Geographic Disparities Associated with Stroke and Myocardial Infarction in East Tennessee

Ashley Pedigo Golden
apedigo@utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Part of the Biostatistics Commons, Cardiovascular Diseases Commons, Epidemiology Commons, Health Services Research Commons, Multivariate Analysis Commons, and the Vital and Health Statistics Commons

Recommended Citation
https://trace.tennessee.edu/utk_graddiss/1182

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
To the Graduate Council:

I am submitting herewith a dissertation written by Ashley Pedigo Golden entitled "Geographic Disparities Associated with Stroke and Myocardial Infarction in East Tennessee." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Comparative and Experimental Medicine.

Agricola Odoi, Major Professor

We have read this dissertation and recommend its acceptance:

Tim E. Aldrich, Karla J. Matteson, Bruce A. Ralston, William L. Seaver

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)
Geographic Disparities Associated with Stroke and Myocardial Infarction in East Tennessee

A Dissertation Presented for the Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Ashley Pedigo Golden
December 2011
Dedication

This dissertation is dedicated to my parents, Ronnie and JoAnn Pedigo, for their many sacrifices, constant love, and support that have enabled me to accomplish my educational goals.
Acknowledgements

This dissertation could not have been completed without the help from many individuals, whom I wish to acknowledge and give thanks. I am sincerely grateful for the opportunity to learn and work under the supervision of Dr. Agricola Odoi. His expertise, guidance, and dedication to excellence have been integral to my educational and professional growth. Thank you, Dr. Odoi, for your patience and wisdom in this journey. I would like to extend my appreciation to the members of my committee: Dr. Tim Aldrich, Dr. Karla Matteson, Dr. Bruce Ralston, and Dr. William Seaver. Each of you has provided both academic and professional support for which I am very thankful. Special thanks are due to Lisa Escalante for her assistance in all things EMS. I am grateful to my colleagues, Jenny and Doreen, for their friendship and support. To my friends and family, thank you for your encouragement and faith in me. I owe my deepest gratitude to my husband, David, whose constant love and understanding provided me strength. Finally, I thank God for his mercy and blessings on me.
Abstract

Stroke and myocardial infarction (MI) are serious conditions whose burdens vary by socio-demographic and geographic factors. Although several studies have investigated and identified disparities in burdens of these conditions at the county and state levels, little is known regarding their geographic epidemiology at the neighborhood level. Both conditions require emergency treatments and therefore timely geographic accessibility to appropriate care is critical. Investigation of disparities in geographic accessibility to stroke and MI care and the role of Emergency Medical Services (EMS) in reducing treatment delays are vital in improving health outcomes. Therefore, the objectives of this work were to: (i) classify neighborhoods based on socio-demographic and geographic characteristics; (ii) investigate spatial patterns of neighborhood level mortality; (iii) identify disparities in geographic accessibility to stroke and MI care; and (iv) identify disparities in EMS transport times for stroke and MI patients in East Tennessee.

Fuzzy cluster analysis was used to classify neighborhoods into peer neighborhoods (PNs) based on their socio-demographic and geographic factors. Neighborhood level spatial patterns of stroke and MI mortality risks were investigated using Spatial Empirical Bayesian smoothing techniques and neighborhoods with high mortality risks identified using spatial scan statistics. Travel times to stroke and cardiac care facilities were computed using network analysis to investigate geographic accessibility. Records of over 3,900 suspected stroke and MI patients, from two EMS providers, were used to investigate disparities in EMS transport delays.
Four distinct PNs were identified. The highest stroke/MI mortality risks were observed in less affluent, urban PNs, and lowest risks in more affluent, suburban PNs. Several significant ($p<0.0001$) stroke and MI high mortality risk spatial clusters were identified. Approximately 8% and 15% of the population did not have timely accessibility to appropriate stroke and MI care, respectively. The disparity was greatest for populations in rural areas. Important disparities in EMS transport delays were identified, with the travel time to a hospital contributing the longest delay.

The identified disparities in neighborhood characteristics, mortality risks, geographic accessibility, and EMS transport delays are invaluable in guiding resource allocation, service provision, and policy decisions to support evidence-based population health planning and policy.
Table of Contents

CHAPTER 1

1.0 Introduction ....................................................................................................................... 1

1.1 Literature Review .............................................................................................................. 4
  1.1.1 Stroke and Myocardial Infarction (MI) ......................................................................... 4
    1.1.1.1 Disease biology, etiology, and symptoms ................................................................. 4
      1.1.1.1.1 Stroke ................................................................................................................... 4
      1.1.1.1.2 Myocardial Infarction ......................................................................................... 5
    1.1.1.2 Diagnoses and treatments .......................................................................................... 7
      1.1.1.2.1 Stroke ................................................................................................................... 7
      1.1.1.2.2 Myocardial Infarction ......................................................................................... 10
    1.1.1.3 Disease Burden ......................................................................................................... 11
      1.1.1.3.1 Stroke ................................................................................................................... 12
      1.1.1.3.2 Coronary Heart Disease and Myocardial Infarction ........................................... 13
    1.1.1.4 Temporal trends ........................................................................................................ 14
      1.1.1.4.1 Stroke ................................................................................................................... 14
      1.1.1.4.2 Coronary Heart Disease and Myocardial Infarction ........................................... 15
    1.1.1.5 Vital Statistics ........................................................................................................... 17
  1.1.2 Socio-demographic determinants of stroke and myocardial infarction ...................... 18
    1.1.2.1 Demographic factors ............................................................................................... 18
    1.1.2.2 Socioeconomic factors ............................................................................................ 19
    1.1.2.3 Geographic clustering of socioeconomic and demographic factors ..................... 21
  1.1.3 Geographic patterns of stroke and myocardial infarction .......................................... 23
    1.1.3.1 Geographic disparities in burdens .............................................................................. 23
    1.1.3.2 Neighborhood level spatial analyses ....................................................................... 24
  1.1.4 Geographic accessibility to stroke and MI care ............................................................ 26
    1.1.4.1 Disparities in access to healthcare ............................................................................. 26
    1.1.4.2 Measuring geographic accessibility to healthcare ....................................................... 28
    1.1.4.3 Improving geographic accessibility .......................................................................... 30
1.1.5 Role of Emergency Medical Services (EMS) in timely access to stroke and MI care ................................................................. 32
  1.1.5.1 Delays in receiving emergency care for stroke and MI ........ 32
  1.1.5.2 Components of pre-hospital delays .................................. 36
    1.1.5.2.1 Decision Delays .................................................. 37
    1.1.5.2.2 Transport Delays ............................................... 39
  1.1.5.3 Improving pre-hospital delays ...................................... 41

CHAPTER 2
2.0 Identifying Unique Neighborhood Characteristics to Guide Health Planning for Stroke and Heart Attack: Fuzzy Cluster & Discriminant Analyses Approaches ....................................................... 43
2.1 Abstract .................................................................................. 44
2.2 Introduction ............................................................................. 45
2.3 Methods .................................................................................. 47
  2.3.1 Study area population ....................................................... 47
  2.3.2 Data acquisition ............................................................... 48
    2.3.2.1 Population characteristics ......................................... 48
    2.3.2.2 Mortality data .......................................................... 48
  2.3.3 Data analysis ..................................................................... 49
    2.3.3.1 Data management ..................................................... 49
    2.3.3.2 Robust principal components analysis (PCA) .......... 49
    2.3.3.3 Fuzzy cluster analysis .............................................. 50
    2.3.3.4 Validation of Peer Neighborhoods (PNs) ................. 52
    2.3.3.5 Comparison of mortality between Peer Neighborhoods (PNs) .................................................. 53
2.4 Results ................................................................................... 54
  2.4.1 Robust principal components analysis ................................ 54
  2.4.2 Fuzzy cluster analysis results ............................................ 54
  2.4.3 Characteristics of identified Peer Neighborhoods (PNs) .... 61
  2.4.4 Evaluation of misclassified neighborhoods ....................... 62
2.4.4 Variables important for classifying Peer Neighborhoods (PNs) .............................................. 65
2.4.5 Disparities in stroke and MI mortality between Peer Neighborhoods (PNs) .............................. 67
2.5 Discussion and Conclusions ........................................................................................................ 67

CHAPTER 3

3.0 Neighborhood Disparities in Stroke and Myocardial Infarction Mortality:
a GIS and Spatial Scan Statistics Approach ...................................................................................... 76
3.1 Abstract ...................................................................................................................................... 77
3.2 Introduction .................................................................................................................................. 78
3.2 Methods ...................................................................................................................................... 80
  3.2.1 Study area and data collection ............................................................................................... 80
  3.2.2 Data analysis .......................................................................................................................... 82
    3.2.2.1 Data management ............................................................................................................ 82
    3.2.2.2 Descriptive analyses, risk standardization and spatial smoothing ............................... 83
    3.2.2.3 Detection and identification of stroke and MI clusters ................................................. 83
    3.2.2.4 Logistic modeling of predictors of high risk stroke or MI clusters ............................... 84
    3.2.2.5 Cartographic displays .................................................................................................... 86
3.3 Results ........................................................................................................................................ 86
  3.3.1 Description of stroke and MI deaths ...................................................................................... 86
  3.3.2 Spatial distribution of mortality risks .................................................................................... 87
    3.3.2.1 Stroke risks ..................................................................................................................... 87
    3.3.2.2 Myocardial infarction risks ............................................................................................ 89
  3.3.3 Spatial clusters of high stroke/MI mortality risks ................................................................. 89
    3.3.3.1 Stroke clusters ............................................................................................................... 89
    3.3.3.2 Myocardial infarction clusters ....................................................................................... 94
  3.3.4 Predictors of high risk stroke and myocardial infarction spatial clusters ............................ 94
    3.3.4.1 Stroke ............................................................................................................................ 94
    3.3.4.2 Myocardial infarction .................................................................................................... 98
3.4 Discussion and Conclusions ......................................................................................................... 100
CHAPTER 4

4.0 Investigation of Disparities in Geographic Accessibility to Emergency Stroke and Myocardial Infarction Care in East Tennessee using GIS and Network Analysis ................................................................. 108

4.1 Abstract ........................................................................................................... 109

4.2 Introduction ..................................................................................................... 110

4.3 Materials and Methods .................................................................................. 111

  4.3.1 Study area population ............................................................................... 111
  4.3.2 Data collection .......................................................................................... 112
    4.3.2.1 Hospital distribution data ................................................................. 112
    4.3.2.2 Street network data .......................................................................... 112
    4.3.2.3 Cartographic boundary files ............................................................ 114
    4.3.2.4 Population data ................................................................................ 114
  4.3.3 Data Analysis ............................................................................................ 114
    4.3.3.1 Network Analysis ............................................................................ 114
    4.3.3.2 Euclidean Distance Model ............................................................... 115
    4.3.3.3 Travel time by air ambulance analysis ............................................ 115

4.4 Results ............................................................................................................ 116

  4.4.1 Hospital distribution .................................................................................. 116
  4.4.2 Ground travel time to accredited stroke centers ....................................... 116
  4.4.3 Ground travel time to cardiac centers ..................................................... 121
  4.4.4 Ground travel time to other hospitals with an ER .................................... 121
  4.4.5 Travel time by Euclidean distance ............................................................ 121
  4.4.6 Travel time by air ambulance .................................................................. 121

4.5 Discussion and Conclusions ......................................................................... 123

CHAPTER 5

5.0 Disparities in Pre-hospital Transport Delays for Acute Stroke and Myocardial Infarction Patients Utilizing Emergency Medical Services .......... 127

5.1 Abstract ........................................................................................................... 128
List of Tables

Table 2.1: Component Loadings from Robust Principal Components Analysis for Socioeconomic and Demographic Variables ....................................................... 55

Table 2.2: Sensitivity Analysis of Fuzzy Cluster Analysis Results for Peer Neighborhoods Based on Socioeconomic and Demographic Population Characteristics ................................................................. 57

Table 2.3: Summary of Degrees of Belonging for Neighborhoods within Peer Neighborhoods as the Fuzzifier changes in Fuzzy Cluster Analysis ......................................................................................................................... 58

Table 2.4: Summary Statistics of Socioeconomic and Demographic Population Characteristics of Peer neighborhoods in East TN ............................................ 63

Table 2.5: Nearest Neighbor Discriminant Analysis Results of Classification of East Tennessee Peer Neighborhoods Based on Socioeconomic & Demographic Characteristics ...................................................... 64

Table 3.1: Spatial Clusters of Age-Adjusted Stroke and Myocardial Infarction Mortality Risks from 1999 to 2007 in East TN Appalachian Region ........................................................................................................... 91

Table 3.2: Univariate Associations of High Risk Stroke and Myocardial Infarction Mortality Clusters with Neighborhood Socioeconomic and Demographic Factor ........................................................................................................... 95

Table 3.3: Final logistic model showing socioeconomic and demographic predictors of high risk stroke mortality cluster ................................................................. 97

Table 3.4: Final logistic model showing socioeconomic and demographic predictors of high risk myocardial infarction mortality clusters ................................................ 99

Table 4.1: Population Distribution within Specified Travel Times to Stroke and Myocardial Infarction Care in East TN Appalachian Region ................................................................................................................. 120

Table 5.1: Definitions of emergency medical services (EMS) transport time intervals for stroke and myocardial infarction patients ............................................. 133

Table 5.2: Pre-hospital transport time intervals for suspected stroke and myocardial infarction (MI) patients from two emergency medical service (EMS) providers in East TN ......................................................... 138
Table 5.3: Characteristics of stroke patients whose time intervals exceeded emergency medical services (EMS) guidelines ........................................141

Table 5.4: Characteristics of stroke and myocardial infarction patients, transported by emergency medical services provider #2 (EMS 2) whose time intervals exceeded EMS guidelines .........................................................143

Table 5.5: Characteristics of myocardial infarction patients whose time intervals exceeded emergency medical services (EMS) guidelines ........................................................................................................145
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Identified Peer Neighborhoods (PN) in East TN based on Socioeconomic and Demographic Population Characteristics using Fuzzy K-means Clustering Algorithm</td>
<td>59</td>
</tr>
<tr>
<td>2.2</td>
<td>Identified Peer Neighborhoods (PN) in East TN based on Socioeconomic and Demographic Population Characteristics using K-means Clustering Algorithm</td>
<td>60</td>
</tr>
<tr>
<td>2.3</td>
<td>Cluster and Regression Tree (CART) Results for Peer Neighborhoods (PNs) in East Tennessee</td>
<td>66</td>
</tr>
<tr>
<td>2.4</td>
<td>Annual Age-Adjusted Stroke and Myocardial Infarction Mortality Risks for Peer Neighborhoods in East Tennessee</td>
<td>68</td>
</tr>
<tr>
<td>3.1</td>
<td>Spatial Empirical Bayes Smoothed Age-Adjusted Stroke Mortality Risk per 100,000 Population from 1999 to 2007 in East TN Appalachian Region</td>
<td>88</td>
</tr>
<tr>
<td>3.2</td>
<td>Spatial Empirical Bayes Smoothed Age-Adjusted Myocardial Infarction Mortality Risk Per 100,000 Population from 1999 to 2007 in East TN Appalachian Region</td>
<td>90</td>
</tr>
<tr>
<td>3.3</td>
<td>Significant Spatial Clusters of High Age-Adjusted Stroke Mortality Risks from 1999 to 2007 in East TN Appalachian Region</td>
<td>92</td>
</tr>
<tr>
<td>3.4</td>
<td>Significant Spatial Clusters of High Age-Adjusted Myocardial Infarction Mortality Risks from 1999 to 2007 in East Tennessee Appalachian Region</td>
<td>93</td>
</tr>
<tr>
<td>4.1</td>
<td>Map of Study Area and Neighboring Counties</td>
<td>113</td>
</tr>
<tr>
<td>4.2</td>
<td>Travel Time to Stroke Centers in East TN Appalachian Region</td>
<td>117</td>
</tr>
<tr>
<td>4.3</td>
<td>Travel Time to Cardiac Centers in East TN Appalachian Region</td>
<td>118</td>
</tr>
<tr>
<td>4.4</td>
<td>Travel Time to Care Centers with an Emergency Room in East TN Appalachian Region</td>
<td>119</td>
</tr>
<tr>
<td>4.5</td>
<td>Comparison of Travel Time by Network Analysis and Euclidean Distance for 50 MPH Travel Speed to Care Centers in East TN Appalachian Region</td>
<td>122</td>
</tr>
</tbody>
</table>
List of Abbreviations

AHA: American Heart Association
ALS: Advanced Life Support
API: Application Programming Interface
BLS: Basic Life Support
CART: Classification and Regression Tree
CDC: Centers for Disease Control and Prevention
CHD: Coronary Heart Disease
CT: Computed Tomography
CTs: Census Tracts
CVA: Cerebrovascular Accident
DA: Discriminant Analysis
DPU: Normalized Average Squared Area
ECG: Electrocardiogram
ED: Emergency Department
EMS: Emergency Medical Services
ER: Emergency Room
ESRI: Environmental Systems Research Institute
FPU: Dunn’s Normalized Partition Coefficient
GIS: Geographical Information Systems
HARN: High Accuracy Reference Network
ICD: International Classification of Diseases
IRB: Institutional Review Board
JCAHO: Joint Commission on the Accreditation of Health Organizations

MI: Myocardial Infarction

MPH: Miles per Hour

NAD: North American Datum

NCHS: National Center for Health Statistics

NHANES: National Health and Nutrition Examination Survey

NIH: National Institutes of Health

NoRLS: No Red Lights and Sirens

P: p-value

PCA: Principal Component Analysis

PCI: Percutaneous Coronary Intervention

PN: Peer Neighborhood

RLS: Red Lights and Sirens

SEB: Spatial Empirical Bayes

SES: Socioeconomic Status

TN: Tennessee

TPA: Tissue Plasminogen Activator

US: United States
CHAPTER 1

1.0 Introduction

Despite recent declines in death risks from stroke and acute myocardial infarction (MI), their burdens remain high and the conditions continue to be of significant public health importance in the United States (US)\(^1\). The burdens of these conditions vary significantly across demographic, socioeconomic, and geographic groups. However, most of the past research has focused on identifying disparities of individual risk or at higher geographic levels, like state or county. Recent studies indicate that finer geographic units are needed to improve our understanding of the geographic distribution of these determinants of health at lower levels so as to better inform local health planning\(^2,3\). Furthermore, research has overwhelmingly found that an individual’s health can be influenced by the characteristics of their neighborhood beyond their individual characteristics\(^4\).

The clustering of demographic and socioeconomic determinants of health into distinct neighborhoods can greatly affect the planning, implementation, and focus of health initiatives that seek to reduce disparities. Given the complex and multidimensional nature of these risk factors, it has been suggested that research should focus on using multivariate methods to identify and better understand population characteristics to guide health planning\(^3\). Thus, since population health planning is typically conducted at the local community level, population studies are needed that investigate important determinants of health, using multivariate methods, and the patterns of stroke and MI risks at the neighborhood level.
Stroke and MI both require specific, time sensitive care. Thus, geographic access to care is an important determinant for obtaining effective emergency care. However, little is known regarding disparities in timely geographic accessibility to stroke and MI care based on actual travel time since most past studies have used distance to the healthcare facility, and not travel time to care. Recently, travel time has been recognized as a better indicator of accessibility compared to distance to healthcare facilities since factors that impede travel (speed limits, road connectivity, and turn restrictions) can be accounted for in the analysis. Therefore, it is important to adjust for these factors in order to give a more accurate and realistic estimate of geographic access to healthcare services. With increased access to GIS, availability of detailed transportation data, and advanced computing, research investigating access to emergency stroke and MI care while taken into account travel impedances is warranted to better identify disparities in geographic accessibility across populations.

Emergency Medical Services (EMS) play a critical role in providing rapid response and transport to emergency stroke and MI care. Patients who utilize EMS transport have higher odds of receiving appropriate treatments within the recommended times. However, transport delays sometimes occur at different time intervals from the time a 911 call is received to the time the patient is delivered to the care facility by the EMS. Thus, it is important to investigate the specific time intervals involved in EMS transport so as to identify when and where delays occur. This information is vital for guiding efforts aimed at reducing these delays. The time intervals to be considered for these types of investigations include the response time, time spent on-scene, and the time required to travel to the hospital. Unfortunately, only a few recent studies have
described the specific time intervals associated with EMS transport for stroke\textsuperscript{15,16} and MI\textsuperscript{17}. Additional studies of different populations, geographic areas, and EMS providers are needed to improve understanding of this component of pre-hospital delays.

With the above issues in mind, the objectives of this study were to: (i) classify neighborhoods into peer neighborhoods based on socio-demographic and geographic characteristics; (ii) investigate spatial patterns of neighborhood level mortality risk of stroke and MI; (iii) identify disparities in timely geographic accessibility to stroke and MI care; (iv) identify disparities in EMS transport times for stroke and MI patients in East Tennessee. The identification of disparities in neighborhood characteristics, mortality risks, geographic accessibility, and EMS transport delays are invaluable in guiding resource allocation, service provision, and policy decisions to support evidence-based population health planning and policy.

This dissertation is divided into six chapters. The first chapter includes the introduction to the study and literature review. Chapters 2, 3, 4, and 5 describe the methods and findings of studies each addressing objectives 1, 2, 3, and 4, respectively. Lastly, chapter 6 provides a summary of the conclusions and recommendations of this study.
1.1 Literature Review

1.1.1 Stroke and Myocardial Infarction (MI)

1.1.1.1 Disease biology, etiology, and symptoms

1.1.1.1.1 Stroke

Stroke, or cerebrovascular accident, is characterized by a sudden loss of consciousness, sensations, or voluntary motions due to rupture (hemorrhagic stroke) or obstruction (ischemic stroke) of a blood vessel, interrupting the blood supply to the brain\textsuperscript{18, 19}. Approximately 87\% of all stroke cases in the US are ischemic\textsuperscript{20}. The most common causes of ischemic strokes include: (i) a thrombosis which is a clot formation in the vessels in the brain or neck; (ii) an embolism, or clot that has moved to the brain from other places in the body; (iii) the narrowing of an artery in the brain, called stenosis\textsuperscript{18}. Transient ischemic attacks, sometimes called mini-strokes, are stroke events that last only a few seconds or minutes and are a major risk factor, along with prior stroke, for future ischemic attacks\textsuperscript{21}.

Other medical conditions are also known to increase the risk of stroke. High blood pressure (BP> 120/80 mmHg), or hypertension, has been found to be a powerful determinant of risk for both ischemic and hemorrhagic stroke, approximately doubling risk of death from stroke\textsuperscript{22, 23}. Additionally, cigarette smoking, high dietary sodium, and excessive alcoholic intake can significantly increase blood pressure and thus are behaviors associated with increased risk of stroke\textsuperscript{1, 18}. High blood cholesterol and coronary artery disease can lead to build up of fatty deposits (plaque) in blood vessels,
which increases the risk of ischemic stroke, specifically, as well as myocardial infarction\textsuperscript{18}. Diabetes and obesity are two other conditions that are associated with increased blood pressure and cholesterol and thus elevated stroke risks\textsuperscript{18,24}. Since hypertension, blood cholesterol, dietary sodium, and diabetes can typically be significantly improved through a healthier diet and increased exercise, these conditions, as well as smoking and alcohol intake, are generally referred to as modifiable risk factors\textsuperscript{18}. These modifiable risk factors synergistically increase stroke risk along with non-modifiable risk factors like genetics and demographic characteristics (age, gender, race)\textsuperscript{25}.

Stroke symptoms typically start suddenly, over seconds to minutes, and result in neurological deficits that are highly dependent on the location and extent of the affected area of the brain\textsuperscript{26}. The five most common symptoms include: (i) sudden numbness or weakness of the face, arm, or leg particularly on one side; (ii) sudden confusion or trouble speaking; (iii) sudden trouble with vision; (iv) sudden dizziness, trouble walking, or loss of balance or coordination; or (v) severe headache of unknown cause\textsuperscript{18}. Several studies have found that facial droop, arm drift, and slurred speech are the signs most significantly associated with correct identification of stroke\textsuperscript{27-29}.

1.1.1.1.2 Myocardial Infarction

Myocardial infarction (MI), or heart attack, occurs when there is insufficient blood supply to the heart, resulting in damage or death to portions of the heart muscle\textsuperscript{30}. If left untreated, MI can cause irregular heart rhythms, called arrhythmias, or the heart to stop. Thus, quick restoration of the blood flow, or reperfusion, minimizes the extent of
damage. The most common etiology of MI is coronary heart disease (CAD), which is a condition in which fatty deposits, or plaque, build up in the arteries. The plaque progressively hardens resulting in narrowed arteries that restrict blood flow, possibly leading to a coronary blockage severe enough that an infarction occurs. Additionally, an area of plaque may rupture, leading to a coronary thrombosis, or clot, that prevents blood from flowing to the heart.

Like stroke, a number of conditions are associated with increased risk of MI: hypertension, high cholesterol levels, obesity, and diabetes mellitus. As well, the same modifiable behavioral risk factors (i.e. diet, exercise, smoking, and alcohol intake) associated with increased stroke risk are determinants of MI risk and co-morbidities. One case-control study estimated that improvements in these risk factors could reduce the risk of an initial MI by approximately 90%. These modifiable risk factors act synergistically with non-modifiable risk factors to increase MI risk, similar to stroke risk.

The major symptoms of MI include: chest pain; discomfort in other areas of the upper body (arm, back, jaw, neck); shortness of breath; or other signs (nausea, lightheadedness, or cold sweat). Chest pain or discomfort, with varying levels of intensity, is the most common symptom. However, some patients have mistaken MI pain as indigestion or heartburn. Additionally, several studies have reported that women are more likely than men to experience symptoms other than chest pain. Thus, recognizing the signs and symptoms of an MI can be quite difficult. Therefore, it is important that individuals (particularly those at higher risk) rapidly seek medical
attention in any instance of these signs and symptoms so that accurate diagnoses can be performed.

1.1.1.2 Diagnoses and treatments

1.1.1.2.1 Stroke

Patient symptoms, medical history, physical examination, neurological evaluation, as well as diagnostic tests, are the foundations for stroke diagnoses. Often, the symptom characteristics (i.e. severity, location, time since onset) along with findings from the physical (i.e. vital signs) and neurological examinations are summarized using a standard score on the National Institutes of Health (NIH) stroke scale. This scale is used worldwide as a clinical assessment tool to provide a quantitative measure of stroke-related neurologic deficit\textsuperscript{37}. This information aids clinicians in the evaluation of the type, location, and severity of the stroke, as well as in decisions for which diagnostic tests are needed\textsuperscript{38}.

The most commonly used, and widely available, diagnostic technique for suspected stroke is computed tomography imaging, which uses low-dose radiation to image the brain. This technique has high sensitivity (89\%) and specificity (100\%) for hemorrhagic stroke, as well as high specificity (96\%) for ischemic stroke\textsuperscript{39}. However, the sensitivity of CT imaging for ischemic stroke is very low (16\%). Thus, additional imaging may be necessary in some cases to confirm the diagnosis. Magnetic resonance imaging utilizes magnetic fields to image the brain and have high sensitivity in detecting the location and extent of brain tissue damage from stroke, particularly in small vessels, but is much more expensive and less available than CT imaging\textsuperscript{39}. Once a diagnosis on
the type (ischemic or hemorrhagic) and location of the stroke has been made, other
diagnostic tests may be performed to determine the underlying etiology. The most
common tests and their purposes include: ultrasound/Doppler and duplex scanning to
assess the flow of blood in the brain in order to detect carotid stenosis;
electrocardiogram (ECG) to identify arrhythmias that are the result of an embolism
-especially from the heart); or an angiogram to determine if the hemorrhage was the
result of an aneurysm or malformation of the vessel$^{40}$.

Treatments for stroke are highly dependent on the diagnosis since key therapies
for ischemic stroke can actually make bleeding associated with hemorrhagic stroke
worse. The primary goal in treatment of ischemic stroke is to restore blood flow to the
brain by removing the obstruction (clot). Anticoagulant drugs (such as aspirin,
Clopidogrel, or dipyridamole) may be given to try to reduce clot enlargement or prevent
new clots from forming by stopping platelets from aggregating$^{41}$. The primary treatment
recommended for ischemic stroke patients is intravenous thrombolytic treatment, such
as tissue plasminogen activator (TPA), given within 180 minutes (3 hours) of the onset
of symptoms$^{42-44}$. It has been reported that, even within the 3 hour window, reducing the
time from symptom onset to treatment increased the odds of a good health outcome$^{43}$. 
Due to the significant risk associated with this therapy, its use is only recommended
when an ischemic stroke diagnosis has been confirmed by a specialist, usually a
neurologist, through evaluation and diagnostic imaging$^{44}$.

Additionally, the American Heart Association states that thrombolytic treatment
should not be given if emergency ancillary care and capabilities to handle bleeding
complications are not readily available\textsuperscript{42}. These ancillary care and capabilities to handle complications are typically available only in stroke centers. Thus, recommendations for the establishment and direct transfer of acute stroke patients to accredited stroke centers have been made\textsuperscript{42, 45}. The Brain Attack Coalition set the following criteria for establishing a hospital as an accredited stroke center: (1) have healthcare personnel with specific expertise in neurosurgery and vascular neurology; (2) have advanced neuro-imaging capabilities such as computed tomography (CT) or magnetic resonance imaging (MRI); (3) can provide surgical and endovascular techniques; and (4) have other specific infrastructure and programmatic elements such as an intensive care unit and a stroke registry\textsuperscript{46}. Several studies have reported decreased time to treatment, increased likelihood of being eligible to receive thrombolytic treatment, and improved health outcomes for stroke patients treated at accredited stroke centers\textsuperscript{45, 47-49}.

