4-2023

Multimorbidity in Diverse Populations: Stratified Analysis of Race/ Ethnicity, Age, Obesity, and Healthcare Costs

Manal Alshakhs
University of Tennessee Health Science Center

Multimorbidity in Diverse Populations: Stratified Analysis of Race/Ethnicity, Age, Obesity, and Healthcare Costs

Abstract

This research aims to fill an essential gap in understanding how Body Mass Index (BMI) cutoffs relate to multimorbidity across races in the United States (US). Given the significant and growing rates of obesity and multimorbidity, as well as the known differences in healthy fat distribution among different races, this is an important area of research. BMI is a widely used but imperfect measure of obesity, as it does not account for differences in body composition. However, it is still used as a diagnostic tool. It is vital to ensure that the cutoffs used to define obesity are appropriate for all populations, particularly given the racial disparities in multimorbidity rates. This proposed framework for evaluating BMI cutoffs across races for multimorbidity considered a range of measures, such as, including incidence rates of prevalent diseases, age, gender, type of patient visits, and type of health insurance to arrive at questioning the current World Health Organization (WHO) BMI cutoffs in the US. This research demonstrated that having the exact BMI cutoffs across all races does not serve all populations ideally through three assessments. First, it assessed differences in the prevalence of multimorbidity by race. It identified disease combinations shared by all races/ethnicities, shared by some, and those unique to one group for each age/obesity level. These findings demonstrated that despite controlling for age and obesity, there are differences in multimorbidity prevalence across races. Second, the study developed models to project total charges for the most common multimorbidity combinations in the US and evaluated the accuracy of these models across different racial and ethnic groups and multimorbidity patterns. The relationship between healthcare costs and multimorbidity varied for each racial group and depended on the specific combination of chronic conditions, age, and obesity status. Third, it assessed the relationship between BMI and healthcare burden across race and healthcare utilization among middle-aged patients in the US. It demonstrated that the relationship between BMI and healthcare burden varied across races within the same healthcare care utilization category. This research can improve health outcomes and reduce the risk of chronic diseases associated with obesity and multimorbidity, particularly among vulnerable populations. It will also be essential to consider the potential implications of any new BMI cutoffs on clinical practice and health policies related to obesity and multimorbidity in serving unique clinical needs. More work must be done to understand how multimorbidity, BMI, age, and healthcare burden associate across races.

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Health Outcomes and Policy Research

Research Advisor

Charisse Madlock-Brown, PhD

Keywords

Healthcare burden, Healthcare costs, Healthcare utilization, Multimorbidity, Race/ethnicity stratification

Subject Categories

Analytical, Diagnostic and Therapeutic Techniques and Equipment | Diseases | Endocrine System

UNIVERSITY OF TENNESSEE HEALTH SCIENCE CENTER

DOCTORAL DISSERTATION

Multimorbidity in Diverse Populations: Stratified Analysis of Race/Ethnicity, Age, Obesity, and Healthcare Costs

---

Author:
Manal Jawad Alshakhs

Advisor:
Charisse Madlock-Brown, PhD

A Dissertation Presented for The Graduate Studies Council of
The University of Tennessee Health Science Center
in Partial Fulfillment of the Requirements for the Doctor of Philosophy degree from
The University of Tennessee

in

Health Outcomes & Policy Research: Health Informatics & Information Management
College of Graduate Health Sciences

April 2023

DEDICATION

To the graduate students who persevere.
Myself…and those who believed in me!

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God, the Almighty, for aiding me in this endeavor.

Words cannot express my gratitude to my advisor, Dr. Charisse Madlock-Brown. Dr. Madlock-Brown constantly guided me throughout the lengthy stages of this degree with patience, kind words of encouragement, and invaluable advice. Her cunning vision and enthusiasm were strong drivers for achieving my goal.

I would also like to thank my committee members, Dr. Jim Bailey, Dr. Simonne Nouer, Dr. Rebecca Reynolds, and Dr. Shelly White-Means for their guidance and support. I could not have made it without them. With their blended specialties, valuable remarks, and splendid smiles, I improved my research.

I am also grateful to the current program director, Dr. Simonne Nouer, and the previous director, Dr. Sajeesh Kumar, for their advice, concern, and reassurance.

Special thanks to Billy Barnett, our IT whiz, the best problem solver with the quickest turnaround. Exceptional thanks to my designer daughter, Huda Aliathan, for her unique slide design “Does One Size Fit All” and constantly being available for slide design consultations.

To my program colleagues at UTHSC, I can’t thank you enough for your support, encouragement, and concern. I could not have survived this without you.

My gratitude extends to my Chair, Dean, colleagues, and friends at Imam Abdulrahman Bin Faisal University in Saudi Arabia for supporting me. I thank the sponsorship and assistance of the Saudi Arabian Cultural Mission in Washington, DC.

To my parents, whom through their prayers I thrived, no words can express my gratitude. My family, immediate and extended, thank you for being my driving force. To my friends, thank you for being there for me. To my Collierville neighbors, thank you for your hospitality and for making me feel at home. To every person I met on this journey, thank you for our interaction.

PREFACE

The body of this dissertation is organized in a way that first introduces readers to our rationale for choosing the research purpose and specific aims—as well as to present an overview of the literature. A concluding chapter relates all research elements back to our final thoughts about the findings and their significance.

For readers to have immediate access to the full presentation of our previously published research studies, the articles are presented in the appendices. This mode of presentation allows for Chapters 2, 3 and 4, which use them as their basis, to focus more narrowly on a summary and discussion of those articles in Appendices A, B and C and to show specifically how they relate to the dissertation's larger goals. References in the chapters to relevant sections, tables, or figures in these appendices look like the following example. The Chapter 2 callout to **Figure A-1** refers to Figure 1 in **Appendix A.** The blue highlight links to the appendix figure. To return to the Chapter 2 callout page, see the PDF navigation note next.

NOTE ON PDF NAVIGATION: Document navigation is greatly facilitated by using Adobe Acrobat's "Previous view" and "Next view" functions. For "Previous view," use quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For "Next view," use Alt/Ctrl+Right Arrow on PC or Command+Right Arrow on Mac. Using these quick keys in tandem allows the reader to toggle between document locations. Since every scroll represents a new view; depending on how much scrolling is done for a specific view destination, more than one press of the back or forward arrows may be needed. For additional navigational tips, click View at the top of the PDF, then Page Navigation. These Adobe Acrobat functions may not be functional for other PDF readers or for PDFs opened in web browsers.

ABSTRACT

This research aims to fill an essential gap in understanding how Body Mass Index (BMI) cutoffs relate to multimorbidity across races in the United States (US). Given the significant and growing rates of obesity and multimorbidity, as well as the known differences in healthy fat distribution among different races, this is an important area of research. BMI is a widely used but imperfect measure of obesity, as it does not account for differences in body composition. However, it is still used as a diagnostic tool. It is vital to ensure that the cutoffs used to define obesity are appropriate for all populations, particularly given the racial disparities in multimorbidity rates. This proposed framework for evaluating BMI cutoffs across races for multimorbidity considered a range of measures, such as, including incidence rates of prevalent diseases, age, gender, type of patient visits, and type of health insurance to arrive at questioning the current World Health Organization (WHO) BMI cutoffs in the US.

This research demonstrated that having the exact BMI cutoffs across all races does not serve all populations ideally through three assessments. First, it assessed differences in the prevalence of multimorbidity by race. It identified disease combinations shared by all races/ethnicities, shared by some, and those unique to one group for each age/obesity level. These findings demonstrated that despite controlling for age and obesity, there are differences in multimorbidity prevalence across races. Second, the study developed models to project total charges for the most common multimorbidity combinations in the US and evaluated the accuracy of these models across different racial and ethnic groups and multimorbidity patterns. The relationship between healthcare costs and multimorbidity varied for each racial group and depended on the specific combination of chronic conditions, age, and obesity status. Third, it assessed the relationship between BMI and healthcare burden across race and healthcare utilization among middle-aged patients in the US. It demonstrated that the relationship between BMI and healthcare burden varied across races within the same healthcare care utilization category.

This research can improve health outcomes and reduce the risk of chronic diseases associated with obesity and multimorbidity, particularly among vulnerable populations. It will also be essential to consider the potential implications of any new BMI cutoffs on clinical practice and health policies related to obesity and multimorbidity in serving unique clinical needs. More work must be done to understand how multimorbidity, BMI, age, and healthcare burden associate across races.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
BMI	Body Mass Index
CCI	Charlson Comorbidity Index
ECI	Elixhauser Comorbidity Index
GVIF	Generalized Variance Inflation Factors
ICD-9-CM	International Classification of Diseases, Ninth Revision, Clinical Modification
ICD-10-CM	International Classification of Diseases, Tenth Revision, Clinical Modification
ROC	Receiver Operating Characteristic
US	United States
WHO	World Health Organization

CHAPTER 1. INTRODUCTION

NOTE: When using Adobe Acrobat, return to the last viewed page using quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command + Right Arrow. See **Preface** for further details.

Background

The standard measure for classifying obesity is Body Mass Index (BMI), defined as a person's weight in kilograms divided by the square of their height in meters[1]. The World Health Organization (WHO) defined obesity as "excessive fat accumulation that presents a risk to health" with a BMI of 30 or above[2]. Increased body fat puts patients at risk for many chronic diseases, such as heart disease, hypertension, lipidemia, and diabetes[3,3,4]. BMI is viewed as an inaccurate indicator of body fat and obesity since it does not account for differences in body composition, such as fat, muscle, bone, and other lean body mass[5,6]. This inaccuracy leads to a delay in the diagnosis of time-sensitive health issues. Despite the widespread use of BMI as a diagnostic tool, there is significant evidence that healthy fat distribution differs among races and that the currently-defined ranges are inappropriate for most non-white, non-male patients[7–10].

Multimorbidity, the co-occurrence of two or more chronic conditions[7], is a growing global healthcare challenge. A significant contributor to multimorbidity is obesity, which affects about 42% of US adults[11]. Multimorbidity impacts around 27% of adults and 58%-77% of the elderly in the US[8]. Adding to the increasing rates of obesity and multimorbidity, the US has an increasingly aging population with obesity[1]. Furthermore, aging is an entrance to multimorbidity and weight gain[12].

Variations in multimorbidity exist by race and ethnicity. Recent research has revealed that multimorbidity rates are greater among African Americans than Caucasians and higher among Caucasians than Asians[13]. Race is also related to quicker rates of multimorbidity progress. In a population of middle-aged adults, one study found that African Americans developed multimorbidity around four years earlier than Caucasians[14].

About 200 years ago, Adolphe Quetelet, a Belgian astronomer, mathematician, statistician, and sociologist, wanted to identify the characteristics of an average man. He came up with Quetelet's index to compare short and tall men[7]. His population was based on French and Scottish participants, and he used height and weight to calculate his index, now known as BMI. This measure was not intended for measuring individual health. It was designed to measure a Western European population statistically to identify the "ideal" man[7,8]. In the 20th century, life insurance companies relied on height and weight tables to determine what to charge policyholders. Even though these tables were unreliable, physicians in the 1950s and 1960s started to rely on them to evaluate patients' health1,3. In 1972, Ancel Keys named Quetelet's index Body Mass Index (BMI). Keys

and his colleagues wanted to find a diagnostic tool to measure body fat to diagnose patients with obesity. BMI was the best tool then, although it did not differentiate between body fat, muscle mass, or bone density. In his study, participants were predominantly from white nations; The United States, Finland, Italy, Japan, and South Africa. In 1985, the National Institute of Health linked BMI to their definition of obesity[15].

