22	7.94	
	Born outside the US	1,674	5.96	3.81	8.83	0.71
Poverty Index	At or above	1,594	11.71	8.96	14.94	
	Below	4,442	5.54	4.19	7.16	<0.01
Education	Completed High School or Greater	1,882	11.39	7.87	15.76	
	Less than High School	4,606	5.35	4.16	6.76	0.01
Smoking	Never	3,363	3.86	2.77	5.22	
	Any	3,107	8.56	6.77	10.63	<0.01
High CRP	*Reference* <0.22 mg/dL	4,721	5.87	4.59	7.38	
	0.22-0.99 mg/dL	1,284	8.15	5.46	11.60	0.16
	>=1.0 mg/dL	494	9.58	5.02	16.20	0.21
Hepatitis A	Seronegative	3,946	6.43	5.06	8.05	
	Seropositive	2,569	6.24	4.40	8.56	0.89
Hepatitis B	Seronegative	6,078	6.05	4.91	7.35	
	Seropositive	437	12.20	5.56	22.28	0.13
Cytomegalovirus	Seronegative	2,145	5.70	4.07	7.74	
	Seropositive	4,370	7.15	5.75	8.77	0.21
Toxoplasma gondii	Seronegative	5,379	6.07	4.88	7.44	
	Seropositive	1,136	8.26	5.27	12.20	0.22
Toxocara	Seronegative	5,493	6.64	5.38	8.09	
	Seropositive	1,022	4.91	3.17	7.22	0.14
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	5.86	4.42	7.60	
	2 Diseases	1,992	7.34	4.89	10.51	
	3 Diseases	853	9.10	5.51	13.94	
	4-5 Diseases	232	5.73	2.25	11.69	0.25[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of dysthymia in each individual group to reference group within each cofactor

c P-values from linear test for trend in dysthymia prevalence with increasing level of cumulative infectious disease burden

Table 5. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Dysthymia with Major Depression by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Dysthymia with Major Depression by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	3.65	2.63	4.92	
Age	20-29	3,229	2.01	1.30	2.97	
	30-39	3,286	5.07	3.26	7.47	0.01
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	3.96	2.64	5.68	
	Non-Hispanic Black	2,068	3.30	2.35	4.48	0.46
	Mexican American	2,162	3.28	1.93	5.18	0.54
Gender	Female	3,482	4.63	3.27	6.35	
	Male	3,033	2.68	1.55	4.30	0.03
Foreign Birth	US Born	4,829	3.88	2.72	5.36	
	Born outside the US	1,674	2.31	0.96	4.63	0.15
Poverty Index	At or above	1,594	6.44	4.29	9.21	
	Below	4,442	3.23	2.06	4.79	0.03
Education	Completed High School or Greater	1,882	5.94	3.14	10.07	
	Less than High School	4,606	3.18	2.17	4.49	0.11
Smoking	Never	3,363	1.95	1.13	3.12	
	Any	3,107	5.08	3.45	7.18	<0.01
High CRP	*Reference* <0.22 mg/dL	4,721	3.28	2.11	4.83	
	0.22-0.99 mg/dL	1,284	5.29	3.11	8.34	0.14
	>=1.0 mg/dL	494	4.53	1.68	9.59	0.56
Hepatitis A	Seronegative	3,946	3.74	2.55	5.28	
	Seropositive	2,569	3.31	1.85	5.42	0.69
Hepatitis B	Seronegative	6,078	3.39	2.35	4.73	
	Seropositive	437	7.90	2.32	18.51	0.24
Cytomegalovirus	Seronegative	2,145	3.90	2.36	6.03	
	Seropositive	4,370	3.38	2.40	4.61	0.59
Toxoplasma gondii	Seronegative	5,379	3.45	2.38	4.81	
	Seropositive	1,136	4.79	2.43	8.36	0.37
Toxocara	Seronegative	5,493	4.08	2.91	5.54	
	Seropositive	1,022	1.07	0.53	1.92	<0.01
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	3.56	2.26	5.30	
	2 Diseases	1,992	4.18	2.07	7.43	
	3 Diseases	853	3.49	1.40	7.11	
	4-5 Diseases	232	2.58	0.23	9.92	0.83[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of dysthymia with major depression in each individual group to reference group within each cofactor

c P-values from linear test for trend in dysthymia with major depression prevalence with increasing level of cumulative infectious disease burden