Recently, the European Cooperative Acute Stroke Study reported a clinical benefit for ischemic stroke patients when the thrombolytic alteplase (or TPA) was administered within 3 to 4.5 hours after the onset of symptoms\textsuperscript{50}. Non-thrombolytic treatments for reperfusion of ischemic stroke patients, particularly used when the time from symptom onset is greater than 3 hours, include: mechanical thrombolysis, intra-arterial thrombolysis, carotid endarterectomy, and angioplasty\textsuperscript{42}. Treatments for hemorrhagic stroke are less defined in the literature, but in general, include surgical procedures to stop bleeding and repair aneurysms\textsuperscript{51}. 
1.1.1.2.2 Myocardial Infarction

Like stroke, patient symptoms and physical evaluation, as well as medical history, form the basis for diagnoses of MI. In fact, one of established criteria for correct diagnosis of MI is clinical history of chest pain due to ischemia (restriction of blood) for at least 20 minutes\(^ {52} \). The other criteria include progressive changes in electrocardiogram (ECG) readings and abnormalities in cardiac enzymes, such as creatine kinase and troponin. The ECG traces the electrical activity of the heart, which during or after a MI, will have abnormal electrical impulses that can be identified through changes in electrical waves recorded by the ECG\(^ {31} \). Elevated levels of cardiac enzymes, along with ECG readings, are strong indicators of MI. Additionally, coronary angiography, which involves special imaging of the heart and arteries, is commonly used to identify the location of arterial blockages. This technique, as well as an echocardiogram, can be used to detect abnormalities in areas of the heart muscle or wall that may indicate past damage from a silent (i.e. MI with no symptoms) or past MI (greater than 48 hours since MI event)\(^ {52} \).

As the first line of defense in managing MI, therapies such as aspirin, nitroglycerin, and oxygen are often given by Emergency Medical Services (EMS) personnel or by physicians in the Emergency Department (ED). Aspirin serves to control pain and thins the blood to prevent or reduce blockages, while nitroglycerin is important in restoring or maintaining the electrical activity of the heart, and oxygen is necessary to relieve shortness of breath often associated with an MI\(^ {31} \). Beyond initial management, the key to stopping the infarct and limiting its size so as to reduce damage is rapid reperfusion through mechanical (percutaneous coronary interventions) or
pharmacological (thrombolytic drugs) removal of the blockage\textsuperscript{53}. Percutaneous coronary interventions (PCI), also referred to as balloon angioplasty or catheterization, involves the insertion of a catheter with an inflatable balloon at the end into the affected artery to relieve the blockage and restore blood flow. Often in conjunction with PCI, a stent (a small mesh tube) is implanted to hold the artery open to reduce the chance of a recurrence\textsuperscript{54}. Several studies have reported that PCI is superior to thrombolytic drugs for treatment of MI\textsuperscript{55-57}. 

Similar to the treatment for stroke, the effectiveness of PCI for MI patients is dependent on time to treatment. Current guidelines from the American Heart Association and American College of Cardiology recommend that the time from first medical contact to PCI be 90 minutes or less\textsuperscript{53}. Reports indicate that health outcomes are improved by up to 50\% when PCI is administered within 60 minutes and 23\% within 180 minutes of the onset of MI symptoms\textsuperscript{58, 59}. However, like the treatment for stroke, PCI requires advanced training of cardiologists and specialized equipment, primarily catheterization and angioplasty laboratories, and therefore not all hospitals are capable of providing these treatments\textsuperscript{60}. Therefore, recommendations have been made for the establishment and direct transport of MI patients to cardiac centers (i.e. hospitals with PCI capabilities) so as to increase the use of PCI and the likelihood of better health outcomes\textsuperscript{33, 53, 61, 62}.

1.1.1.3 Disease Burden

Cardiovascular diseases (CVD), and the acute coronary (MI) and cerebrovascular (stroke) events associated with it, are major causes of disability and premature death throughout the world\textsuperscript{63}. 

11
1.1.1.3.1 Stroke

Stroke is the third most common cause of death, behind CHD and cancer, in developed countries\textsuperscript{64}. Of the estimated 15 million people worldwide who suffer a stroke each year, 5.5 million (3 million women and 2.5 million men) die and another 5 million will be permanently disabled\textsuperscript{64}. Stroke is the biggest single cause of debilitation in many developed countries, including the United Kingdom (UK)\textsuperscript{64} and the US\textsuperscript{24}. Age standardized (standardized to the European Union population) international death risks (per 100,000 population) are highest in China (men: 365.4; women: 239.1), the Russian Federation (men: 351.4; 189.2), Bulgaria (men: 239.8; women: 137.5), and Romania (men: 204.3; women: 122.6)\textsuperscript{1}. On the other hand, the lowest risks were observed in high income, developed countries, including (but not limited to): Switzerland (men: 16.6; women: 12.4), Australia (men: 22.0; women: 16.3), Canada (men: 24.2; women: 17.3), France (men: 26.5; women: 14.0), the Netherlands (men: 26.6; women: 16.2), and Austria (men: 27.1; women: 16.2). It has been reported that 85% of all stroke deaths occur in lower income countries\textsuperscript{65}. For instance in North America, the stroke death risk is about two times higher in Mexico (men: 60.5; women: 52.4) compared to the US (men: 31.6; women: 24.3) and Canada (men: 24.2; women: 17.3)\textsuperscript{1}.

It is estimated that, on average, an American dies of a stroke every 4 minutes\textsuperscript{66}. Approximately 1 of every 18 adult deaths in the US in 2007 were attributed to stroke\textsuperscript{66}. There is evidence that mortality from stroke has recently declined in the US. During the period 1999-2007, there was a 34.3% decrease in the annual stroke death risk and 18.8% decrease in actual number of stroke deaths\textsuperscript{1}. However, the estimated overall stroke prevalence (3.0%) remained relatively unchanged during the same period. The
annual stroke incidence continues to be high with an estimated 795,000 stroke events in the US population, of which 610,000 were first attacks\(^1\). Thus, stroke is a serious burden to the US health system with annual costs estimated at $40.9 billion in 2007\(^1\).

**1.1.1.3.2 Coronary Heart Disease and Myocardial Infarction**

Even though the current study is focused only on MI, data on the burden of MI are typically not separated from that of CHD (particularly from countries other than the US), so the burden of MI is presented here as that for CHD, except when separately available. Globally, more deaths are attributable to CHD, including MI, than any other cause of death. The highest number of deaths from CHD in 2002 occurred in India (1,531,534 deaths), China (702,925 deaths), and the Russian Federation (674,881 deaths)\(^64\). Age-adjusted CHD death risks (per 100,000 population) are lowest for many high income, developed countries including (but not limited to): Japan (men: 47.6; women: 13.8), South Korea (men: 51.4; women: 20.0); France (men: 58.4, women: 12.2); the Netherlands (men: 66.2; women: 22.8); and Switzerland (men: 78.2; women: 19.4)\(^1\). Similar to stroke, it has been reported that the greatest burden of CHD occurs in low income countries, particularly those in Eastern Europe (Croatia, Romania), Central Asia (Kazakhstan, Ukraine, and Turkey), and North Africa (Egypt, Libya)\(^64\). In North America, some differences exist in CHD death risks, although less distinct than disparities in stroke risks, between the higher income countries, like the US (men: 153.3; women: 60.4) and Canada (men: 130.8; women: 42.8), and Mexico (men: 136.8; women: 73.2)\(^1\).
Coronary heart disease (CHD) accounted for approximately 1 of every 6 US deaths in 2007\textsuperscript{66}. The annual incidence of MI is 610,000 new attacks and 325,000 recurrent attacks each year and it is estimated that approximately 15\% of those who have an MI in the US will die of it\textsuperscript{1}. The estimated average number of years of life lost due to MI is 16.6 years\textsuperscript{1}. From 2005 to 2008, the overall prevalence of MI among US adults ≥20 years of age was 3.1\%\textsuperscript{1}. The burdens of CHD and MI constitute huge expenditures in the US healthcare system with annual costs estimated at $177.5 billion\textsuperscript{1}.

1.1.1.4 Temporal trends

Globally, the burdens of stroke and CHD have declined in most developing countries over recent decades\textsuperscript{64}. The decline has been attributed, in part, to improved awareness, prevention, diagnosis, and treatment of stroke and MI patients. Additionally, significant portions of the trend of decreasing burdens are due to favorable changes in modifiable risk factors, such as decreased smoking, improvements in physical activity and diet, and better control of hypertension and high blood cholesterol\textsuperscript{65}. However, temporal patterns in stroke and CHD burdens vary for different populations in the world, with some increasing trends in mortality for Eastern Europe and other low income countries.

1.1.1.4.1 Stroke

Both the incidence and mortality risk of stroke have decreased in North America, most Western European countries, China, and Japan\textsuperscript{67, 68}. The decline was rapid, particularly in mortality risk, from the 1960s through the next two decades and continued
to decline in recent years\textsuperscript{64}. In Japan, the age-adjusted mortality risk declined almost 70\% in Japan between the time periods of 1965-1969 and 1995-1997\textsuperscript{69}. Similarly an almost 60\% decline was observed in Canada over the same period. However, increasing trends in mortality risks were observed for most Eastern European countries, with a 15\% increase in Albania\textsuperscript{69}.

Similar to the decreases observed in other developed countries, the stroke mortality risk in the US declined almost 60\% between the time periods of 1965-1969 and 1995-1997\textsuperscript{69}. Data from the National Center for Health Statistics (NCHS) indicate that the overall US stroke mortality risk decreased an additional 34.3\% from 1997 to 2007\textsuperscript{1}. However, this decline was not consistent across population sub-groups with greater declines observed for men compared to women and people $\geq 65$ years of age compared with younger age groups. Reports from the Framingham Heart Study indicate that incidence rate of stroke (first stroke per 1000 person-years) has been declining over time from the period 1950-1977 (men: 7.6; women: 6.2) to the period 1990-2004 (men: 5.3; women: 5.1)\textsuperscript{68}. However, others have reported that stroke incidence in 2005 was decreasing for whites but not blacks in US, especially among younger blacks\textsuperscript{1}.

\subsection*{1.1.1.4.2 Coronary Heart Disease and Myocardial Infarction}

The risks of death from CHD have decreased in the last several decades for most developed countries, particularly North America and western European countries\textsuperscript{69}. The greatest declines in mortality for the period 1988-1998 were observed in Denmark (men: -49\%; women: -46\%), Sweden (men: -43\%; women: -40\%), and the Netherlands (men: -39\%; women: -29\%)\textsuperscript{64}. Of countries that had an increase in CHD
mortality risk for the same period, the greatest increases were observed in Croatia (men: +62%; women: +61%), Kazakhstan (men: +56%; women: +36%), and Ukraine (men: +49%; women: +38%). Others have also reported that CHD incidence and mortality risks in Eastern and Central European countries were either increasing or not declining as rapidly as risks in higher income European countries\textsuperscript{70}. The increasing burdens of CHD, including MI, in developing and low income countries has been attributed, at least in part, to increasing longevity, urbanization, and lifestyle changes of those populations\textsuperscript{64}.

Like trends in stroke mortality, CHD mortality in the US declined by 64% (from 470 to 169.2 deaths per 100,000 population) during 1950-1999\textsuperscript{71}. Additionally, a 26.3% decline in mortality for the period 1998-2007 has been reported\textsuperscript{66}. However, the decreasing trends in CHD mortality are not consistent across population sub-groups, with slower declines for blacks than whites\textsuperscript{72, 73}. There has been less evidence of decline in incidence of CHD, compared with stroke. Data from the National Health and Nutrition Examination Survey (NHANES), for the periods 1988-1994 and 1999-2002, indicate that the 10-year risk of developing CHD changed little for adults 20 to 74 years of age\textsuperscript{74}. However, the rate of hospitalization for MI was found to have decreased by 23.4% from 2002 (1131 per 100,000 person-years) to 2007 (866 per 100,000 person-years) for US Medicare beneficiaries, with the degree of reduction greater for whites than other races\textsuperscript{75}. Thus, despite favorable declines, stroke and MI remain significant burdens in the US population, particularly for some populations.
1.1.1.5 Vital Statistics

Vital statistics refer to the information on vital, or life, events of the population, including: births (natality), deaths (mortality), disease (morbidity), marriages, and divorces. In the US, these statistics are maintained by the National Center for Health Statistics and can be freely obtained for specific geographic units (i.e. state, county, etc.) from the CDC, state or local health departments, and some not-for-profit agencies.

Mortality data are an important source of health-related data that can be analyzed for a variety of characteristics and at varying geographic levels. Although mortality data are useful and commonly used in epidemiological studies to assess health and its patterns, they are not without limitations. First, the accuracy of the cause of death given on a death certificate can be affected by errors made by physicians or in coding, differences in diagnostic criteria, issues arising when there are multiple causes of death, or errors in data entry. Lloyd-Jones et al. (1998) reported that death certificates overrepresented coronary heart disease as cause of death, particularly for older populations, and cautioned that its use in etiologic studies could potentially lead to a bias towards the null value. There is also concern that mortality data reflects past, rather than current, health needs.

However, mortality is often the most commonly available data for observational, population-based studies. These data are useful in investigating the characteristics of those dying from disease, which is important in identifying potential predictors of disease that aid in the understanding of the disease etiology. Mortality data are also
widely used as indicators of health status and healthcare needs in population health planning.

1.1.2 Socio-demographic determinants of stroke and myocardial infarction

1.1.2.1 Demographic factors

In the US, disparities in the burdens of stroke and MI exist across different demographic population sub-groups. Considerable evidence exists of higher stroke and MI risks for persons who are ≥ 65 years of age. It is estimated that 81% of those who die of CHD are ≥ 65 years of age. In 2002, the mean age at death from stroke was 79.6 years and only 11.9% of all stroke deaths occurred in persons < 65 years of age. The association between age and stroke or MI risk is modified by gender, with men dying at younger ages than women. The average age at first MI in the US is 64.5 years for men compared to 70.3 years for women, while the average age of first stroke onset was 71 years for men and 75 years for women. Women accounted for 60.6% of stroke deaths in 2007 due, in part, to the large number of elderly women in the US. Age-adjusted risks overwhelmingly show higher risk in stroke and MI prevalence, incidence, and mortality for men compared to women. Since women are typically older at first MI or stroke, the survival rate for women tends to be lower than for men until 75 years of age, after which median survival rates for MI (3.2 years) and stroke (2.2 years) are the same for both genders.

Population studies have found that race is an important determinant of stroke or MI risk with increased risks among blacks compared to other races, even after adjusting for gender and age. In 2007, the overall age-adjusted CHD mortality risk
was 126.0 per 100,000 population while race-specific risks were 191.6 for blacks, 165.6 for whites, 122.3 for Hispanics, 112.2 for American Indians or Alaska Natives, and 91.7 for Asians or Pacific Islanders\(^6\). Age-adjusted race-specific mortality risks for females were 121.5 for blacks, 94.2 for whites, 77.8 for Hispanics, 65.6 for American Indians or Alaska Natives, and 55.0 for Asians or Pacific Islanders. The overall age-adjusted stroke mortality risk was 53.5 deaths per 100,000 persons in 2006 while the race-specific risks were 48.1 for white males, 74.9 for black males, 47.2 for white females, and 65.5 for black females\(^2\). Age-adjusted, gender- and race-specific mortality risks for other races are scarce due to relatively large standard errors\(^8\). The gender-specific prevalence of both MI and stroke in 2008 were higher for blacks compared to other races\(^1\). Additionally, whites had higher mean age for both first stroke onset and death from stroke than other races\(^8\). The risks observed for blacks compared to other races were probably due to the higher prevalence of medical conditions (high blood pressure, high cholesterol, diabetes) that increase stroke and MI risk reported for blacks\(^1,2,83,88\).

### 1.1.2.2 Socioeconomic factors

It has become widely accepted that socioeconomic factors can impact health and interact with demographic characteristics to produce disparities in health across population groups. The associations of stroke and MI risks with income and education, the two most commonly assessed factors, are predominantly described as inverse, with increasing risk being associated with decreasing levels of income\(^18,89-93\) and education\(^8\). Composite measures of socioeconomic status (SES) or deprivation have also been reported to be inversely associated with stroke and MI risk\(^3,88-90,97-100\). These measures typically include a combination of factors, including: income, education
employment, occupation, poverty status, single parenthood, marital status, family size, housing value or housing ownership. However, these studies have primarily focused on socioeconomic characteristics at the individual level and stroke or MI risk. Research has overwhelmingly found that an individual’s health can be influenced by the socioeconomic and demographic characteristics of their neighborhood beyond their individual characteristics. 

The associations between neighborhood level socioeconomic factors and individual level stroke and MI (or CHD) risk have been investigated with some studies finding that median neighborhood income was positively associated with decreased mortality risk, even after adjustment for individual characteristics. However, many other studies investigated the association between socioeconomic factors and individual risk using composite measures of SES or neighborhood deprivation. Overall, lower SES or deprivation of the neighborhood was associated with increased incidence or mortality. These studies investigated neighborhood characteristics as contextual effects in multilevel models that sought to explain individual level risk. Thus, ecological studies are needed to investigate disparities in risk with the neighborhood as the unit of analysis since this is important in identifying high risk communities and targeting resources to address health disparities and improve population health. The association of socioeconomic characteristics with MI risk at the neighborhood level has been described by only a few studies using census tracts as the geographic unit of analysis. The studies found that MI risks were greater for low income or deprived neighborhoods. However, only one (income) or composite SES measures were used to represent all socioeconomic factors and their potential
association with MI risk. Evidence from recent research indicates that many socioeconomic factors are not interchangeable, and so the use of one measure or a composite measure ignores the complex relationships between the factors\textsuperscript{94,101}. Additionally, marked demographic (age, gender, race) differences exist for socioeconomic measures, particularly income and education\textsuperscript{90,101,113-115}. These complex interrelationships among socioeconomic and demographic factors imply that as many risk factors as realistically possible are needed to understand the relationships with stroke and MI risk, as well as identify disparities in risk for different populations.

1.1.2.3 Geographic clustering of socioeconomic and demographic factors

Although socioeconomic and demographic characteristics are known to be important determinants of stroke and MI, little is known regarding neighborhood clustering of these risk factors. Several studies from the US, Canada, and Europe have shown that socioeconomic factors can interact with demographic characteristics, such as race and sex, to produce different effects across populations for a number of health outcomes\textsuperscript{115-119}. Additionally, some studies have indicated that those characteristics most associated with poorer health outcomes tend to cluster, geographically, into socially deprived neighborhoods\textsuperscript{120-122}. Therefore, it has been recommended that research should focus on identifying disparities among sub-groups to better understand health needs at the neighborhood level and guide health programs geared toward reducing/eliminating these disparities\textsuperscript{83,101}.

Identifying unique neighborhood profiles in relation to the socioeconomic and demographic determinants of health is useful in understanding population health needs.
of neighborhoods. However, given the complex and multidimensional nature of socioeconomic and demographic factors, multivariate approaches for investigating the associations of these determinants with health outcomes have been recommended\textsuperscript{83, 94, 101, 120, 123}. Multivariate analyses, like principal component analyses (PCA) and cluster analyses, have been used in classifying neighborhoods based on their socioeconomic and demographic characteristics\textsuperscript{120}. These techniques identify groups of neighborhoods that have similar socio-demographic characteristics; however, standard PCA does not account for high correlations among variables, different variable scales (like income in dollars versus proportion of males), or outliers. Robust PCA, which assigns weights to outliers, has been shown to reduce the classification bias associated with the aforementioned issues\textsuperscript{124-126}.

Furthermore, some neighborhoods may be similar to more than one group of neighborhoods given the complex interrelationships between the socioeconomic and demographic risk factors and the multi-factorial nature of causation of stroke and MI. However, standard methods of cluster analysis, like k-means, assign neighborhoods/observations to only one group\textsuperscript{127}. Recent reports have recommended a generalization of the k-means algorithm (called fuzzy k-means cluster analysis) that allows observations (neighborhoods, in this case) to have a degree of belonging to two or more clusters, giving additional insight into the grouping of neighborhoods\textsuperscript{128-130}. Therefore, the use of multivariate strategies enables health professionals to utilize a needs-based approach to planning and service provision, based on unique neighborhood profiles and health needs, instead of using a “one-size-fits-all” strategy.
1.1.3 Geographic patterns of stroke and myocardial infarction

1.1.3.1 Geographic disparities in burdens

Geographic disparities in the burdens of CVD, including stroke and MI, have been observed for different geographic areas throughout the world. Globally the highest stroke and CHD burdens were observed in low income countries, like those in Eastern Europe, Central Asia, and North Africa. However, other studies have reported important regional variations even within high income countries. In Finland, a Northeast to Southwest gradient in decreasing risk was observed for both ischemic stroke and MI. Several other studies have found that northern parts of the country have higher burdens of CVD than southern areas in a number of countries, including: UK, France, Spain, and China. Geographic disparities in CHD and stroke have also been reported in Canada, Australia, and Japan. In most countries, including the US, greater disparities in stroke and MI burdens have been observed in poor populations for rural residents compared to urban residents.

The highest burdens of stroke and MI in the US have been reported in the southeastern states and in populations living in rural areas, particularly in the Appalachian region. Many areas of Appalachia, including parts of Tennessee, form a portion of the US “stroke belt”. It has been estimated that the overall average stroke mortality is 20% greater in the stroke belt compared to the rest of the US. While the burdens of stroke and MI have declined slightly in the Southeastern US and Appalachian region, including the stroke belt, these communities are still experiencing higher burdens of these conditions compared to the rest of the US.
population\textsuperscript{80, 81, 86}. Additionally, it has been reported that persons in the southeastern states and Appalachia had lower education levels, higher prevalence of diabetes and high blood pressure, and less health-care coverage than other regions\textsuperscript{81}.

In 2007, Tennessee had the highest age-adjusted stroke mortality risk in the US with 58.1 deaths per 100,000 persons compared to the national risk of 44.1 deaths per 100,000 persons\textsuperscript{1}. For coronary heart disease including MI, Tennessee ranks 4\textsuperscript{th} highest in the US with an age-adjusted mortality risk of 171.1 deaths per 100,000 persons compared to the national risk of 135.1 death per 100,000 persons\textsuperscript{1}. For MI only, the 2006 estimated annual age-adjusted mortality risk in Tennessee was 85.5 deaths per 100,000 persons, compared to the national risk of 58.9 per deaths per 100,000 persons\textsuperscript{149}. In addition to high risks for stroke and MI, rural areas, like Appalachia and many other regions of TN, have been reported to utilize health care services less than urban areas\textsuperscript{143, 150}. Rural areas may also experience unequal distribution of health care facilities compared to urban areas\textsuperscript{56} as well as geographical barriers to accessing care, including distance, travel time, and road conditions.

1.1.3.2 Neighborhood level spatial analyses

The geographic distributions of stroke and MI risks have been investigated at country, state, and county levels\textsuperscript{1, 69, 86, 142}. However, geographic disparities have been shown to exist even after adjusting for variations in common risk factors like demographic factors (race, age), socioeconomic measures (income, education), behaviors (smoking, physical activity), and other conditions (diabetes, hypertension)\textsuperscript{141, 145, 146, 151}. Therefore, it has been recommended that research should focus on
identifying disparities at lower geographic levels, like neighborhoods, to better understand health needs and thus, provide needs-based health services. Several US and Canadian studies have evaluated the performance of the different definitions of neighborhoods in the US (block groups, census tracts, and zipcodes) and Canada (dissemination areas, census tracts) and found them to be similar and concluded that census tracts accurately represent natural neighborhood boundaries. In the US, census tracts (CTs) are statistical subdivisions of a county that have between 2,500 and 8,000 persons, do not cross county boundaries, and are homogenous with respect to population characteristics, economic status, and living conditions. These characteristics enable census tracts to be good proxies of natural neighborhoods and therefore useful in describing neighborhood health disparities and population characteristics. Additionally, census tract level socioeconomic and demographic data are available to all states in the US through the US Census Bureau, as well as for populations in other countries like Canada (census tracts), the United Kingdom (postcode sectors), and Italy (census tracts) that approximately correspond to US census tracts. Thus, census tracts are an excellent choice to use as the unit of analysis in spatial analyses that seek to identify disparities in risk and their determinants at the neighborhood level.

Many studies have considered socioeconomic or demographic characteristics of populations in relation to stroke or MI, but historically most of these analyses have been done at state or county geographic levels. Recent studies have defined neighborhoods as census tracts or similar geographic units in the US.
Canada\textsuperscript{3, 108}, Sweden\textsuperscript{92, 93, 97}, and Italy\textsuperscript{157}. However, neighborhoods were not used as the unit of analysis in these studies. Rather, neighborhood characteristics were investigated as contextual effects in multilevel models that sought to explain individual level risk. A study from Spain was the only one identified that investigated neighborhood (census tract) level distributions of stroke and CHD mortality risks and the associations with neighborhood deprivation and environmental variables\textsuperscript{109}. They concluded that neighborhood deprivation was ecologically associated with neighborhood cardiovascular disease mortality. Thus, more ecological studies are needed to investigate potential socioeconomic and demographic predictors of spatial patterns and clustering of high stroke and MI risk with the neighborhood as the unit of analysis.

1.1.4 Geographic accessibility to stroke and MI care

1.1.4.1 Disparities in access to healthcare

Access to healthcare is a complex concept with multiple dimensions. There is a distinction between potential care (where the potential for receiving care exists) and realized care (where health services are actually utilized by the patient)\textsuperscript{160}. Accessibility to care describes the ability to get potential care which may be impeded by both spatial, like travel impedances and locations of healthcare services, as well as aspatial factors like the ability to pay\textsuperscript{160, 161}. The time sensitive nature of stroke and MI treatments implies that geographic accessibility to an appropriate healthcare center is critical in achieving favorable health outcomes. Additionally, the requirement for specialized equipment and training of health care professionals implies that not every healthcare facility can appropriately provide care for these patients\textsuperscript{46, 60}. Thus, the geographic
distribution of stroke and MI specialty centers in relation to population distribution will also impact access to outcome improving treatments.

In 2007, it was reported that 70% of CHD deaths and 54% of stroke deaths in the US occurred out of the hospital\(^1\). Long travel times to appropriate care have been implicated as a factor that contributes to high risk of out of hospital deaths in the US\(^{60}\), Canada\(^{140}\), and UK\(^{162}\). Likewise, it has been found that as the distance travelled increased, the utilization of healthcare services tended to decrease in the US\(^{163}\). Other studies from the UK\(^{56}\) and Canada\(^{89}\) have shown that MI patients living in closer proximity to cardiac centers are more likely to receive PCI than those living further away. In 2006, it was estimated that only 25% of US hospitals were capable of performing PCI\(^{61,164}\). Nallamothu and others (2006) found that almost 80% of the US population lived within 60 minutes driving time of a cardiac center (with PCI capabilities), and greater than 40% lived in areas in which the closest hospital was a cardiac center\(^{164}\). In Canada, travel time to stroke care was within 60, 90, and 120 minutes for 67.3%, 78.2%, and 85.3% of the population, respectively\(^{140}\). However, timely access was found to vary substantially across urban, suburban, and rural neighborhoods, with the greatest disparities in timely access found for rural populations.

Other US studies have reported similar disparities in rural areas with distance and travel time comprising the greatest determinants of access to cardiovascular health services\(^{143,163,165}\). Moreover, the unequal geographic distribution of healthcare centers compounds the disparities in access for rural populations\(^{56,160}\). The Joint Commission for the Accreditation of Healthcare Organizations (JHACO) has reported that nearly 800 hospitals (out of the approximately 5000 acute-care hospitals in the US) in 49 states
have achieved and maintain primary stroke center designation\textsuperscript{166}. Since the majority of stroke centers are large hospitals located primarily in urban areas, the limited availability of stroke centers across the US places an additional geographic barrier in receiving appropriate, timely care\textsuperscript{48, 167, 168}. Thus, it is important for population health planners to be aware of neighborhoods that experience disparities in geographic accessibility, so as to better target health service and research programs.