Eventually, BMI became a standard health measure for all racial groups and genders. However, for over 200 years, it was mainly based on white/male participants at a time when populations had a negligible prevalence of obesity. Now, BMI, a scale based on a questionable representative sample of a population, determines the patient's health risks. As BMI increases, so do the health risks. Generalizing BMI to all racial groups will not accurately measure their health risks.

In 2004, WHO published a consultation regarding redefining BMI cutoff points for Asians. They acknowledged that many Asian people are at high risk for developing type 2 diabetes and cardiovascular disease at lower BMI cutoffs for overweight and obesity. They also noted that BMI cutoffs should be interpreted with morbidity and mortality, not just height and weight. WHO also proposed that countries could decide their BMI cutoffs based on population risk factors[15]. A similar scenario has occurred in China, Malaysia, and India, where excessive fat accumulation and increased risk factors started to show at lower BMI cutoffs, indicating a further need to re-examine cutoff values for obesity[16–18].

Assessing the BMI cutoffs while associating them with multimorbidity and healthcare burden across races based on the US population is essential to improve health outcomes. My research aims to build a framework that acquires an evaluation of BMI cutoffs across races while focusing on multimorbidity. Multimorbidity positively correlates with age and obesity and varies by race[19–21]. Various studies have been done on BMI cutoffs across races for specific populations[16,22,23]. Adiposity measures, all-cause mortality, waist circumference, belly-hip ratio, and incidence rates of certain prevalent diseases are some measures used to reach the new BMI cutoffs[24]. However, a multimorbidity base understanding of BMI across races in the United States (US) is missing.

In a recently published paper, researchers redefined disease-specific BMI cutoffs by sex and race based on association with metabolic disease. They linked obesity to the future prediction of medical morbidity and mortality based on a study done in 2004 and considering hypertension, dyslipidemia, and diabetes as risk factors[25]. Each of these three diseases had a different obesity BMI cutoff. That study had data for weight circumference, which we do not have. This limitation, different types of data collected in the dataset, led me to seek other measures to address a multimorbidity base understanding of BMI across races.

To my knowledge, I will be the first to study the following from a generalizable US population:

1. Identify multimorbidity patterns by race/ethnicity as stratified by age and obesity.
2. Identify differences in multimorbidity-related healthcare costs by race as stratified by age and obesity.
3. Assess the association of BMI with multimorbidities and healthcare burden across races.

Purpose

Understanding different multimorbidity sets across age and race groups will help manage joint and complex healthcare plans, set guidelines on how these sets interact, and understand how management features can work together. Integrated care could be developed for specific groups with multimorbidity after profiling these patients properly according to their education, income, social support, health literacy, and self-management capabilities[26]. Patient profiling via risk segmentation will improve the quality of care and help identify patients who can be targeted with preventative care to reduce morbidity, mortality, and healthcare costs. Traditional healthcare management focuses on treating each morbidity separately and is not suited to managing multiple chronic conditions. While studying interventions in the context of multimorbidity and disease sets, researchers found that managing one disease often had a positive effect on managing the other[27]. Multimorbidity negatively impacts work productivity, mental health, healthcare costs, quality of life, and coordinated healthcare management[9].

For over 20 years, research has indicated that standard BMI cutoff values should be revised because of the difference in fat distribution among races.[28] Re-examining and questioning BMI cutoffs was applied in certain countries[29–31]. Identifying a race-based BMI scale will determine health risk factors in the appropriate stage. It will also lead to better outcomes resulting in preventative care and early intervention. Patient profiling via patient segmentation will improve the quality of care, help identify patients likely to incur high costs, and can be targeted with preventative care early on. This knowledge extends to policymakers when allocating resources considering that race is one of the social determinants of health. Obesity needs to have a new conceptualization other than only relying on BMI. An optimal BMI measure must be redefined according to age, race, and multimorbidity[32]. To avoid delays in diagnosing time-sensitive health issues, we must understand BMI's influence by age and race and its correlation to multimorbidity.

This collection of new data and a new understanding of BMI will be significant for policymakers and patient healthcare outcomes. It can lay a baseline for physicians to target patients with the most influencing multimorbidities and improve physician-patient communication for those patients who do not understand the prolonged effect of their lifestyle.

Aims

Aim 1: Identify multimorbidity patterns across races adjusted for age and obesity

Multimorbidity is associated with age, race/ethnicity, gender, and weight, which magnifies managing it on both sides: physician and patient. I will analyze data from Cerner HealthFacts® Database, a national dataset, for middle-aged (45-64) and elderly (65+) adults. This database provides electronic health records for about 70 million patients from hospitals around the United States from 2001 to 2017. I will present subgroups of multimorbidity patterns across races that are homogenous in obesity class and age segment under a certain threshold. This research will enable the development of applicable preventive interventions to address the issue of multimorbidity across race/ethnicity.

Aim 2: Identify differences in multimorbidity total charges across races adjusted for age and obesity using a comorbidity index

Patients with more multimorbidities have a higher severity index, yielding a higher healthcare cost[33]. However, these costs are elevated by multimorbidity burden rather than the number of morbidities[20,34]. In this aim, I will determine whether the relationship between multimorbidity disease severity and total healthcare charges is consistent across races. The results of this investigation will highlight the total healthcare charges across all major racial-ethnic groups in the dataset. These results can set a baseline for proper intervention methods for each racial/ethnic group.

Aim 3: Assess the association of BMI with multimorbidities and healthcare burden across races

This study examines the relationship between BMI and healthcare burden among middle-aged patients in the US. Additionally, the study intends to explore how this relationship varies based on race and healthcare utilization. Various factors, including race and access to healthcare, may influence the relationship between BMI and health outcomes. The number of emergency room visits and insurance type will be considered potential confounders. By stratifying the study population based on race and healthcare utilization, we aim to understand better how these factors may impact the relationship between BMI and healthcare burden. This information could help inform public health initiatives and healthcare policies.

CHAPTER 2. AIM 1 STUDY[1]

NOTE: This chapter refers frequently to content in **Appendix A.** When using Adobe Acrobat, after going there, return to the last viewed page using quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command + Right Arrow. See **Preface** for further details.

Introduction

This chapter summarizes the research to test Aim 1: "Identify multimorbidity patterns across races adjusted for age and obesity,**"** as shown in the article "Multimorbidity Patterns Across Race/Ethnicity Stratified by Age and Obesity: A Cross-sectional Study of a National US Sample" (**Appendix A)**. My summary of the article's content, presented next, includes a discussion of the findings as they relate to my ETD Aim 1.

Summary

The prevalence of multimorbidities increases as people age and is associated with increased mortality, hospitalization, polypharmacy, and adverse health outcomes. Obesity exacerbates multimorbidity prevalence and is associated with increased healthcare utilization and poor health outcomes. Identifying homogeneous subgroups with specific multimorbidity patterns can help develop targeted interventions, but previous studies have been limited and not stratified by factors that impact the multimorbidity pattern. Disparities in multimorbidity exist by race/ethnicity and sociodemographic factors such as age and obesity. Patients need proactive, precise, patient-centered care plans that explicitly address their most critical needs with the most prevalent multimorbidity combinations.

This study aimed to identify prevalent multimorbidity patterns within each racial/ethnic category stratified by age and obesity status. Multimorbidity is the presence of two or more diseases within an individual. Patients were classified with obesity based on an average BMI of 30+ and without obesity if their BMI was less than this cutoff. Racial categories were based on those present within the Cerner HealthFacts data warehouse. The study used frequent itemset detection to find combinations of diseases above the threshold of 5% prevalence.

---

[1] Final submission reproduced with open access permission. Alshakhs, M., Jackson, B., Ikponmwosa, D. *et al*. Multimorbidity patterns across race/ethnicity as stratified by age and obesity. *Sci Rep* **12**, 9716 (2022). https://doi.org/10.1038/s41598-022-13733-w [28] (**Appendix A**).

This study was a cross-sectional analysis of patient encounter records from the Cerner HealthFacts data warehouse, which includes over 70 million patients treated at hospitals and clinics throughout the United States between 2001 and 2017. The study focuses on patients aged 45 and above with a BMI value between 18.5 and 206, assigned race/ethnicity, and at least one ICD-10-CM diagnosis code indicating a medical condition.

The study examined two cohorts, middle-aged and elderly, stratified by race/ethnicity, with most patients being Caucasian, followed by African American. Obesity was prevalent among patients of all races/ethnicities except for Asians/Pacific Islanders. African American patients had the highest number of total and distinct multimorbidities, with the lowest number found in middle-aged cohorts without obesity. The study provided a table with the number of total multimorbidities for each race/ethnicity cohort by age and weight class (**Table A-2**).

The study identified several multimorbidity patterns shared across all race/ethnic groups analyzed within each age/obesity cohort (**Figure A-1**). The patterns included one or more of the following ICD-10-CM codes: I10: Hypertension, E78: Lipidemia, or E11: Diabetes. The prevalence of the patterns varied by race/ethnicity and age/obesity cohort. African American patients had the highest number of total multimorbidities and the most distinct multimorbidities for each age/weight group. The elderly cohort patterns included three new clinical categories, D: Disease of blood and blood-forming organs, N: Diseases of the genitourinary system, and G: Diseases of the nervous system.

Some races shared some multimorbidity patterns (**Figure A-2**). The study identified ten multimorbidity patterns for patients without obesity and 23 for patients with obesity in the middle-aged cohort. The most prevalent diseases include diabetes, hypertension, lipidemia, dorsalgia, and other joint disorders. In the elderly cohort, there were 47 multimorbidity patterns for patients without obesity and 90 for patients with obesity. The most prevalent diseases were heart disease, hypertension, chronic obstructive pulmonary disease, and chronic kidney disease. The study found that certain combinations of diseases were more prevalent among certain racial and ethnic groups, such as African Americans and Native Americans having higher prevalence of nicotine dependence and hypertension.

The study identified distinct multimorbidities that are unique to specific racial/ethnic groups, both in middle-aged and elderly cohorts (**Figures A-3**, **A-4**, and **A-5**). The African American cohort had the most distinct multimorbidity patterns, and the Native American cohort had distinct patterns including diabetes. The study also found that the median BMI values differed significantly among the racial/ethnic groups within each age and weight class. For example, all patient cohorts without obesity exhibited median BMI values of 25-26, which fall into the overweight weight class.

Even after stratification, we found differences in multimorbidity prevalence across races/ethnicities, and some combinations were distinct to certain racial/ethnic groups. Common morbidities present in disease combinations across all races/ethnicities

were lipidemia, hypertension, and diabetes, regardless of age or obesity level. The study showed that multimorbidity increased with age and was highest among African Americans and lowest among Asian/Pacific Islanders. The disease composition of multimorbidity also varied by race/ethnicity, and African Americans presented with the most distinct multimorbidities at an earlier age than other races/ethnicities. The study's findings can be used to inform public policy and to develop patient guidelines for various obesity levels and ages.

Conclusion

Our analysis revealed disparities in the prevalence of multimorbidity across races, even after controlling for factors such as age and obesity. African Americans were found to experience distinct multimorbidities at an earlier age compared to other racial groups. This information on prevalent multimorbidity patterns across races provides crucial insight into their specific clinical needs.

CHAPTER 3. AIM 2 STUDY[2]

NOTE: This chapter refers frequently to content in **Appendix B.** When using Adobe Acrobat, after going there, return to the last viewed page using quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command + Right Arrow. See **Preface** for further details.

Introduction

This chapter summarizes the research to test Aim 2: "Identify differences in multimorbidity total charges across races adjusted for age and obesity using a comorbidity index," as shown in the article "Racial Differences in Healthcare Expenditures for Prevalent Multimorbidity Combinations in the U.S." (**Appendix B)**. My summary of the article's content, presented next, includes a discussion of the findings as they relate to my ETD Aim 2.