Table 6. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Bipolar Disorder by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Bipolar Disorder by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	0.97	0.59	1.49	
Age	20-29	3,229	0.96	0.43	1.85	
	30-39	3,286	0.97	0.57	1.53	0.99
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	0.93	0.45	1.70	
	Non-Hispanic Black	2,068	1.91	1.21	2.87	0.07
	Mexican American	2,162	0.96	0.22	2.69	0.96
Gender	Female	3,482	0.86	0.39	1.65	
	Male	3,033	1.07	0.55	1.87	0.62
Foreign Birth	US Born	4,829	1.09	0.65	1.71	
	Born outside the US	1,674	0.24	0.09	0.49	<0.01
Poverty Index	At or above	1,594	2.24	1.33	3.53	
	Below	4,442	0.74	0.34	1.40	0.02
Education	Completed High School or Greater	1,882	1.31	0.56	2.58	
	Less than High School	4,606	0.90	0.51	1.46	0.36
Smoking	Never	3,363	0.68	0.43	1.02	
	Any	3,107	1.21	0.61	2.13	0.14
High CRP	*Reference* <0.22 mg/dL	4,721	0.89	0.44	1.59	
	0.22-0.99 mg/dL	1,284	1.38	0.52	2.96	0.45
	>=1.0 mg/dL	494	0.94	0.54	1.52	0.88
Hepatitis A	Seronegative	3,946	1.04	0.60	1.69	
	Seropositive	2,569	0.67	0.32	1.22	0.23
Hepatitis B	Seronegative	6,078	0.97	0.58	1.53	
	Seropositive	437	0.88	0.24	2.23	0.84
Cytomegalovirus	Seronegative	2,145	0.90	0.36	1.87	
	Seropositive	4,370	1.03	0.52	1.84	0.79
Toxoplasma gondii	Seronegative	5,379	0.91	0.51	1.49	
	Seropositive	1,136	1.31	0.77	2.09	0.24
Toxocara	Seronegative	5,493	0.96	0.56	1.52	
	Seropositive	1,022	1.03	0.43	2.05	0.87
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	0.96	0.46	1.77	
	2 Diseases	1,992	1.12	0.54	2.04	
	3 Diseases	853	0.86	0.33	1.80	
	4-5 Diseases	232	0.16	0.00	0.90	0.70[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of bipolar disorder in each individual group to reference group within each cofactor

c P-values from linear test for trend in bipolar disorder prevalence with increasing level of cumulative infectious disease burden

Table 7. Age Adjusted[a] and Age Specific Prevalence (in Percent) of Atypical Bipolar Disorder by Levels of All Cofactors, among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Sample Size	Prevalence of Atypical Bipolar Disorder by Level of each Cofactor	Lower 95% Confidence Interval	Upper 95% Confidence Interval	P-value T-Test [b]
Total		6,515	1.45	0.94	2.14	
Age	20-29	3,229	1.52	0.83	2.55	
	30-39	3,286	1.40	0.69	2.52	0.84
Race and Hispanic Origin	*Reference* non-Hispanic White	2,013	1.69	1.01	2.65	
	Non-Hispanic Black	2,068	1.10	0.56	1.93	0.25
	Mexican American	2,162	0.95	0.56	1.52	0.15
Gender	Female	3,482	1.31	0.72	2.19	
	Male	3,033	1.60	0.80	2.85	0.64
Foreign Birth	US Born	4,829	1.64	1.05	2.45	
	Born outside the US	1,674	0.35	0.09	0.91	<0.01
Poverty Index	At or above	1,594	1.60	0.57	3.53	
	Below	4,442	1.41	0.84	2.22	0.80
Education	Completed High School or Greater	1,882	1.07	0.43	2.19	
	Less than High School	4,606	1.54	0.92	2.41	0.43
Smoking	Never	3,363	0.65	0.25	1.34	
	Any	3,107	2.13	1.23	3.42	0.02
High CRP	*Reference* <0.22 mg/dL	4,721	1.35	0.77	2.20	
	0.22-0.99 mg/dL	1,284	2.06	0.75	4.44	0.45
	>=1.0 mg/dL	494	1.28	0.26	3.72	0.92
Hepatitis A	Seronegative	3,946	1.52	0.95	2.29	
	Seropositive	2,569	1.21	0.26	3.43	0.69
Hepatitis B	Seronegative	6,078	1.42	0.89	2.15	
	Seropositive	437	1.98	0.40	5.74	0.65
Cytomegalovirus	Seronegative	2,145	1.47	0.79	2.51	
	Seropositive	4,370	1.43	0.75	2.49	0.95
Toxoplasma gondii	Seronegative	5,379	1.59	0.98	2.43	
	Seropositive	1,136	0.67	0.15	1.87	0.09
Toxocara	Seronegative	5,493	1.63	1.06	2.40	
	Seropositive	1,022	0.39	0.01	2.06	0.01
Cumulative Infectious Disease Burden	0-1 Diseases	3,438	1.53	0.92	2.39	
	2 Diseases	1,992	1.77	0.60	4.01	
	3 Diseases	853	0.38	0.02	1.80	
	4-5 Diseases	232	0.00	0.00	97.50	0.22[c]