1.1.4.2 Measuring geographic accessibility to healthcare

Historically, most studies of geographic access to care have mainly used straight-line (or Euclidean) distances\textsuperscript{56, 169, 170}. Recently, travel time has been recognized as a better indicator of accessibility than distance to the healthcare facility since factors that impede travel (speed limits, road connectivity, and turn restrictions) can be accounted for in the analysis\textsuperscript{4-8}. Furthermore, it has been reported that people relate more easily to travel time than distance when making care seeking decisions\textsuperscript{171}. Some studies, though, have reported that Euclidean distance is often a good estimate of travel time\textsuperscript{6, 172}. However, others have found that at lower travel speeds it underestimated the travel time while overestimating at higher speeds\textsuperscript{173}. The weak correlations between distance and travel time have been attributed to the inability of Euclidean models (as estimates of access) to capture the complexity of street networks and associated travel impedances\textsuperscript{8, 173}. Thus, it is important to adjust for these factors in order to give a more accurate and realistic estimate of geographic access to healthcare services.
With increased availability of Geographic Information Systems (GIS), travel time estimates have become widely used in assessing geographic accessibility to care\textsuperscript{4,174,175}. However, studies assessing travel times have used average values for all segments of the road, regardless of actual impedances of each segment\textsuperscript{164,165,176}. Schuurman and others (2006) found that segment-by-segment information on speed limits and other travel impedances, such as turn impedances (like no u-turns) and connectivity, produced a more realistic model for calculation of travel time\textsuperscript{7}. Connectivity refers to the degree to which a street segment is accessible and connected to other streets, i.e. an interstate has higher connectivity than a rural state or county highway\textsuperscript{177}.

Network analysis is a methodology that models how resources or travel flows along a network, like roads, typically with the goal to reduce cost (which can be distance or time) of travel for a defined route while taking into account these travel impedances\textsuperscript{178}. The nodes of the network are defined as origins and destinations through which travel occurs along a route, such as between patients and hospitals. Assessing timely geographic accessibility to emergency care necessitates the use of a routing technique that seeks the shortest path, or least travel time between patients and hospitals by minimizing travel impedances. The most widely used method for computing shortest paths is Dijkstra’s algorithm\textsuperscript{177}, which minimizes travel time by favoring hierarchical routing techniques for road segment travel impedances\textsuperscript{179}. Hierarchical routing prioritizes road impedances by assigning lower cost (travel time) to roads with higher speed limits, fewer turn restrictions, and high connectivity (like interstate highways)\textsuperscript{179}.
Wang and others (2008) investigated access to health care services for late stage breast cancer patients in Illinois using network analysis to calculate travel time between the geographic zip code of where the patient lived and the hospital\(^8\). However this method assumes all patients living within a zipcode begin their travel at the geographic centroid, and thus does not consider the population distribution which can be sparse for low populations. Onega and others (2008) found that the population weighted centroid of a zipcode gave more reliable travel estimates for accessibility to cancer care across the US, particularly for sparsely population areas\(^5\). Similarly, Berke and others (2009) found that population weighted centroids were superior to geographic centroids\(^180\). However, they also found that polygon based methods, such that service area rings are created around a healthcare center based on travel time, were the best choices for estimating travel time without geographic or population restraints. Thus, network analysis, particularly when employed with the service area methodology, provides realistic, accurate, and efficient estimation of travel time. Given that the eligibility and efficacy of stroke and MI treatments is dependent on time-to-receipt of medical care, accurate estimations of travel time to appropriate hospitals are quite critical.

1.1.4.3 Improving geographic accessibility

The identification of disparities in timely geographic accessibility warrant strategies for improvement. Some studies have found that air transport of stroke and MI patients, particularly those in rural areas, improves access to time sensitive treatments\(^168, 181\). Thus, recommendations have been made that if ground travel exceeds 60 minutes, air transport should be considered for priority dispatch patients,
including those experiencing a stroke or MI\textsuperscript{42, 53}. Although many factors (such as cost, availability, maintenance issues, weather, terrain and protocols of transport) influence the use of air ambulances\textsuperscript{11}, several studies have reported that air transport is cost effective and can increase geographic accessibility to care\textsuperscript{181-183}.

The use of strategic telemedicine linkages between specialty centers (like stroke and MI), local hospitals with emergency rooms, and/or EMS providers has become increasingly popular for improving accessibility to outcome improving care\textsuperscript{48, 167, 184}. Telemedicine is defined as “the use of medical information exchanged from one site to another via electronic communications to improve patients health”\textsuperscript{185}. Thus, it encompasses all aspects of medicine that are practiced from a distance, including: phone, fax, electronic record sharing, electronic image and video distribution, and other interactive protocols that bring together patients and emergency care providers that are separated by distance. It has been estimated that 71% of the US population has geographic accessibility to a hospital with an emergency room within 30 minutes travel time and that 98% are within 60 minutes travel time\textsuperscript{176}. Thus, telemedicine between these hospitals and stroke/MI specialty centers could reduce the time to diagnoses, expedite decisions on transfers (by ground or air ambulance) to specialty centers, allow for increased provisions of triage or supplementary care for the patient, and improve the likelihood of receiving outcome improving treatments within the medically recommended times for a large majority of the US population\textsuperscript{11, 62}. Furthermore, pre-notification of suspected stroke and MI diagnoses by EMS has been reported to significantly reduce in-hospital delays associated with diagnosis, imaging, and treatment, as well as improve the odds of receiving time sensitive treatments\textsuperscript{45, 186, 187}. Given this evidence,
telemedicine linkages between stroke/MI specialty centers, ERs, and EMS have been overwhelmingly recommended as reliable and cost effective methods for improving geographic accessibility to stroke and MI care\textsuperscript{11, 53}.

Timely geographic accessibility to appropriate stroke and MI care is vital for receipt of effective treatment, but other factors are also important\textsuperscript{188}. Other studies have shown that many factors potentially affect the timely access to appropriate medical care, including: (a) socioeconomic factors\textsuperscript{84, 189, 190}, (b) the recognition of symptoms and decision to seek emergency care by patients or caregivers/bystanders\textsuperscript{26, 35}, (c) issues involved in the use of emergency medical services\textsuperscript{9} (like correct diagnoses\textsuperscript{16} and transportation policies related to pre-hospital triage\textsuperscript{164, 186}), and (d) hospital arrival to treatment delays\textsuperscript{44, 46, 49}. Thus, complete analysis of timely access to care must assess not only geographic accessibility but also factors associated with pre-hospital delays, use of EMS, and EMS transport delays.

1.1.5 Role of Emergency Medical Services (EMS) in timely access to stroke and MI care

1.1.5.1 Delays in receiving emergency care for stroke and MI

The emergency nature of treatments and severity of stroke and MI events requires that patients receive the care they need in appropriate health facilities in a timely manner so as to achieve good health outcomes. Current North American and European guidelines recommend intravenous thrombolytic treatment within 3 hours of the onset of stroke symptoms\textsuperscript{42, 44, 158}, while time from first medical contact to PCI for MI patients should be 90 minutes or less\textsuperscript{53}. Additionally, recommendations for the
establishment and direct transfer of stroke and MI patients to specialty centers stress the importance of rapid transport and treatment\textsuperscript{42,45}. However, transport delays to appropriate care can hinder stroke/MI patients from receiving treatments with the recommended therapeutic windows.

There are two general types of delays that affect the receipt of time sensitive stroke and MI treatments: pre-hospital and in-hospital delays. The pre-hospital delay refers to the time period from the onset of stroke/MI symptoms to arrival at an appropriate hospital, incorporating the time required to recognize symptoms, decide to seek medical attention, and transport to the hospital\textsuperscript{11}. The in-hospital delay generally refers to delays during the time from arrival at the hospital to receipt of treatment and includes components, such as the times for: triage, diagnosis, imaging, and treatment administration\textsuperscript{12}.

Given the 3 hour window required for stroke treatment, it has been recommended that pre-hospital delay should not exceed 2 hours, while in-hospital delays should not exceed one hour\textsuperscript{11,12}. Several studies have reported median length of pre-hospital delays for stroke patients to range from 1.5 – 3 hours in a number of countries, including: Switzerland\textsuperscript{191}, Germany\textsuperscript{192,193}, Italy\textsuperscript{194}, Turkey\textsuperscript{195}, Japan\textsuperscript{190}, Canada\textsuperscript{196} and the US\textsuperscript{15,189,197,198}. However, fewer than half of all stroke patients arrive within 2 hours of the onset of symptoms in each of the studies. In the US, data from the Paul Coverdell National Acute Stroke Registry (including 142 hospitals in Georgia, Illinois, Massachusetts, and North Carolina) indicated that 65% of the patients who arrived at the hospital within 2 hours of the onset of symptoms received imaging within
one hour, significantly more than those (39%) whose pre-hospital delay was > 2 hours\textsuperscript{189}. Other studies throughout the world have reported that stroke patients with extended pre-hospital delays (2 or more hours) were more likely to have extended in-hospital delays as well\textsuperscript{162, 186, 191, 194}.

While there are no clear recommendations for the acceptable length of pre-hospital delays for MI patients, it has been reported that MI outcomes are improved by up to 50% when PCI is administered within 60 minutes and 23% within 180 minutes of the onset of symptoms\textsuperscript{58, 59}. Thus, as the length of pre-hospital delay decreases, the odds of a better outcome increase. Unfortunately, the median length of pre-hospital delays for MI patients has been reported to be 2.0 – 3.5 hours for studies in Turkey\textsuperscript{58}, Finland\textsuperscript{199}, Canada\textsuperscript{72, 187}, and the US\textsuperscript{9, 10, 13, 17, 200}. It has been reported that MI patients with prolonged pre-hospital delay (>2 hours) were less likely to receive PCI within the recommended 90 minutes of hospital arrival\textsuperscript{201}. Thus, for both stroke and MI, pre-hospital delays have a significant impact on in-hospital delays.

Patient demographic factors have been reported by some studies, both in the US\textsuperscript{36, 104, 200, 202} and abroad\textsuperscript{29, 58, 72, 190}, to be associated with increased delays in receiving MI care. These demographic factors include older age (>65 years), female gender, and minority race/ethnicity. However, the relationships seem to be less clear for stroke patients with some studies finding longer delays for these characteristics\textsuperscript{15, 189}, but others not\textsuperscript{190, 191, 197}. The observed differences are possibly due to the population dynamics such that a greater proportion of stroke patients are female and older and that minority populations experience language barriers. Other factors associated with delays
(particularly pre-hospital delays) include: history of past stroke/MI\textsuperscript{200,203}; co-morbidities (i.e. diabetes, hypertension, or other CHD related illnesses)\textsuperscript{191,200,202}; type or severity of the attack\textsuperscript{58,190,191}; being/living alone at onset\textsuperscript{19,190,197}; awakening with symptoms\textsuperscript{19,190,197}; and transfer from another hospital\textsuperscript{19,190,197}. However, like demographic characteristics, significant associations of these factors with delays were not consistent. However, the utilization of EMS has consistently and overwhelmingly been shown to reduce both pre-hospital and in-hospital delays in receiving treatments for stroke\textsuperscript{12,189-191,197,203} and MI\textsuperscript{58,72,104,200,202}.

Utilization of EMS as the first medical contact and transport to specialty centers is a critical component of North American and European recommendations in which the goals are reduce delays in receiving stroke and MI care\textsuperscript{13,204}. Approximately 60% of out-of-hospital cardiac deaths in the US are treated by EMS personnel\textsuperscript{1}. Thus, EMS is a key contributor to improved outcomes and so EMS personnel must be proficient in their ability to recognize, assess, manage, triage, and transport stroke and MI patients\textsuperscript{11}. Because of the time sensitive nature of treatments for stroke and MI, priority medical dispatch has been designated for all suspected stroke and MI emergency calls\textsuperscript{53,204} and has been associated with significantly lower risks of death before arrival to a hospital\textsuperscript{199}. Therefore, it is important for emergency dispatchers to correctly identify suspected stroke and MI cases to avoid inappropriate priority assignment\textsuperscript{16}. Patient assessment, management of symptoms, and triage at the scene by EMS personnel have also been associated with both shorter delays and improved outcomes\textsuperscript{9,58,200}. It has been suggested that the most important part of assessment by EMS in relation to stroke and MI is the determination of the onset of symptoms, in addition to patient medical
history\textsuperscript{11}. This key information is critical in determining the eligibility for time sensitive treatments and developing effective care protocols to improve patient outcomes.

The final role of EMS in emergency stroke and MI care is the rapid transport of the patient to an appropriate hospital. In addition to medical attention and assessment, transport by EMS is advantageous over self-transportation since, when using emergency lights and sirens, they are allowed to exceed legal speed limits and have the right-of-way through intersections\textsuperscript{204, 205}. Additionally, many studies have reported that patients arriving at hospital emergency rooms by ambulance are assessed by physicians significantly faster than those arriving by personal transportation\textsuperscript{205}. This may be in part due to the pre-notification protocol followed by EMS that allow emergency room physicians and staff to be prepared for the patient on arrival\textsuperscript{11}.

1.1.5.2 Components of pre-hospital delays

Several studies have suggested that the pre-hospital delay is the source of the longest delay for acute patients\textsuperscript{12, 197, 206}. Pre-hospital delays for stroke and MI patients have been routinely found to have significant impacts on timeliness and eligibility of patients to receive time sensitive, emergency treatments associated with better health outcomes\textsuperscript{12, 197, 206}. Many studies have investigated the impact of pre-hospital delays on receiving treatment for stroke\textsuperscript{12, 189-191, 197, 203} and MI\textsuperscript{58, 72, 104, 200, 202} and found that utilization of Emergency Medical Services (EMS), among other factors, is associated with reduced pre-hospital delay. These studies considered pre-hospital time as one interval, from the onset of symptoms to arrival at the hospital; however, it has been suggested that pre-hospital delays should be further sub-divided into decision delays
and transport delays so as to better identify disparities and target interventions to reduce these delays\textsuperscript{12, 104}.

1.1.5.2.1 Decision Delays

Decision delays include delays in the lengths of time for stroke or MI symptom recognition and the decision to seek medical care, including choosing to utilize EMS for transport. Decision delays have been suggested to be the longest and most variable component of pre-hospital delays\textsuperscript{12, 15, 202, 206}. The ability to recognize symptoms has been implicated as the most important factor in delay for both stroke and MI\textsuperscript{59, 207}. According to the Behavioral Risk Factor Surveillance Survey (BRFSS), only about 27% of people were aware of the 5 most common MI warning signs and symptoms\textsuperscript{35}. The proportion of the population that were aware of symptoms and understood the need to call 911 was higher among non-Hispanic whites (30.3%), women (30.8%), and those with higher levels of education (33.4%) compared to non-whites (15.3%), men (22.5%), and those with less than a high school education (15.7%). Similarly for stroke, awareness of the 5 warning symptoms for stroke was 38% among all respondents and highest for non-Hispanic whites (41.3%) versus non-whites (28.2%), women (41.5%) versus men (34.5%), those with college degrees (47.6%) versus those with less than a high school education (22.5%)\textsuperscript{26} and those < 65 years of age (47%) versus those ≥ 65 years (28%)\textsuperscript{208}. These findings are similar to other studies that have assessed knowledge of stroke and MI symptoms\textsuperscript{26, 53}.

Some recent studies have reported that knowledge of stroke and MI symptoms was not significantly associated with seeking medical care\textsuperscript{28, 205, 209}. Instead, the studies
found that the most commonly reported reasons for not seeking medical attention were related to self-efficacy (belief in one’s capability to perform a task successfully) and included: not believing the symptoms were serious, thinking the symptoms would just go away, and being concerned about the cost of care, particularly calling EMS. Additionally, it has been found that the severity of symptoms and not living/being alone at the onset of symptoms were important factors in the decision to seek medical care. One study estimated that someone other than the patient made the decision to seek medical care in 66% of stroke cases. A number of studies have reported that women and older patients are more likely to utilize EMS for stroke and MI emergencies compared to men and patients < 65 years of age. However, other studies have found no associations with patient demographic characteristics. Interestingly, patients who were employed and had higher levels of education were found by one study to be less likely to utilize EMS. The reasons for this association were unclear. Using data from the National Registry of Myocardial Infarction 2, Canto and others (2002) found that, in addition to female gender and older age, the following variables were significantly associated with increased EMS use: white race, Medicare/Medicaid payer status, living in Southern US, higher number of co-morbidities (diabetes, hypertension, smoker, hyperlipidemia), and prior history of MI. Additionally, several studies have reported that more severe symptoms and critical presentations for stroke and MI are significantly more likely to be taken to the hospital by EMS. Interestingly, many factors that increase the likelihood of using EMS are also reported to be associated with reduced decision delays, and overall reduced pre-hospital delays. Thus, use of EMS may possibly be the
most, or one of the most, important factors in reducing pre-hospital delays. It has been suggested that if the interventions targeted to reduce decision delays are successful over time, then delays associated with EMS transport will become increasingly important to identify and remediate\textsuperscript{198}.

1.1.5.2.2 Transport Delays

Transport delays describe delays in the length of time from when the patient calls 911 to when the patient arrives at the hospital. Since EMS play a critical role in providing rapid transport of acute stroke and MI patients, it is important to investigate the specific time intervals involved in transport in order to better recognize when and where delays occur so that they might be more effectively addressed. These intervals include the response time, time spent on-scene, and the time required to travel to the hospital. These intervals are generally defined as follows: response time describes the time from EMS dispatch to arrival on-scene where the patient is located; on-scene time refers to the time from EMS arrival at the scene to EMS departure with patient from the scene; and travel time is the time elapsed from leaving the scene to arrival at the hospital\textsuperscript{16, 17, 198}. Some studies have defined response time as the time from the 911 call to arrival on-scene\textsuperscript{15, 211}. However, several studies have reported that these data are often missing for nearly half of the records\textsuperscript{72, 212}. One study found that the median elapsed time between 911 call and EMS dispatch was 2 minutes, and suggested that the difference in using the time of 911 call or time of EMS dispatch would not significantly impact estimates for the response or total times\textsuperscript{198}. 

39
Unfortunately, only a few recent studies have described the specific time intervals associated with EMS response and transport for stroke and MI. Some studies have reported median response times to stroke or MI emergency calls to be 5 minutes, 5.5 minutes, 6 minutes, 7.5 minutes, and 8 minutes. Shorter response times have been associated with both priority dispatch and use of lights and sirens in both urban and rural settings. Other studies have found both higher call volumes and traffic congestion to be important for describing longer response delays in urban systems. Median on-scene times have been reported for stroke (13 minutes) and MI (14.5 minutes). However, these were lower than median on-scene times, ranging 18 to 20 minutes, reported for suspected stroke patients by other studies. It was unclear whether these differences were related to population variation or diversity in EMS protocols related to assessment, triage, or treatment of patients on the scene. Differences in EMS on-scene times have also been suggested to possibly reflect varying levels of efficiency, experience, or attitude of EMS personnel serving different populations. One study also observed longer on-scene times for more serious patients that required transport by Advanced Life Support (ALS) services which have intubation and ventilator capabilities, compared to Basic Life Support (BLS). Moreover, once at the hospital, ALS patients were seen by a physician twice as fast. Conversely, it has been reported that the reasons for longer on-scene times are not only due to the seriousness of the patient condition, but other reasons are unclear. The median travel time from the scene to the hospital was reported by two studies in urban areas as 11 minutes for stroke patients. Reports of specific travel times for MI patients were not identified in the literature. Given the limited literature
available on specific EMS transport intervals, additional studies of different populations, geographic areas, and EMS providers are needed to improve understanding of this component of pre-hospital delays.

1.1.5.3 Improving pre-hospital delays

Despite the clear evidence of the benefit of utilizing EMS for transport to emergency stroke and MI care, up to half of those experiencing MI\(^9\) and stroke\(^203\) choose to self transport (including transport by family members, coworkers, or others). Thus, it has been suggested that, in addition to increasing awareness/knowledge of symptoms, community interventions should also focus on ways of increasing patient or family/bystander self-efficacy and motivation for seeking medical care, with strong encouragement to utilize EMS\(^{28,198,214}\). Furthermore, it has been recommended that educational interventions also include stressing the benefits, which include reduced pre-hospital delays and increased likelihood of receiving timely treatments, of direct transport to a stroke or cardiac specialty center\(^{45,187}\). While some studies have indicated that direct transport to a specialty hospital may not significantly reduce the overall pre-hospital delay (symptom onset to arrival at the hospital), in-hospital delays (imaging, diagnosis, and treatment) are significantly reduced so that total time to treatment is shorter and patients are more likely to receive time sensitive treatment that improve health outcomes\(^{17,46,204}\). Thus, recommendations for pre-hospital protocols to incorporate EMS bypassing non-specialty centers have been made\(^{11,42,58,187}\). Based on these recommendations, patient/family member education on the benefits of choosing specialty centers, in addition to reviews of EMS transport protocols to include bypass options have been shown to be associated with improved outcomes\(^{11,42,204,215}\).
Distance and travel time have been reported to be the greatest determinants of access to health services for populations, particularly those living in rural areas\textsuperscript{143}. Several studies have found that use of air transport for patients in rural areas decreased delays and increased the likelihood of receiving time sensitive treatments\textsuperscript{168, 181, 182}. It has been recommended that air transport should be used if travel time by ground ambulance to the specialty center exceeds 1 hour\textsuperscript{11}. Additionally, strategic telemedicine linkages between stroke or cardiac centers and EMS providers have been recommended to improve access for patients living far away from care and/or in rural areas. Telemedicine programs have been shown to be feasible, reliable and improve outcomes, particularly for rural patients\textsuperscript{184, 216}. Additionally, with guidance from specialty centers, EMS personnel may be able to reduce assessment and treatment times of patients at the scene.

Identifying the characteristics of patients with extended EMS delays can also inform both local public health officials and EMS for efforts to reduce delays and improve the likelihood of receiving timely emergency treatments. Thus, studies that identify delays in specific EMS transport intervals, and the factors affecting them, are needed to provide vital pieces of information for local health initiatives that seek to improve accessibility to health services and outcomes for stroke and MI patients.
CHAPTER 2

2.0 Identifying Unique Neighborhood Characteristics to Guide Health Planning for Stroke and Heart Attack: Fuzzy Cluster & Discriminant Analyses Approaches

Ashley Pedigo, MS^1; William Seaver, PhD^2; and Agricola Odoi, BVM, M.Sc., PhD^1

^1 The University of Tennessee, Department of Comparative Medicine, 2407 River Dr, Knoxville, Tennessee 37996-4543

^2 The University of Tennessee, Department of Statistics, Operations and Business Science, 342 Stokely Management Center, 916 Volunteer Blvd., Knoxville, TN 37996

This chapter is a manuscript that has been published in PLoS One journal. (PLoS ONE 2011, 6(7): e22693. doi:10.1371/journal.pone.0022693).

My contributions to this paper included data preparation, analysis, interpretation of results, gathering and reviewing of literature, formulation of discussion topics, as well as drafting and editing the manuscript.
2.1 Abstract

Socioeconomic, demographic, and geographic factors are known determinants of stroke and myocardial infarction (MI) risk. Clustering of these factors in neighborhoods needs to be taken into consideration during planning, prioritization and implementation of health programs intended to reduce disparities. Given the complex and multidimensional nature of these factors, multivariate methods are needed to identify neighborhood clusters of these determinants so as to better understand the unique neighborhood profiles. This information is critical for evidence-based health planning and service provision. Therefore, this study used a robust multivariate approach to classify neighborhoods and identify their socio-demographic characteristics so as to provide information for evidence-based neighborhood health planning for stroke and MI.

The study was performed in East Tennessee Appalachia, an area with one of the highest stroke and MI risks in USA. Robust principal component analysis was performed on neighborhood (census tract) socioeconomic and demographic characteristics, obtained from the US Census, to reduce the dimensionality and influence of outliers in the data. Fuzzy cluster analysis was used to classify neighborhoods into Peer Neighborhoods (PNs) based on their socioeconomic and demographic characteristics. Nearest neighbor discriminant analysis and decision trees were used to validate PNs and determine the characteristics important for discrimination. Stroke and MI mortality risks were compared across PNs. Four distinct PNs were identified and their unique characteristics and potential health needs described. The highest risk of stroke and MI mortality tended to occur in less affluent PNs located in urban areas, while the suburban most affluent PNs had the lowest risk.
Implementation of this multivariate strategy provides health planners useful information to better understand and effectively plan for the unique neighborhood health needs and is important in guiding resource allocation, service provision, and policy decisions to address neighborhood health disparities and improve population health.

2.2 Introduction

Stroke is the third most common cause of death and leading cause of debilitation in the US\textsuperscript{24}. Coronary heart disease, including myocardial infarction (MI), accounts for nearly 1 out of every 6 deaths in the US\textsuperscript{217}. These health conditions are serious burdens to the US health system with prevalence estimates of 2.9% and 3.6 % and annual costs estimated at $73.7 and $177.1 billion for stroke and MI, respectively\textsuperscript{217}.

These burdens vary by demographic, socioeconomic, and geographic factors. Several studies have reported geographic variations in prevalence and mortality of stroke and MI with the highest risks being reported in southeastern US\textsuperscript{24, 81, 159} and in populations living in rural areas\textsuperscript{81, 143, 159}. Tennessee ranks 3rd and 4th highest in the US for stroke and coronary heart disease including MI, respectively\textsuperscript{217}. The 2006 annual age standardized mortality risks of stroke and MI in Tennessee were 67.5 and 85.5 deaths per 100,000 persons, compared to the national risks of 53.5 and 58.9 deaths per 100,000 persons, respectively\textsuperscript{149}. Many rural areas of Tennessee, including the Appalachian Region, form part of the “stroke belt” of the US\textsuperscript{81, 83, 159}. Populations that are male\textsuperscript{24, 81, 84, 99}, black\textsuperscript{81, 83, 88}, or 60-65 years of age and older\textsuperscript{24, 25, 81, 94} have higher stroke or MI prevalence and mortality than other demographic groups. The relationships with socioeconomic factors are predominantly described as inverse with increasing risk
of stroke or MI being associated with decreasing levels of income, education, and composite measures of socioeconomic status (SES) or deprivation that include factors like employment, occupation, single parenthood, marital status, housing value or housing ownership, to mention but a few.

Although socioeconomic, demographic, and geographic factors are known to be important determinants of stroke and MI, little is known regarding the clustering of these risk factors in neighborhoods. Research has overwhelmingly found that an individual’s health can be influenced by the socioeconomic and demographic characteristics of their neighborhood beyond their individual characteristics. Clustering of these determinants of health across neighborhoods inevitably impacts health outcomes and thus health planning. Therefore, research should focus on identifying disparities among sub-groups to better understand health needs at the neighborhood level and guide health programs geared toward reducing/eliminating these disparities. Moreover, the multi-factorial nature of disease determinants implies that as many risk factors as reasonably possible need to be included for the most realistic analyses. Thus, the analysis of the complex and multidimensional nature of socioeconomic, demographic, and geographic risk factors requires the use of multivariate approaches.

With these issues in mind, the objective of this study was to classify neighborhoods in East Tennessee (using multivariate techniques) based on demographic, socioeconomic, and geographic risk factors for stroke and MI to better identify and understand population characteristics and health needs at the neighborhood level to support population health planning and policy. Many of these risk
factors are expected to be interdependent, such that clusters based on these characteristics will not be mutually exclusive. Thus, this study uses multivariate methods (robust principal components analysis, fuzzy cluster analysis, discriminant analysis, and classification trees) to address this issue.

2.3 Methods

2.3.1 Study area population

This study was performed in the East Tennessee Appalachian region, an area that includes eleven counties: Claiborne, Cocke, Grainger, Greene, Hamblen, Hancock, Hawkins, Jefferson, Knox, Sevier, and Union. These counties were chosen because of their high risk of stroke and/or MI. This area has a population of just over 857,000 and includes 168 census tracts (CTs). Census tracts are statistical subdivisions of a county that have between 2,500 and 8,000 persons, do not cross county boundaries, and are homogenous with respect to population characteristics, economic status, and living conditions\textsuperscript{155}. The US Census Bureau further describes the design of CTs to provide a relatively stable set of geographic units that allow statistical comparisons of population characteristics between decennial censuses. Additional information on how the boundaries of the CTs are determined can be found at the US Census Bureau\textsuperscript{218}. Census tracts have been shown to be good proxies of natural neighborhood boundaries and are thus useful in describing neighborhood population characteristics, as well as health needs\textsuperscript{120,219}. Furthermore, other studies of socioeconomic characteristics in the US have used census tracts to represent neighborhoods\textsuperscript{119,153}. Given these characteristics, CTs were used in this study to represent neighborhoods as the
geographical unit of analysis and therefore all analyses, results, and inferences were made at this population level.

2.3.2 Data acquisition

2.3.2.1 Population characteristics

Census tract level socioeconomic, demographic, and population data for the study area were obtained from the census 2000 summary file 3\textsuperscript{155}. Since these data are available in the US only through the decennial census, the 2000 data was deemed best suited to match the disease data (1999-2007). The variables considered in the study were those that have been reported in the literature\textsuperscript{89, 90, 97, 99} to be associated with risk of stroke and MI either independently or as part of a composite measure. They include: race, gender, age (40-49, 50-59, 60-64, 65 years and older), marital status (for population 15 years and older), population living below poverty, per capita income, educational attainment (less than high school, high school graduate, some college, bachelor degree, or graduate degree), single parent households, housing ownership, housing value, and the urban/rural classification of each neighborhood.

2.3.2.2 Mortality data

Mortality data covering the period 1999-2007 were obtained from the Tennessee Department of Health and were used for comparison of mortality risks across neighborhoods. Stroke and MI mortality cases were defined using ICD 10 codes I60-I69 and I21-I22, respectively. Mortality case addresses were geo-coded using Batch Geocode\textsuperscript{220} and imported into ArcGIS 9.3\textsuperscript{221}. Point-in-polygon join was used to connect
the mortality data to the census tract cartographic boundary files obtained from the U.S. Census Bureau\textsuperscript{155}.