Summary

This study addressed the need for economic models to better understand the burden of multimorbidity, which is the presence of two or more diseases in an individual. The current models do not consider the varying costs of different disease combinations, making it difficult to target the highest-cost patients for intensive interventions.
It is essential to accurately project patient costs for the most prevalent multimorbidity combinations by race/ethnicity, particularly for the elderly population with two or more chronic conditions, and account for a significant portion of Medicare spending. This study aimed to identify the expected total charges associated with the most prevalent multimorbidity combinations by race/ethnicity, assess differences in expenditures for these combinations, and assess differences in model accuracy by race/ethnicity.
We employed a cross-sectional design using de-identified data from the Cerner HealthFacts® data warehouse, which includes over 490 million patient encounters for over 70 million patients treated at hospitals and clinics at 792 non-affiliated healthcare systems throughout the United States between 2001 and 2017. The inclusion criteria for patients were age 45+, BMI value present and between 18.5 and 206, assigned race category, assigned gender, encounters with total charges greater than $0, and an encounter with an ICD-10-CM diagnosis code that is included in one or more of the 38 broad diagnoses that make up the most prevalent multimorbidities in the U.S.

This study analyzed the financial burden across races in middle-aged and older adults, with demographic, multimorbidity, and healthcare utilization variables as the primary independent variables. The study used total charges as the primary outcome

[2] Article reused from prepared manuscript with authors' permission. Alshakhs, M., Goedecke, P., Bailey, J., and Madlock-Brown, C. "Racial Differences in Healthcare Expenditures for Prevalent Multimorbidity Combinations in the U.S." (2023) (**Appendix B**).

variable, categorizing patient encounters into inpatient, outpatient, and emergency visit categories.

The study excluded hospitals with missing census divisions or rural/urban status information and imputed missing teaching facility information. The study excluded encounters with $0 listed for total charges. We removed patient records if treated in two different census divisions or rural and urban hospitals for easy interpretation.
We used statistical analysis to examine the distribution of the outcome variable, total charges, and tested for multicollinearity among the independent variables. Regression analysis was then used to compare total charges of different multimorbidity combinations by race/ethnicity. A generalized linear model with gamma distribution and log link function was applied. A 3-way analysis of variance (ANOVA) test was conducted to determine interaction effects between BMI and race and Elixhauser Comorbidity Index (ECI) ranks on total charges.

The final population included middle-aged and elderly patients, with most being female. The breakdown of demographics by race is shown in **Tables B-1** and **B-2**. Outcomes, including visit type, emergency room visits, ECI score, admission days, and charges, are shown in **Table B-3**. After eliminating outliers and testing for collinearity, the total number of morbidities variable was removed. A generalized linear model with a Gamma distribution and log link function was used to assess model performance. The exponential model was selected as optimal, exhibiting the least sum of square error in both cohorts.

We found that shared multimorbidities were significant predictors for healthcare charges across races and that mean total charges were higher in the elderly cohorts than in the middle-aged cohorts. The study also showed that the model overestimated total charges for specific multimorbidity patterns and underestimated them for others, depending on race and obesity status (**Figure B-1**). The variability of mean model residuals varied by race and obesity status as well (**Figure B-2**).

This study aimed to identify the total charge trends for the most prevalent multimorbidity combinations and to assess the accuracy of the model predictions across races. It revealed that the relationship between cost and multimorbidity was inconsistent for each racial group. Specific multimorbidity patterns were more inaccurate for some groups, and some racial groups could drive the overall inaccuracy of cost estimates for specific multimorbidity combinations.

Conclusion

This study found that model accuracy varied across obesity status and race. A well-fit cost model helped understand the relationship between cost and multimorbidity, which has been missing from the literature. The relationship between cost and multimorbidity was inconsistent for each racial group, highlighting the need for accurate models for all racial groups to address inequalities.

CHAPTER 4. AIM 3 STUDY[3]

NOTE: This chapter refers frequently to content in **Appendix C.** When using Adobe Acrobat, after going there, return to the last viewed page using quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command + Right Arrow. See **Preface** for further details.

Introduction

This chapter summarizes the research to test Aim 3: "Assess the association of BMI with multimorbidities and healthcare burden across races," as shown in the article "Assessing the Relationship Between Healthcare Burden and BMI as Stratified by Race in the U.S." (**Appendix C**). My summary of the article's content, presented next, includes a discussion of the findings as they relate to my ETD Aim 3.

Summary

There are limitations to using BMI to assess weight status, and it is essential to consider individual differences in body shape and composition. The World Health Organization has recommended rescaling BMI for Asian individuals, and other countries have done the same based on various measures. Assessing the relationship between BMI and healthcare burden across races in the U.S. is crucial in determining overall health status and informing a tailored treatment plan. Higher BMIs are often associated with a greater risk of obesity-related health conditions that strain the healthcare system. This study aimed to assess the relationship between BMI and multimorbidity as stratified by race to address health disparities and design target interventions to improve health outcomes.

This study used a cross-sectional design and data from the Cerner HealthFacts® data warehouse for 2016-2017, comprising encounter data from over 70 million patients across the U.S. who received treatment at hospitals and clinics from 792 non-affiliated healthcare systems. Patients aged 45-64 with BMI values between 18.5 and 75 and assigned race and gender were included in the study, while those with cancer diagnoses, pregnancy, and ICD-9-CM diagnostic codes were excluded. The patients were stratified by race and divided into three subgroups based on healthcare utilization patterns, defined as a minimum of two outpatient visits within the two-year study period. We also examined the impact of insurance and visit types on the burden of multimorbidity across races.

[3] Article reused from prepared manuscript with authors' permission. Alshakhs, M., Goedecke, P., Chinthala, L., Weiskopf, N., and Madlock-Brown, C. "Assess the association of BMI with multimorbidities and healthcare burden across races" (2023) (**Appendix C**).

We used logistic regression to model the relationship between independent variables (demographics, payer type, smoking, alcohol use, urbanism, and healthcare utilization) and the outcome variable (healthcare burden measure). The primary independent variables included race, age, gender, marital status, BMI, and 2014-2015 healthcare utilization. The study used the CCI score to measure the degree of illness in patients, calculated from the ICD-10-CM diagnosis codes. We applied the Kolmogorov-Smirnov test to compare the distribution of the outcome variable among different cohorts. We used the generalized variance inflation factor analysis to remove variables with high GVIF scores and analyzed the ROC curve to determine the best BMI cutoff for each group.

The final patient population was 1,271,697 patients used to investigate the impact of BMI on healthcare burden across different races and healthcare utilization cohorts. The logistic regression model showed that BMI was the most significant predictor of healthcare burden across all races and utilization cohorts (**Figure C-2**). The top three predictors varied depending on healthcare utilization and race. The model best predicted the healthcare burden for the Caucasian racial group, followed by the Asian/Pacific Islander group, and the African American group was predicted least well by the model. The average BMI for the most significant AUC in the healthcare utilizer cohorts was 34 for the African American cohort, 27 for the Asian/Pacific Islander cohort, 32 for the Caucasian cohort, and 35 for the Native American cohort (**Figure C-3**).

This study found that the relationship between BMI and healthcare burden varied across race and healthcare utilization, with BMI and age being the most important variables impacting the prediction of healthcare burden. The findings highlight the complexity of the relationship between BMI, healthcare burden, and healthcare utilization across races. The research can aid in treatment plans and resource allocation, identifying factors associated with increased or decreased utilization, evaluating the effectiveness of interventions targeted at specific utilization subgroups, and identifying potential areas for cost savings. It can also help researchers and healthcare providers to understand and address health disparities.

Conclusion

This research demonstrated that the relationship between BMI and CCI varied across races within same healthcare care utilization cohorts. Some of this variation could be driven by access to healthcare resources. Understanding how multimorbidity accumulates over time across populations is essential. More work must be done to understand how multimorbidity, BMI, and healthcare burden associate across races.

CHAPTER 5. DISCUSSION AND FUTURE WORK

NOTE: When using Adobe Acrobat, return to the last viewed page using quick keys Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command + Right Arrow. See **Preface** for further details.

Discussion

This research identified multimorbidity patterns across race/ethnicity-specific while considering age and obesity status. It sheds light on the prevalence of multimorbidity patterns in these populations, revealing that African Americans had a higher prevalence of multimorbidity, irrespective of age or obesity status and that this condition tends to develop earlier in this group. Furthermore, it showed that certain combinations of multimorbidity patterns are more common in specific racial/ethnic groups, mainly middle-aged patients with or without obesity. This knowledge is critical in understanding the underlying mechanisms contributing to disease co-occurrence in patients of different races/ethnicities and ages. These findings can guide the development of patient-centered care models.

This research also evaluated the accuracy of model predictions for total charges of the most prevalent multimorbidity combinations across racial groups. Additionally, it was the first to identify trends in total charges for these multimorbidity combinations. Our findings revealed that African Americans had the highest mean total charges, while members of the Hispanic race had the lowest charges in both cohorts. However, our model exhibited inconsistency in accurately predicting total charges by race based on the various multimorbidity patterns. We observed that the total charges were often over- or underestimated across different multimorbidity patterns, and the estimates were extreme in some cases. This highlighted the challenges of accurately modeling total charge estimates for diseases that may interact in a multimorbidity, as their effects are not simply additive. Therefore, it is imperative to develop more robust models to ensure that the healthcare system can better serve all populations. Improved modeling of underserved populations, considering both race and multimorbidity, is necessary to address this critical issue.

In addition, this research assessed the relationship between BMI and healthcare burden across race and healthcare utilization. This research demonstrated that the relationship between BMI and healthcare burden varied across races within the same healthcare care utilization cohorts. This variation showed that multimorbidity impacted patients across races differently, indicating that having the exact BMI cutoffs across races is questionable in the US. Some of this variation could be driven by access to healthcare resources. Most of the research regarding multimorbidity focused on a specific point in time and on a few multimorbidities simultaneously. Understanding how multimorbidity accumulates over time across populations is not addressed. More work must be done to understand how multimorbidity, BMI, age, and healthcare burden associate across races.

Future Work

This research can be the baseline for exploring the potential impact of interventions targeting specific multimorbidity patterns in different racial/ethnic groups. Researchers can further investigate the potential role of cultural factors and healthcare access in shaping multimorbidity patterns among different racial/ethnic groups. Understanding how cultural beliefs and healthcare utilization patterns impact multimorbidity could lead to more targeted and effective interventions.

More research needs to be done to develop accurate and robust models for predicting total charges among patients with multimorbidity, particularly for underserved populations. This could involve using machine learning techniques, such as deep learning, to capture the complex interactions between multiple diseases and patient demographics. Conduct longitudinal studies to monitor changes in total charges over time for different racial groups with multimorbidity. This could help identify trends and potential interventions to reduce healthcare disparities.

Longitudinal studies could also be conducted to monitor changes in the relationship between BMI and healthcare severity over time across different racial groups. This could provide insights into the dynamic nature of this relationship and how it evolves. Investigating the impact of access to healthcare resources on the relationship between BMI and healthcare burden across different racial groups could be promising. This could include examining how differences in healthcare utilization and access impact the development and progression of multimorbidity. This proposed research could lead to more effective and targeted interventions for addressing healthcare disparities among racial/ethnic groups.