a Age-adjusted, except for age cofactor

b P-value from t-test comparing prevalence of atypical bipolar disorder in each individual group to reference group within each cofactor

c P-values from linear test for trend in atypical bipolar disorder prevalence with increasing level of cumulative infectious disease burden

Table 8. Summary of Regression Analyses for Variables Predicting Increasing Cumulative Infectious Disease Burden[a, 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
Gender	Female		
	Male	0.00	0.85
Foreign Birth	U.S. born		
	Born outside the U.S.	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[1]	Not Obese		
	Obese	0.09	<0.01
Diabetes	No Diagnosis		
	Diagnosis	0.19	<0.01
Total Cholesterol[2]	Normal		
	High	0.08	0.03
Low HDL[3]	Normal		
	Low	0.10	<0.01

High Triglycerides[4]	Normal		
	High	0.09	0.01
Hypertension/ Hypertension Medication[5]	*Reference* Neither diagnosis nor medication		
	Hypertension diagnosis with no hypertension medication	0.07	0.04
	Hypertension diagnosis and on hypertension medication	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[6]	*Reference* Lowest tertile		
	Middle tertile	-.07	0.02
	Highest tertile	-.11	<0.01

a. Age-adjusted, except for age cofactor

b. Cumulative infectious disease burden is a measure of burden of infection based on the number of infections for which a sample person tested positive out of 5 organisms (Toxoplasma gondii, Ttoxocarr spp, cytomegalovirus, Hepatitis A and Hepatitis B virus)

1. Obesity – body mass index greater than or equal to 30
2. High total cholesterol – greater than or equal to 240 mg/dL
3. Low HDL – less than 40 mg/dL for Males and less than 50 for Females
4. High triglycerides – greater than or equal to 150 mg/dL
5. Hypertension/ Hypertension Medication - those with a physician diagnosis of hypertension or high blood pressure and are currently taking medication for it compared with those with a physician diagnosis of hypertension or high blood pressure or who measured greater than or equal to 140 systolic or greater than or equal to 90 diastolic blood pressure at the NHANES exam but who are NOT currently taking medication compared to those who had neither been diagnosed with hypertension/high pressure and who's blood pressure was normal at exam.
6. Lowest tertile is <0.5 mg/dL, middle tertile 0.5-0.6 and highest tertile >=0.7 mg/dL

Table 9. Association of increasing Cumulative Infectious Disease Burden on Stroke Outcome from Complex Stepwise Models, among adults age 20 years and older, NHANES III.

	Cofactors[1]	CIDB Beta	P-Value	OR for CIDB	OR Lower Limit, 95% Confidence	OR Upper Limit, 95% Confidence
Model 1	**CIDB Age, Race and Hispanic origin, Education, Smoking, Alcohol Use**	0.25	0.03	1.29	1.02	1.63
Model 2	CIDB Age, Race and Hispanic origin, Education, Smoking, Alcohol Use, **Diabetes**	0.24	0.05	1.27	1.00	1.61
Model 3	CIDB Age, Education, Smoking, Diabetes, **Low HDL, Hypertension/High Blood Pressure Medication Use or Diagnosis**	0.31	0.01	1.37	1.08	1.73
Model 4	CIDB Age, Education, Smoking, Diabetes Low HDL, Hypertension/High Blood Pressure Medication Use or Diagnosis, **High CRP, Bilirubin level**.	0.32	0.01	1.38	1.09	1.74