2.3.3 Data analysis

2.3.3.1 Data management

With the exception of income and housing value, all variables were analyzed as the proportion of the population in each neighborhood. One neighborhood in Knox county, that had a population of 232 and included a mental health facility, was removed from the analysis due to missing data values for many variables.

2.3.3.2 Robust principal components analysis (PCA)

When the ultimate goal of the analysis is to identify group structure within data using cluster analysis based on many variables, principal components analysis (PCA) can be used to reduce the dimensionality of the data\textsuperscript{125}. This process reduces bias in clustering since substantial interdependencies, or high correlations, often exist among the many variables being considered. However, outliers can also bias the orthogonal linear combinations, as well as the cluster formation. Thus in this study, robust PCA in NCSS\textsuperscript{222} was performed to reduce the dimensionality of the 22 strongly interdependent socioeconomic and demographic variables and to decrease the influence of outliers prior to subsequent cluster analysis\textsuperscript{127, 223}. This method uses weights that are inversely proportional to the degree to which an observation is outlying\textsuperscript{124}. The robust PCA was performed on the correlation matrix, which has values standardized by variance for the whole dataset, instead of just one variable, since major differences in variability and scale were expected amongst these variables\textsuperscript{224}. Kaiser’s eigenvalue cutoff of 1.0 was
used to retain five components that accounted for 80% of the variation. The five retained component scores (with a mean of zero and variance of 1.0) were multiplied by the square root of their eigenvalues to retain maximum-ordered variances. This was done to ensure that principal components with high variances would have more weight in subsequent cluster analysis.

2.3.3.3 Fuzzy cluster analysis

Clustering techniques can be used on the robust PCA scores to find groups or clusters that contain observations with similar socioeconomic and demographic characteristics. Typically, there is a hard or crisp assignment of observations into clusters, such as with k-means. However, a generalization of the k-means clustering algorithm (called fuzzy k-means clustering) allows observations to have a non-crisp assignment to clusters. This non-crisp assignment allows observations to have a degree of belonging to two or more clusters, i.e., some observations may partly belong to other clusters.

The fuzzy K-means clustering algorithm is based on minimizing the following objective function:

$$ J = \sum_{i=1}^{n} \sum_{g=1}^{C} u_{ig}^m d_{ig}^2 $$

where \( u_{ig} \) is the degree of belonging of the \( i \)th observation to the \( g \)th cluster, \( m \) is the fuzzifier (\( m \geq 1 \): \( m=1 \) or close to 1 gives a crisp solution; and as \( m \) increases greater than 1, the solution becomes more and more fuzzy with each increment); and \( d_{ig}^2 \) is a Euclidean measure of distance based on the robust principal component scores. With the computation of the degrees of belonging, there is a re-estimate of the cluster centroids in a fuzzy way according to the following relationship:
In this case, \( i = 1, 2, \ldots, n \) observations, \( g = 1, 2, \ldots, r \) clusters, and \( y_i \) is the robust principal component score in this study. There is an iterative computation of Euclidean distances relative to the cluster centroids. New values of \( u_{ig} \), which minimize \( J \) (equation (1)) for given distance measures, can be computed by:

\[
u_{ig} = \left( \sum_{g=1}^{c} \frac{d_{ij}/d_{ig}}{d_{ij}} \right)^{2/(m-1)}^{-1}
\]

(3)

where \( i = 1, 2, \ldots, n \) observations, \( j = 1, 2, \ldots, n \) observations, \( g = 1, 2, \ldots, r \) clusters.

The minimization of equation (1) with respect to the centroids (equation 2) and the degree of belonging (equation 3) continues until the differences between successive membership matrices are less than some pre-assigned value (in this study the value is 0.001).

The fuzzy clustering strategy allows a sensitivity analysis on cluster structure as well as assessment of the uniqueness of each observation to a particular cluster by varying the fuzzifier and the number of clusters. The fuzzifier is increased typically by small amounts from 0.10 up to 0.25. Some data sets will be extremely sensitive to changes in the fuzzifier and others not\(^{130}\). The tremendous amount of information provided by the degree of belonging information can be summarized using either (a) Dunn’s normalized partition coefficient (FPU), with values closer to one reflecting hard partition and values closer to zero fuzzy solutions; or (b) the normalized average squared error (DPU), where values closer to zero indicate hard solutions and values near one are fuzzy solutions\(^{128, 226, 227}\). The solution that will provide the best insight to the cluster structure of the data, in this case the population profiles of neighborhoods (observations), should be neither too hard nor too fuzzy\(^{127}\). This is addressed with the
fuzzy indices, FPU and DPU, and with the validation of classification into each cluster with discriminant analysis with the original variables. A more comprehensive discussion on the selection of a fuzzy solution (i.e. number of clusters and fuzzifier), is available in Seaver, et al (2004)\textsuperscript{129}.

In this study, fuzzy cluster analysis was performed in NCSS\textsuperscript{222} using the principal component scores from robust PCA of the population characteristics to identify peer neighborhoods (PNs). In order to identify the solution with the most distinction between PNs, a sensitivity analysis was performed by varying the fuzzifier from 1.0 to 1.6 and the number of clusters from 3 to 6, based on the suspected group structure of the study area.

2.3.3.4 Validation of Peer Neighborhoods (PNs)

After identifying PNs, it was important to assess accuracy of PN identity, identify misclassified neighborhoods, and determine the characteristics most important for separating the neighborhoods. This was done using: (a) non-parametric nearest neighbor discriminant analysis (DA) with two neighbors (k=2) in SAS 9.2\textsuperscript{228} and (b) classification and regression tree (CART) in AnswerTree 3.0\textsuperscript{229}. The performance of the DA was evaluated by estimating error rates (or probabilities of misclassification) in the classification of neighborhoods using cross validation (or jack-knife) method where \( n-1 \) neighborhoods were used to predict the classification of the neighborhood held out\textsuperscript{230}.

The means of socioeconomic and demographic variables were compared in each PN between misclassified and non-misclassified neighborhoods using Hotelling’s two sample t-test to investigate characteristics of the misclassified neighborhoods.
Randomization tests of significance were used since the assumption of multivariate normality was not met\textsuperscript{124, 231}.

When distributional assumptions are uncertain and more flexibility is needed, classification (decision) trees can be used to predict the assignment of observations into discrete groups based on one or more predictor variables. One particular advantage of classification trees is that they readily lend themselves to being displayed graphically, making them easier to interpret and use. Classification trees construct hierarchical decision rules in the form of binary trees starting with the original classification for the data and ending with somewhat homogeneous groups of observations. Computationally, decisions must be made on: the criteria for predictive accuracy, selecting splits, stopping point for splitting, and selecting the "right-sized" tree. However, the goal in this study was simplicity of the tree and ease in comparison with the traditional nearest neighbor results to validate the uniqueness of identified PNs. Thus, CART\textsuperscript{229} with binary splits at four levels was used.

2.3.3.5 Comparison of mortality between Peer Neighborhoods (PNs)

Annual age-adjusted mortality risks of PNs for stroke and MI were calculated by the direct standardization method in Stata 10\textsuperscript{232} using the 2000 Tennessee population as the standard population. A two sample test of equality of proportions for each PN pair was performed and the p-values adjusted for multiple testing using the Simes method\textsuperscript{233}. Spatial distribution of identified PNs were displayed in ArcGIS 9.3\textsuperscript{221}.
2.4 Results

2.41 Robust principal components analysis

The five retained components from robust PCA explained 80% of the variation in the data. The first component explained 34% of the variation and was primarily composed of socioeconomic (education, income, housing value, employment) and geography (urban versus rural) variables (Table 2.1). Socio-demographic variables (race, single parent families, married population, and home ownership) were heavily loaded onto component 2, which explained the next largest portion (26%) of the variation. Component 3 was also a demographic perspective of the data, with average family size and age primarily loaded on this component. Variables for race and gender were also important for component 3. Components 4 and 5 have less clear interpretations. Component rotation, using Varimax rotation (results not shown), did not change the loadings or interpretation, except to make a few variables more distinct for components 4 (race and age) and 5 (gender and rural geography). A regular PCA on the current data (results not shown) yielded the same percentage of total variation explained (80%); however, the first three components explained less variation individually compared to the robust (Table 2.1). By adjusting for outliers in the robust PCA, the variation is more distinctly partitioned in the components, allowing for better interpretation.

2.4.2 Fuzzy cluster analysis results

In sensitivity analyses, the solution that will provide the most insight into the data is one that has a higher FPU value and lower DPU value without being too close to a
Table 2.1: Component Loadings from Robust Principal Components Analysis for Socioeconomic and Demographic Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of variation explained</td>
</tr>
<tr>
<td>Living in urban area</td>
<td>34%</td>
</tr>
<tr>
<td>Living in rural area</td>
<td>26%</td>
</tr>
<tr>
<td>White race</td>
<td>9%</td>
</tr>
<tr>
<td>Black race</td>
<td>6%</td>
</tr>
<tr>
<td>Male</td>
<td>5%</td>
</tr>
<tr>
<td>Age 40-49 years</td>
<td></td>
</tr>
<tr>
<td>Age 50-59 years</td>
<td></td>
</tr>
<tr>
<td>Age 60-65 years</td>
<td></td>
</tr>
<tr>
<td>Age over 65 years</td>
<td></td>
</tr>
<tr>
<td>Single parent families</td>
<td></td>
</tr>
<tr>
<td>Average family size</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td></td>
</tr>
<tr>
<td>Per capita income</td>
<td></td>
</tr>
<tr>
<td>Homeowners</td>
<td></td>
</tr>
<tr>
<td>Less than high school degree</td>
<td></td>
</tr>
<tr>
<td>High school degree</td>
<td></td>
</tr>
<tr>
<td>Some college education</td>
<td></td>
</tr>
<tr>
<td>Bachelor degree</td>
<td></td>
</tr>
<tr>
<td>Graduate degree</td>
<td></td>
</tr>
<tr>
<td>Below poverty</td>
<td></td>
</tr>
<tr>
<td>Median housing value</td>
<td></td>
</tr>
</tbody>
</table>
completely fuzzy solution (where FPU=1 and DPU=0). Thus, the results along with later validation revealed that the best clarity in neighborhood structure was achieved with the four PN solution at fuzzifier of 1.4 (Table 2.2). The optimum number of PNs could have been three, with very similar values for four PNs; however, indication from the fuzzy indices, a stronger classification rate, as well as, a priori knowledge of the study area, particularly the location of urban centers, indicated four PNs was the most sensible solution. The three PN solution tended to group small to medium sized urban centers (like those in Greene, Jefferson, and Sevier counties) with more rural neighborhoods, while the four PN solution separated them into different PNs (Figure 2.1). This is similar to, but not as exaggerated as, results from preliminary analyses of the data using hard clustering methods (K-means), where every neighborhood outside of Knox County was grouped into one PN (Figure 2.2). Due to the known demographic diversity and socioeconomic variability of small to medium sized cities compared to rural neighborhoods in the study area, it was clear that those solutions (from standard k-means) were not providing good insight into the structure of neighborhood characteristics in the study area.

In the sensitivity analysis, one not only looks at the fuzzy indices, but also the patterns in membership belonging for neighborhoods in each PN as the fuzzifier changes. A summary of degrees of belonging for neighborhoods within each PN at different fuzzifiers is presented in Table 2.3. A stable neighborhood would have a primary (the PN to which it is classified) degree of belonging that is greater than 0.75. Fuzzy neighborhoods were described as having secondary and tertiary degrees of
Table 2.2: Sensitivity Analysis of Fuzzy Cluster Analysis Results for Peer Neighborhoods Based on Socioeconomic and Demographic Population Characteristics

<table>
<thead>
<tr>
<th>Fuzzifier (m)</th>
<th>Three PNs</th>
<th>Four PNs</th>
<th>Five PNs</th>
<th>Six PNS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FPU*</td>
<td>DPU*</td>
<td>FPU</td>
<td>DPU</td>
</tr>
<tr>
<td>1.01</td>
<td>0.999</td>
<td>0.000</td>
<td>0.999</td>
<td>0.000</td>
</tr>
<tr>
<td>1.1</td>
<td>0.924</td>
<td>0.023</td>
<td>0.914</td>
<td>0.027</td>
</tr>
<tr>
<td>1.2</td>
<td>0.722</td>
<td>0.067</td>
<td>0.706</td>
<td>0.097</td>
</tr>
<tr>
<td>1.3</td>
<td>0.413</td>
<td>0.266</td>
<td>0.456</td>
<td>0.241</td>
</tr>
<tr>
<td>1.4</td>
<td>0.471</td>
<td>0.202</td>
<td>0.465</td>
<td>0.227</td>
</tr>
<tr>
<td>1.5</td>
<td>0.264</td>
<td>0.354</td>
<td>0.292</td>
<td>0.357</td>
</tr>
<tr>
<td>1.6</td>
<td>0.119</td>
<td>0.640</td>
<td>0.091</td>
<td>0.722</td>
</tr>
</tbody>
</table>

DPU = Normalized average square error, values close to 1 are hard solutions; FPU = Dunn’s normalized partition coefficient, values close to 1 are fuzzy solutions; PN = Peer Neighborhood

*One wants to identify a solution that has a high FPU index and low DPU index without being too close to a completely fuzzy solution (where FPU=1 and DPU=0)
Table 2.3: Summary of Degrees of Belonging for Neighborhoods within Peer Neighborhoods as the Fuzzifier changes in Fuzzy Cluster Analysis

<table>
<thead>
<tr>
<th>PN</th>
<th>Stable 1 M=1.1</th>
<th>Fuzzy 2 (%) M=1.1</th>
<th>Stable 1 M=1.3</th>
<th>Fuzzy 2 (%) M=1.3</th>
<th>Stable 1 M=1.4</th>
<th>Fuzzy 2 (%) M=1.4</th>
<th>Stable 1 M=1.5</th>
<th>Fuzzy 2 (%) M=1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>4 (5.9)</td>
<td>52</td>
<td>16 (23.9)</td>
<td>45</td>
<td>21 (31.8)</td>
<td>27</td>
<td>38 (58.5)</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>3 (13.0)</td>
<td>13</td>
<td>6 (31.6)</td>
<td>12</td>
<td>7 (36.8)</td>
<td>0</td>
<td>20 (100.0)</td>
</tr>
<tr>
<td>3</td>
<td>51</td>
<td>8 (15.0)</td>
<td>28</td>
<td>26 (48.1)</td>
<td>10</td>
<td>40 (80.0)</td>
<td>0</td>
<td>45 (100.0)</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>1 (5.9)</td>
<td>13</td>
<td>14 (53.8)</td>
<td>7</td>
<td>25 (78.1)</td>
<td>0</td>
<td>37 (100)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>151</strong></td>
<td><strong>16 (9.6)</strong></td>
<td><strong>105</strong></td>
<td><strong>62 (36.9)</strong></td>
<td><strong>74</strong></td>
<td><strong>93 (55.7)</strong></td>
<td><strong>27</strong></td>
<td><strong>140 (83.8)</strong></td>
</tr>
</tbody>
</table>

M= fuzzifier in fuzzy cluster analysis; PN = peer neighborhood

1The number of neighborhoods within the PN that are stable, i.e. have secondary or tertiary degrees of belonging to other PN(s) less than 0.25

2The number (%) of neighborhoods within the PN that are fuzzy, i.e. have secondary or tertiary degrees of belonging to other PN(s) greater than 0.25
Figure 2.1: Identified Peer Neighborhoods (PN) in East Tennessee based on Socioeconomic and Demographic Population Characteristics using Fuzzy K-means Clustering Algorithm
Figure 2.2: Identified Peer Neighborhoods (PN) in East Tennessee based on Socioeconomic and Demographic Population Characteristics using K-means Clustering Algorithm.
belonging greater than 0.25 to other PN(s) than the one in which it is classified. At m=1.1, there are only 16 neighborhoods (9.6% of the total sample) with a secondary and tertiary degree of belonging of at least 0.25 or more. This indicates that these neighborhoods have a tendency to move elsewhere, i.e. have characteristics similar to another PN. At m=1.3, 36.9% of the sample is showing this tendency, but more so the neighborhoods in PNs 3 and 4. At m=1.4, the neighborhoods in PNs 3 and 4 are moving quickly toward diffused (or equal) degrees of belonging across all PNs, while PNs 1 and 2 are moving in that direction slowly. At m=1.5, there is too much fuzziness since only a few neighborhoods in PN 1 (41.5%) have a strong degree of belonging to that PN. If there were no fuzziness in the clustering structure, these changes would not have occurred so quickly\textsuperscript{127,129}. Given that the desired solution should not be too fuzzy nor too hard, the suitable choices for the fuzzifier were m=1.3 or 1.4. It would be expected that the fuzzy neighborhoods would form their own PN if the number of PNs was increased to 5 or 6 if the neighborhoods were uniquely different than the already established PNs, but this was not seen. Thus, the fuzzy observations actually lie in the space between the PNs, such that they are similar to more than one based on some characteristics. The fuzzifier m=1.4 was chosen for the final solution because of the additional information it gave for some of the fuzzy neighborhoods, i.e. that they actually had similar characteristics to one or more other PNs, and because of the later strong validation with discriminant analysis and classification trees.

\textbf{2.4.3 Characteristics of identified Peer Neighborhoods (PNs)}

Peer neighborhood 1 was located primarily in rural, including the mountainous, areas (Figure 2.1) and was characterized by higher proportions of married people and
homeowners, medium levels of income and housing value, but lower levels of education (Table 2.4). The most urbanized was PN 2, located in the downtown portions of cities with significantly lower median housing values, per capita income, education levels, proportion of homeowners, and proportion of married people compared to other PNs (Figure 2.1 & Table 2.4). This PN also had the highest proportions of single parent households, minorities, and younger populations. Peer neighborhood 3 was located in semi-urban areas and had the highest proportion of population ≥ 65 years, as well as the second highest levels of economic and higher education variables. Located in the suburban areas, PN 4 was the most affluent with significantly higher per capita income, housing value, employment, homeownership, and higher education (bachelor and graduate degrees) than other PNs (Figure 2.1 & Table 2.4).

2.4.4 Evaluation of misclassified neighborhoods

Both nearest neighbor DA and CART resulted in 86% correct classification of the four PNs (Table 2.5). This was by far the highest classification for any number of clusters (results not presented). The misclassified neighborhoods were often located along geographic borders of PNs (Figure 2.1). Thus, it was not surprising that these neighborhoods had degrees of belonging split between the PNs they bordered geographically. Additionally, the misclassified neighborhoods tended to be located just outside urban areas or areas that may have developing industry and/or transitioning population. For example, PN 1 had nine misclassified neighborhoods. The cross validation results in DA indicated that six of those nine were predicted to be in PN 3, while the other three were in PN 4. According to Hotelling’s test, the six neighborhoods predicted for PN 3 had a significantly higher urban population while the three
<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Living in urban areas (%)</td>
<td>11.4C</td>
<td>100.0A</td>
<td>87.5B</td>
<td>88.8BA</td>
</tr>
<tr>
<td>Below poverty (%)</td>
<td>17.1B</td>
<td>41.7A</td>
<td>15.1B</td>
<td>8.16C</td>
</tr>
<tr>
<td>Housing median value ($)</td>
<td>70741BC</td>
<td>36616C</td>
<td>83466B</td>
<td>128997A</td>
</tr>
<tr>
<td>Living in rural areas (%)</td>
<td>5.64A</td>
<td>0.00B</td>
<td>0.21B</td>
<td>0.26B</td>
</tr>
<tr>
<td>White (%)</td>
<td>97.4A</td>
<td>53.7B</td>
<td>91.4A</td>
<td>92.4A</td>
</tr>
<tr>
<td>Black (%)</td>
<td>1.11B</td>
<td>42.4A</td>
<td>5.13B</td>
<td>3.73B</td>
</tr>
<tr>
<td>Male (%)</td>
<td>49.6A</td>
<td>47.1B</td>
<td>48.1AB</td>
<td>49.0A</td>
</tr>
<tr>
<td>Population 40-59 yrs (%)</td>
<td>15.5AB</td>
<td>13.4C</td>
<td>14.8BC</td>
<td>16.7C</td>
</tr>
<tr>
<td>Population 50-59 yrs (%)</td>
<td>13.3A</td>
<td>8.69C</td>
<td>11.6B</td>
<td>13.0AB</td>
</tr>
<tr>
<td>Population 60-65 yrs (%)</td>
<td>5.38A</td>
<td>3.00C</td>
<td>4.43B</td>
<td>4.30B</td>
</tr>
<tr>
<td>65 yrs and over (%)</td>
<td>12.7B</td>
<td>11.4B</td>
<td>15.6A</td>
<td>12.6B</td>
</tr>
<tr>
<td>Single parent families (%)</td>
<td>6.89BC</td>
<td>17.3A</td>
<td>7.90B</td>
<td>5.03C</td>
</tr>
<tr>
<td>Average family size (#)</td>
<td>2.95A</td>
<td>2.99A</td>
<td>2.88A</td>
<td>2.93A</td>
</tr>
<tr>
<td>Married (%)</td>
<td>64.3A</td>
<td>34.1C</td>
<td>54.2B</td>
<td>62.0A</td>
</tr>
<tr>
<td>Employed (%)</td>
<td>55.3B</td>
<td>45.9C</td>
<td>58.3B</td>
<td>65.3A</td>
</tr>
<tr>
<td>Per capita income ($)</td>
<td>14795B</td>
<td>10735C</td>
<td>17654B</td>
<td>27859A</td>
</tr>
<tr>
<td>Homeowner (%)</td>
<td>81.8A</td>
<td>36.2C</td>
<td>63.1B</td>
<td>75.8A</td>
</tr>
<tr>
<td>Less than high school education (%)</td>
<td>36.5A</td>
<td>31.9A</td>
<td>25.7B</td>
<td>10.7C</td>
</tr>
<tr>
<td>High school graduate (%)</td>
<td>36.5A</td>
<td>29.2B</td>
<td>30.2B</td>
<td>18.7C</td>
</tr>
<tr>
<td>Some college (%)</td>
<td>18.8B</td>
<td>28.4A</td>
<td>26.6A</td>
<td>29.9A</td>
</tr>
<tr>
<td>Bachelor degree (%)</td>
<td>5.18C</td>
<td>5.98C</td>
<td>10.8B</td>
<td>22.8A</td>
</tr>
<tr>
<td>Graduate degree (%)</td>
<td>2.78C</td>
<td>3.32CB</td>
<td>5.29B</td>
<td>14.7A</td>
</tr>
</tbody>
</table>

A,B,C,D Mean separation based on Tukey (p<0.05) adjustment method. Means of the variable between peer neighborhoods that have the same letter are not significantly different.
Table 2.5: Nearest Neighbor Discriminant Analysis Results of Classification of East Tennessee Peer Neighborhoods Based on Socioeconomic & Demographic Characteristics

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual Peer Neighborhood</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>57</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>6</td>
<td>3</td>
<td>40</td>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>66</td>
<td>19</td>
<td>50</td>
<td>32</td>
<td>167</td>
</tr>
</tbody>
</table>
neighborhoods predicted for PN 4 had significantly higher housing values than the rest of the neighborhoods in PN 1. Similar results were found for misclassified neighborhoods in other PNs. The three misclassifications in PN 2 had significantly lower urban populations than the rest of the PNs and had equal degrees of belonging to PNs 2 and 3. PN 3 had the most misclassifications with 10 neighborhoods predicted to be either in PN 1 (if they had a significantly lower proportion of urban population) or in PN 4 (if they had significantly higher median housing values and lower proportions of the population living below poverty). The least number of misclassified neighborhoods occurred in PN 4 where two neighborhoods were predicted to belong in PN 3. However, no differences in socioeconomic and demographic characteristics from Hotelling’s test were found.

2.4.4 Variables important for classifying Peer Neighborhoods (PNs)

CART results show that the first split was on percent urban population ≤ 38.183% leading to 97% correct classification in PN 1 (Figure 2.3). The second split occurred with percent urban population > 38.183 and housing value > $105,850. This resulted in 87% correct classification in PN 4. The third split occurred when percent urban population was greater than 38.183% and housing value ≤ $105,850. This produced two groups with percent below poverty level ≤ 27.276% yielding a 74% correct classification in PN 3. When the percentage below poverty level was > 27.276%, 89% of the neighborhoods were correctly classified in PN 2.

Given that the CART and DA yielded similar classification results, the uniqueness of the four identified PNs was supported. The percent of population living in urban areas, the median housing value, and the percent of population living below poverty in a
Figure 2.3: Cluster and Regression Tree (CART) Results for Peer Neighborhoods (PNs) in East Tennessee
neighborhood were the most important variables in determining correct classification of neighborhoods.

2.4.5 Disparities in stroke and MI mortality between Peer Neighborhoods (PNs)

Peer neighborhood 4, the most affluent PN and located in the suburbs, had significantly lower \((p=0.01)\) risks for stroke and MI mortality than all other PNs (Figure 2.4). Conversely, the most urban and least affluent neighborhood, PN 2, tended to have higher risks of stroke and MI mortality, although these were not significantly \((p=0.6)\) different from the risks for both PN 1 and PN 3. Only the MI mortality risk for PN 2 was greater than the state risk of 85.5/100,000, while the risk for PN 4 was the only one below the US risk (58.9/100,000). The stroke mortality risks in PNs 2 and 3 exceeded both the state (67.5/100,000) and US (53.5/100,000) risks.

2.5 Discussion and Conclusions

To our knowledge, this is the first study to investigate the clustering of population characteristics that are risk factors associated with stroke or MI at the neighborhood level. Based on knowledge of the study area, the four PNs identified are a unique and sensible classification of neighborhoods based on socioeconomic, demographic, geographic characteristics for East Tennessee. The geographic distribution of identified PNs revealed that the most affluent neighborhoods are located in suburban areas, while the least affluent neighborhoods were located in the downtown areas. These findings are consistent with those from other studies that have investigated neighborhood level socioeconomic and demographic determinants of health\(^{120, 234}\).
Figure 2.4: Annual age-adjusted stroke and myocardial infarction mortality risks for peer neighborhoods in East Tennessee
Several studies have considered socioeconomic or demographic characteristics of populations in relation to stroke or MI, but historically most of these analyses have been done at state\textsuperscript{24, 81, 100, 158, 159} or county\textsuperscript{1, 2} geographic levels. Recent studies indicate that finer geographic units are needed to increase the clarity of distributions of determinants of health\textsuperscript{83, 101} and to better guide local health planning and targeted health programs. To address this issue, the current study was performed at the census tract level, which have been found to be good proxies of natural neighborhoods\textsuperscript{120, 219}. Additionally, census tract level socioeconomic and demographic data are available to all states in the US through the US Census Bureau, as well as for populations in other countries like Canada (census tracts)\textsuperscript{156} and the United Kingdom (postcode sectors) that approximately correspond to US census tracts\textsuperscript{102, 154}. Given the lack of socioeconomic information provided in US vital records, population studies must rely on census data in order to investigate population characteristics at a neighborhood level. Comprehensive data at the census tract level is also limited by the decennial nature of the US census, such that the data may be outdated or not accurately reflect neighborhood composition due to population growth and migration. To address this issue, it has been recommended that only data from the closest census falling within five years of the study period should be used\textsuperscript{235}. Thus, the 2000 census data were best suited to match the disease data (1999-2007) for this study. Furthermore, the 2010 census data were not available at the time of this study’s analyses. Since census tract level was the best available data for the current study, robust multivariate methods were utilized to be able to include many socioeconomic and demographic variables in order to reduce bias and get the most comprehensive insight into neighborhood
characteristics of the study area. As this was a population health planning approach and the goal was to better understand neighborhood effects, individual level risk factors (like genetics, co-morbidities, medical history, or modifiable behaviors) that may affect stroke or MI patterns\(^{105,153}\) were not included in the analyses. Although census data are useful and are currently the best available data for these types of analyses to address these types of research and health planning questions, they are not without limitations. Some of the limitations associated with census data include both sampling (e.g. missing street address) and non-sampling errors (e.g. phrasing of questions which may influence the response) and hence the data obtained\(^{236,237}\).

The association of socioeconomic and demographic characteristics with survival after MI at the neighborhood level has been described by other studies using census tracts as the geographic unit of analysis\(^{94,112}\). However, these studies included only one or a few demographic factors and measures of socioeconomic status. Other studies have found that neighborhood SES is important in determining risk using composite socioeconomic and demographic measures\(^{42,99}\). Evidence from recent research indicates that many socioeconomic and demographic characteristics are not interchangeable, and so the use of one measure or a composite measure ignores the complex relationships between the factors\(^{94,101}\). The results from the robust PCA in this study also indicated that, despite high correlations between variables, additional information existed that would be lost if some variable(s) were removed. For instance, while many of the variables that heavily loaded on component 1 were highly correlated, their loadings differed across the other components. Thus, the variables were explaining different pieces of information or variation across those components. These complex
interrelationships among socioeconomic and demographic factors imply that as many risk factors as realistically possible are needed for the most holistic analysis.