LIST OF REFERENCES

1. Products - Data Briefs - Number 360 - February 2020. Published June 26, 2020. Accessed November 2, 2020. https://www.cdc.gov/nchs/products/databriefs/db360.htm

2. Obesity. Accessed November 20, 2022. https://www.who.int/health-topics/obesity

3. Blair D, Habicht Jp, Sims Eah, Sylwester D, Abraham S. Evidence For An Increased Risk For Hypertension With Centrally Located Body Fat And The Effect Of Race And Sex On This Risk. *American Journal of Epidemiology*. 1984;119(4):526-540. https://doi.org/10.1093/oxfordjournals.aje.a113770

4. Changes in Body Weight and Body Fat Distribution as Risk Factors for Clinical Diabetes in US Men | American Journal of Epidemiology | Oxford Academic. Accessed November 20, 2022. https://academic.oup.com/aje/article/159/12/1150/86091

5. Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How Useful Is Body Mass Index for Comparison of Body Fatness across Age, Sex, and Ethnic Groups? *American Journal of Epidemiology*. 1996;143(3):228-239. https://doi.org/10.1093/oxfordjournals.aje.a008733

6. Yusuf S, Hawken S, Ôunpuu S, et al. Obesity and the risk of myocardial infarction in 27 000 participants from 52 countries: a case-control study. *The Lancet*. 2005;366(9497):1640-1649. https://doi.org/10.1016/S0140-6736(05)67663-5

7. Eknoyan G. Adolphe Quetelet (1796 1874) the average man and indices of obesity. *Nephrology Dialysis Transplantation*. 2007;23(1):47-51. https://doi.org/10.1093/ndt/gfm517

8. The Editors of Encyclopaedia Britannica. Adolphe Quetelet. In: Encyclopædia Britannica; 2020. https://www.britannica.com/biography/Adolphe-Quetelet

9. Your Fat Friend. The Bizarre and Racist History of the BMI. Published October 15, 2019. https://elemental.medium.com/the-bizarre-and-racist-history-of-the-bmi-7d8dc2aa33bb

10. Ward BW, Schiller JS. Prevalence of multiple chronic conditions among US adults: estimates from the National Health Interview Survey, 2010. *Prev Chronic Dis*. 2013;10:E65. https://doi.org/10.5888/pcd10.120203

11. The Lancet. Making more of multimorbidity: an emerging priority. *The Lancet*. 2018;391(10131):1637. https://doi.org/10.1016/S0140-6736(18)30941-3

12. Alexis McKee, MD and John E Morley, MB, BCh. *Obesity in the Elderly*.; 2018. https://www.ncbi.nlm.nih.gov/books/NBK532533/

13. Quiñones AR, Markwardt S, Botoseneanu A. Multimorbidity combinations and disability in older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*. 2016;71(6):823-830.

14. Quiñones AR, Botoseneanu A, Markwardt S, et al. Racial/ethnic differences in multimorbidity development and chronic disease accumulation for middle-aged adults. Ginsberg SD, ed. *PLoS ONE*. 2019;14(6):e0218462. https://doi.org/10/gmxd3d

15. Blackburn H, Jacobs D. Commentary: Origins and evolution of body mass index (BMI): continuing saga. *International Journal of Epidemiology*. 2014;43(3):665-669. https://doi.org/10/gmxd3t

16. He W, Li Q, Yang M, et al. Lower BMI cutoffs to define overweight and obesity in China: Chinese BMI Cutoffs. *Obesity*. 2015;23(3):684-691. https://doi.org/10/f3pg2s

17. Deurenberg P, Yap M, van Staveren W. Body mass index and percent body fat: a meta analysis among different ethnic groups. *Int J Obes*. 1998;22(12):1164-1171. https://doi.org/10/fsdv2m

18. Deurenberg-Yap M, Chew S, Lin V, Tan B, van Staveren W, Deurenberg P. Relationships between indices of obesity and its co-morbidities in multi-ethnic Singapore. *Int J Obes*. 2001;25(10):1554-1562. https://doi.org/10/fs5fg6

19. King DE, Xiang J, Pilkerton CS. Multimorbidity Trends in United States Adults, 1988-2014. *J Am Board Fam Med*. 2018;31(4):503-513. https://doi.org/10.3122/jabfm.2018.04.180008

20. Agborsangaya CB, Ngwakongnwi E, Lahtinen M, Cooke T, Johnson JA. Multimorbidity prevalence in the general population: the role of obesity in chronic disease clustering. *BMC Public Health*. 2013;13(1):1161. https://doi.org/10/f5kdtw

21. Afshar S, Roderick PJ, Kowal P, Dimitrov BD, Hill AG. Multimorbidity and the inequalities of global ageing: a cross-sectional study of 28 countries using the World Health Surveys. *BMC Public Health*. 2015;15. https://doi.org/10.1186/s12889-015-2008-7

22. Heymsfield SB, Peterson CM, Thomas DM, Heo M, Schuna JM. Why are there race/ethnic differences in adult body mass index-adiposity relationships? A quantitative critical review: BMI and race/ethnicity. *Obesity Reviews*. 2016;17(3):262-275. https://doi.org/10.1111/obr.12358

23. Albrecht SS, Mayer-Davis E, Popkin BM. Secular and race/ethnic trends in glycemic outcomes by BMI in US adults: The role of waist circumference. *Diabetes Metab Res Rev*. 2017;33(5):e2889. https://doi.org/10/f9p9fh

24. Burkhauser RV, Cawley J. Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*. 2008;27(2):519-529. https://doi.org/10.1016/j.jhealeco.2007.05.005

25. Stanford FC, Lee M, Hur C. Race, Ethnicity, Sex, and Obesity: Is It Time to Personalize the Scale? *Mayo Clinic Proceedings*. 2019;94(2):362-363. https://doi.org/10.1016/j.mayocp.2018.10.014

26. Rijken M, van der Heide I. Identifying subgroups of persons with multimorbidity based on their needs for care and support. *BMC Fam Pract*. 2019;20(1):179. https://doi.org/10/gmxd3s

27. Flegel K. What we need to learn about multimorbidity. *CMAJ*. 2018;190(34):E1001-E1001. https://doi.org/10/gmxd3k

28. Alshakhs M, Jackson B, Ikponmwosa D, Reynolds R, Madlock-Brown C. Multimorbidity patterns across race/ethnicity as stratified by age and obesity. *Sci Rep*. 2022;12(1):9716. https://doi.org/10.1038/s41598-022-13733-w

29. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *The Lancet*. 2004;363(9403):157-163. https://doi.org/10/fm3fgw

30. Lauderdale D, Rathouz P. Body mass index in a US national sample of Asian Americans: effects of nativity, years since immigration and socioeconomic status. *Int J Obes*. 2000;24(9):1188-1194. https://doi.org/10.1038/sj.ijo.0801365

31. Kuczmarski RJ, Flegal KM. Criteria for definition of overweight in transition: background and recommendations for the United States. *The American Journal of Clinical Nutrition*. 2000;72(5):1074-1081. https://doi.org/10.1093/ajcn/72.5.1074

32. Müller MJ, Braun W, Enderle J, Bosy-Westphal A. Beyond BMI: Conceptual Issues Related to Overweight and Obese Patients. *Obes Facts*. 2016;9(3):193-205. https://doi.org/10/f8tdhv

33. Sambamoorthi U, Tan X, Deb A. Multiple chronic conditions and healthcare costs among adults. *Expert Rev Pharmacoecon Outcomes Res*. 2015;15(5):823-832. https://doi.org/10.1586/14737167.2015.1091730

34. Booth HP, Prevost AT, Gulliford MC. Impact of body mass index on prevalence of multimorbidity in primary care: cohort study. *Family Practice*. 2014;31(1):38-43. https://doi.org/10.1093/fampra/cmt061

APPENDIX A. CHAPTER 2 ARTICLE

NOTE: Navigation with Adobe Acrobat Reader or Adobe Acrobat Professional: To return to the last viewed page, use key commands Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For "Next view," use Alt/Ctrl+Right Arrow on PC or Command+Right Arrow on Mac. See **Preface** for further details. If needed, use this link to return to **Chapter 2** after navigating within this appendix.

Introduction

Final submission reproduced with open access permission. Alshakhs, M., Jackson, B., Ikponmwosa, D. *et al.* Multimorbidity patterns across race/ethnicity as stratified by age and obesity. *Sci Rep* **12**, 9716 (2022).
https://doi.org/10.1038/s41598-022-13733-w

Article

Multimorbidity Patterns Across Race/Ethnicity Stratified by Age and Obesity: A Cross-sectional Study of a National US Sample

Manal Alshakhs, MS[1]
Bianca Jackson, MPH[1]
Davina Ikponmwosa, MHIIM[1]
Rebecca Reynolds, EdD, MHA[1,2,3]
*Charisse Madlock-Brown, PhD, MLS[1,2,3]

[1]Health Outcomes and Policy Program, University of Tennessee Health Science Center, Memphis, Tennessee, USA
[2]Health Informatics and Information Management Program, University of Tennessee Health Science Center, Memphis, Tennessee, USA
[3]Center for Health System Improvement, University of Tennessee Health Science Center, Memphis, Tennessee, USA

*Corresponding Author:
Charisse Madlock-Brown, Health Informatics and Information Management Program, University of Tennessee Health Science Center, 66 North Pauline St. rm 221, Memphis, TN 38163, USA. Email: cmadlock@uthsc.edu

Abstract
The objective of our study is to assess differences in prevalence of multimorbidity by race. We applied the FP-growth algorithm on middle-aged and elderly cohorts stratified by race, age, and obesity level. We used 2016-2017 data from the Cerner HealthFacts® Electronic Health Record data warehouse. We identified disease combinations that are shared by all races/ethnicities, those shared by some, and those that are unique to one group for each age/obesity level. Our findings demonstrate that even after controlling for age and obesity, there are differences in multimorbidity prevalence across races. There are multimorbidity combinations distinct to some racial groups—many of which are understudied. Some multimorbidities are shared by some but not all races. African Americans presented with the most distinct multimorbidities at an earlier age. The identification of prevalent multimorbidity combinations amongst subpopulations provides information specific to their unique clinical needs.

Introduction
Multimorbidities become more prevalent as individuals age, and they have been associated with substantial burden and increased mortality (1). The treatment of multimorbidity can be complicated by the involvement of multiple medical specialties (2). These patients are also at risk of repeated hospitalization, polypharmacy, adverse drug events, and increased care dependence (3). Obesity exacerbates multimorbidity prevalence and further contributes to increased healthcare utilization and adverse health outcomes. Internationally, the medical community has called for a greater understanding of multimorbidity patterns (4). Developing the means to identify and study homogeneous multimorbidity subgroups has been theorized to be a way to develop and target interventions more effectively (5). Studies have also shown that specific racial and ethnic groups are at greater risk of poor health outcomes (6).

To minimize the adverse effects of multimorbidity, patients need proactive, precise, and patient-centered care plans that explicitly address patients' most critical needs with the most prevalent multimorbidity combinations (7). Further, providers should be aware of the most likely multimorbidity combinations to develop and coordinate appropriate care plans when treating patients with these disease groupings (8). It is crucial to identify specific multimorbidity patterns because the impact of multimorbidity on health-related quality of life varies for different combinations(9).

Although previous studies have identified factors that may impact multimorbidity, most relevant studies were not conducted in the US, and few studies adjusted for obesity or age (2,10). Additionally, there remains a need to identify homogeneous combinations among race/ethnic groups. Most previous research identified multimorbidity patterns using counts or cluster analysis and are disease specific (11). These approaches are limited as there exists "tremendous diagnostic heterogeneity, variation in the number of chronic conditions, and the severity of illness characterized the multiple chronic conditions population making identification of multimorbidity trends difficult" (12). One research team using clustering for multimorbidity pattern identification state "though recognition of general patterns of disease co-occurrence is useful for policy planning, the heterogeneity of persons with significant multimorbidity (>=3 conditions) defies neat

classification. A simple count of conditions may be preferable for predicting usage" (13). Chong et al. say the benefits of segmenting patients by multimorbidity patterns will be the "facilitation of healthcare service planning, promotion of the evaluation of health service innovations, and improving care integration" (14). Understanding which diseases cluster together most frequently will lead to the understanding of which disease clusters have the "most significant impact on essential patient outcomes" (15). Though it may be tempting to assume patient groups need care plans tailored to each disease, research has shown that fragmented care can lead to ineffective and potentially harmful interventions (16).