1. Significant cofactors in model - those added are bold and those retained in model are regular font

Table 10. Final Simplified Backwards Stepwise Model of Stroke Outcome for Total Population, among adults age 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*	1.29, 4.10						
	>=70	5.55*	3.43, 8.99						
Sex	Male	1.55*	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*	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*	1.11, 2.14	1.81	0.75, 4.38	1.11	0.49, 2.49	1.72*	1.13, 2.61
	Never (Ref)	1.00		1.00		1.00		1.00	
Diabetes	Diagnosis	2.66*	1.87, 3.78	2.40	0.91, 6.31	3.38*	1.37, 8.32	2.16*	1.49, 3.14
	No Diagnosis (Ref)	1.00		1.00		1.00		1.00	
Total Cholesterol[1]	High	1.52*	1.07, 2.17	2.67*	1.19, 6.00	0.65	0.33, 1.28	1.56*	1.07, 2.28
	Normal (Ref)	1.00		1.00		1.00		1.00	
HDL[2]	Low	1.57*	1.05, 2.35	1.24	0.47, 3.32	1.66	0.75, 3.65	1.89*	1.29, 2.76
	Normal (Ref)	1.00		1.00		1.00		1.00	
Hypertension[3]	Hypertension diagnosis and on hypertension medication	2.76*	1.77, 4.29	2.77*	1.07, 7.15	1.56	0.77, 3.16	2.95*	1.78, 4.90
	Hypertension diagnosis with no medication	1.41	0.84, 2.35	1.89	0.72, 5.01	0.96	0.40, 2.29	1.27	0.81, 2.01
	Neither diagnosis nor medication (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*	1.22, 3.16	2.48	0.90, 6.84	2.26	0.90, 5.63	1.40	0.86, 2.29
Bilirubin[4]	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*	0.44, 0.86	0.27*	0.09, 0.87	0.75	0.33, 1.71	0.93	0.59, 1.48
Cumulative Infectious Disease Burden		**1.37***	**1.09, 1.73**	**1.96***	**1.24, 3.11**	**1.06**	**0.69, 1.62**	**1.12**	**0.84, 1.51**

* Significant at P<0.05; Ref=Reference group

1. High total cholesterol – greater than or equal to 240 mg/dL
2. Low HDL – less than 40 mg/dL for Males and less than 50 for Females
3. Hypertension/ Hypertension Medication - those with a physician diagnosis of hypertension or high blood pressure and are currently taking medication for it compared with those with a physician diagnosis of hypertension or high blood pressure or who measured greater than or equal to 140 systolic or greater than or equal to 90 diastolic blood pressure at the NHANES exam but who are NOT currently taking medication compared to those who had neither been diagnosed with hypertension/high pressure and who's blood pressure was normal at exam.
4. Lowest tertile is <0.5 mg/dL, middle tertile 0.5-0.6 and highest tertile >=0.7 mg/dL

Table 11. Odds ratios and p-values from logistic models adjusting for age, CIDB and each individual cofactor as well as testing for interaction of CIDB and each cofactor on stroke outcome among adults age 20 and older from NHANES III.

Cofactor	Levels of Cofactor	OR[a]	95% Confidence Interval		P-Value	P-Value for interaction of CIDB 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
CIDB		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
CIDB		1.38	1.09,	1.73	0.01	
Gender	Female					
	Male	1.16	0.81,	1.67	0.42	0.81
CIDB		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
CIDB		1.51	1.23,	1.84	<0.01	
Metropolitan Residence	No					
	Yes	0.82	0.59,	1.13	0.22	0.35
CIDB		1.49	1.22,	1.82	<0.01	
Poverty Index	At or above					
	Below	1.68	1.20,	2.36	<0.01	0.47
CIDB		1.46	1.16,	1.83	<0.01	
Education	*Reference* Completed High School or Greater					
	Less than High School	1.69	1.21,	2.37	<0.01	0.57
CIDB		1.39	1.14,	1.69	<0.01	
Smoking	Never					
	Any	1.73	1.32,	2.28	<0.01	0.87
CIDB		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
CIDB		1.46	1.18,	1.81	<0.01	
Obesity[1]	Not Obese					
	Obese	1.39	0.91,	2.13	0.12	0.16
CIDB		1.48	1.21,	1.81	<0.01	
Diabetes	No Diagnosis					
	Diagnosis	3.80	2.70,	5.35	<0.01	0.05
CIDB		1.45	1.18,	1.78	<0.01	