When using a high number of risk factors to classify neighborhoods into similar groups, issues with interdependencies among variables, different variable scales, and outliers are likely to arise. A major strength of this study was the use of robust PCA to account for these issues and reduce their bias on cluster analysis\textsuperscript{127}. Furthermore, the fuzzy cluster strategy was utilized to allow neighborhoods to have associations with more than one PN, giving insight into the structure of the data when groups may not be mutually exclusive\textsuperscript{129}. The drastic difference in results (Figures 2.1 and 2.2) revealed that the fuzzy clustering approach provided more insight into the true structure of the neighborhoods, while the traditional k-means approach seemed to be more influenced by outliers in Knox county, masking the characteristics of some neighborhoods in other counties. The complex interrelationships between the risk factors and the multi-factorial nature of causation of stroke and MI indicate that some overlap between groups could be expected. These areas of overlap are particularly important when considering neighborhood health needs since the identified unique population profiles are valuable in the development of population health programs. Information on the tendency of a neighborhood to move toward another PN from the sensitivity analysis of the fuzzy method is very useful when developing population health programs since every neighborhood is important. This allows health initiatives to be targeted at the neighborhood level based on the population characteristics and health needs, instead of a larger area that has more diverse characteristics. The implication of this is that, within an administrative unit (such as a county), health professionals are able to use a needs-
based approach to planning and service provision, based on unique neighborhood profiles and health needs, instead of using a "one-size-fits-all" strategy. Thus, within an administrative unit, different programs can be designed to meet the distinct needs of the different neighborhood types based on their unique profiles.

In order to get the most comprehensive idea of the structure of the neighborhoods and to fully understand the uniqueness of those misclassified observations where overlap between PNs could be expected, it was important to explore the cluster solution using several validation methods. The majority of misclassified neighborhoods were found in PNs 3 and 1. This was expected given that these PNs had levels of socioeconomic and demographic variables somewhere in between the distinct high and low extremes of PNs 4 and 2, respectively (Table 2.2). The fuzzy analysis allows the overlap of the misclassified neighborhoods with fuzzy degrees of belonging across another PN to be highlighted. This implies that it may be necessary to consider some neighborhoods in more than one PN in the population health planning of those different areas. For example, when designing a targeted health program for improving heart attack mortality risk for PN 3, one would also want to consider those neighborhoods classified as PN 2 but had high degrees of belonging (i.e. similar characteristic) to PN 3. Though these neighborhoods were classified in PN 2 because of their urban locations, their demographic and socioeconomic characteristics were more consistent with PN 3. Thus we would expect health needs for these neighborhoods to be similar to PN 3. Practically, heart health education campaigns, such as diet and exercise recommendations, geared toward less diverse and higher income populations like PN 3, might be additionally presented to those neighborhoods
in PN 2 that were similar to PN 3. Therefore, in addition to statistical analyses, visual evaluation of the grouping of neighborhood characteristics into PNs and the prior knowledge of relationships between the variables and health outcome of interest are important in recognizing patterns that are useful in aiding resource allocation and service provision.

Several studies have found that risks of stroke and MI are inversely related to socioeconomic factors like education and income and positively associated with demographic factors like proportion of males, blacks, and population over 65\textsuperscript{81,100,158}. In this study, these characteristics were clustered in neighborhoods located in the most urbanized downtown areas. Similar results have been reported by a Canadian study\textsuperscript{120}. In addition to urbanicity, the current study also found that median housing value and the proportion of the population living in poverty were the key factors in classifying PNs. While urban populations have not been directly reported to have increased stroke and MI risk, they tend to have socioeconomic and demographic factors consistent with increased risk, i.e. tend to be the less affluent segments of the population.

Indeed, this study found that a significant disparity exists in both stroke and MI mortality between less affluent, urbanized neighborhoods and more affluent, suburban neighborhoods. This is very concerning since recent reports indicate that the disparity in cardiovascular death risks is widening between lower and higher socioeconomic status groups\textsuperscript{83}. This study provides information on the unique socioeconomic and demographic profiles of neighborhoods that can aid in understanding disparities in health outcomes by identifying the unique challenges and health needs between
neighborhoods. For instance, although PNs 1 and 3 seem to have similar socioeconomic characteristics, close evaluation reveals that these PNs greatly differ. PN 3 has a significantly more urban, older, and educated population than PN 1. If only socioeconomic characteristics are considered, these populations would incorrectly be considered similar. From a health planning perspective, it is clear that older populations, like PN 3, would have different health needs than other segments of the population. Additionally, PNs 1 and 4 have similar stroke risks (47 and 46.6 annual deaths per 100,000 population, respectively). Both PNs have higher levels of income; however, PN 4 is a somewhat younger, more urban and more ethnically diverse than PN 1. Thus, different characteristics at the neighborhood level must be considered in targeting health education and outreach activities in order to improve outcomes and reduce disparities.

The neighborhood focused approach of this study is applicable to health planning in areas other than East Tennessee. The generalizability is not specifically in the study findings, but in the application of the methodology to provide insight into the unique population characteristics and potential health needs of other communities based on empirical evidence. The findings of this study serve as examples of the type of information that can be obtained from this approach and its usefulness from a population health planning perspective. It would be expected that a different number of PNs with different sets of unique profiles would be identified using this methodology in different populations. However, the health outcome improvement programs and health disparity reduction strategies could then be specifically tailored to the results and specific needs of neighborhoods of interest.
In conclusion, the robust and fuzzy multivariate techniques utilized in this study to classify neighborhoods based on socioeconomic or demographic characteristics identified four unique population profiles in the study area. Stroke and MI mortality risk differed between the identified PNs. The PNs with highest risk also had the highest levels of socioeconomic variables known or suspected to be associated with higher risk of stroke or MI and were located in the urbanized downtown areas. The lowest mortality risk was associated with the most affluent PN. These findings provide population health planners a unique opportunity to better understand and effectively plan for the unique neighborhood health needs. Thus, implementation of these methodologies and the careful integration of its findings in health planning activities will be useful in guiding health resource allocation, service provision, and policy decisions at the local level. Moreover, this information is important for addressing neighborhood health disparities not only in the East Tennessee Appalachian Region, but also for other health planning regions throughout the US and other countries given the availability of socioeconomic and demographic data.
CHAPTER 3

3.0 Neighborhood Disparities in Stroke and Myocardial Infarction Mortality: a GIS and Spatial Scan Statistics Approach

Ashley Pedigo, MS¹; Tim Aldrich, PhD²; and Agricola Odoi, BVM, M.Sc., PhD¹

¹The University of Tennessee, Department of Comparative Medicine, 2407 River Drive Knoxville, Tennessee 37996-4543, United States.

²East Tennessee State University, Department of Biostatistics and Epidemiology, P.O. Box 70259 Johnson City, TN 37614-1709

This chapter is a manuscript that has been published in BMC Public Health journal (BMC Public Health 2011, 11:644; doi: 10.1186/1471-2458-11-644).

My contributions to this paper included data preparation, analysis, interpretation of results, gathering and reviewing of literature, formulation of discussion topics, as well as drafting and editing the manuscript.
3.1 Abstract

Stroke and myocardial infarction (MI) are serious public health burdens in the US. These burdens vary by geographic location with the highest mortality risks reported in the southeastern US. While these disparities have been investigated at state and county levels, little is known regarding disparities in risk at lower levels of geography, such as neighborhoods. Therefore, the objective of this study was to investigate spatial patterns of stroke and MI mortality risks in the East Tennessee Appalachian Region so as to identify neighborhoods with the highest risks.

Stroke and MI mortality data for the period 1999-2007, obtained free of charge upon request from the Tennessee Department of Health, were aggregated to the census tract (neighborhood) level. Mortality risks were age-standardized by the direct method. To adjust for spatial autocorrelation, population heterogeneity, and variance instability, standardized risks were smoothed using Spatial Empirical Bayesian technique. Spatial clusters of high risks were identified using spatial scan statistics, with a discrete Poisson model adjusted for age and using a 5% scanning window. Significance testing was performed using 999 Monte Carlo permutations. Logistic models were used to investigate neighborhood level socioeconomic and demographic predictors of the identified spatial clusters.

There were 3,824 stroke deaths and 5,018 MI deaths. Neighborhoods with significantly high mortality risks were identified. Annual stroke mortality risks ranged from 0 to 182 per 100,000 population (median: 55.6), while annual MI mortality risks ranged from 0 to 243 per 100,000 population (median: 65.5). Stroke and MI mortality risks exceeded the state risks of 67.5 and 85.5 in 28% and 32% of the neighborhoods,
respectively. Six and ten significant (p<0.001) spatial clusters of high risk of stroke and MI mortality were identified, respectively. Neighborhoods belonging to high risk clusters of stroke and MI mortality tended to have high proportions of the population with low education attainment.

These methods for identifying disparities in mortality risks across neighborhoods are useful for identifying high risk communities and for guiding population health programs aimed at addressing health disparities and improving population health.

### 3.2 Introduction

On average, every 34 and 40 seconds, myocardial infarction (MI) and stroke events occur in the US, respectively\(^{217}\). Stroke ranks third in causes of death and is the leading cause of debilitation among Americans\(^{24}\). It is estimated that approximately 15% of those who have an MI will die of it\(^{217}\). These health conditions are serious economic burdens to the US health system with annual costs estimated at $73.7 billion for stroke and $177.1 billion for MI\(^{217}\).

Place of residence is an important determinant of cardiovascular health and disparities in the burdens of stroke and MI have been observed for different geographic areas\(^{24, 83, 217}\). The highest risks of mortality have been reported in the southeastern US\(^{86, 141, 142, 217}\) and in populations living in rural areas\(^{81, 143, 159}\), particularly in the Appalachian region\(^{145, 146}\). Many areas of the Appalachian region, including parts of Tennessee, form a portion of the US “stroke belt”. Tennessee ranks 3\(^{rd}\) highest in the US for stroke\(^{217}\), and had an annual age-adjusted stroke mortality risk for the period 2000-2006 of 67.5 deaths per 100,000 persons compared to the national risk of 53.5
deaths per 100,000 persons. For coronary heart disease including MI, Tennessee ranks 4th highest in the US with an annual age-adjusted mortality risk for the period 2000-2006 of 85.5 deaths per 100,000 persons compared to the national risk of 58.9 death per 100,000 persons.

The geographic distributions of stroke and MI mortality have been investigated at state and county levels. However, geographic disparities have been shown to exist even after adjusting for variations in common risk factors like demographic factors (race, age), socioeconomic measures (income, education), behaviors (smoking, physical activity), and other conditions (diabetes, hypertension). These findings suggest that geographic variation in stroke and MI mortality could be due to more localized distributions of neighborhood risk factors. The clustering of determinants of stroke and MI at the neighborhood level can greatly affect the planning, implementation, and focus of health initiatives that seek to reduce disparities. Therefore, research should focus on identifying disparities at the neighborhood level to better understand health needs and thus, provide needs-based health services. While many studies have defined neighborhoods as census tracts or smaller geographic units, the neighborhoods have not been used as the unit of analysis for many past studies investigating cardiovascular disease and stroke. Rather, these studies have investigated neighborhood characteristics as contextual effects in multilevel models that seek to explain individual level risk. Thus, ecological studies are needed to investigate the spatial patterns and clustering of high mortality risk with the neighborhood as the unit of analysis since this is important in identifying high risk
communities and targeting resources to address health disparities and improve population health.

When investigating disease patterns in small geographic areas like neighborhoods, however, there are some challenges that must be addressed. Due to population heterogeneity, mortality risks from areas of low population will likely have higher variances and therefore be more unstable than those from areas of high population\textsuperscript{238}. This variance instability of small geographic areas is referred to as the small number problem\textsuperscript{239}. Spatial smoothing of risks is used to mitigate this issue by reducing the “noise” from areas with low population and therefore high variances\textsuperscript{240}.

With these issues in mind, the objective of this study was to investigate spatial patterns and detect local neighborhood clusters of high risk of stroke and MI mortality in the East Tennessee Appalachian Region. The identification of neighborhoods with high risks is expected to aid local health planners in understanding the specific neighborhood health needs to guide health planning and provision of health services. Thus, identified clusters of high risks of stroke and MI mortality will be useful in guiding resource allocation, service provision, and policy decisions at the local/neighborhood level that are crucial for addressing neighborhood health disparities.

3.2 Methods

3.2.1 Study area and data collection

The study area included eleven counties of the East Tennessee Appalachian Region that have some of the highest risks of stroke and/or MI in the state: Claiborne,
Cocke, Grainger, Greene, Hamblen, Hancock, Hawkins, Jefferson, Knox, Sevier, and Union counties. This area had a population of just over 780,000 persons in 2000 and included 168 census tracts. Census tracts (CTs) are statistical subdivisions of a county that have between 2,500 and 8,000 persons, do not cross county boundaries, and are homogenous with respect to population characteristics, economic status, and living conditions\textsuperscript{155}. Since they are good proxies of natural neighborhood boundaries and are therefore useful in describing neighborhood population characteristics and health disparities\textsuperscript{120, 241}, CTs were chosen as the geographical unit of analysis and were used to represent neighborhoods in this study.

Mortality data from 1999 to 2007 were obtained free of charge, upon request, from the Tennessee Department of Health. Thus, although these data are freely available on request from the responsible authorities, they are not currently openly available for internet downloads. Stroke and MI deaths were identified by ICD 10 codes I60-I69 and I21-I22, respectively. For the 8,842 mortality records obtained, complete street address data were available for 94%, while the other 6% had missing or inadequate (such as post office box) address data. The addresses were geo-coded using BatchGeo\textsuperscript{220}, an online geo-coding service which implements the Google Maps geocoding application programming interface (API) that has some of the highest quality geocoding databases available\textsuperscript{242, 243}. Exact, or roof top, address matches were obtained for 67% of the data, while 30% were range interpolated between two points on the street and 3% were matched to the zipcode. The geographic coordinates were imported into ArcGIS 9.3\textsuperscript{221} where point-in-polygon join was used to link the mortality
data to the openly available census tract level cartographic boundary files downloaded from the U.S. Census Bureau website\textsuperscript{237}.

Census tract level socioeconomic, demographic, and population data for the study area were obtained from the openly available census 2000 summary file \textsuperscript{3}\textsuperscript{244}. Since these data are available in the US only through the decennial census, the 2000 data was deemed best suited to match the disease data (1999-2007). The neighborhood variables chosen to be assessed as potential predictors of the geographic distribution of MI and stroke high risk mortality clusters were based on current knowledge in the literature. They include: black race\textsuperscript{80, 81, 83, 86}, gender\textsuperscript{2, 24, 81, 84}, age 65 years and older\textsuperscript{24, 25, 81, 94}, household income\textsuperscript{94, 105, 106, 112, 153}, education less than high school\textsuperscript{81, 89, 96, 112}, population below poverty\textsuperscript{107, 112, 153}, median housing value\textsuperscript{89, 97, 99}, geography (urban versus rural)\textsuperscript{81, 83, 143, 148}, and factors like employment, single parent families, marital status, and housing ownership that have been used in composite measures of socioeconomic status (SES) or deprivation\textsuperscript{89, 90, 97, 99}.

3.2.2 Data analysis

3.2.2.1 Data management

One neighborhood in Knox county, that had a population of 232 and included a mental health facility, was removed from the analysis due to missing data values for most of the variables. With the exception of median household income, median housing value, and family size, all variables were analyzed as the proportion of the population in each neighborhood.
3.2.2.2 Descriptive analyses, risk standardization and spatial smoothing

All descriptive analyses were done in SAS 9.2. Significance of the difference in median age between genders was assessed using the Wilcoxon rank sum test. Mortality risks for neighborhoods were age-adjusted using direct standardization in Stata 11. All risks were expressed as the annual number of deaths per 100,000 population.

The raw (unsmoothed) age-adjusted risks were expected to have high variances due to the small number problem since there were areas of low population and some neighborhoods with only a few cases of stroke/MI in the study area. To address this issue, as well as adjust for spatial autocorrelation and population heterogeneity, the raw age-adjusted risks were smoothed using Spatial Empirical Bayes (SEB) smoothing using 2nd order queen weights in GeoDa. In this smoothing method, the risks for low population neighborhoods in areas without clear spatial patterns are shrunk toward the global mean of the study area. Conversely, in areas where obvious spatial patterns exist, the less reliable estimates from low population areas are adjusted towards a local mean. Thus, the SEB smoothed risks are more stable than raw (unsmoothed) risks.

3.2.2.3 Detection and identification of stroke and MI clusters

To detect the presence of high risk stroke and MI clusters and identify their locations, the spatial scan statistic, implemented in SaTScan, was used. The technique uses circular windows of variable radius that move across the study area to compare the number of deaths in the window with what would be expected if the deaths
were distributed randomly in space\textsuperscript{248}. The window radius varies from zero up to a specified maximum. Each window defines a set of different neighboring CTs, such that if the geographic centroid of a census tract is contained in the window, then the deaths and population from that whole census tract are included. Clusters are identified based on a likelihood ratio test\textsuperscript{249} with a p-value obtained using Monte Carlo replications\textsuperscript{250}. The primary cluster, with the highest significant likelihood, is interpreted such that there is an increased risk of stroke/MI mortality within the window compared to outside\textsuperscript{251}.

Non-overlapping, spatial clusters of high risk of stroke/MI mortality were identified using a purely spatial, discrete Poisson model\textsuperscript{249} adjusted for age distribution. Since the results of this analysis can be sensitive to model parameters, particularly window size, care must be taken in its choice. The goal of the current analyses was to identify local clusters of high mortality risks among neighborhoods. Thus, similar to another study\textsuperscript{252}, the window size of 5\% of the total population was chosen based on the population of the largest neighborhood so that potentially one single neighborhood could constitute a distinct high risk cluster.

3.2.2.4 Logistic modeling of predictors of high risk stroke or MI clusters

The outcome of interest in this modeling was binary, reflecting whether a neighborhood belonged to a cluster or not. Univariate associations of continuous variables with the outcomes were assessed using Wilcoxon rank sum test for non-normally distributed data, while chi-square and exact tests were used for categorical variables. Variables with significant associations based on a liberal p-value (p=0.20)
were considered in the modeling process along with some non-significant variables that had been shown in literature to be strongly associated with the outcome.

Multiple logistic models were used to investigate potential associations between log odds of a neighborhood being in a high risk stroke or MI cluster and a number of neighborhood level socioeconomic and demographic characteristics. The assumption of linearity of continuous variables with the log odds of the outcome (belonging to a stroke or MI cluster) for logistic modeling were assessed using graphical methods. Only the proportions of the population ≥ 65 years and of single parent families met this assumption for stroke cluster, while the proportions of population with less than high school education, those living below poverty and median housing value met the assumption for the MI outcome. Therefore, these variables were modeled as continuous variables. The variables not meeting the linearity assumption were transformed into categorical variables using either *a priori* considerations or quartile cutpoints from the distribution of the variable.

The model was built by starting with the full model and then removing variables based on the following criteria: (1) the highest non-significant p-value (with significance set to p=0.05); (2) a likelihood ratio test of the model with and without the variable that was non-significant; and (3) the variable was not an important confounder of other variables in the model. Variables were considered important confounders if their removal from the model resulted in a large (greater than 20%) change in the coefficients of any of the remaining variables in the model. Categorical variables were analyzed as regular dummy variables. The significance in the model of each group of dummy
variables (belonging to one categorical variable) was analyzed using a likelihood ratio test. Two-way interaction terms between gender, race, age, income, education, poverty, and geography were assessed for statistical significance\textsuperscript{2, 81, 90, 111}. Model fits were assessed using the Pearson and Hosmer-Lemeshow goodness of fit tests and residual diagnostics. The predictive abilities of the models were evaluated using sensitivity, specificity, and overall correct classifications.

3.2.2.5 Cartographic displays

All cartographic manipulations and displays were done in ArcGIS 9.3\textsuperscript{221}. The intervals for displaying the age-adjusted SEB smoothed mortality risks of stroke and MI in the choropleth maps were determined using Jenk’s optimization classification scheme. Since SEB risks are more appropriate for mapping in small areas compared to unsmoothed risks\textsuperscript{239, 240}, only the former are presented. Significant spatial clusters were displayed in ArcGIS 9.3\textsuperscript{221}.

3.3 Results

3.3.1 Description of stroke and MI deaths

There were 3,824 stroke deaths in the study area from 1999 to 2007. No stroke deaths were reported in 18 of the 168 neighborhoods. Women accounted for 2,435 (63.7%) of the stroke deaths. The median age was significantly (p<0.001) lower for men (median 78; range 4-103), than women (median 81; range 3-103). Persons dying from stroke or MI in the study were primarily white (94%) and had less than a high school education (45%). It is worth noting that 92% of the population in the study area was
white, while 25% of the population older than 18 years had less than high school education.

Myocardial infarction was the cause of 5,018 deaths during the study period. No deaths were reported in 17 neighborhoods; 15 of these neighborhoods also had no reported stroke deaths. More MI deaths occurred in men (2,745 deaths, 54.6%) than women (45.4%). Again, the median age of death was significantly (p<0.001) lower for men (median 71; range 21-102), than women (median 81; 27-106).

3.3.2 Spatial distribution of mortality risks

3.3.2.1 Stroke risks

The annual median age-adjusted raw (unsmoothed) stroke risk for the study area was 55.6 deaths/100,000 population (range: 0-182), with 28% of the neighborhoods exceeding the state stroke mortality risk of 67.5. Similarly, the annual median SEB smoothed stroke risk was 56.1 deaths/100,000 population (range: 0.1-174). The annual median risk for the study area remained constant from 1999 to 2007. The highest stroke risks (greater than 110 deaths/100,000) were observed in three neighborhoods in Knox county and one neighborhood each in Jefferson and Hamblen counties (Figure 3.1). It appeared that the neighborhoods with stroke risks higher than the state risk were concentrated across neighborhoods in the northwest portions of Cocke and Greene counties, in addition to a few neighborhoods in Grainger, Hamblen, and Jefferson counties, as well as in the downtown area of Knox county. These neighborhoods are primarily located in or near city centers in the study area.
Figure 3.1: Spatial Empirical Bayes Smoothed Age-Adjusted Stroke Mortality Risk per 100,000 Population from 1999 to 2007 in East Tennessee Appalachian Region
3.3.2.2 Myocardial infarction risks

The annual median raw (unsmoothed) age-adjusted MI mortality risk was 65.5 deaths/100,000 population (range: 0-243), while the median SEB smoothed risk was 63.5 (range: 0.5-235). Myocardial infarction mortality risks in the study area were higher than the state risk of 85.5 in 32% of the neighborhoods. The spatial distribution of neighborhood risks revealed patterns of high risks across the study area (Figure 3.2). The areas with the highest MI risks (greater than 140 deaths/100,000) included all neighborhoods in Claiborne county and all but one neighborhood in Cocke county. In addition to these counties, neighborhoods with risks above the state risk were also located in Greene, Jefferson, Hamblen, Grainger, and Knox counties in a pattern very similar to that for stroke risks.

3.3.3 Spatial clusters of high stroke/MI mortality risks

Table 3.1 displays results of identified significant spatial stroke and MI mortality clusters. For each cluster, the table gives the number of census tracts in the cluster, the total population, the observed number of stroke or MI deaths in the cluster area, the expected number of deaths based on the Poisson model, the estimated annual number of cases per 100,000 persons, and the significance level (p-value) obtained from the likelihood ratio test with Monte Carlo permutations. Figures 3.3 and 3.4 display geographic distributions of the significant spatial clusters of stroke and MI, respectively.

3.3.3.1 Stroke clusters

Six significant (p<0.001) spatial clusters of high risk of stroke mortality were identified (Table 3.1 and Figure 3.3). The smallest cluster, which was also the primary
Figure 3.2: Spatial Empirical Bayes Smoothed Age-Adjusted Myocardial Infarction Mortality Risk Per 100,000 Population from 1999 to 2007 in East Tennessee Appalachian Region.
Table 3.1: Spatial Clusters of Age-Adjusted Stroke and Myocardial Infarction Mortality Risks from 1999 to 2007 in East Tennessee Appalachian Region

<table>
<thead>
<tr>
<th>Cluster</th>
<th># of Census Tracts (Neighborhoods)</th>
<th>Population</th>
<th>Observed # of Deaths</th>
<th>Expected # of Deaths</th>
<th>Annual # of Deaths /100,000 Persons</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>5,447</td>
<td>136</td>
<td>37.76</td>
<td>195.6</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>17,243</td>
<td>174</td>
<td>91.95</td>
<td>102.8</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>34,887</td>
<td>270</td>
<td>174.24</td>
<td>84.2</td>
<td>0.001</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>30,158</td>
<td>266</td>
<td>187.63</td>
<td>77.0</td>
<td>0.001</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>24,711</td>
<td>180</td>
<td>120.65</td>
<td>81.0</td>
<td>0.001</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>12,008</td>
<td>107</td>
<td>67.27</td>
<td>86.4</td>
<td>0.004</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>36,945</td>
<td>608</td>
<td>243.77</td>
<td>177.7</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>24,596</td>
<td>334</td>
<td>159.80</td>
<td>148.9</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>13,856</td>
<td>213</td>
<td>88.78</td>
<td>171.0</td>
<td>0.001</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>28,823</td>
<td>333</td>
<td>197.74</td>
<td>120.0</td>
<td>0.001</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>30,158</td>
<td>363</td>
<td>236.98</td>
<td>109.2</td>
<td>0.001</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>9,568</td>
<td>124</td>
<td>61.61</td>
<td>143.4</td>
<td>0.001</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>35,548</td>
<td>325</td>
<td>231.94</td>
<td>99.9</td>
<td>0.001</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2,818</td>
<td>47</td>
<td>20.94</td>
<td>160.0</td>
<td>0.001</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>8,566</td>
<td>88</td>
<td>54.98</td>
<td>114.1</td>
<td>0.009</td>
</tr>
</tbody>
</table>

#: Number
Figure 3.3: Significant Spatial Clusters of High Age-Adjusted Stroke Mortality Risks from 1999 to 2007 in East Tennessee Appalachian Region
Figure 3.4: Significant Spatial Clusters of High Age-Adjusted Myocardial Infarction Mortality Risks from 1999 to 2007 in East Tennessee Appalachian Region.
cluster, was comprised of only 1 neighborhood in Hamblen county. The relative risk of this cluster was 3.7 (Figure 3.3), implying that the risk of death from stroke was 3.7 times higher within cluster 1 than other neighborhoods in the study area. Relative risks for the secondary clusters ranged from 1.5 to 1.9. Cluster 3 accounted for the highest number of stroke deaths and was composed of 6 neighborhoods in Cocke and Hamblen counties. The second largest cluster (cluster 4) included 6 neighborhoods in Greene county. The majority of the high risk stroke clusters were located in or near city centers.

### 3.3.3.2 Myocardial infarction clusters

There were nine significant \( p<0.009 \) spatial clusters of high risk of MI mortality (Table 3.1 and Figure 3.4). The primary cluster was the largest cluster in both the number of MI deaths and geographic size, and included neighborhoods in Cocke and Hamblen counties. The populations in cluster 1 neighborhoods had a risk of death from MI that was 2.7 times greater than other neighborhoods in the study area. Relative risks for the secondary clusters ranged from 1.4 to 2.5. Cluster 7 was the second largest and included neighborhoods in Jefferson, Hamblen, and Grainger counties. Neighborhoods in Claiborne, Greene, and Knox counties were also parts of significant high risk MI clusters. The majority (76%) of neighborhoods in significant high risk stroke clusters also belonged to significant high risk MI clusters.

### 3.3.4 Predictors of high risk stroke and myocardial infarction spatial clusters

#### 3.3.4.1 Stroke

The univariate associations of socioeconomic and demographic variables with the outcome of belonging to a high risk stroke cluster are presented in Table 3.2.
Table 3.2: Univariate Associations of High Risk Stroke and Myocardial Infarction (MI) Mortality Clusters with Neighborhood Socioeconomic and Demographic Factors

<table>
<thead>
<tr>
<th>Neighborhood level socioeconomic and demographic variables</th>
<th>Significance value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stroke cluster</td>
</tr>
<tr>
<td>Geography (rural, suburban, urban)</td>
<td>0.10§</td>
</tr>
<tr>
<td>Proportion of black population</td>
<td>0.02§</td>
</tr>
<tr>
<td>Proportion population age ≥ 65 years</td>
<td>0.02§</td>
</tr>
<tr>
<td>Proportion of single parent families</td>
<td>0.04§</td>
</tr>
<tr>
<td>Proportion of owner occupied housing units</td>
<td>0.08§</td>
</tr>
<tr>
<td>Median household income ($)</td>
<td>0.15§</td>
</tr>
<tr>
<td>Proportion of population with &lt; high school education</td>
<td>0.18§</td>
</tr>
<tr>
<td>Proportion of married persons</td>
<td>0.20§</td>
</tr>
<tr>
<td>Average family size</td>
<td>0.34</td>
</tr>
<tr>
<td>Proportion of population living below poverty</td>
<td>0.41</td>
</tr>
<tr>
<td>Median housing value ($)</td>
<td>0.41</td>
</tr>
<tr>
<td>Gender</td>
<td>0.60§</td>
</tr>
<tr>
<td>Proportion of population employed</td>
<td>0.67</td>
</tr>
</tbody>
</table>

§ Variables assessed in subsequent multivariable logistic regression model
Variables with significant associations, based on a liberal p-value=0.20 were further assessed in the multivariable logistic model. Even though gender was non-significant it was included because disparities in stroke risk and mortality by gender have been reported in literature [2, 8, 40, 41]. The other non-significant variables were not included because they were each highly correlated ($r>0.70$) with median household income. The final model had a highly significant ($p=0.0002$) likelihood. The proportion of the population with less than a high school education ($p=0.015$) and that were black ($p=0.019$) were significant variables in the model (Table 3.3). Neighborhood geography (rural, suburban, urban) was not significant ($p=0.1$), but was included in the final model because it was an important confounder of race such that its removal resulted in a 30% change for coefficients for proportion of blacks. No interaction terms were significant at the $p<0.05$ level. Neighborhoods with higher proportion of population with less than a high school education had significantly higher odds of belonging to a stroke cluster compared to those with low proportion of the population with less than high school education.