Several research groups have identified the prevalence of specific multimorbidity patterns of two or more diseases. Our previous work identified the most prevalent multimorbidity patterns of two or more diseases for a cohort with obesity (17). Held et al. used association rule analysis mining to find dyads and triads of diseases for a cohort of men of older age in Australia (18). Van den Bussche et al identified combinations of 3 multimorbidities in a sample of German Elderly patients (19). While these studies shed considerable light on how diseases tend to combine, they are limited in that they include only a special population and are not stratified by factors that are likely to impact the pattern of multimorbidity.

Disparities in multimorbidity exist by race and ethnicity. Recent studies have found that African Americans tend to have higher rates of multimorbidity compared with Caucasians, while rates are higher among Caucasians compared with Asians (6). Race is also associated with fasters rates of multimorbidity development. One study demonstrated that African Americans developed multimorbidity about four years earlier than Caucasians in a middle-aged adult population (6). An analysis of multimorbidity networks similarly found that African Americans have the most densely connected network at the organ-level, followed by Caucasians and Native Americans20. There is an urgent need to understand the specific prevalence of multimorbidities between races/ethnicities. These gaps motivate the need for similar research to be conducted within a US patient population, and further emphasize the need to evaluate the effect of race/ethnicity on multimorbidity patterns.

Researchers have also demonstrated that multimorbidity can vary by sociodemographic factors. In a Canadian study among middle-aged adults, the prevalence of multimorbidity increased from 29.7% in individuals that were 45-49 years of age to 52% for the group 60-64 (2). Differences in multimorbidity patterns among men and women have also been observed in a Spanish population of adults over 65 years of age (2,20). Further, a cross-sectional study conducted in Singapore reported an association between multimorbidity and increasing age, lower socio-economic status, female sex, and mental disorders (10).

Although many associations with sociodemographic characteristics and multimorbidity have been found, among the most significant are age and obesity (22,23). Increasing obesity rates contribute to the growing prevalence of multimorbidity worldwide. The prevalence of multimorbidity in persons with obesity exceeds 60% in the US (17,24). Compared with patients of normal weight, patients with obesity have an increased risk of

developing multimorbidity (1,25). Multimorbidity combinations containing obesity, specifically, may result in increased social isolation and vulnerability (26), poorer outcomes, increased hospitalization and healthcare costs, compared with other multimorbidity combinations (2,27). Interestingly, more middle-aged adults than older adults tend to be afflicted with obesity and multimorbidity. They tend to live longer with multimorbidity and associated complications, warranting special attention to obesity associated multimorbidity within this specific population (25).

Our work will address different multimorbidity patterns across race/ethnicity in the United States. Although we are not the first to use a frequent item-set algorithm to assess multimorbidity (18), our study is the first to classify distinct prevalent multimorbidities by race/ethnicity and stratified for obesity and age. Disadvantaged populations are disproportionately affected by multiple chronic conditions (28), which may affect their multimorbidity patterns. Therefore, we aim to identify patterns of morbidity that may be prevalent for only some groups.

The objectives of the research are to:

- Identify prevalent multimorbidity patterns by race/ethnicity category within middle-aged and elderly cohorts
  which will be further subdivided into with and without obesity groups.
- Compare individual disease prevalence of multimorbidities found in all race/ethnicity.
- Identify multimorbidities found in more than one but not all race/ethnicity.
- Identify multimorbidities that are distinct to one race/ethnicity.
- Assess potential multimorbidity disease burden across cohorts.

Methods

Design

This cross-sectional study employed data collected in 2016–2017 and stored the Cerner HealthFacts data warehouse, which includes patient encounter records for over 70 million patients treated at hospitals and clinics throughout the United States between 2001 and 2017. This dataset includes medical histories, diagnoses, laboratory information, prescriptions, patient demographics, clinic type, procedures, and surgical data documenting over 490 million encounters from 792 non-affiliated health care systems. It includes outpatient, inpatient, emergency, and other encounter types. The inclusion criteria for patients were: (1) Age 45+, (2) at least one clinical encounter during 2016–2017 with Body Mass Index (BMI) values between 18.5 and 206 present, (3) assigned race/ethnicity, (4) ≥ 1 International Classification of Diseases, 10th revision, Clinical Modification (ICD-10-CM) diagnosis code that indicates a medical condition, and (5) having no ICD-9-CM classification documented during this timeframe. We chose a two-year period for the study of multimorbidity to increase the likelihood that each patient had each morbidity concurrently. For each patient, diagnoses were aggregated across all visits for this period. We did not require patients to have more than one encounter as research shows healthcare use and access vary by race (29), and being too restrictive could limit the representativeness of our sample.

Ethical Considerations
The data are de-identified and exclude the 16 identifiable variables that necessitate Internal Review Board (IRB) approval for access. Because of the de-identified nature of the data, this study is not considered human subjects research. The University of Tennessee Health Science Center IRB determined that this research is exempt from IRB according to the National Institutes of Health Office of Human Subjects Research policy. This research was performed in accordance with all relevant guidelines/regulations.

Variables
Our outcome variables were multimorbidity patterns prevalent within each racial/ethnic category stratified by age and obesity status. There is no international consensus defining multimorbidity (30). We defined multimorbidity as the presence of two or more chronic conditions within one individual (31). For different definitions, the duration a patient must have a chronic condition varies from 3 months to a year (or may not list a specific duration) (32). Datasets derived from EHRs do not typically capture the duration that a patient may suffer from a diagnosis which may extend until the end of life. Several different organizations maintain lists of chronic conditions that are used in research, but each one omits some chronic conditions (33). Our aim is to include a broad array of conditions to better reflect patient disease states by using all ICD-10-CM codes representing conditions that can be chronic. The ICD-10-CM codes we considered include those from the following clinical categories: E00–E89, endocrine, nutritional, and metabolic diseases; F01–F99, mental, behavioral, and neurodevelopmental disorders; G00–G99, diseases of the nervous system; I00–I99, diseases of the circulatory system; J00–J99, diseases of the respiratory system, K00–K95, diseases of the digestive system; M00–M99, diseases of the musculoskeletal system and connective tissue; and N00–N99, diseases of the genitourinary system. The specific subcodes we considered are described in the text and enumerated in **Supplementary Table 2**. Patients were considered positive for individual diseases or comorbidities if their record included one or more ICD-10-CM diagnostic code(s) within the respective broad disease category during 2016–2017. For example, a patient with ICD-10-CM codes E11 and I10 would be positive for type II diabetes and essential (primary) hypertension, respectively. All sub-classifications of diseases were included under the umbrella of their broad disease code. For example, code E11.0 (type 2 diabetes mellitus with hyperosmolarity) would be categorized under the broader parental code E11 (type 2 diabetes mellitus).

Our analysis includes multimorbidities based on prevalent codes as prevalence-based selection of ICD-10-CM codes in multimorbidity research has been shown to be robust (34). Patients were classified with obesity if they had an average BMI of 30+ during the study period, and without obesity if their BMI was less than this cutoff. Patients were further stratified into middle-aged if they were between 45 and 64, and elderly if they were 65+. We considered a BMI value to be valid if it was less $\leq$ 206 as the highest recorded BMI values are in the low 200s between 206 and 224 (35,36). We recognize that due to the size of our sample, outliers are unlikely to have a noticeable impact on our results. Since BMI was used to stratify our results, we excluded it as an outcome variable. Racial categories were based on those present within the Cerner HealthFacts data warehouse. Because we were unable to separate patients listed as Asian/Pacific Islander

into either Asian or Pacific Islander categories, we decided to merge Asian, Pacific Islander, and Asian/Pacific Islander categories into one category. To ensure that we could identify the groups from which any category was drawn, we excluded patients from the study if their race/ethnicity was missing or listed as other. One category, Mid-Eastern Indian, was removed as the authors were unclear as to its meaning. The final racial categories used in the analysis were African American, Caucasian, Asian/Pacific Islander, Biracial, Hispanic, and Native American.

Identifying Multimorbidity Patterns

We used a Spark distributed cluster for our analysis. We used frequent itemset detection to find combinations of diseases above the threshold of 5% prevalence. This algorithm identified groups of 2 + diagnoses that appear together in the dataset for at least 5% of patients within each race/ethnicity age obesity-level cohort. Our rationale for setting support to 5% is as follows: the tremendous diagnostic heterogeneity and variation in the number of chronic conditions for the multimorbidity population suggests there will be many combinations of low frequency. Therefore, we did not want to set support too high. Frequent itemset detection can reveal disease clusters in which all patients share all group attributes. Below is an example problem where the minimum support (threshold of prevalence) is 60%.

Patient lists of ICD-10-CM diagnosis codes where each row represents one patient:
I10,G47,E78
I10,E78
G47
I10,R06
I10,R06,E78

The frequent patterns above the 60% threshold of minimum support are:
I10
E78
I10,E78

Finding frequent itemsets can be computationally expensive on large datasets. However, the parallel frequent pattern growth (FP-growth) algorithm is an efficient distributed frequent itemset mining algorithm that reduces the number of candidate itemsets when used on a distributed cluster to find patterns on large datasets (37).

The FP-growth algorithm is a scalable frequent itemset algorithm frequently used on large datasets38. The parameters for the FP-growth algorithm are as follows: Let I = {P1, P2…Pm} represent a set of m diagnoses. Each patient's diagnoses list L for the study period contains a set of diagnoses such that $L \subseteq I$. The support (occurrence frequency) of a pattern A, where A is a set of diagnoses is the number of patient lists containing A. A pattern is considered frequent if A's support is greater than or equal to a pre-defined minimum support threshold, ξ. We used the SparkR implementation of this algorithm (39). We applied the FP-growth algorithm on each cohort stratified by race/ethnicity, age, and obesity level. We compared results across races/ethnicities for each age range by

obesity level. We compared the disease combinations shared by all races/ethnicities, those shared by some, and those unique to one group for each age/obesity level.

We calculated confidence intervals for multimorbidity prevalence for diseases shared across races/ethnicities to identify which races/ethnicities may have similar prevalence. We used the Clopper–Pearson method to generate binomial proportion confidence intervals (40). To compare prevalence rates by race, we used the g-test of independence. These statistical analyses were performed using the R DescTools package (41). Sensitivity analysis included using the Kruskal–Wallis ANOVA for medians of BMI across race for each age/weight class to determine if our findings could be impacted by one race having a higher median BMI within the weight class, which could lead to skewed results. We used the WRS2 R package for this test (42). The g-test were used to determine statistical significance. P values < 0.05 were considered to be statistically significant.

Results

Sample: Middle Aged and Elderly Cohorts Stratified by Race/Ethnicity

A total of 1,212,956 patients matched our criteria in the middle-aged cohort and 1,003,498 patients in the elderly cohort. **Supplementary Table 1** shows the patient population for this study. Our patient population's average number of visits was seven, and 76% of patients had 2 + encounters. In each cohort, most patients were Caucasian (77% and 87% respectively), followed by African American (19% and 10% respectively), as shown in **Table 1**. Patients listed as Asian/Pacific Islander accounted for 2% of the Middle-aged cohort and 2% of the Elderly cohort. Native Americans accounts for 1% in the middle-aged cohort, and < 1% in the elderly cohort. Both the Biracial and Hispanic cohorts account for < 1% of the samples in both age groups. Except for Asians/Pacific Islanders, most middleaged patients for each race/ethnicity were in the with obesity category. For the elderly cohorts, there were more patients with obesity for all three categories.