High Total Cholesterol[2]	Normal					
	High	1.54	1.03,	2.28	0.03	0.47
CIDB		1.48	1.22,	1.81	<0.01	
Low HDL[3]	Normal					
	Low	1.92	1.37,	2.68	<0.01	0.22
CIDB		1.47	1.20,	1.81	<0.01	
High Triglycerides[4]	Normal					
	High	2.66	1.81,	3.90	<0.01	0.29
CIDB		1.98	1.43,	2.74	<0.01	
Hypertension/ Hypertension Medication[5]	Hypertension diagnosis and on hypertension medication	3.49	2.22,	5.49	<0.01	0.02
	Hypertension diagnosis with no hypertension medication	1.55	0.91,	2.64	0.10	0.84
	Reference Neither diagnosis nor medication					
CIDB		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
CIDB		1.48	1.21,	1.81	<0.01	
Bilirubin[6]	*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
CIDB		1.48	1.21,	1.81	<0.01	

a. ORs for covariates refer to the increase in odds of stroke for each increase in level of the cofactor when interaction is not in the model; ORs for CIDB refer to the increase in odds of stroke for each increase in level of CIDB

1. Obesity – body mass index greater than or equal to 30
2. High total cholesterol – greater than or equal to 240 mg/dL
3. Low HDL – less than 40 mg/dL for Males and less than 50 for Females
4. High triglycerides – greater than or equal to 150 mg/dL
5. Hypertension/ Hypertension Medication - those with a physician diagnosis of hypertension or high blood pressure and are currently taking medication for it compared with those with a physician diagnosis of
hypertension or high blood pressure or who measured greater than or equal to 140 systolic or greater than or equal to 90 diastolic blood pressure at the NHANES exam but who are NOT currently taking medication compared to those who had neither been diagnosed with hypertension/high pressure and who's blood pressure was normal at exam.
6. Lowest tertile is <0.5 mg/dL, middle tertile 0.5-0.6 and highest tertile >=0.7 mg/dL

Figure 1. DAG Illustrating Possible Causal relationships Between Covariates, Cumulative Infectious Disease Burden (CIDB) and Stroke, among adults age 20 and older from NHANES III.