Goodness of fit tests showed no evidence ($p=0.389$) that the model was not fitting the data well. The model had very high specificity (97.8%) (i.e. the ability to correctly predict no cluster given the neighborhood was not in a cluster). However, it had a relatively low (20%) sensitivity (i.e. the ability to predict being in a stroke cluster given that the neighborhood was truly in a cluster). The positive predictive value, or the probability of being in a cluster given the model predicted cluster, was 62.5%. The negative predictive value, or the probability of not being in a cluster given that the model predicted no cluster, was 87.4%. Overall, the model has a correct classification rate of
Table 3.3: Final logistic model showing socioeconomic and demographic predictors of high risk stroke mortality clusters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>LRT* p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.036</td>
<td></td>
<td>-8.467, -3.605</td>
</tr>
<tr>
<td>Geography</td>
<td></td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>Referent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suburban</td>
<td>1.299</td>
<td>-0.170, 2.769</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.351</td>
<td>-0.340, 3.042</td>
<td></td>
</tr>
<tr>
<td>Proportion of Blacks</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 0.02</td>
<td>Referent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;0.02 - ≤ 0.05</td>
<td>1.179</td>
<td>-0.127, 2.486</td>
<td></td>
</tr>
<tr>
<td>&gt;0.05 - ≤ 0.10</td>
<td>1.631</td>
<td>-0.095, 3.357</td>
<td></td>
</tr>
<tr>
<td>&gt; 0.10</td>
<td>-0.629</td>
<td>-2.589, 1.35</td>
<td></td>
</tr>
<tr>
<td>Proportion of Pop with &lt; High School education</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤0.17</td>
<td>Referent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;0.17 - ≤ 0.30</td>
<td>2.913</td>
<td>0.699, 5.127</td>
<td></td>
</tr>
<tr>
<td>&gt;0.30 - ≤ 0.37</td>
<td>3.022</td>
<td>0.740, 5.304</td>
<td></td>
</tr>
<tr>
<td>&gt;0.37</td>
<td>3.898</td>
<td>1.527, 6.268</td>
<td></td>
</tr>
</tbody>
</table>

*LRT (Likelihood ratio test) p-value = test of significance of each group of dummy variables (belonging to one categorical variable). Thus, this tests the statistical significance of the variable as a whole (all parameter estimates of the categories of variable in the model).
86.2%. There were a few outliers, with large positive residuals in the model. These neighborhoods were primarily urban, with the lowest proportion of population of blacks, and the lowest levels of population without high school education.

3.3.4.2 Myocardial infarction

The univariate associations of the socioeconomic and demographic variables of interest with the outcome of belonging to a high risk MI mortality cluster are presented in Table 3.2. Variables with significant associations, based on a liberal p-value=0.20 were further assessed in the multivariable logistic model. The proportion of the neighborhood population of blacks was non-significant, but it was included in the analyses because disparities in MI risk and mortality by race have been reported in the literature.\(^{80, 81, 83, 86}\) The final model, based on the prescribed criteria for removal of variables, had a highly significant likelihood (p<0.001) (Table 3.4). The proportion of the population with less than high school education, modeled as a continuous variable, was the strongest predictor of the odds of being in a MI cluster. Geography (p=0.05) and gender (p=0.03) were significant based on the likelihood ratio test of their respective dummy variables as a group. Suburban and urban neighborhoods had significantly higher odds of belonging to an MI cluster compared to rural neighborhoods. Neighborhoods with a higher proportion of males versus females also had higher odds of being in a cluster. The proportion of the population of black race was not significant (p=0.1), but was included in the final model because it was an important confounder for both geography and gender such that its removal resulted in a more than 20% change for their coefficients. No interaction terms were significant at the p<0.05 level.
Table 3.4: Final logistic model showing socioeconomic and demographic predictors of high risk myocardial infarction mortality clusters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>LRT* p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.541</td>
<td></td>
<td>-8.865, -4.220</td>
</tr>
<tr>
<td>Proportion of Pop with &lt; High School education</td>
<td>14.562</td>
<td></td>
<td>8.963, 20.610</td>
</tr>
<tr>
<td>Geography</td>
<td></td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Rural Referent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suburban</td>
<td>1.558</td>
<td>0.205, 2.911</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.544</td>
<td>-0.033, 3.122</td>
<td></td>
</tr>
<tr>
<td>Proportion of Blacks</td>
<td></td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>&lt; 0.02 Referent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;0.02 - ≤ 0.05</td>
<td>0.306</td>
<td>-0.844, 1.456</td>
<td></td>
</tr>
<tr>
<td>&gt;0.05 - ≤ 0.10</td>
<td>-0.991</td>
<td>-2.950, 0.968</td>
<td></td>
</tr>
<tr>
<td>&gt; 0.10 Referent</td>
<td>-1.494</td>
<td>-3.231, 0.244</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Proportion of Male Population ≤0.50 Referent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of Male Population &gt;0.50</td>
<td>1.024</td>
<td>0.086, 1.962</td>
<td></td>
</tr>
</tbody>
</table>

LRT (Likelihood ratio test) p-value = test of significance of each group of dummy variables (belonging to one categorical variable). Thus, this tests the statistical significance of the variable as a whole (all parameter estimates of the categories of variable in the model).
Goodness of fit tests showed no evidence (p=0.521) that the model was not fitting the data well. The model had very high specificity (90.2%). However, it had a relatively low (51.1%) sensitivity. The positive predictive value was 65.7% while the negative predictive value was 83.3%. Overall, the model had a correct classification rate of 80%. There were only three neighborhoods that the model did not fit well. These were rural neighborhoods that had the most extreme high levels of the proportions of the population without high school education.

3.4 Discussion and Conclusions

The results show that spatial patterns of high risk of stroke and MI exist in the study area. These findings are consistent with those from other studies that have reported that southern states like Tennessee\textsuperscript{80,142,148,159,217}, and specifically Appalachian counties\textsuperscript{145,146,253}, have excess risk of stroke and MI. The excess risk has mostly been attributed to variations in the distribution of stroke and MI risk factors such as race, socioeconomic status, geography (urban vs. rural), and prevalence of other chronic diseases, such as diabetes and hypertension\textsuperscript{83,91,142,159}. However, other studies have reported that geographic disparities exist even after adjusting for variations in these risk factors\textsuperscript{141,145,146,151}. The apparent inconsistency in the association between high risks of stroke/MI and risk factors at the state and county levels suggests that disparities may be due to more localized distributions of risk factors.

To our knowledge, this is the first study to investigate spatial patterns and clusters of stroke and MI risk to better understand observed disparities and identify specific health needs at the neighborhood level to aid population health planning. The
results of the current study provide evidence that the risk of stroke and MI can be highly variable within a county and therefore studies that perform analyses at the county level fail to identify these disparities at lower (neighborhood) levels. For example, Knox and Hamblen counties are often reported to have lower risks of stroke and MI and are not considered economically distressed/disadvantaged when compared to other counties in the area\textsuperscript{145, 146}. However, it is evident from the findings here that a few neighborhoods in these counties have very high risks and are part of significant spatial clusters for stroke and MI. If analyses, research, and planning activities to address disparities in risk are conducted at county or higher levels as is often done, these spatial disparities within the counties would be missed. Therefore, neighborhoods would likely be erroneously ignored in programs geared towards addressing disparities in MI and stroke risk. The implication is that for health research and planning activities to be most effective, the focus must be on neighborhood level characteristics and specific needs to alleviate the variation seen at higher geographic levels.

Other studies have used multilevel analyses, including both neighborhood and individual characteristics, to describe disparities in MI risk for individuals\textsuperscript{93, 94, 106, 110, 112, 153, 157}. One study, using data from the Atherosclerosis Risk in Communities Study, categorized neighborhoods (CTs) into tertiles by neighborhood median household income and found that greater incidence risk of MI was associated with living in lower income neighborhoods\textsuperscript{105}. Diez Rouz, et al. (2001) also found that living in a disadvantaged neighborhood was associated with increased incidence of coronary heart disease, including MI, while adjusting for individual income, education, and occupation and defining neighborhoods as census block groups\textsuperscript{106}. However, some
differences in incidence remained between neighborhoods after adjusting for common socioeconomic factors. The failure of individual level risk factors to substantially explain risk at aggregated levels is a common finding in multilevel studies. Some authors have suggested that neighborhood level socioeconomic variables capture information above and beyond the individual level, and so do not serve only as proxies for individual risk factors. Similar to reports from other studies, we found that neighborhoods with a high proportion of the population with low education had higher stroke and MI risks. However, we did not find significant association between median household income and risk of MI or stroke. This is contrary to findings from previous studies and is likely because these were individual level studies while ours is a population/group (neighborhood) level study. In addition to the level of education, the confounding identified between the geography (urban versus rural), race, and gender distribution of each neighborhood is potentially important to understanding how geographic disparities arise in the study area. The influence of neighborhood socioeconomic and social conditions on health may be related, in part, to availability and accessibility to health care services, the built environment and infrastructure (i.e. quality schools, recreational facilities, stores and restaurants with healthy foods), neighborhood based attitudes towards health and related behaviors (i.e. smoking, physical activity, and diet), and the degree of social support. Since health planning is performed at the population level, identifying geographic disparities for neighborhoods can provide insight into the social conditions, structures, and mechanisms that influence health outcomes in the population to better provide effective population based education campaigns and prevention strategies. Thus, studies, such as this one, that investigate neighborhood
level patterns in risk should be considered in addition to those multilevel studies that assess risk of individuals in neighborhoods to ensure community health resources, services, and other efforts are best targeted to the populations at greatest risk.

Although mortality data are useful and commonly used in epidemiological studies to assess health and its patterns, they are not without limitations. First, the accuracy of the cause of death given on a death certificate can be affected by errors made by physicians or in coding, differences in diagnostic criteria, issues arising when there are multiple causes of death, or errors in data entry\textsuperscript{77}. Lloyd-Jones et al. (1998) reported that death certificates overrepresented coronary heart disease as cause of death, particularly for older populations, and cautioned that its use in etiologic studies could potentially lead to a bias towards the null value\textsuperscript{78}. There is also concern that mortality data reflects past, rather than current, health needs. However, mortality is often the most commonly available data for observational, population-based studies since (in the US) it is freely available through organizations, like health departments and the Centers for Disease Control and Prevention\textsuperscript{77}. Unfortunately, the mortality data in this study contained only decedent’s residential address for geo-coding to the census tract level and gave no information on whether the address was a place other than a private home, such as nursing homes or prisons, thus limiting the ability to assess any effect such issues would have on the results of the study. However, we did identify to the best of our ability, the addresses known to be nursing homes and found that no more than 15 deaths occurred at any given address. Thus, we do not believe these issues would significantly affect the spatial patterns observed.
From a methodological standpoint, while neighborhood level analyses provide the advantage of better insight and understanding of health disparities and needs, they are not without limitations. Due to the small number problem, visualization of raw risks from areas with low population or small number of deaths can be misleading. In this study, this problem was overcome using SEB smoothing of risks that reduces noise associated with population heterogeneity and variance instability by borrowing strength from neighbors. While the removal of noise from low populations with unstable risks eases visual interpretation, it may possibly introduce artifacts into the map and therefore these risks should only be used for visualization and not statistical analyses. Additionally, many smoothing techniques, including the SEB used in this study, are prone to edge effects such that neighborhoods on the edges of the study area have fewer neighbors than those in the interior, so there is less information to borrow from neighbors in smoothing. Thus the risks are shrunk toward a global instead of the local mean. Despite these disadvantages, spatial smoothing of risks minimizes erroneous visual interpretations associated with raw risks by reducing noise, making spatial patterns more evident, and reducing attention to outliers by focusing on the overall geographic pattern of the study area. In this study, the smoothed risks did not change the raw pattern very much, except to make localized patterns more visually obvious for both stroke and MI. This result indicates that extreme values (very high and low risks) in the wide mortality risk range were composed of neighborhoods with stable risks, i.e. risks with low variance. Since the SEB has a larger impact on unstable risks and little to no impact on stable risks (i.e. those with low variances), it is not
unexpected that there were minimal differences between the raw (unsmoothed) and SEB risks.

The visual interpretation of spatial patterns can be strongly affected by the number and width of class intervals used to represent risk values\textsuperscript{239, 258}. To reduce this potential bias, it has been suggested that intervals should be based on the overall shape of the distribution and not statistical frequency\textsuperscript{258}. Thus, this study employed the Jenks, or natural breaks, classification method which defines intervals based on the natural distribution of breaks or groupings in the data\textsuperscript{259}. The visualization of spatial patterns of disease is an important component in identifying geographic disparities. However, it is standard epidemiology practice not to rely on one’s visual interpretation of a map of disease risks to differentiate significant spatial clusters from what may seem to be a cluster visually but is not statistically significant\textsuperscript{240, 257}. Furthermore, interpretations of spatial patterns from visual investigations become even more difficult when the population is heterogeneously distributed throughout the study area, resulting in differences in variances of disease risks across different areas in the map. Thus, statistical comparisons are needed to identify areas where statistically significant clusters of stroke and MI mortality exist, while taking into account population distribution, to better understand disease disparities. This explains the need to use SEB risk maps as well as spatial scan statistics to identify significant high risk spatial clusters. Moreover, other studies have also indicated that interpreting the results of cluster detection along with the spatial distribution of risk, especially with Bayesian smoothing, can strengthen findings of spatial analysis\textsuperscript{117, 260, 261}. 
Spatial scan statistics were used to identify and assess the statistical significance of areas with high risk of stroke and MI clusters. This methodology, implemented in SaTScan 8.0, has many advantages over other cluster detection methods: it corrects for multiple comparisons, adjusts for population heterogeneity in the study area, identifies clusters without a priori specification of their suspected location or size and thus limits pre-selection bias, and allows for adjustment for covariates. Using visualization of spatial patterns of SEB smoothed risk in conjunction with the results of spatial scan statistics in this study, the neighborhoods with the highest risks were consistent and easy to identify. Detection of spatial clusters of disease allows health planners to effectively identify and plan for the specific characteristics and health needs of the populations with the highest risks of disease. For instance, median levels of stroke and MI mortality risk were observed for Knox County in the smoothed risk maps, but cluster detection highlighted just a few neighborhoods with statistically significant higher risk than surrounding neighborhoods in the county. The implication is that health planning and programs can be focused to specific neighborhoods of high risk to better meet their health needs instead of using a one-size-fits-all strategy for all neighborhoods within a county. Thus, neighborhood level analysis allows limited resources and efforts to be targeted to the highest risk communities.

In conclusion, spatial clusters of high risks were identified at the neighborhood level, indicating that not all of the population within counties in the study area experience similar risks of stroke and MI. The implication is that from an effective health planning standpoint, a neighborhood/community level approach is important to ensure that resources and efforts are targeted to the populations most in need. This study also
demonstrated that the use of spatial statistics, cluster detection methods, and GIS can aid health planners in appropriately assessing and identifying spatial disparities in risk in populations so as to understand the unique characteristics and needs and inform evidence based health planning.
CHAPTER 4

4.0 Investigation of Disparities in Geographic Accessibility to Emergency Stroke and Myocardial Infarction Care in East Tennessee using GIS and Network Analysis

Ashley S. Pedigo, MS¹ and Agricola Odoi, BVM, M.Sc., PhD¹

¹The University of Tennessee, Department of Comparative Medicine, 2407 River Drive Knoxville, Tennessee 37996-4543, United States.

This chapter is a manuscript that has been published in Annals of Epidemiology journal (Ann Epidemiol 2010, 20:924–930).

My contributions to this paper included data preparation, analysis, interpretation of results, gathering and reviewing of literature, formulation of discussion topics, as well as drafting and editing the manuscript.
4.1 Abstract

Stroke and myocardial infarction (MI) require timely geographic accessibility to emergency care. Historically, studies used straight line distances as measures of geographic accessibility. Recently, travel time has been recognized as a better indicator of accessibility since travel impedances can be considered. This study utilized finer grained transportation data and network analysis to investigate neighborhood disparities in travel time to emergency stroke and MI care.

Travel times to stroke and cardiac centers were computed using network analysis, while considering distance, speed limit, road connectivity, and turn impedances. Neighborhoods within 30, 60, or 90 minutes travel were identified. Travel time by air ambulance was calculated and adjusted for flying speed and some delays.

Approximately 8% and 15% of the study population did not have timely geographic accessibility to emergency stroke and MI care, respectively. Populations with poor access were located in rural areas. The entire study population had timely access by air ambulance.

This study identified disparities in geographic accessibility to emergency stroke and MI care in East Tennessee. Use of air ambulance or telemedicine could play a vital role in addressing these disparities. This information is important for evidence-based health planning and resource allocation.
4.2 Introduction

Stroke and myocardial infarction (MI) are serious burdens to the US health system with prevalence estimates of 2.9% and 3.6% and annual estimated costs of $73.7 and $177.1 billion, respectively.217 The burdens of these conditions in the US vary by geographical location with the highest risks being reported in the southeastern states.158 Tennessee ranks 3rd and 4th highest in the US with age-adjusted mortality risks of 54.6 and 167.8 per 100,000 population compared to national rates of 43.6 and 135.0 per 100,000 for stroke and coronary heart disease including MI, respectively.217 The Appalachian Region of Tennessee is part of the “stroke belt” and has some of the highest risks of stroke and MI in the state.253

Both stroke and MI have time sensitive treatments. Improved outcomes have been observed for ischemic stroke patients when intravenous thrombolytic treatment is received within 180 minutes of the onset of symptoms. Myocardial infarction outcomes are improved by up to 50% when percutaneous coronary interventions, such as balloon angioplasty, are administered within 60 minutes and 23% when performed within 180 minutes of the onset of symptoms.58,59 It is therefore evident that the earlier patients receive treatment, the greater the likelihood of more favorable outcomes. In addition to being time sensitive, these treatments require specialized equipment and medical expertise that are not available in all hospitals.43,46,60,158 This implies that travel time to an appropriate hospital is a critical component of access to effective treatment.190,262

Access to healthcare is a complex concept with multiple dimensions. Distinction has been made between potential care (where the potential for receiving care exists) and realized care (where health services are actually utilized by the patient).160,161,165
Accessibility to care describes the ability to get potential care which may be impeded by both spatial, like travel impedances, and aspatial factors like the ability to pay\textsuperscript{160, 161}. The focus of this study is spatial or geographic accessibility to emergency care for stroke and MI in the East Tennessee Appalachian Region.

Historically, studies have used straight line (or Euclidean) distances as measures of geographic accessibility\textsuperscript{169, 263}. Recently, travel time has been recognized as a better indicator of accessibility since travel impedances such as speed limits can be considered\textsuperscript{4-6}. With increasing use of Geographic Information Systems (GIS), travel time estimates are becoming widely used in assessing geographic accessibility to care\textsuperscript{4, 5, 161, 175, 264}. The current study sought to enhance these strategies by utilizing advanced GIS methodologies and finer grained transportation data for travel time estimation to investigate disparities in geographic accessibility to emergency stroke and MI care in the East Tennessee Appalachian region.

4.3 Materials and Methods

4.3.1 Study area population

This study was carried out in eleven counties of the East Tennessee Appalachian region that, except for Knox County, all have stroke and MI rates above the national rates and many have rates at or above the state stroke and MI rates (Figure 4.1). Knox County was included because it has some of the best hospitals in the state and patients from other counties in the study area often seek medical care in Knox county. This area is comprised of 168 census tracts with a total population of approximately 857,000. A county “buffer” around the study area was included in the analysis to ensure that all
hospitals nearest to the study area were included since some residents may choose to travel out of the study area for care (Figure 4.1). Factors affecting utilization of care are not considered in the scope of this study, thus a county buffer is sufficient for investigating potential geographic accessibility.

4.3.2 Data collection

4.3.2.1 Hospital distribution data

Distribution of stroke, cardiac, and all other hospitals with an emergency room (ER) in the study area and surrounding counties, were obtained from the Joint Commission on the Accreditation of Health Organizations\textsuperscript{166}. Accredited stroke centers were defined as those that met criteria set by the Brain Attack Coalition\textsuperscript{46}. Hospitals that offered percutaneous coronary intervention services, were considered cardiac centers. Hospitals with the capability of giving emergency stroke or MI care were identified as those with ER. The hospital addresses were geo-coded using GPS Visualizer\textsuperscript{265} and imported into Arc GIS 9.3\textsuperscript{221}. Hospitals were also contacted to determine presence/absence of a helipad and access to air ambulances.

4.3.2.2 Street network data

The street network dataset was obtained from Street Map USA\textsuperscript{266}. This dataset provided information on driving restrictions and connectivity of all street segments in the United States, where a segment refers to sections of the street that have the same characteristics. The reference system used for the data was the North American Datum
Figure 4.1: Map of Study Area and Neighboring Counties

Study Area Counties
1 = Claiborne
2 = Hancock
3 = Hawkins

Neighboring (Buffer) Counties
4 = Union
5 = Grainger
6 = Hamblen
7 = Knox
8 = Jefferson
9 = Greene
10 = Cocke
11 = Sevier
(NAD) 1983 State Plane coordinate system with the high accuracy reference network (HARN) for Tennessee. This coordinate system uses the Lambert conformal conic projection which minimizes scale distortion within the state to allow more accurate distance computations, and is the official projection of the Tennessee Department of Transportation. Travel time for each street segment was computed from the distance and speed limit using the field geometry calculator in Arc GIS 9.3.

4.3.2.3 Cartographic boundary files

Census tract geographic boundary files, used to overlay the street network dataset and hospital locations in ArcGIS 9.3, were downloaded from the US Census Bureau. Census tracts were used to represent neighborhoods and thus chosen as the geographical unit of analysis because they are good proxies of natural neighborhood boundaries and are therefore useful in describing neighborhood characteristics and health disparities.

4.3.2.4 Population data

Census tract and block group level population data were obtained from the 2000 Census summary file 3 and the 2007 population estimates. Block group population data was used to obtain population weighted centroids for census tracts to account for unequal distribution of the population within the census tract.

4.3.3 Data Analysis

4.3.3.1 Network Analysis

Ground travel time to hospitals was calculated with the service area solver of the network analyst extension in Arc GIS 9.3 using the street network dataset. This tool
utilizes Dijkstra’s algorithm for computing shortest paths and minimizes travel time by favoring hierarchical routing techniques for travel impedances. The algorithm calculates travel time along the road network and its travel impedances, without regard to census tract centroids or boundaries, until the travel time exceeds the given limits, which in this study were 30, 60, and 90 minutes. The following algorithm specifications were used: (1) no U-turns to prevent the route from doubling back on to the same street and including dead end streets; (2) segment by segment connectivity values; and (3) the cost (or impedance) set as distance in feet to ensure the shortest travel time route in an emergency situation. Analysis of the 30, 60, and 90 minute travel time buffers was performed separately for stroke, cardiac, and ERs. Neighborhoods (CTs) were assigned to a travel time if their population weighted centroid was within the buffer boundary.

4.3.3.2 Euclidean Distance Model

A Euclidean distance model of travel time to hospitals was calculated using the Buffer Wizard tool in Arc GIS 9.3. Travel time buffers of 30, 60, and 90 minutes were computed from each care center based on assumed average travel speeds of 20, 30, 40, 50, and 65 miles per hour (mph). These buffers were overlayed on the network travel time buffers for comparison.

4.3.3.3 Travel time by air ambulance analysis

Air ambulance flight times from the centroids of CTs to hospitals were calculated using Euclidean distance with a cruising speed of 130 mph, 20-minute patient preparation time, and 3.5-minute warm-up delay. Travel times by air ambulance were analyzed for roundtrip and one-way travel. The roundtrip analysis is indicative of air
ambulances being stationed at the hospital, while the one-way analysis represents situations where air ambulances are distributed throughout the study area so that there is minimal to no travel time required to reach the patient.

4.4 Results

4.4.1 Hospital distribution

There were two accredited stroke centers (Figure 4.2) and nine cardiac centers (Figure 4.3) in the study area. The stroke centers, located in Knox county, were also cardiac centers. In addition to the two stroke/cardiac centers and seven cardiac only centers, there were six other ERs (Figure 4.4). Grainger and Union counties did not have any care center.

Three primary stroke centers, located in Washington and Sullivan counties of Tennessee and in Asheville, North Carolina, were included in the study because they were within the county buffer (Figures 4.2 and 4.4). Similarly, seven cardiac only centers and eight other ERs were located in the county buffer (Figure 4.4).

4.4.2 Ground travel time to accredited stroke centers

Neighborhoods that were within 30 minutes travel time to a stroke center included 93 census tracts and contained 56% of the study area population (Table 4.1 and Figure 4.2). Nearly seven percent of the study population could not reach a stroke center within 60 to 90 minutes. Additionally, one percent of the study population did not have access to a primary stroke center within 90 minutes. The proportion of senior citizens who did not have access to a stroke center within 90 minutes was similar to that of the total population.
Figure 4.2: Travel Time to Accredited Stroke Centers in East Tennessee Appalachian Region
Figure 4.3: Travel Time to Cardiac Centers in East Tennessee Appalachian Region
Figure 4.4: Travel Time to Care Centers with an Emergency Room in East Tennessee Appalachian Region
Table 4.1: Population Distribution within Specified Travel Times to Stroke and Myocardial Infarction Care in East Tennessee Appalachian Region

<table>
<thead>
<tr>
<th>Travel Time (minutes)</th>
<th># Census Tracts</th>
<th>2000 Total Population* (%)</th>
<th>2007 Total Population† (%)</th>
<th>2000 Population≥65 years</th>
<th>2007 Population≥65 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stroke Centers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>93</td>
<td>439,109 (56.2%)</td>
<td>484,814 (56.9%)</td>
<td>54,941 (54.3%)</td>
<td>55,077 (54.3%)</td>
</tr>
<tr>
<td>60</td>
<td>56</td>
<td>281,403 (36.0%)</td>
<td>305,945 (35.7%)</td>
<td>38,365 (37.9%)</td>
<td>38,356 (37.8%)</td>
</tr>
<tr>
<td>90</td>
<td>16</td>
<td>53,549 (6.9%)</td>
<td>59,327 (6.9%)</td>
<td>7,029 (6.9%)</td>
<td>6,990 (6.9%)</td>
</tr>
<tr>
<td>&gt;90</td>
<td>3</td>
<td>6715 (0.9%)</td>
<td>7052 (0.8%)</td>
<td>928 (0.8%)</td>
<td>975 (1.0%)</td>
</tr>
<tr>
<td><strong>Cardiac Centers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>138</td>
<td>664,732 (85.1%)</td>
<td>729,995 (85.2%)</td>
<td>86,092 (85.0%)</td>
<td>86,259 (85.1%)</td>
</tr>
<tr>
<td>60</td>
<td>24</td>
<td>104,041 (13.3%)</td>
<td>114,120 (13.3%)</td>
<td>13,694 (13.5%)</td>
<td>13,588 (13.4%)</td>
</tr>
<tr>
<td>90</td>
<td>6</td>
<td>12,003 (1.5%)</td>
<td>13,023 (1.5%)</td>
<td>1487 (1.5%)</td>
<td>1,551 (1.5%)</td>
</tr>
<tr>
<td>&gt;90</td>
<td>0</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>ERs‡</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>159</td>
<td>748,386 (95.9%)</td>
<td>820,833 (95.8%)</td>
<td>97,618 (96.4%)</td>
<td>97,710 (96.4%)</td>
</tr>
<tr>
<td>60</td>
<td>9</td>
<td>32,390 (4.2%)</td>
<td>36,305 (4.2%)</td>
<td>3,655 (3.6%)</td>
<td>3,688 (3.6%)</td>
</tr>
<tr>
<td>90</td>
<td>0</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>&gt;90</td>
<td>0</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>168</td>
<td>780,776</td>
<td>857,138</td>
<td>101,273</td>
<td>101,398</td>
</tr>
</tbody>
</table>

*Population numbers are based on U.S. Census Bureau 2000 summary file 3
†Population numbers are based on US Census Bureau 2007 population estimates
‡ER = Care centers with an Emergency Room
4.4.3 Ground travel time to cardiac centers

Neighborhoods within 30 minutes travel time to a cardiac center accounted for 85% of the total study population (Figure 4.3, Table 4.1). Only about 2% of the study population could not reach a cardiac care center within 60 minutes. However, the entire study population could travel to cardiac care within 90 minutes.