Table 1. Demographic breakdown across age and weight class

Race/Ethnicity	Middle Aged			Elderly		
	Total % of Population (Total Population)	Without obesity	With obesity	Total % of Population (Total Population)	Without obesity	With obesity
African American	19% (235,612)	42% (98,759)	58% (136,859)	10% (102,138)	55% (55,745)	45% (46,393)
Asian/Pacific Islander	2% (21,013)	76% (15,947)	24% (5,066)	2% (15,427)	84% (12,982)	16% (2,445)
Biracial	<1% (1,220)	49% (596)	51% (624)	<1% (506)	57% (288)	43% (218)
Caucasian	77% (936,472)	48% (448,299)	52% (488,173)	87% (876,538)	59% (520,821)	41% (355,717)
Hispanic	<1% (7008)	46% (3200)	54% (3808)	<1% (3,227)	59% (1,895)	41% (1,332)
Native American	1% (11,578)	38% (4443)	62% (7135)	<1% (5,662)	52% (2,963)	43% (2,699)

Table 2 shows the number of total multimorbidities at the 5% threshold for prevalence across weight class for each race/ethnicity cohort with the number of distinct and overall multimorbidities for each by age and weight cohort. We found African Americans patients have the highest number of total multimorbidities and the most distinct multimorbidities for each age/weight group. Middle-aged without obesity cohorts have the lowest number of multimorbidities for each race/ethnicity, and Elderly with obesity have the most. **Supplementary Table 2** describes the diagnosis codes and corresponding diagnoses included in the observed multimorbidity patterns.

Table 2. The number of multimorbidity patterns by race/ethnicity at 5% threshold across weight class

Race	Middle Aged Without obesity		Middle Aged With obesity		Elderly Without obesity		Elderly With obesity	
	Overall	Distinct	Overall	Distinct	Overall	Distinct	Overall	Distinct
African American	17	6	50	20	112	45	157	37
Asian/Pacific Islander	4	0	15	0	41	1	98	2
Biracial	2	0	11	0	58	3	65	1
Caucasian	5	0	24	2	58	8	95	9
Hispanic	6	0	12	0	44	0	63	0
Native American	13	1	22	3	60	4	107	3

Multimorbidities Shared Across all Races/Ethnicities

Figure 1a–d show the confidence intervals for multimorbidity patterns shared across all race/ethnic groups analyzed within each age/obesity cohort. Gray sectors indicate the combination was not prevalent above the 5% threshold for a give group. The multimorbidity patterns shared across all racial and ethnic groups included one or more of the following ICD-10-CM codes with the corresponding diagnoses: I10: Hypertension, E78: Lipidemia, or E11: Diabetes. As described in **Supplementary Table 2**. There were only two shared patterns across all races/ethnicities in the middle-aged cohort without obesity (**Figure 1a**). Caucasians had the lowest prevalence for the E11: Diabetes, I10: Hypertension combination, and Biracial had the lowest prevalence for the E78: Lipidemia, I10: Hypertension combination. In the middle-aged patients with obesity cohort, there were six multimorbidity patterns. In this cohort, the E78: Lipidemia + I10: Hypertension and E11: Diabetes + I10: Hypertension combinations had the highest prevalence for each race/ethnicity. There were only two clinical categories (E: Endocrine, nutritional and metabolic, and I: Circulatory system) represented in the middle-aged cohort without obesity, while that increased to 7 in patients with obesity (including K: Digestive system and M: Musculoskeletal system and connective tissue).

For five of the multimorbidity patterns observed for middle-aged patients with obesity (**Figure 1b**), the 95% confidence interval for the African American patients did not overlap with any other racial/ethnic group with one exception, E11: Diabetes + I10: Hypertension. For the E78: Lipidemia and E11: Diabetes diagnosis codes, the 95% confidence interval overlaps with that of the Asian/Pacific Islander cohort. For patients without obesity, there is an increase from two patterns in the middle-aged to 26 patterns in the elderly.

Results for the elderly cohorts appear in **Figure 1c,d** and Supplementary **Figure 1**. There are 26 patterns for the without obesity cohort and 37 for the with obesity cohort. Due to the difficulty in interpreting so many patterns, we have limited the set that appear in **Figure 1d** to only those with prevalence equal to or above 0.08%. In the without obesity cohort (**Figure 1c**), five patterns include I25: Heart Disease, and two include N18: Chronic kidney disease (CKD). For the E78: Lipidemia + I10: combination, the Asian/Pacific Islander and African American cohort have the highest estimates. For three patterns, the Caucasian cohort estimate error bars do not overlap with any other group and are the lowest.

In the elderly patients with obesity cohort (**Figure 1d** and **Supplementary Figure 1**), 25 multimorbidity patterns were present. The following diagnoses appeared in 3 or more multimorbidity patterns: I10: Hypertension (19 patterns), E78: Lipidemia (13patterns), E11: Diabetes (12 patterns), I25: Heart disease (8 patterns), N18: Chronic kidney disease (6 patterns), K21: GERD (4 patterns), I50: Heart failure (3 patterns), and E87: Other disorders of fluid, electrolyte and acid–base balance (3 patterns). The elderly cohort patterns include three new clinical categories, D: Disease of blood and blood-forming organs, N: Diseases of the genitourinary system, and G: Diseases of the nervous system.

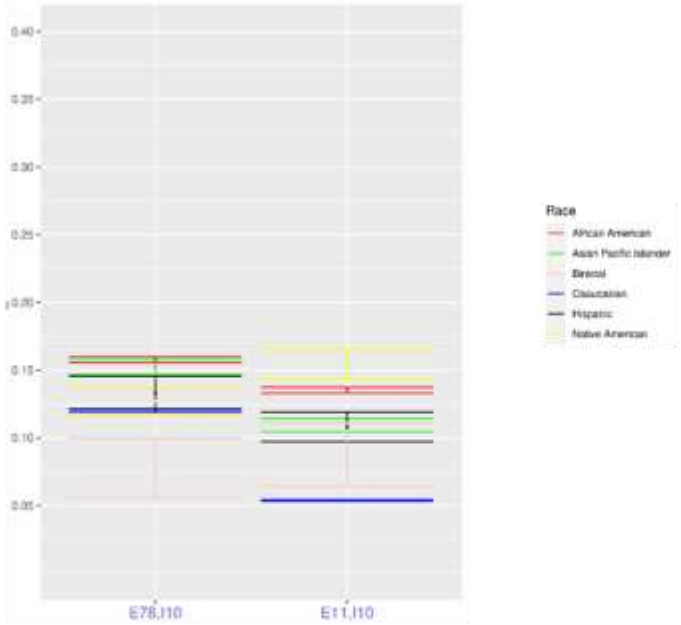


Figure 1a *Confidence Interval Overlap for Middle Aged Patients Without Obesity Shared by All Races*

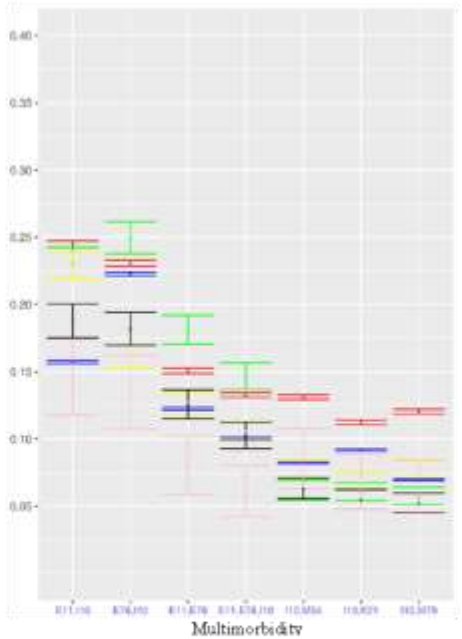


Figure 1b *Confidence Interval Overlap for Middle Aged Patients with Obesity Shared by All Races*

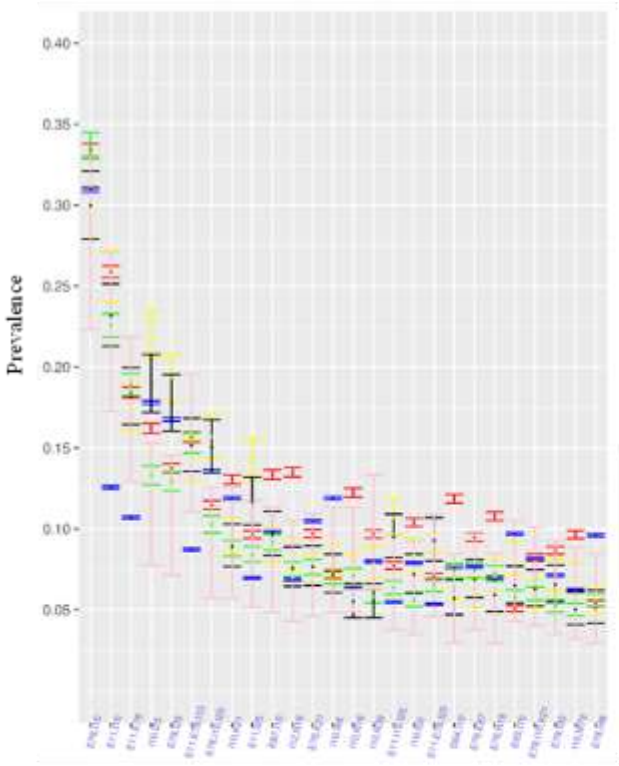


Figure 1c *Confidence Interval Overlap for Elderly Patients Without Obesity Shared by All Races*

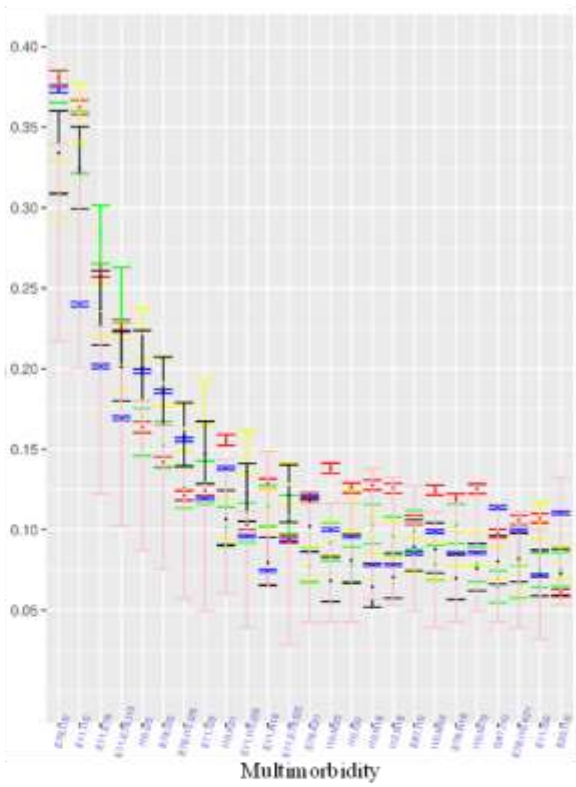


Figure 1d *Confidence Interval Overlap for Elderly Patients with Obesity Shared by All Races (Top Average Prevalent Multimorbidities)*

Figure 1.

Figures 1a and **1b** demonstrate the 95% confidence interval overlap for shared multimorbidities among middle aged patients without obesity and patients with obesity, respectively, at 5% prevalence threshold. The g-test of independence was calculated resulting in all significant results with $p < .05$.