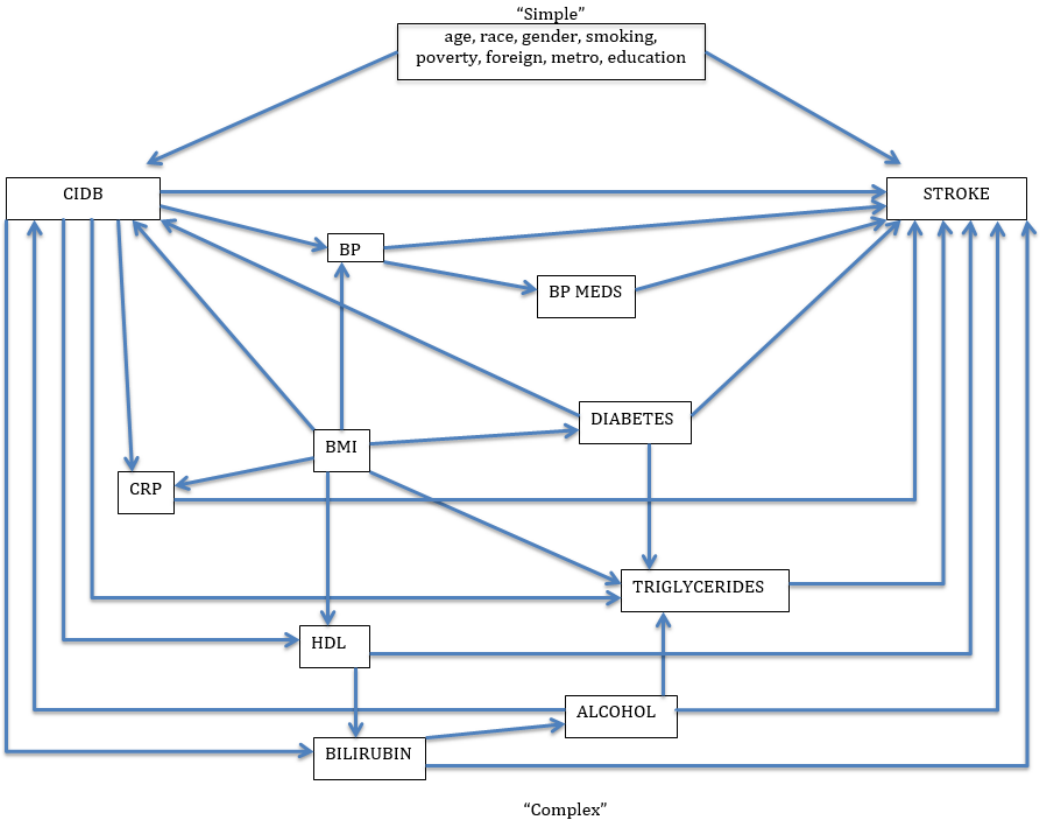
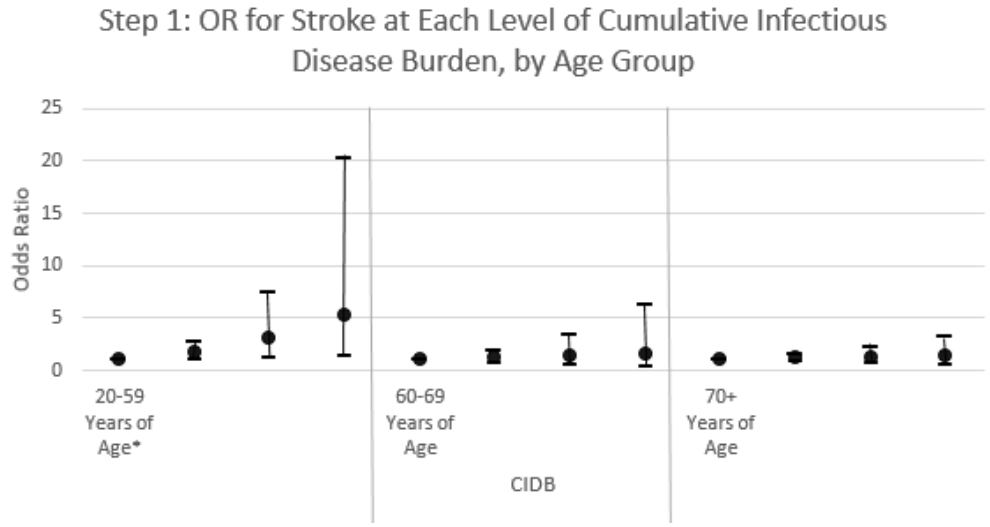
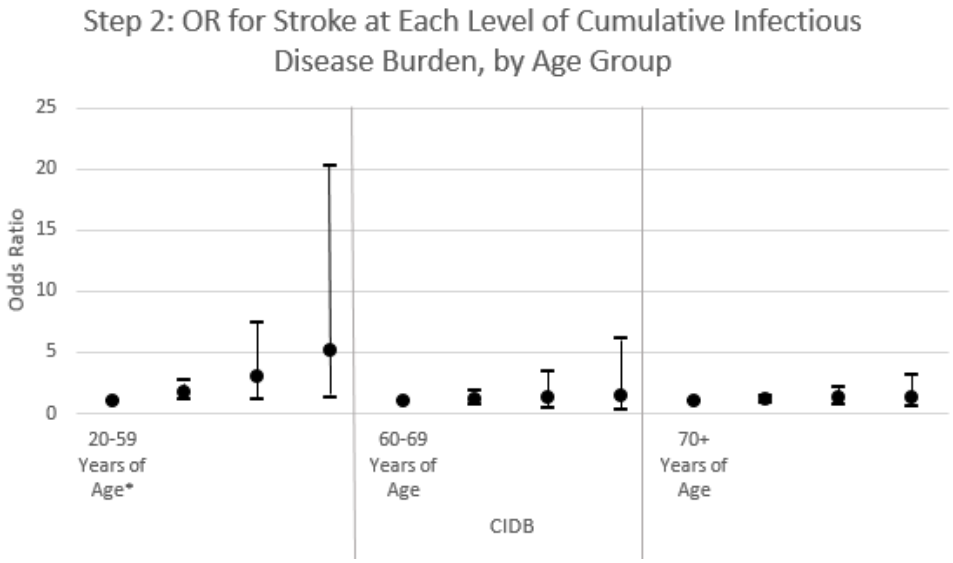


Figure 2. a-e) Graphs depicting ORs at each modeling step to illustrate interaction, among adults age 20 and older from NHANES III. ORs are the increase in odds of stroke for that level of cumulative infectious disease burden compared to the 0-1 CIDB level reference group. * Indicates significant age group ORs for the age group at α=0.05.