4.4.4 Ground travel time to other hospitals with an ER

The majority (Table 4.1), of the study area population was within 30 minutes driving time to an ER with the following exceptions: some of the rural areas near the Kentucky border and a few neighborhoods in the mountainous region along the North Carolina border (Figure 4.4). However, all these neighborhoods were within 60 minutes driving time to an ER.

4.4.5 Travel time by Euclidean distance

At 50 mph or higher speeds, the Euclidean distance model tended to overestimate travel time compared to the network model to the point where there was no 90 minute travel buffer in the study area. However, for lower speeds, the model underestimated travel time. Due to space limitations, only 50 mph travel speed was presented (Figure 4.5). It serves as an estimate of the average travel speed in an emergency situation.

4.4.6 Travel time by air ambulance

All hospitals included in the analyses supported air ambulances (had helipad or landing site). Thus, all neighborhoods in the study area could reach a stroke center, cardiac center, or ER within 30 minutes of flight time. When roundtrip travel was
Figure 4.5: Comparison of Travel Time by Network Analysis and Euclidean Distance for 50 MPH Travel Speed to Care Centers in East Tennessee Appalachian Region. (a) Stroke Centers. (b) Cardiac Centers. (c) Hospitals with an Emergency Room.
considered, all neighborhoods in the study area could reach a cardiac center or an ER within 30 minutes and a stroke center within 60 minutes.

4.5 Discussion and Conclusions

It has been shown that for every 15 minutes beyond one hour that treatment is delayed, risk of mortality is increased 1.6 times\textsuperscript{59}. Likewise, it was reported that even within the 180 minute recommended time for treatment of stroke, outcomes improve as time from onset of symptoms to treatment decreases\textsuperscript{44}. Moreover, after arrival to an appropriate care facility, delays of up to 1 hour may be experienced during diagnosis, imaging, or initiation of treatment\textsuperscript{44, 189, 190}. Therefore, shorter travel times to stroke or cardiac centers provide better opportunities for receiving the most effective treatments and improving outcomes.

The results of this study indicate that populations in some neighborhoods of the East Tennessee Appalachian Region are spending half or more of the time within which treatment should be given, travelling to a hospital. Most of these neighborhoods are in the very rural areas. Other studies have reported similar disparities in rural areas with distance and travel time comprising the greatest determinants of geographic accessibility to health services\textsuperscript{140, 143}. Moreover, the unequal geographic distribution of healthcare centers compounds the disparities in access by rural populations\textsuperscript{56}. In this study area, the stroke centers and most of the cardiac centers are clustered in or near the urban areas. Thus, it is important for population health planners to be aware of neighborhoods that lack geographic accessibility so as to better target health service and research programs.
A very positive finding of this study is the improved geographic accessibility to emergency stroke and cardiac care, even for the most rural areas, when air ambulances are used. Although it may not be cost effective and practical to have more stroke or cardiac centers in rural areas, it may be quite beneficial to consider emergency transport of certain patients by air ambulance. Silliman, et al (2003), found that use of air ambulance for rural residents could improve access to thrombolytic treatment with 38% of ischemic stroke patients in their study receiving treatment within the 180 minute recommended time\textsuperscript{168}. They suggested that the increased cost of air transportation ($4623 per patient) would be small compared to the potential savings related to the costs of poorer outcomes from delayed treatments\textsuperscript{168}. Although many factors (such as cost, availability, maintenance issues, weather, terrain and protocols of transport) influence the use of air ambulances\textsuperscript{11}, studies have reported that air transport is cost effective\textsuperscript{181,182}.

Since approximately 96% of the study area population can travel to an ER within 30 minutes, the majority could benefit from strategic telemedicine linkages between stroke or cardiac centers and ERs and/or emergency service providers. Telemedicine programs for stroke care are feasible, reliable and improve outcomes\textsuperscript{184,216}. To our knowledge, there are currently no telemedicine programs for stroke or MI in the study area.

The methodology in this study used travel time estimates to identify neighborhoods lacking timely geographical accessibility to stroke and MI care. While it has been reported that travel time is a better indicator of geographic access than distance to healthcare since travel impedances can be taken into account\textsuperscript{5,7}, other
studies have reported that Euclidean distance is a good estimate of travel time in some cases\textsuperscript{6,172}. Thus, Euclidean distance model for travel time estimates was included in the current study. At lower travel speeds, it underestimated the travel time compared to the network model. Shadid et al. (2009) found the same result in comparing Euclidean distance to travel time in an urban area of Canada\textsuperscript{173}. At higher travel speeds, the Euclidean model overestimated the travel. Studies have reported weak correlations between travel time and distance, likely due to the inability of the Euclidean model to capture the complexity of street networks and associated travel impedances\textsuperscript{5,173}. The studies that suggested distance as a proxy for travel time report that the relationship may not hold in areas where the geographic units are small\textsuperscript{172} or in areas with many travel impedances\textsuperscript{6}.

Several studies have used network data to estimate travel time to care; however, they utilized average values for all segments of the street, regardless of actual impedances of each segment\textsuperscript{5,140,164,173,268}. Such estimates are generally inaccurate since impedances, especially speed limits, change along different segments of the street. This study has shown that finer grained, segment by segment information on travel impedances are readily available in products like StreetMap USA and enable more accurate estimates of travel time. Wang et al. (2008) estimated travel time based on real-world road network data, but the travel time was calculated between the zip code centroids and hospitals\textsuperscript{5}. The current study utilized the service area solver of the network analyst extension in Arc GIS 9.3 where travel time was not calculated between two points but along the street network without boundaries set by geographic polygons. This allows travel time to be compared at lower geographic units of analysis, like CTs.
This study is not without limitations. First, the travel time estimates were based on the legal speed limits. These may be conservative, since emergency vehicles may travel above the legal limit and are exempt from other travel restrictions (e.g. traffic lights). However, these factors are difficult to include in the model since official guidelines on these issues were unavailable. On the other hand, up to half of those experiencing MI and stroke choose to self transport and would thus be subject to the standard traffic laws used in this analysis. Second, this study lacked data on travel impedances such as traffic congestion, weather conditions, and time of day for adjustment of travel time estimates. These factors are difficult to model due to their inherent variability and would require significant assumptions on the average or median values. A potential solution for this issue in future research could be the use of ambulance data to attempt to capture variability in travel time after readily available impedance factors are taken into account.

The above limitations notwithstanding, the advanced GIS methodologies and finer grained transportation data used to accurately estimate travel time in this study identified disparities in geographic accessibility to emergency stroke and MI care in East Tennessee. With specific time sensitive treatments, travel time to appropriate care is an important part of improving health outcomes. Use of air ambulances, telemedicine and/or other outcome improving programs in rural areas could play a critical role in addressing these disparities. This information is invaluable for evidence-based population health planning initiatives that seek to address disparities in geographic accessibility to care for stroke and MI.
CHAPTER 5

5.0 Disparities in Pre-hospital Transport Delays for Acute Stroke and Myocardial Infarction Patients Utilizing Emergency Medical Services

Ashley Pedigo Golden, MS¹ and Agricola Odoi, BVM, M.Sc., PhD¹

¹The University of Tennessee, Department of Biomedical and Diagnostic Sciences, 2407 River Drive, Knoxville, TN 37996-4543, United States.

This chapter is a manuscript that will be submitted for publication in Circulation. My contributions to this paper included data preparation, analysis, interpretation of results gathering and reviewing of literature, formulation of discussion topics, as well as drafting and editing the manuscript.
5.1 Abstract

Pre-hospital delays in receiving emergency stroke and myocardial infarction (MI) care have significant impacts on outcomes. However, use of Emergency Medical Services (EMS) has been shown to reduce delays. Therefore, the objective of this study was to identify disparities in EMS transport times for suspected stroke and MI patients.

Over 3,900 records of suspected stroke and MI patients, from 2006 to 2009, were obtained from two EMS providers (EMS 1 & EMS 2) in East Tennessee. Summary statistics of transport time intervals were computed and their associations with patient characteristics investigated.

Transport times for stroke and MI were similar. Most (80-83%) emergency calls had response times of ≤10 minutes. Over 1/3 of the calls had on-scene times exceeding the recommended 15 minutes. Patients served by EMS 2 were mainly from rural communities and experienced significantly (p<0.05) longer travel times (median=23, range=0-94) than those served by EMS 1 (median=14.5, range=0-74). Almost all suspected MI patients (>96%) from both EMS 1 and 2 were taken to cardiac centers, but only 10% of EMS 2 suspected stroke patients were taken to a stroke center compared to 66% of EMS 1. Older age and being taken to a specialty center were associated with exceeding recommended times.

Since use of EMS are critical for timely access to stroke and MI treatments, the findings of this study are important for guiding local health initiatives that seek to improve health services and outcomes for stroke and MI patients.
5.2 Introduction

Despite recent declines in death rates from stroke and acute myocardial infarction (MI), the burdens of these diseases remain high in the US with direct and indirect costs estimated at $40.9 and $177.5 billion in 2007 for stroke and CHD, respectively\(^1\). Both stroke and MI require time sensitive treatments and therefore transport time to appropriate health facilities is critical for good health outcomes. Improved outcomes have been observed for ischemic stroke patients when intravenous thrombolytic treatment, such as tissue plasminogen activator (tPA), is received within 3 hours of the onset of symptoms\(^42\). Thus, rapid transport and treatment of acute stroke patients is important and therefore it has been recommended that these patients be transported directly to primary stroke centers\(^42\). Current guidelines from the American Heart Association (AHA) and American College of Cardiology recommend that the time from first medical contact to percutaneous coronary intervention (PCI) be 90 minutes or less for MI\(^53\). Reports indicate that health outcomes are improved by up to 50% when PCI is administered within 60 minutes and by 23% if given within 180 minutes of the onset of symptoms\(^58,59\).

While there are two general types of delays (pre-hospital and in-hospital) that can affect timely receipt of stroke and MI treatments, some studies have suggested that the pre-hospital interval, from onset of symptoms to arrival at the hospital, is the source of the longest delay\(^12\). Studies have reported that utilization of Emergency Medical Services (EMS), among other factors, are associated with reduced delays on the time to receipt of treatment for stroke\(^12\) and MI\(^58\). These studies considered pre-hospital time as one time interval, that is, from the onset of symptoms to arrival at the hospital.
However, to better identify disparities and target interventions to reduce delays, pre-hospital delays should be further sub-divided into decision delays and transport delays\textsuperscript{12}. Since EMS play a critical role in providing rapid transport of acute stroke and MI patients, it is important to investigate the specific time intervals involved in transport in order to better recognize when delays occur so that specific delays can be effectively reduced. The specific time intervals involved in EMS transport include: response time, on-scene time, and time required to travel to the hospital. Unfortunately, only a few recent studies have investigated and described the specific time intervals associated with EMS transport for stroke\textsuperscript{15,16} and MI\textsuperscript{17}. Thus, additional studies of different populations, geographic areas, and EMS providers are needed to improve our understanding of this component of pre-hospital delays. Therefore, the objective of this study was to identify disparities in EMS transport time delays, and the factors affecting them, for suspected stroke and MI patients.

5.3 Methods

5.3.1 Study area population

This case study was performed within the East Tennessee Appalachian region, an area that has some of highest mortality risks of stroke and MI in the country\textsuperscript{1}. Thus, although the results are generalizable only to the population within this region, the methods and implications of the study are important and useful to other EMS and populations. The study population included those in two counties each served by one of two participating EMS providers. One EMS (EMS 1) provided service to a mostly urban
county, while the other, EMS 2, served a less populous county within many rural and some mountainous areas.

5.3.2 Data collection and management

Suspected stroke and MI cases reported between 2006 and 2009 were extracted from the EMS dispatch databases. Cases were defined as stroke or MI cases when at least one of the following was true: 1) the emergency caller mentioned one or more symptoms for stroke and MI as defined by the AHA\(^\text{18}\); 2) EMS observed one or more of the defined symptoms; 3) EMS had the clinical impression of a suspected MI or stroke. Confidential patient data were removed before records were released to investigators.

Records provided by EMS 1 were manually entered into an electronic database. Entries were checked for accuracy. Suspected stroke cases from EMS 1 were available for the period 2006-2008, while suspected MI cases from the same EMS provider were only available for 2006 and 2008. Records from EMS 2 covering the period 2007 to 2009 for stroke and MI were received as an electronic database. Data preparations and cleaning were done in SAS 9.2\(^\text{228}\). A total of 4,411 records matched the case definitions. Records were excluded (495 cases, 11.2%) from analyses under the following conditions: 1) patient was <18 years (165 cases, 3.7%); 2) duplicate records (33 cases, 0.75 %); 3) dispatch did not result in transport (48 cases, 1.1%); or 4) the record was missing all time components (249, 5.6%).

The EMS transport time was divided into time intervals deemed useful for identifying time points at which delays could potentially occur: total time, response time, on-scene time, and travel time (Table 5.1). The time of the emergency/911 call was
missing in over half of the records and thus was not used in calculating EMS response. However, for records with 911 call time recorded, the median elapsed time between receiving the call and ambulance dispatch was 1 minute (range: 0-2).

Other variables of interest in describing disparities in EMS transport times included: patient age and gender, year, season (winter: December-February; spring: March-May; summer: June-August; fall: September-November), dispatch reason, and whether the patient was taken to a specialty (stroke or cardiac) center. The hospital was considered a specialty stroke center if it met criteria set by the Brain Attack Coalition, while for MI, the hospital was considered a cardiac center if it provided percutaneous coronary intervention (PCI) services, as per the Joint Commission on the Accreditation of Health Organizations website JCAHO.

Additional variables that were collected by EMS 2, but not EMS 1 included: residency status in the study area, red lights/sirens (RLS) mode of the ambulance to the scene, RLS mode to the hospital, and who made the choice of hospital to which the patient was taken (EMS or patient/family member).

5.3.3 Data analysis

Descriptive analyses for all variables, time intervals, and proportion of patients exceeding the literature or EMS compliance guidelines were performed in SAS 9.2. Cases were considered as exceeding the guidelines if: response time was >10 minutes, on-scene time was >15 minutes, or total time was >60 minutes.

Continuous variables (time intervals and age) were assessed for normality using

Table 5.1: Definitions of emergency medical services (EMS) transport time intervals for stroke and myocardial infarction patients
<table>
<thead>
<tr>
<th>Transport time interval</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total time</td>
<td>Time from EMS dispatch to arrival at hospital</td>
</tr>
<tr>
<td>Response time</td>
<td>Time from EMS dispatch to EMS arrival at scene</td>
</tr>
<tr>
<td>On-scene time</td>
<td>Time from EMS arrival at scene to EMS departure from scene</td>
</tr>
<tr>
<td>Travel time</td>
<td>Time from EMS departure from scene to arrival at hospital</td>
</tr>
</tbody>
</table>
Shapiro-Wilks test and were found to be non-normally distributed (p<0.05). Thus, significant differences in medians of time intervals between categorical variables were assessed using Wilcoxon rank sum (for binary variables) or Kruskal-Wallis (for variables with multiple categories) tests. Associations between categorical variables and the binary outcomes of exceeding recommended times were investigated using chi-square test, or Fisher’s exact test for low cell counts, with significance set at p=0.05. For categorical variables that were significantly associated with exceeding a time guideline, differences between levels/categories were assessed using two-sample test of proportions, adjusted for multiple comparisons using Simes method.

5.4 Results

5.4.1 Patient characteristics

5.4.1.1 Stroke

There were 1,075 suspected stroke cases for EMS 1. Significantly (p=0.01) more cases occurred during 2008 (41.8%) than 2006 (21.9) or 2007 (36.1%). Significantly (p=0.04) more stroke cases occurred in the fall (31.3%) compared to spring (21.0%), summer (21.4%), or winter (26.3%). The most common dispatch reason for EMS 1 stroke patients were ‘Cerebrovascular accident (CVA)/Stroke’ (92.5%), followed by ‘Unconscious/Syncope’ (2%). More vague dispatch reasons (weakness, fall, dizziness, numbness, headache) tended to be reported for female than male patients (p=0.06). The majority of patients were females (63.7%) who were significantly (p<0.0001) older (median: 77, range: 22-100) than males (median: 69, range 22-99). Significantly (p=0.0001) more stroke patients were taken by EMS 1 to an accredited stroke center.
(66.3%) than to a non-stroke center (36.3%). There were no differences (p=0.3) in the use of stroke centers between genders (males: 31.5%, females: 34.8%), nor was there an association between age and the use of stroke centers (p=0.18).

A total of 511 suspected stroke cases were reported by EMS 2. Significantly (p=0.02) more cases occurred in 2009 (38%) compared to 2008 (32.6%) and 2007 (28.6%). Like stroke cases from EMS 1, significantly (p=0.03) more EMS 2 stroke cases occurred in the fall (35.4%) than spring (14.9%), summer (29.9%), or winter (19.8%). The most common dispatch reasons for EMS 2 stroke patients were 'Unconscious/Fainting/Syncope' (21.1%), followed by 'Convulsions/Seizures' (16.8%), and 'CVA/Stroke' (9.6%). Suspected stroke patients with dispatch reasons 'Unconscious/Fainting/Syncope' or 'CVA/Stroke' were more likely (p=0.005) to be transported to a stroke center by EMS 2 than patients with other dispatch reasons. More stroke patients were female (53%) compared to male and the median patient age was 56 (range: 18-98). Unlike stroke patients from EMS 1, there was no difference (p=0.09) in median age between genders for EMS 2 stroke patients. Only 56% of EMS 2 patients were residents of the study area. The distribution of stroke cases by age, gender, year, season, or being taken to a stroke center were not significantly (p>0.05) different between study area residents and non-residents. Only 10% of suspected stroke patients were taken to a stroke center by EMS 2. There were no differences in the use of stroke centers across genders (p=0.11) or ages (p=0.13).

The ambulance response mode was RLS for 93% of suspected stroke patients. The ambulance travel mode from the scene to the hospital was RLS for only 33% of
patients. Use of the RLS mode to the hospital was significantly (p=0.008) higher when the dispatch reasons were more serious (unconscious, convulsions, stroke) compared to other reasons (headache, dizziness, sick person, weakness). However, use of the RLS mode was not significantly (p=0.08) associated with being taken to a stroke center. The hospital to which the patient was taken was selected by EMS personnel for 80% of suspected stroke patients, of which the majority (76%) were taken to the closest facility. When the dispatch reasons were more serious, the choice of the hospital tended to be made by EMS personnel (p=0.05).

5.4.1.2 Myocardial Infarction (MI)

There were 1,754 suspected MI cases for EMS 1 in 2006 (36%) and 2008 (64%). Similar to stroke, significantly (p=0.03) more cases occurred in the fall (36%) season than spring (24.7%), summer (21.7%), or winter (17.5%). The majority of patients were female (53.0%) who were significantly (p<0.0001) older (median: 61, range: 18-99) than male patients (median: 56, range 18-101). The most commonly reported dispatch reason was ‘Chest pain’ (98.7%). The majority (96%) of MI patients were taken to cardiac centers by EMS 1.

A total of 576 suspected MI cases were recorded by EMS 2. There were no differences in the volume of cases across seasons or years. The majority of MI patients transported by EMS 2 were female (52%) and were significantly (p<0.0001) older (median age: 62; range: 19-96) than males (median age: 54; range: 19-94). ‘Chest pain’ (51.8%) and ‘Shortness of breath’ (28.5%) were the two most common dispatch reasons for EMS 2 MI patients. A significant (p=0.02) difference was observed between
genders for dispatch reasons with more males (55.1%) reporting ‘Chest pain’ and more females (64%) reporting ‘Shortness of breath’. As for EMS 1 MI patients, almost all (98.8%) EMS 2 MI patients were taken to cardiac centers.

Similar to EMS 2 stroke patients, only 59% of MI patients transported by EMS 2 were study area residents and their characteristics (age, gender, year, season, or being taken to a MI center) did not differ from those of non-residents. Response mode to the scene was RLS for 96% of patients and was significantly (p=0.04) higher when the dispatch reason was ‘Chest pain’ compared to other dispatch reasons. In contrast, RLS were used for 36% of patients during travel to the hospital. The EMS personnel selected the hospital for 75% of suspected MI patients, of which the majority (79%) were taken to the closest facility. Significantly (p=0.02) more females (61%) and study area residents (85%; p<0.0001) had the choice of hospital made by the patient/family member.

5.4.2 Pre-hospital transport time intervals

5.4.2.1 Stroke

The total time was significantly (p<0.0001) longer for EMS 2 patients compared to EMS 1 (Table 5.2). Patients served by EMS 2 had significantly (p<0.0001) shorter response times than those served by EMS 1. However, the on-scene times of the two EMS providers were not significantly different. The longest interval was travel times to the hospital which were significantly (p<0.0001) longer for EMS 2 compared to EMS 1 patients (Table 2). Patients taken to stroke centers by EMS 1 had significantly shorter (p<0.0001) travel times (median: 15; range: 0-77) than those taken to non-stroke
Table 5.2: Pre-hospital transport time intervals for suspected stroke and myocardial infarction (MI) patients from two emergency medical service (EMS) providers in East Tennessee

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>EMS 1</th>
<th>EMS 2</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median</td>
<td>IQR*</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1075</td>
<td>35</td>
<td>28.5-41.5</td>
</tr>
<tr>
<td>Total</td>
<td>1041</td>
<td>35</td>
<td>28.5-41.5</td>
</tr>
<tr>
<td>Response</td>
<td>1047</td>
<td>6</td>
<td>3.5-8.5</td>
</tr>
<tr>
<td>On-scene</td>
<td>1052</td>
<td>13</td>
<td>9-17</td>
</tr>
<tr>
<td>Travel</td>
<td>1048</td>
<td>15</td>
<td>11-19</td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1717</td>
<td>34</td>
<td>27-41</td>
</tr>
<tr>
<td>Response</td>
<td>1738</td>
<td>6</td>
<td>4-8</td>
</tr>
<tr>
<td>On-scene</td>
<td>1739</td>
<td>13</td>
<td>9.5-16.5</td>
</tr>
<tr>
<td>Travel</td>
<td>1728</td>
<td>14</td>
<td>9.5-18.5</td>
</tr>
</tbody>
</table>

*IQR = Interquartile range

*P-value for difference between median times of EMS 1 & 2
centers (median: 20; range: 0-70). More serious dispatch reasons were associated with significantly shorter response (median: 5; range: 0-29; p=0.02) and travel (median: 21; range: 0-64; p=0.001) times compared to the response (median: 4; range 0-46) and travel (median: 23; range: 0-77) times for less serious dispatch reasons.

Patients taken to stroke centers by EMS 2 had significantly (p=0.01) longer on-scene times (median: 17; range: 0-101) and shorter travel times (median: 14; range: 0-60) compared to on-scene (median: 13; range: 0-122) and travel (median 23: range 0-77) times of patients taken to non-stroke centers. While no significant (p>0.05) associations existed between time intervals and dispatch reasons, patients whose hospital choices were made by the patient/family members had significantly (p=0.003) longer travel times than those for whom the choice was made by EMS personnel. As expected, when the ambulance used RLS mode to the hospital, the median travel time (median 19; range 0-68) was significantly (p=0.0001) lower than when no RLS was used (median 26; range 0-77).

5.4.2.2 Myocardial Infarction (MI)

Like stroke patients, EMS transport times for MI patients differed between the EMS providers (Table 5.2). Patients served by EMS 2 had significantly (p<0.0001) shorter response times, but significantly (p<0.0001) longer travel times and total times compared to EMS 1 patients. The transport times for EMS 1 MI patients were not significantly associated with any patient characteristics.

On the other hand, EMS 2 female patients had significantly (p=0.002) longer travel times (median: 24; range: 0-94) than male patients (median: 22; range: 0-73).
Study area residents experienced both significantly (p=0.002) longer on-scene (median: 14; range 0-51) and travel (median: 24; range: 0-94) times compared to non-residents. Patients whose hospital choice was made by the patient/family members had significantly (p<0.0001) longer travel times (median: 25; range: 0-94) than those whose choices were made by EMS personnel (median: 22; range: 0-76). As for stroke patients, use of RLS mode to the hospital resulted in significantly (p=0.0002) shorter median travel times (median 21; range 0-74) than when RLS was not used (median 24; range 0-94).

5.4.3 Exceeding time guidelines

5.4.3.1 Stroke

The guidelines for EMS response (≤ 10 minutes), on-scene (≤15 minutes), and total (≤ 60 minutes) times were exceeded in 21.5%, 34.9%, and 2.2% of suspected stroke cases from EMS 1 and in 17.6%, 41.5%, and 16.4% of suspected stroke cases from EMS 2, respectively. For EMS 1 patients, significantly more females and those taken to stroke centers exceeded the total time guideline (Table 5.3). Patients whose on-scene time exceeded the guideline were significantly older than those whose on-scene time did not exceed the guideline.

Stroke patients transported by EMS 2 whose total, response, and on-scene times exceeded the EMS guidelines were more likely to be older (Table 5.3). Cases in fall and summer were more likely to exceed the response time guideline than winter cases. Significantly fewer patients with critical dispatch reasons exceeded the response time guidelines compared to those with “Other” dispatch reasons. Study area residents were
Table 5.3: Characteristics of stroke patients whose time intervals exceeded emergency medical services (EMS) guidelines

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>EMS 1</th>
<th>EMS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total &gt;60 min (n=154)</td>
<td>Response &gt;10 min (n=231)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median¹</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Median²</td>
<td>74</td>
<td>74</td>
</tr>
<tr>
<td>Gender: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>21 (91.3)**</td>
<td>150 (65)</td>
</tr>
<tr>
<td>Male</td>
<td>2 (8.7)</td>
<td>81 (35)</td>
</tr>
<tr>
<td>Specialty center: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>20 (87)*</td>
<td>147 (63.6)</td>
</tr>
<tr>
<td>No</td>
<td>3 (13)</td>
<td>84 (36.4)</td>
</tr>
<tr>
<td>Year: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>6 (26.1)</td>
<td>55 (23.8)</td>
</tr>
<tr>
<td>2007</td>
<td>5 (21.7)</td>
<td>86 (37.2)</td>
</tr>
<tr>
<td>2008</td>
<td>11 (47.8)</td>
<td>88 (38.1)</td>
</tr>
<tr>
<td>2009</td>
<td>1 (4.4)</td>
<td>2 (0.9)</td>
</tr>
</tbody>
</table>

¹Median age of patients exceeding time guidelines
²Median age of patients not exceeding time guidelines
³Other dispatch reasons: altered mental status, weakness, dizziness, numbness, headache, unknown
*P ≤ 0.05
**P ≤ 0.01
***P ≤ 0.001
A,B,C,D Categories with different letters are significantly (p<0.05) different
Table 5.3 continued: Characteristics of stroke patients whose time intervals exceeded emergency medical services (EMS) guidelines

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>EMS 1</th>
<th>EMS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total &gt;60 min (n=154)</td>
<td>Response &gt;10 min (n=231)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Season: # (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td>7 (30.4)</td>
<td>66 (28.6)</td>
</tr>
<tr>
<td>Winter</td>
<td>9 (39.1)</td>
<td>59 (25.5)</td>
</tr>
<tr>
<td>Spring</td>
<td>2 (8.7)</td>
<td>50 (21.7)</td>
</tr>
<tr>
<td>Summer</td>
<td>5 (31.7)</td>
<td>56 (24.2)</td>
</tr>
<tr>
<td><strong>Dispatch reason: # (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVA/Stroke</td>
<td>21 (91.3)</td>
<td>212 (92.6)</td>
</tr>
<tr>
<td>Unconscious</td>
<td>0 (0.0)</td>
<td>3 (1.3)</td>
</tr>
<tr>
<td>Seizure</td>
<td>0 (0.0)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Fall</td>
<td>0 (0.0)</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Other&lt;sup&gt;3&lt;/sup&gt;</td>
<td>2 (8.7)</td>
<td>11 (4.8)</td>
</tr>
</tbody>
</table>

<sup>1</sup>Median age of patients exceeding time guidelines
<sup>2</sup>Median age of patients not exceeding time guidelines
<sup>3</sup>Other dispatch reasons: altered mental status, weakness, dizziness, numbness, headache, unknown

*P ≤ 0.05
**P ≤ 0.01
***P ≤ 0.001

A, B, C, D Categories with different letters are significantly (p<0.05) different
Table 5.4: Characteristics of stroke and myocardial infarction patients, transported by emergency medical services provider #2 (EMS 2) whose time intervals exceeded EMS guidelines

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>Stroke</th>
<th>Myocardial Infarction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total &gt;60 min (n=84)</td>
<td>Response &gt;10 min (n=90)</td>
</tr>
<tr>
<td>Study area resident: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residents</td>
<td>52 (66.7)*</td>
<td>47 (70.2)**</td>
</tr>
<tr>
<td>Non-residents</td>
<td>26 (33.3)</td>
<td>20 (29.9)</td>
</tr>
<tr>
<td>Ambulance mode to scene: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lights/Sirens</td>
<td>72 (92.3)</td>
<td>58 (86.6)</td>
</tr>
<tr>
<td>No Lights/Sirens</td>
<td>6 (7.7)</td>
<td>9 (13.4)</td>
</tr>
<tr>
<td>Ambulance mode to hospital: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lights/Sirens</td>
<td>49 (62.8)</td>
<td>24 (35.8)</td>
</tr>
<tr>
<td>No Lights/Sirens</td>
<td>29 (37.2)</td>
<td>43 (64.2)</td>
</tr>
<tr>
<td>Hospital choice made by: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMS</td>
<td>30 (38)**</td>
<td>47 (70.2)*</td>
</tr>
<tr>
<td>Patient/family</td>
<td>48 (62)</td>
<td>20 (29.9)</td>
</tr>
</tbody>
</table>

*P ≤ 0.05  
**P ≤ 0.01  
***P ≤ 0.001
significantly more likely to have total time and on-scene time that exceeded guidelines compared to non-residents among EMS 2 stroke patients (Table 5.4). Significantly more patients, whose choices of hospital were made by the patient/family, had total time exceeding guidelines than those for whom EMS personnel chose the hospital. The ambulance mode to the hospital was RLS for significantly more EMS 2 stroke patients whose on-scene times exceeding guidelines (Table 5.4).