Figures 1c and **1d** demonstrate the 95% confidence interval overlap for shared multimorbidities among elderly patients without obesity and patients with obesity, respectively, at 5% prevalence threshold. The chi square test was calculated resulting in all significant results with $p < .05$ except for the multimorbidity patterns circled in red. Disease Clinical Categories; **E**: Endocrine, nutritional and metabolic, **F**: Mental, behavior and neurodevelopmental, **G**: Nervous System, **I**: Circulatory system, **J**: Respiratory system, **K**: Digestive system, **N**: Genitourinary system, **M**: musculoskeletal system and connective tissue. (1d) Demonstrates shared multimorbidities with average prevalence ≥ 0.08.

Multimorbidities Shared Among some Races/Ethnicities

Multimorbidities shared across two or more races/ethnicities, but not all, in the middle-aged and elderly cohorts are shown in **Figure 2a–d**. Gray sectors indicate the combination was not prevalent above the 5% threshold for a given group. A total of 10 multimorbidity patterns for patients without obesity and 23 for patients with obesity appeared in the middle-aged patient cohort (**Figure 2a,b**). Clinical categories not represented in the middle-aged cohort shared by all but appear in the set of those shared by some are F: Mental, Behavioral and Neurodevelopmental disorders and J: Diseases of the respiratory system. All combinations for the middle-aged included at least one of the following diseases: E11: Diabetes, I10: Hypertension, E78: Lipidemia, M54: Dorsalgia, or M25: other joint disorders. In the middle-aged cohort without obesity (**Figure 2a**), African Americans and Native Americans shared all multimorbidity patterns except for the E78: Lipidemia I25: Heart disease combination. African Americans and Native Americans had higher prevenances of the F17: Nicotine dependence-I10: Hypertension combination than the Caucasian cohorts in both the with and without obesity middle-aged samples (**Figure 2a–d**). In the middle-age patients with obesity cohort (**Figure 2b**), the multimorbidities with the highest prevalence for each group include I10: Hypertension, I25: Chronic ischemic heart disease, G47: Sleep disorders, and K21: Gastro-esophageal reflux disease appeared in several combinations. Patterns including musculoskeletal system diseases and connective tissue (those that start with M) were prevalent for the African American, Native American, Asian/Pacific Islander, and Biracial cohorts only. Middle-aged with obesity African Americans and Native American groups (**Figure 2b**) have a similar prevalence for the following combinations: E11: Diabetes + M25: Other joint disorders; E11: Diabetes + M79: Other and unspecified soft tissue disorders; M25: Other joint disorders + M79: Other and unspecified soft tissue disorders, and M25: Other joint disorders + M54: Dorsalgia (**Figure 2b**). The only two combinations present for the Asian/Pacific Islander without obesity (**Figure 2a**) include both diabetes E11: Diabetes and E78: Lipidemia.

Results for the elderly cohorts appear in **Figure 2c–d** and **Supplementary Figs. 2, 3**. A total of 47 multimorbidity patterns for patients without obesity and 90 for patients with obesity appeared in the elderly patient cohort. Due to the difficulty in interpreting so many patterns, we have limited the set that appear in **Figure 2c–d** to only those with prevalence equal to or above 0.065. The rest of the patterns for the without obesity cohort appear in **Supplementary Figure 2**, and those for the with obesity appear in **Supplementary Figure 3**. No new clinical categories are represented. Combinations in the cohort without obesity (**Figure 2c** and **Supplementary Figure 2**) include diagnoses not found in the middle-aged cohort, including N17: Acute Kidney Failure and F17: Nicotine dependence. In the elderly cohort without obesity (**Figure 2c** and **Supplementary Figure 2**), there are 11 patterns comprised of three morbidities and eight comprised of four morbidities. The patients in the elderly with obesity cohort (**Figure 2d** and **Supplementary Figure 3**) exhibited 46 multimorbidity combinations comprised of three multimorbidities and five comprised of four multimorbidities. In the cohort without obesity, the three most prevalent patterns comprised three diagnoses (**Figure 2c**) all include one to two diagnoses in the I: Diseases of the circulatory system category. These

patterns are not prevalent for the Caucasian cohort**.** The highest prevalent pattern with four diagnoses (**Figure 4c**) includes I25: Heart Disease and is prevalent for all groups except Caucasians. In the elderly cohort, the same highest prevalent pattern of four morbidities, E11: Diabetes + E78: Lipidemia + I10: Hypertension + I25: Heart Disease, is prevalent for all groups except Biracial. In the cohort without obesity, all patterns of three or four are prevalent for the African American cohort except for E03: Other Hypothyroidism + E78: Lipidemia + I10: Hypertension (**Supplementary Figure 2**). That combination is prevalent for the Caucasian and Native American cohorts only. Five of the top prevalent patterns of three morbidities in the elderly cohort with obesity include N18: Chronic Kidney Disease. The combination of I10: Hypertension + J44: Other chronic obstructive pulmonary disease (COPD) was prevalent for all groups except Asian/Pacific Islander and Biracial in the without obesity patients (**Figure 2c**) and for all except Asian/Pacific Islander in the elderly with obesity patients (**Figure 2d**). F17: Nicotine dependence appeared in the African American and Native American elderly cohort without obesity (**Figure 2c**), yet it did not appear in any shared pattern in elderly patients with obesity (**Figure 2d**).

Figure 2a *Multimorbidities Shared by Some Races for Middle Aged Without Obesity Patients*

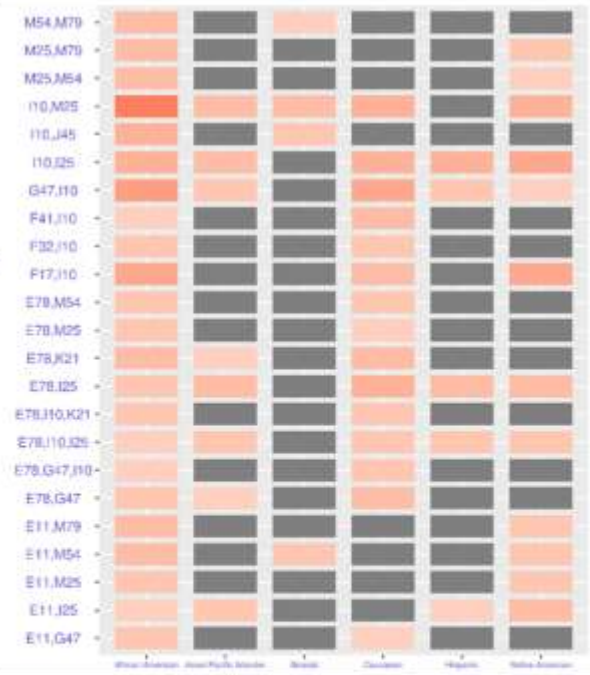


Figure 2b *Multimorbidities Shared by Some Races for Middle Aged Patients with Obesity*

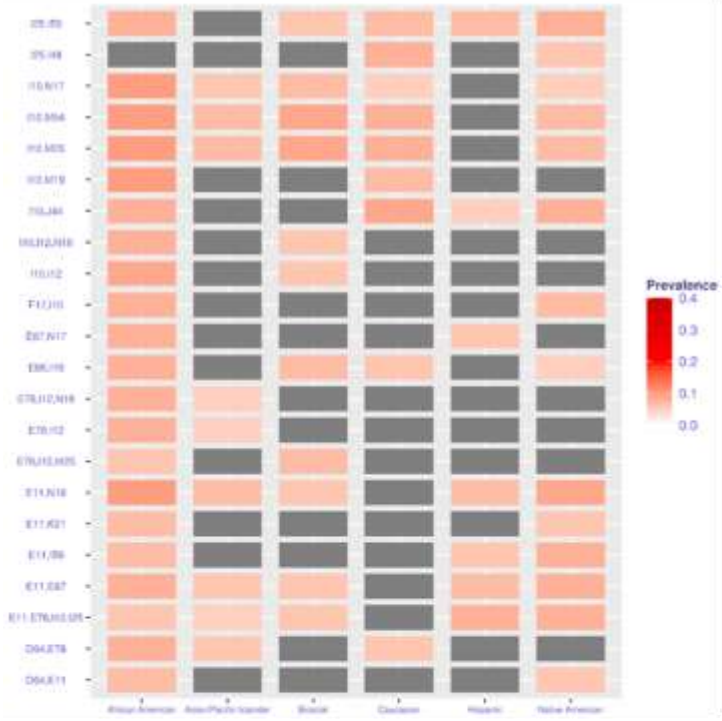


Figure 2c *Multimorbidities Shared by Some Races for Elderly Without Obesity Patients & Average Prevalence ≥ 0.065*

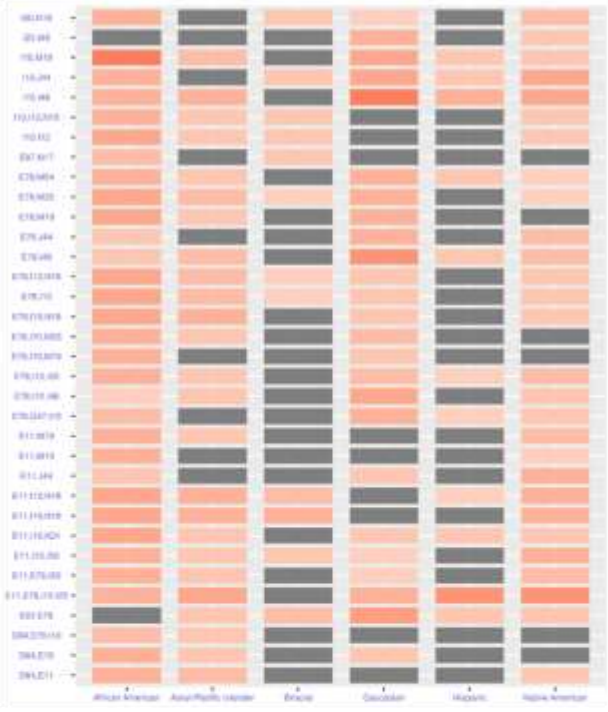


Figure 2d *Multimorbidities Shared by Some Races for Elderly Patients with Obesity & Average Prevalence ≥ 0.065*

Figure 2.

Figures 2a and **2b** show shared multimorbidities across two or more races in middle aged patients without and with obesity, respectively, at 5% prevalence threshold. The g-test of independence was calculated resulting in all significant results with $p<.05$. Gray sectors indicate the combination was not prevalent above the 5% threshold for a given group. The g-test of independence was calculated resulting in all significant results with $p < 0.05$ except for red boxed multimorbidity pattern(s) indicating non-significant result.

Figures 2c and **2d** Demonstrate the shared multimorbidities among two or more races in the elderly patients without and with obesity, respectively, at 5% prevalence threshold. The g-test was calculated resulting in all significant results with $p < 0.05$ except for the multimorbidity patterns circled in red. Disease Clinical Categories; **E**: Endocrine, nutritional and metabolic, **F**: Mental, behavior and neurodevelopmental, **G**: Nervous System, **I**: Circulatory system, **J**: Respiratory system, **K**: Digestive system, **N**: Genitourinary system, **M**: musculoskeletal system and connective tissue.

Multimorbidities Distinct to One Race/Ethnicity

We also identified distinct multimorbidities unique to a specific racial/ethnic group in both middle-aged and elderly cohorts (**Figure 3a–d**). Gray sectors indicate the combination was not prevalent above the 5% threshold. The African American cohort had the most distinct multimorbidity patterns in both weight groups in the middle-aged cohort (**Figure 3a,c**). They also had distinct multimorbidity patterns appear with F17: Nicotine dependence or N18: Chronic kidney disease (CKD) in both cohorts. In the middle-aged patients with obesity cohort (**Figure 3c**), African Americans had K21: GERD, D64: Other anemias, I50: Heart failure, and musculoskeletal disorders (those starting with M) appearing in several distinct multimorbidity patterns. The cohort without obesity had ten distinct patterns of 3 morbidities. The Native American middle-aged cohort (**Figure 3b,e**) had F17: Nicotine dependence appear in a distinct multimorbidity pattern in both cohorts. Each pattern for the Native American includes E11: Diabetes in both with and without obesity groups. Additionally, in the cohorts with obesity, Native Americans had two distinct patterns of three diagnoses (**Figure 3e**). For the middle-aged cohort, the Caucasian group had two distinct patterns among the patients with obesity (**Figure 3d**). Both include a diagnosis in the E: Endocrine, nutritional and metabolic. One includes F41: Other anxiety disorders.