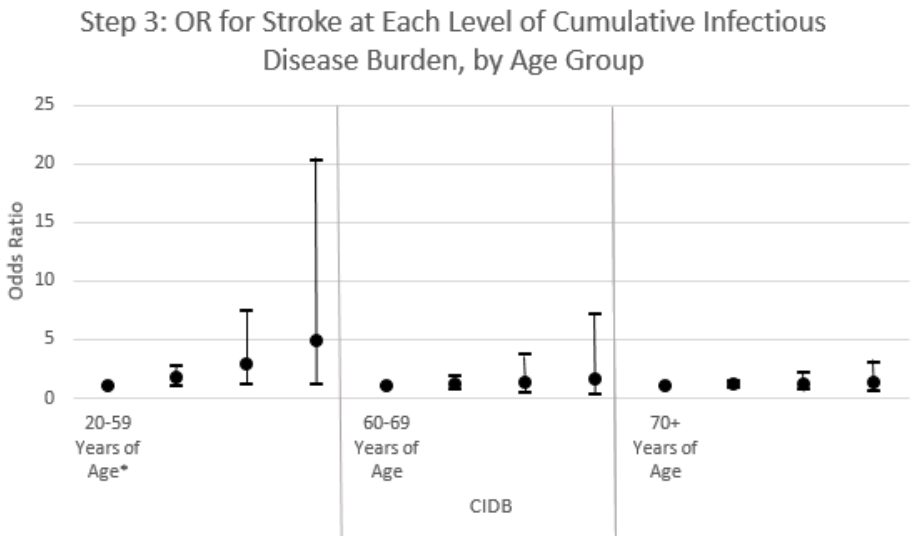
a)



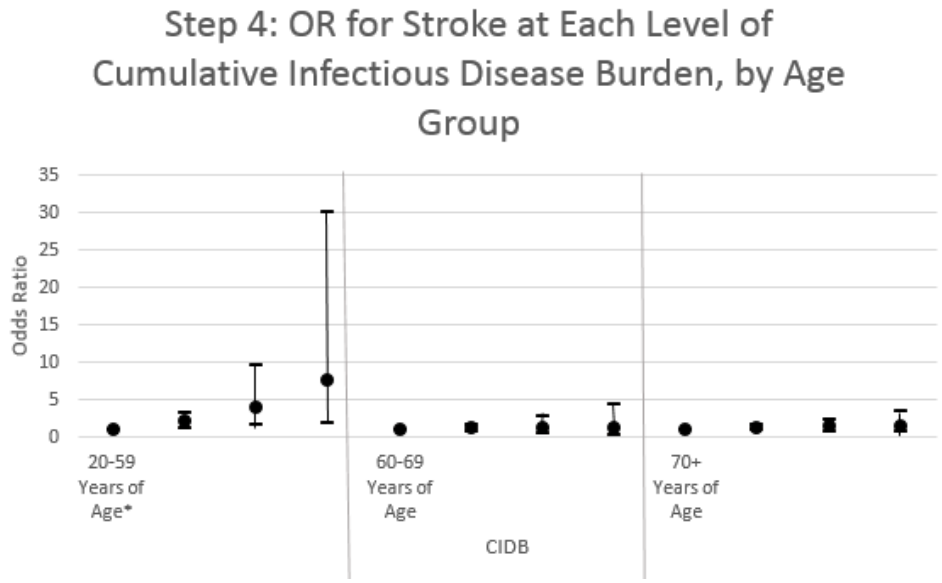
b)



c)


Step 3: OR for Stroke at Each Level of Cumulative Infectious Disease Burden, by Age Group
Odds Ratio
25
20
15
10
5
0
20-59 Years of Age*
60-69 Years of Age
70+ Years of Age
CIDB

d)


Step 4: OR for Stroke at Each Level of Cumulative Infectious Disease Burden, by Age Group
Odds Ratio
35
30
25
20
15
10
5
0
20-59 Years of Age*
60-69 Years of Age
70+ Years of Age
CIDB

e)