5.4.3.2 Myocardial Infarction (MI)

The recommend guidelines for EMS response, on-scene, and total times were exceeded in 18.0%, 34.6%, and 2.4% of suspected MI patients from EMS 1 and in 16.5%, 37.9%, and 14.6% of those from EMS 2, respectively. More EMS 1 MI cases that exceeded the total time guideline occurred in the year 2008 compared to 2006 (Table 5.5). Older patients and those taken to cardiac centers accounted for significantly higher proportions of EMS 1 MI patients that exceeded the on-scene time guideline than younger patients and those not taken to cardiac centers.

Older MI patients, who were transported by EMS 2, were more likely to have both total and on-scene times that exceeded the guidelines (Table 5.5). Significantly more patients that were not taken to cardiac centers by EMS exceeded the total time guideline than patients that were taken to cardiac centers. More MI patients with the dispatch reason ‘chest pain’ had on-scene times exceeding guidelines than patients with ‘shortness of breath’ or ‘other’ dispatch reasons. As was seen for stroke patients, study area residents were significantly more likely to have total and on-scene times exceeding guidelines than non-residents among EMS 2 MI patients (Table 5.4).
Table 5.5: Characteristics of myocardial infarction patients whose time intervals exceeded emergency medical services (EMS) guidelines

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>EMS 1</th>
<th>EMS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total &gt;60 min (n=42)</td>
<td>Response &gt;10 min (n=315)</td>
</tr>
<tr>
<td>Age (median)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median¹</td>
<td>60.5</td>
<td>59</td>
</tr>
<tr>
<td>Median²</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>Gender: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>25 (59.5)</td>
<td>166 (52.7)</td>
</tr>
<tr>
<td>Male</td>
<td>17 (40.5)</td>
<td>149 (47.3)</td>
</tr>
<tr>
<td>Specialty center: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3 (7.1)</td>
<td>15 (4.8)</td>
</tr>
<tr>
<td>No</td>
<td>39 (92.9)</td>
<td>300 (95.2)</td>
</tr>
<tr>
<td>Year: # (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>7 (16.7)**</td>
<td>119 (37.8)</td>
</tr>
<tr>
<td>2007</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2008</td>
<td>35 (83.3)</td>
<td>196 (62.2)</td>
</tr>
<tr>
<td>2009</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

¹Median age of patients exceeding time guidelines
²Median age of patients not exceeding time guidelines
³Other dispatch reasons: altered mental status, weakness, dizziness, numbness, headache, unknown

*P ≤ 0.05
**P ≤ 0.01
***P ≤ 0.001
Table 5.5 continued: Characteristics of myocardial infarction cases whose time intervals exceeded emergency medical services (EMS) guidelines

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>EMS 1 Total &gt;60 min (n=42)</th>
<th>EMS 1 Response &gt;10 min (n=315)</th>
<th>EMS 1 On-scene &gt;15 min (n=606)</th>
<th>EMS 2 Total &gt;60 min (n=84)</th>
<th>EMS 2 Response &gt;10 min (n=95)</th>
<th>EMS 2 On-scene &gt;15 min (n=218)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Season</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall: # (%)</td>
<td>19 (45.2)</td>
<td>111 (35.3)</td>
<td>255 (42.1)</td>
<td>18 (22.5)</td>
<td>21 (28.4)</td>
<td>64 (29.4)</td>
</tr>
<tr>
<td>Winter: # (%)</td>
<td>7 (16.7)</td>
<td>55 (17.5)</td>
<td>108 (17.8)</td>
<td>21 (26.3)</td>
<td>23 (31.2)</td>
<td>52 (23.9)</td>
</tr>
<tr>
<td>Spring: # (%)</td>
<td>8 (19.1)</td>
<td>71 (22.5)</td>
<td>135 (22.3)</td>
<td>16 (20.0)</td>
<td>12 (16.2)</td>
<td>36 (16.5)</td>
</tr>
<tr>
<td>Summer: # (%)</td>
<td>8 (19.1)</td>
<td>78 (24.8)</td>
<td>108 (17.8)</td>
<td>25 (31.3)</td>
<td>18 (24.3)</td>
<td>66 (30.3)</td>
</tr>
<tr>
<td><strong>Dispatch reason</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chest pain: # (%)</td>
<td>41 (97.6)</td>
<td>310 (98.4)</td>
<td>599 (98.5)</td>
<td>38 (47.5)</td>
<td>36 (48.7)</td>
<td>113 (51.9)^A</td>
</tr>
<tr>
<td>Short of Breath: # (%)</td>
<td>0 (0.0)</td>
<td>1 (0.3)</td>
<td>1 (0.2)</td>
<td>24 (30)</td>
<td>18 (24.3)</td>
<td>58 (26.6)^B</td>
</tr>
<tr>
<td>Other*: # (%)</td>
<td>1 (2.4)</td>
<td>4 (1.3)</td>
<td>5 (0.8)</td>
<td>18 (22.5)</td>
<td>20 (27.0)</td>
<td>47 (21.6)^B</td>
</tr>
</tbody>
</table>

*Median age of patients exceeding time guidelines
**Median age of patients not exceeding time guidelines
*Other dispatch reasons: altered mental status, weakness, dizziness, numbness, headache, unknown

*P ≤ 0.05
**P ≤ 0.01
***P ≤ 0.001
the hospital choice was made by the patient/family, more patients had total times exceeding guidelines than those whose hospital choice was made by EMS personnel. Patients that had on-scene times exceeding the guidelines were more likely to have had RLS used as the ambulance mode to the hospital (Table 5.4).

5.5 Discussion and Conclusions

Pre-hospital delays for stroke and MI patients have a significant impact on eligibility for and timeliness to receive emergency treatments. Utilization of EMS has been shown to reduce treatment delays for stroke and MI patients. However, only a few studies have investigated the specific time intervals related to EMS response and transport as a component of pre-hospital delay. The EMS times for stroke and MI were similar, which was expected given that both conditions are priority emergency dispatches. The median response times in this study were similar to those from other studies: 5 minutes, 5.5 minutes, 6 minutes, 7.5 minutes, and 8 minutes. Median on-scene times for the current study were also consistent with other studies that reported on-scene time of 13 minutes for stroke and 14.5 minutes for MI. However, these results were lower than others that found median on-scene times for suspected stroke patients ranging from 18 to 20 minutes. It was unclear whether these differences were related to population variation or diversity in EMS assessment, triage, or treatment protocols. Differences in EMS on-scene times may reflect varying levels of efficiency, experience, or attitude of EMS personnel serving different populations.

While the on-scene time accounted for the longest EMS time component in one study, travel time to the hospital comprised the longest time delay in this study. The
median travel time to the hospital was reported by two studies in urban areas as 11 minutes for stroke patients\textsuperscript{15,198}. In this study, the travel times for both EMS providers were longer than 11 minutes. This is most likely related to both factors that affect travel time and the distribution of hospitals in urban (EMS 1) versus rural (EMS 2) areas.

Distance and travel time have been reported to be the greatest determinants of access to health services for populations living in rural areas\textsuperscript{143}. In our previous research that included the current study area, a GIS network analysis model estimated longer travel times (>90 minutes travel to the nearest stroke center) for the most rural areas, but also in less rural areas that had limited access to high speed and/or highly connected roads (like interstate highways)\textsuperscript{270}. Additionally, results from that work indicated that specialty centers were clustered in urban centers. The unequal geographic distribution of health facilities has been reported to compound the disparities in access\textsuperscript{56}.

For both EMS providers, travel times for stroke patients were significantly lower when the patients were taken to stroke centers. Other studies have reported that direct transport to a stroke center was associated with short symptom-onset-to-treatment time, and better outcomes\textsuperscript{45}. Given the distance, travel time, and geographic distribution of stroke centers in the study area, it was expected that EMS 2 patients transported to a stroke center would experience longer travel times. The unexpected lower travel times were attributed to the fact that air transport was used for 63% of patients transported to stroke centers. However, the travel time reported in these cases most likely represents the travel time by ambulance to the helicopter landing site (median: 11 minutes, range: 9-15.2) which was significantly (p=0.0001) lower than ground travel time to a stroke center (median: 25.5 minutes; range: 22.4-46.5). Therefore, the association between
travel time and transportation to a stroke center for EMS 2 may have misclassification bias since the definition of travel time included arrival at the hospital and data on travel time from the helicopter landing site to arrival at the stroke centers were not available. In our previous research, which included this study area, we found that patients transported by air ambulance from a helipad could reach a stroke center could within 15 minutes270. Whether the association would be attenuated or not with the addition of air transport time to the travel time component is difficult to tell. However, the association between shorter travel time and use of stroke centers was strong for EMS 1 patients, although none used air transport. These findings suggest that use of stroke centers, regardless of mode of transportation, was associated with shorter travel times in this study. Additionally, it has been found that use of air transport for stroke patients in rural areas decreased travel delays and increased the likelihood of receiving life-saving treatments168.

Recommendations on how EMS should optimize pre-hospital time intervals are limited in the literature17. This study identified patients with long delays based on guidelines for the response (<10 minutes)17, on-scene (<15 minutes)17, and total (<60 minutes)211 times. Guidelines for travel time could not be found. Characteristics of patients with these long delays, as well as information on the factors affecting them, can inform both local public health officials and EMS on how to target initiatives related to education or protocols to reduce delays and improve the likelihood of stroke and MI patients receiving appropriate treatments.
Patient demographic factors reported by some studies to be associated with increased pre-hospital delay for MI include: older age, females, and black ethnicity\textsuperscript{200}. However, the relationships seem to be less clear for stroke patients with some reporting significant associations\textsuperscript{15}, while others reported no associations\textsuperscript{190}. The current study found that age and gender were significantly associated with longer delays for some EMS intervals for both stroke and MI. This is probably due to the population dynamics since a greater proportion of patients were female and older, particularly for stroke. A study reported that EMS is more likely to be utilized by older and/or female patients for stroke and MI\textsuperscript{209}. In this study, older patients were consistently more likely to have on-scene times exceeding the recommended 15 minutes. This is probably because longer times are needed to stabilize the patient. Similar to another study\textsuperscript{200}, males were more likely to have serious dispatch reasons for both stroke and MI compared to females. These dispatch reasons were also associated with increased use of RLS to the hospital and use of specialty centers, resulting in shorter response, travel, and total times for more critical patients in this study. Similar results for serious symptoms have been reported by other studies\textsuperscript{190,198,200}. Race could not be investigated due to lack of variation in the population. Consistent with the demographic profile of our study area, more than 96% of the patients were white, non-Hispanic race.

Over 40% of stroke and MI patients from EMS 2 were non-residents and despite similarity in patient demographics between study area residents and non-residents, significant differences on delay intervals were observed. Residents were more likely to have longer response times because they are generally spread throughout the study area, including in the more rural areas, compared to non-residents that tended to be
concentrated in the more urban areas. As well, residents were more likely to have the hospital chosen by the patient/family member as opposed to EMS. Choosing a hospital further away contributes to the longer delays observed for study area residents compared to non-residents. Similar to findings from another study\textsuperscript{15}, the longer delays in response time during the fall season probably reflects increased call volume and traffic congestion resulting from increased numbers of visitors in the area for the season.

Patients transported to specialty centers were more likely to have longer delays in on-scene time probably because the patients were more critical and therefore required more time to be stabilized. This is evidenced by the associations between more severe dispatch reasons and being taken to specialty centers with longer on-scene delays. Moreover, patients with high priority, requiring use of RLS to the hospital, had significantly longer on-scene and travel times to a specialty center. One study also reported longer on-scene times for more serious patients; moreover, once at the hospital, these patients were seen by a physician twice as fast\textsuperscript{208}. Conversely, a study reported that the implications of longer on-scene times are unclear\textsuperscript{198}. Despite increased delays in on-scene time, the results from this study indicated that patients taken to specialty centers had shorter travel time to the hospital. Other studies have indicated that direct transport to a specialty hospital may not significantly decrease the overall pre-hospital delay, but that in-hospital delays are significantly reduced and therefore total time to treatment is shorter resulting in better health outcomes\textsuperscript{17, 46}. Thus, recommendations for pre-hospital protocols to incorporate EMS bypassing non-specialty centers are advocated\textsuperscript{42, 58}.
When the choice of hospital was made by the patient/family, the travel time was significantly higher than when EMS personnel made the choice. The policy observed by both EMS in this study was ‘informed decision’, implying that the patient/family’s wishes must be observed for transport after medical information has been given by the EMS personnel. In instances when the EMS personnel make the decision, the policy is generally to take the patient to the nearest facility with exceptions of level 1 trauma. Thus, the shorter travel times observed in the current study for patients whose choice of hospital was made by EMS personnel probably reflect this nearest facility policy. These patients were also more likely to have serious dispatch reasons and to have had the ambulance use RLS *en route* to the hospital. For the majority of patients in both EMS, the closest hospital was a cardiac center, which would explain the high percentages (96-98%) of transport to cardiac centers observed. It has been reported that the closest hospital for 40% of the US population is a cardiac center\textsuperscript{164}. However, there were only two specialty stroke centers in the study area. When patients were transported to stroke centers, the decision was more often made by EMS. Given the policy of ‘informed decision’, the study area population could greatly benefit from targeted education that encourages patients/family to choose specialty centers when stroke or MI is suspected.

The decision delay component (including recognition of symptoms, decision to seek care, and use of EMS) and the factors affecting it that have been investigated by many other studies\textsuperscript{205, 209} was beyond the scope of this study which sought to characterize only EMS associated delays. Thus, the current study contributes to the body of evidence for only a portion of the total pre-hospital delay. Others have suggested, however, that if public education interventions targeted to reduce decision
delays are successful over time, delays associated with EMS transport, identified in this study, will become increasingly important.  

This study was limited by the unavailability of data on some patient related factors that have been previously reported to be associated with pre-hospital delays, including: history of past stroke/MI, co-morbidities; type or severity of the attack, being/living alone at onset, awakening with symptoms, and transfer from another hospital. Furthermore, since all confidential patient data were removed from the database, identified delays for patients could not be linked to their medical records, hospital discharge, or personal outcomes. Future studies would benefit from such data linkages so as to investigate impacts of delays on health outcomes.

The time of the emergency/911 call was missing in over half of the records and thus was not used in calculating EMS response for this study. However, of the records with the 911 call time recorded, there was a median elapsed time of 1 minute between receiving the call and ambulance dispatch and therefore would have negligible impact on our findings.

Misclassification of stroke and MI could have occurred since cases were selected based on symptoms mentioned/observed by callers, dispatchers, or EMS personnel. However, since the goal was to assess timeliness of EMS transport when MI and stroke are suspected, regardless of final diagnosis, this should not have had a significant impact on our findings. Moreover, a study reported that pre-hospital times for suspected stroke patients were not different between final diagnoses and another suggested that
pre-hospital times are not likely to be affected by final diagnosis since it is not rendered by EMS but by physicians at hospitals\textsuperscript{212}.

Notwithstanding these limitations, this study identified important disparities in EMS transport delays. Long delays were identified for on-scene and travel times. Further investigations are needed to ascertain the factors that contribute to longer on-scene times for more serious patients and how these delays can be shortened through professional education, training, or improved protocols. “Load and go” approaches should be evaluated for appropriateness based on specific patient characteristics to reduce on-scene times\textsuperscript{198}.

Since, the longest delay involved travel time to the hospital, strategic telemedicine linkages between stroke or cardiac centers and EMS would be beneficial. Telemedicine programs have been shown to be feasible, reliable and improve outcomes\textsuperscript{216}. Moreover, with guidance from specialty centers via telemedicine linkages, EMS personnel may be able to reduce on-scene time. To our knowledge, there are currently no telemedicine programs for stroke or MI in the study area. Long delays for rural EMS patients might also be addressed through increased use of air transport for rural patients\textsuperscript{270}.

While almost all patients were taken to cardiac centers, possibly because of the high availability of such centers in the study area, a lot fewer were taken to stroke centers. Efforts to increase the number of suspected stroke patients being taken to stroke centers in the study area are warranted since patients taken to specialty centers in this study were found to have shorter travel times to care. Further, patient/family member education on the benefits of choosing specialty centers, in addition to reviews
of EMS transport protocols to include bypass options are recommended for local health planners and EMS personnel to increase the use of specialty centers for suspected stroke and MI patients in East Tennessee. These methods have been investigated and shown to be linked to improved outcomes by other studies\textsuperscript{42, 59}.

Efforts aimed at improving the EMS portion of delay, including implementing priority dispatches, dispatcher and paramedic training, rapid triage and assessment, transport, and prior notification have been reported to be successful in decreasing delays\textsuperscript{199} and should be considered. The delays in specific EMS transport intervals, and the factors affecting them, identified here are vital pieces of information for local health initiatives that seek to improve health services and outcomes for stroke and MI patients.
6.0 Summary, Discussions, and Conclusions

To my knowledge, this was the first study to investigate the clustering of population characteristics that are associated with stroke or MI at the neighborhood level. Based on knowledge of the study area, the four PNs identified were a unique and sensible classification of neighborhoods based on socioeconomic, demographic, and geographic characteristics of East Tennessee. The geographic distribution of identified PNs revealed that the most affluent neighborhoods were located in suburban areas, while the least affluent neighborhoods were located in the downtown areas. Thus, it was not surprising to find that neighborhoods in the urban PN had higher stroke and MI mortality risks compared to those in the suburban, affluent PNs. The primarily rural PNs also had higher mortality risks than the more affluent PNs, but had different risk factors than the urban PNs despite similar mortality risks. These results indicate that the use of one, a few, or composite measures of socio-demographic factors is not sufficient for capturing unique neighborhood characteristics, challenges, and health needs that aid local health planners in understanding disparities in health outcomes. Thus, implementation of the multivariate methodologies used in this study and the careful integration of the findings in health planning activities will be useful in guiding health resource allocation, service provision, and policy decisions at the local level.

One of the strengths of this study was the use of robust PCA to account for issues of interdependencies among variables, different variable scales and outliers that are likely to arise when using a large number of risk factors to classify neighborhoods.
into similar groups with similar characteristics. Furthermore, the fuzzy cluster strategy allowed neighborhoods to have associations with more than one PN, giving insight into the structure of neighborhood characteristics and allowing overlaps between neighborhoods to be identified. The observed overlaps implied that it was necessary to consider some neighborhoods in more than one PN in population health planning for these areas. This allows health initiatives to be targeted at the neighborhood level based on the unique population characteristics. The implication of this is that health professionals are able to use a needs-based approach to planning and service provision instead of using a “one-size-fits-all” strategy.

High risk spatial clusters of stroke and MI mortality were identified at the neighborhood level, indicating varying mortality risks within counties in the study area. The implication is that from an effective health planning standpoint, a neighborhood/community level approach is necessary to ensure that resources and efforts are targeted to the populations most in need. To my knowledge, this is the first study to investigate spatial patterns and clusters of stroke and MI risk to better understand observed disparities and identify specific health needs at the neighborhood level. The results of this study provide evidence that the risk of stroke and MI can be highly variable within a county and therefore studies that perform analyses at the county level fail to identify these disparities at lower (neighborhood) levels. If analyses, research, and planning activities to address disparities in risk are conducted at county or higher levels, as is often done, the identified spatial disparities at the neighborhood level would be missed and erroneously ignored in programs geared towards addressing disparities in MI and stroke risk. Thus, for health research and planning activities to be
most effective, the focus must be on neighborhood level characteristics and specific needs to alleviate the variation seen at higher geographic levels.

Neighborhoods with high proportions of blacks and those with low education attainment had higher stroke mortality risks, while urban neighborhoods and those with high proportion of females had higher MI risks. Significant confounding was identified between the geography (urban versus rural), race, and gender distribution of neighborhoods. This finding is potentially important to understanding how geographic disparities arise in the study area. Since health planning is performed at the population level, identifying geographic disparities for neighborhoods can provide insight into the social conditions, structures, and mechanisms that influence health outcomes in the population to better provide effective population based education campaigns and prevention strategies. Thus, studies, such as this one, that investigate neighborhood level patterns in risk should be performed in addition to multilevel studies that assess risks of individuals in neighborhoods to ensure community health resources, services, and other efforts are best targeted to the populations at greatest risk. This study also demonstrated that the use of spatial statistics, cluster detection methods, and GIS can aid health planners in appropriately assessing and identifying spatial disparities in risk in populations to better understand the unique characteristics and needs and inform evidence based health planning.

Due to the time sensitive nature of stroke and MI treatments, timeliness of access to appropriate care is an important part of improving health outcomes. Some neighborhoods of the East Tennessee Appalachian Region are so far from the nearest
stroke or cardiac center that patients from these neighborhoods have to spend half or more of the time within which treatments for stroke or MI should be given, travelling to a hospital. Most of these neighborhoods are in the very rural areas. Moreover, the stroke centers and most of the cardiac centers are clustered in or near the urban areas. Furthermore, this study provided evidence that use of an air ambulance for emergency transport improved geographic accessibility to emergency stroke and cardiac care, even for the most rural areas. Additionally, since approximately 96% of the study area population was found to be within 30 minutes travel time of an emergency room (ER), strategic telemedicine linkages between stroke or cardiac centers, ERs, and/or emergency service providers would also likely improve geographic accessibility to care. However, to my knowledge, there are currently no telemedicine programs for stroke or MI in the study area.

Given that the eligibility and efficacy of stroke and MI treatments is dependent on timeliness of treatment, accurate estimations of travel time to appropriate hospitals are quite critical. The current study enhanced previous methods for assessing travel time to care by utilizing advanced GIS methodologies and segment level transportation data on travel impedances. This methodology provided realistic, accurate, and efficient estimation of travel time that were important for identifying neighborhoods without timely geographical accessibility to stroke and MI care. This information is invaluable for evidence-based population health planning initiatives that seek to address disparities in geographic accessibility to care for stroke and MI patients.

Emergency medical services transport times comprised an important portion of the critical, time sensitive treatment windows for suspected stroke and MI patients in the
study population. Both EMS providers in the study provided rapid response to the majority of suspected stroke and MI patients; however, long delays in on-scene and travel times were identified. Over 1/3 of the patients experienced long on-scene delays. These delays were often associated with older patients and those with more serious symptoms, suggesting that long on-scene delays are highly influenced by the additional time required to assess, stabilize, and extract these patients. Delays for these patients could be reduced through EMS personnel education, training, or improved protocols. Additionally, telemedicine support could aid EMS personnel in reducing on-scene assessment and treatment times of patients. The longest delay in EMS transport time was travel time to the hospital, with patients living in the more rural areas experiencing longer delays than those in the more urban areas. This finding was not unexpected given that stroke and cardiac centers were located in urban areas. Thus, similar to recommendations for increasing geographic accessibility for patients in rural areas, use of air transport, compared to ground ambulance, could reduce delays in travel times. While air transport was used for a few stroke patients in the study area, efforts aimed at increasing the rate of its use for both stroke and MI patients and those living in rural areas are needed to reduce travel time delays, as well as use of specialty centers.

Stroke and MI patients taken to specialty centers were found to have shorter travel times in this study. Unfortunately, the number of stroke patients taken to stroke centers was very low. Stroke patients were more likely to be taken to a stroke center when EMS personnel chose the hospital compared to when the patients themselves or a family member made the hospital choice. Given these findings, efforts by local health planners and EMS personnel to improve the use of specialty centers in East Tennessee
should include patient/family member education on the benefits of choosing specialty centers and reviews of EMS transport protocols to include bypass options of non-specialty centers. Since use of EMS is important in reducing delays to stroke and MI treatments, the identified disparities in delays for specific EMS transport times, and the factors affecting them, are vital pieces of information for local health initiatives that seek to improve health services and outcomes for stroke and MI patients in East Tennessee.

The findings of this work, as well as some of the challenges encountered in its execution, allow the identification of potential avenues through which future research may address disparities and continue to enhance current methodologies for investigating health determinants and health outcomes at the neighborhood level. For instance, there is evidence from the findings of this work that future research and health planning approaches could benefit from integrating the analytical approaches used here so as to identify and better understand neighborhood disparities in health outcomes for improved service provision.

Populations, and thus their characteristics, are dynamic due to migration, urbanization and changes in economic conditions, housing and social developments, as well as changing social attitudes. This implies that the findings of studies investigating population characteristics, such as this one, represent only a snapshot of the disparities. Therefore, such studies need to be supported by temporal analyses to assess changes over time to better support population health planning. Fortunately, the methods utilized in the current studies lend themselves to addressing temporal patterns very well and so future investigations would benefit from including a temporal component.
An on-going challenge during this work was the lack of some data at the neighborhood level. It is known that co-morbidities (particularly obesity, diabetes, and hypertension) are important determinants of stroke and MI health. These, in addition to behavioral risk factors (lack of physical activity, poor diet, smoking, and alcohol consumption), and data on the built environment are important in explaining observed disparities but are not readily available at the neighborhood (census tract) level. Assessment of the distribution of these factors at the neighborhood level would provide additional insights into determinants of these neighborhood health disparities. Thus, future neighborhood level studies should include these factors to investigate their impact, along with demographic and socioeconomic influences, on geographic disparities of neighborhoods in order to better provide effective evidence-based disease control strategies.

Additionally, this work investigated disparities in only stroke and MI mortality. Since not all patients that experience stroke and MI die from it, disparities observed for mortality may differ from those observed for populations having new stroke/MI (incidence) and recurrent disease (prevalence). Further, the focus of health planning initiatives should take into account the impact of stroke and MI on quality of life measures, including: healthy years of life lost (HYLL), an estimate of the length of time a person would have lived if they had not died from stroke/MI; potential years of life lost (PYLL), similar to HYLL but more weight is given to deaths among younger persons and so it measures the impact of premature death; and disability adjusted life years (DALY), which estimates the number of healthy years lost to both premature death and disability. Thus, it is worthwhile, from a health planning standpoint, for future studies to identify
factors associated with not only incidence and prevalence of disease, but also measures of quality of life to aid in the development/improvement of disease control strategies at the neighborhood level. In addition to health outcome measures, hospitalization data are useful in assessing who, when, why, and how patients are treated for stroke and MI. This is particularly important for studies investigating timely access to care and its potential impact on receiving time sensitive stroke and MI treatments.

Future studies addressing EMS delays should consider using prospective methods that would allow: (a) integration of data from 911 dispatch centers (for better collection of the time of the 911 call); (b) improved collection of data by EMS on the time of onset of patient symptoms; and (c) inclusion of data on in-hospital delays (delays in receipt of care) in order to identify and better understand the role EMS delays play in access to stroke and MI care. Additionally, future studies should follow patients transported by EMS to assess health outcomes (death, recurrent attack, re-hospitalization, complications, and quality of life measures) so as to investigate potential associations between health outcomes and timeliness of access to treatment.

In conclusion, this work identified important disparities in neighborhood characteristics, stroke and MI mortality risks, geographic accessibility to stroke and MI care, and delays in EMS transport of stroke and MI patients. These findings are invaluable in guiding resource allocation, service provision, and policy decisions to support evidence-based population health planning and policy that aim to reduce these disparities in the East Tennessee study area. Furthermore, the novel and advanced
techniques utilized to investigate these disparities enhance current strategies for integrating the broad determinants health in research and planning. Additionally, these methods can be employed in future studies for the identification of important health disparities for a wide variety of health outcomes (mortality, hospitalization, HYLL, PYLL, DAYL, etc.), while allowing effective and appropriate consideration for both fundamental and temporal changes in the population at risk and modes of access (travel time, EMS) to care. Overall, the neighborhood focused approach and findings of this study are important for focusing efforts aimed at reducing/eliminating health inequities at the local level with the ultimate goal of improving health of the entire population.
References


178


172. Yamashita T, Kunkel SR. The association between heart disease mortality and geographic access to hospitals: County level comparisons in Ohio, USA. *Social Science & Medicine* 2010;70:1211-1218.


229. SPSS. Answer Tree 3.1. In. 3.1 ed. Chicago, IL; 2002.


245. StataCorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.


Vita

Ashley Pedigo Golden was born on June 10, 1985 in Westmoreland, Tennessee. She earned a Bachelor of Science (2006) and Masters of Science (2008) in Food Science and Technology with a minor in Statistics from the University of Tennessee, Knoxville. She received the Doctorate of Philosophy in Comparative and Experimental Medicine with a concentration in Epidemiology in 2011 from the University of Tennessee, Knoxville. Ashley married David Golden on March 15, 2011.