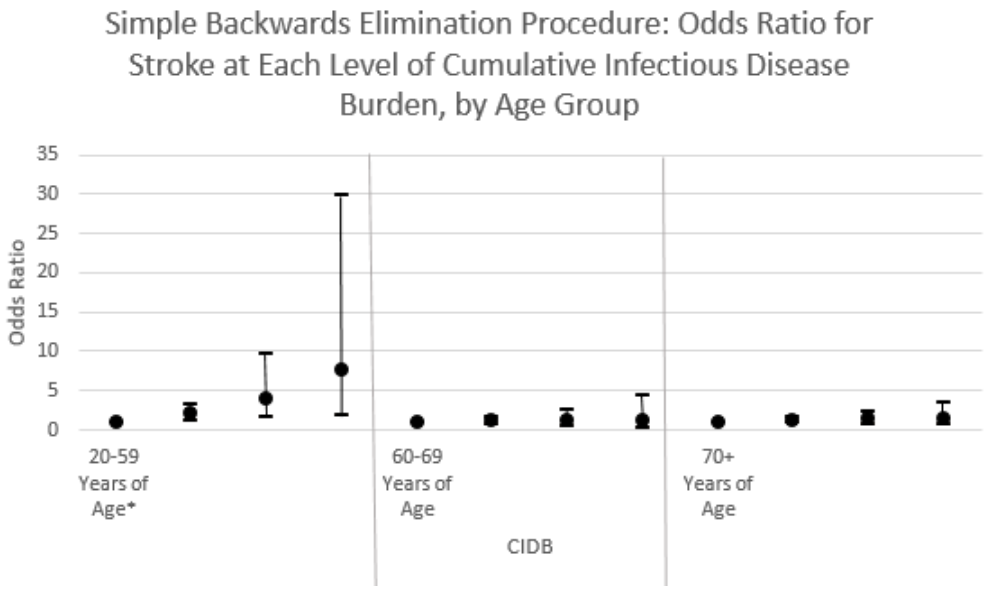

Simple Backwards Elimination Procedure: Odds Ratio for Stroke at Each Level of Cumulative Infectious Disease Burden, by Age Group
Odds Ratio
35
30
25
20
15
10
5
0
20-59 Years of Age*
60-69 Years of Age
70+ Years of Age
CIDB

Table 12. Summary of Regression Analyses for Variables Predicting Increasing Cumulative Infectious Disease Burden[a], among adults age 20 years to 39 years, NHANES III.

Cofactor	Level of Cofactor	Beta	P-value
Age	20-29		
	30-39	0.10	0.01
Race and Hispanic Origin	*Reference* Non-Hispanic White		
	Non-Hispanic Black	0.45	<0.01
	Mexican American	0.81	<0.01
Gender	Female		
	Male	0.04	0.17
Foreign Birth	U.S. born		
	Born outside the U.S.	1.00	<0.01
Poverty Index	At or above		
	Below	0.37	<0.01
Education	Completed High School or Greater		
	Less than completed High School	0.60	<0.01
Smoking	Never		
	Any	-0.05	0.23
High CRP	*Reference* <0.22 mg/dL		
	0.22-0.99 mg/dL	-0.06	0.27
	>=1.0 mg/dL	0.04	0.62

a. Age-adjusted, except for age cofactor

b. Cumulative infectious disease burden is a measure of burden of infection based on the number of infections for which a sample person tested positive out of 5 organisms (Toxoplasma gondii, Toxocara spp, cytomegalovirus, Hepatitis A, and Hepatitis B virus)

Table 13. Multivariate Models Obtained from Simple Backwards Stepwise Elimination with Cumulative Infectious Disease Burden[a] Retained in the Model, among those ages 20-39 in NHANES III

Outcome	Covariates Retained	Odds Ratio for CIDB	95% Confidence Interval for CIDB Odds Ratio	P-Value for CIDB
Major Depression	Age Race Sex Poverty Smoking			
	CIDB	0.77	0.53, 1.11	0.16
Severe Major Depression	Age Race Sex Poverty Smoking			
	CIDB	0.77	0.52, 1.16	0.21
Dysthymia	Age Sex Poverty Education Smoking			
	CIDB	0.99	0.73, 1.35	0.95
Dysthymia with Major Depression	Age Sex Poverty Smoking			
	CIDB	0.91	0.58, 1.41	0.66
Bipolar Disorder	Poverty			
	CIDB	0.83	0.54, 1.28	0.39
Atypical Bipolar Disorder	Foreign Birth Smoking			
	CIDB	0.86	0.46, 1.61	0.63

a. Cumulative infectious disease burden is a measure of burden of infection based on the number of infections for which a sample person tested positive out of 5 organisms (Toxoplasma gondii, Toxocara spp, cytomegalovirus, Hepatitis A, and Hepatitis B virus)