In the elderly, the African American cohort without obesity (**Figure 3c**) had many combinations that combined with I10: Hypertension, E78: Lipidemia, or E11: Diabetes, including N18: Chronic kidney disease (CKD), I12: Hypertensive chronic kidney disease, E87: Other disorders of fluid, electrolyte and acid–base balance, F17: Nicotine dependence, K21: GERD, D64: Other anemias, and I50: Heart failure. In the cohort without obesity (**Figure 4a**), there are 21 multimorbidities comprised of 3 or 4 diseases. In the cohort without obesity (**Figure 4a**), there are 21 multimorbidities comprised of three or four diseases and 19 in the cohort with obesity (**Figure 5a**). Caucasians had F41: Anxiety Disorders and F32: Depressive episode appearing in distinct multimorbidity patterns in the elderly cohorts with and without obesity (**Figs. 4d, 5d**). This group also had I48: Atrial fibrillation and flutter appear in distinct multimorbidity patterns in the elderly cohorts (**Figs. 4d, 5d**). In the cohort without obesity (**Figure 4d**), there are four multimorbidities comprised of three or four diseases as five in the cohort with obesity (**Figure 5a**). For the Asian/Pacific Islander, without obesity group (**Figure 4b**), the single pattern displayed includes I10: Hypertension and M81: Osteoporosis without current pathological fracture. For the cohort with obesity, the Asian/Pacific Islander cohort (**Figure 5b**) had one pattern of size four. The two patterns for this cohort are composed of I10: Hypertension and diseases in the E: Endocrine, nutritional, and metabolic clinical category. For the Biracial without obesity cohort (**Figure 4c**), all patterns include musculoskeletal disorders (those starting with M). The only distinct pattern for the Biracial cohort with obesity (**Figure 5c**) includes E03: Other hypothyroidism. In the Native American without obesity cohort (**Figure 4e**), the combinations include I50: Heart Failure, J18: Pneumonia, and I65: Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction. In the cohort with obesity (**Figure 5c**), patterns include F17: Nicotine dependence and J44: Other chronic obstructive pulmonary disease (COPD). For both the Biracial and Native American samples in the cohort without

obesity (**Figs. 4c, e**), there is one pattern of three morbidities and one in the cohort with obesity (**Figure 5c, e**).

To determine whether any one racial/ethnic group might have an excessively high or low median BMI within a given weight class, we performed Kruskal–Wallis ANOVA to compare median BMI values across race/ethnicity for each age-group and weight class. Median values by race/ethnicity appear in Table 3. The results of the ANOVA analysis showed that the difference in median BMI values for each age or weight class was statistically significant ($p < 0.05$). For the cohorts without obesity, the difference between each racial/ethnic group was just one point (**Table 3**). All patient cohorts without obesity exhibited median BMI values of 25–26, which fall into the overweight weight class. For middle-aged patients with obesity, the median BMI value for each racial group was 33–36 (**Table 3**). For the elderly cohort without obesity, all median values are between 24 and 25. For the elderly with obesity cohort, all medians are between 32 and 34 (**Table3**).

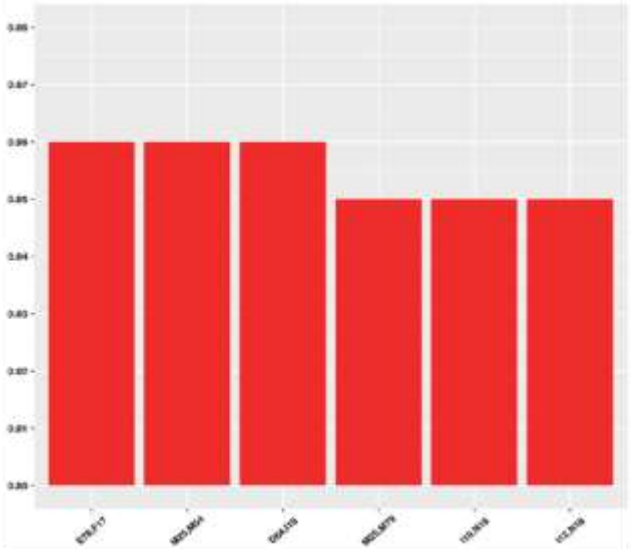


Figure 3a *African American Distinct Multimorbidities for Middle Aged Patients Without Obesity.*

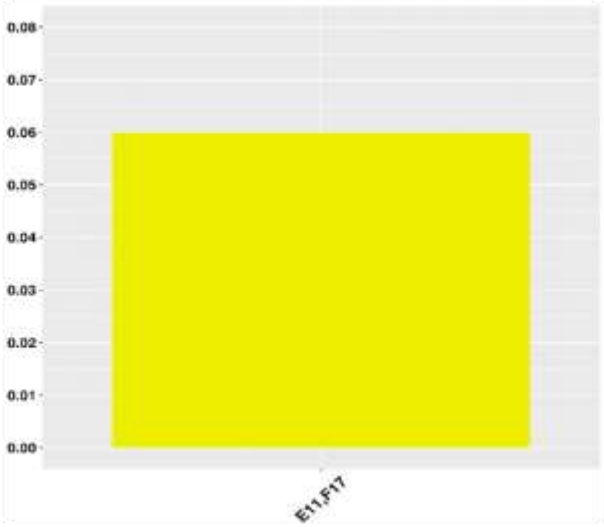


Figure 3b *Native American Distinct Multimorbidities for Middle Aged Patients Without Obesity.*

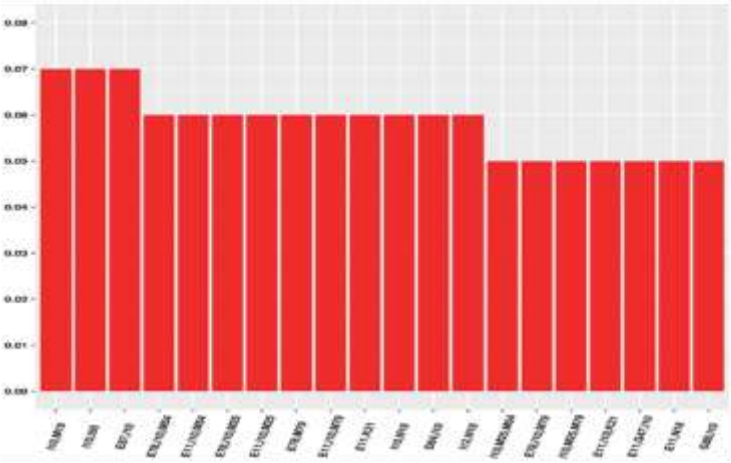


Figure 3c *African American Distinct Multimorbidities for Middle Aged Patients with Obesity.*

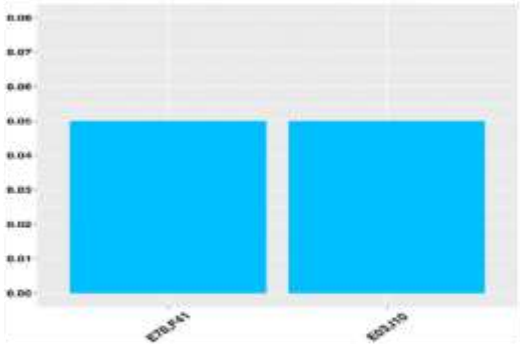


Figure 3d *Caucasian Distinct Multimorbidities for Middle Aged Patients with Obesity.*

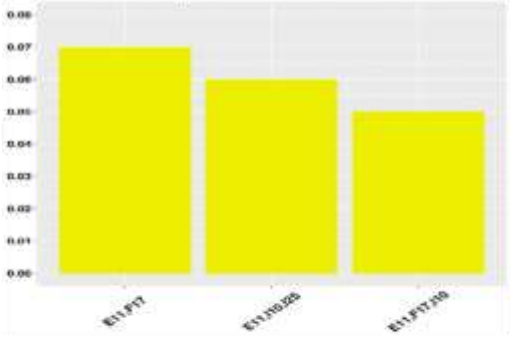


Figure 3e *Native American Distinct Multimorbidities for Middle Aged Patients with Obesity.*

Figure 3.

Figures 3a and **3b** demonstrates distinct multimorbidities across races in middle-aged without obesity at 5% prevalence threshold. Gray sectors indicate the combination was not prevalent above the 5% threshold for a given group.

Figures 3c – 3e demonstrate the distinct multimorbidities for each race in middle aged patients at 5% prevalence threshold. Disease Clinical Categories; D: Blood and immune mechanism disorders, E: Endocrine, nutritional and metabolic, F: Mental, behavior and neurodevelopmental, G: Nervous System, I: Circulatory system, J: Respiratory system, K: Digestive system, N: Genitourinary system, M: musculoskeletal system and connective tissue. The g-test of independence was calculated resulting in all significant results with $p < 0.05$.

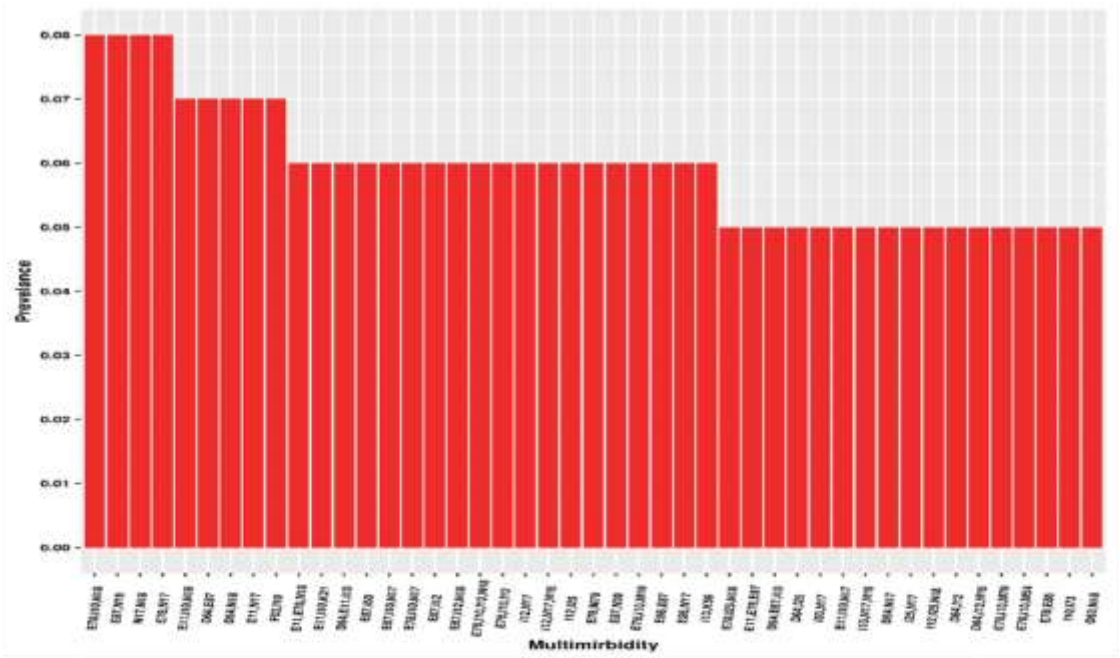


Figure 4a *Distinct Multimorbidities for African American Elderly Cohort Without Obesity*

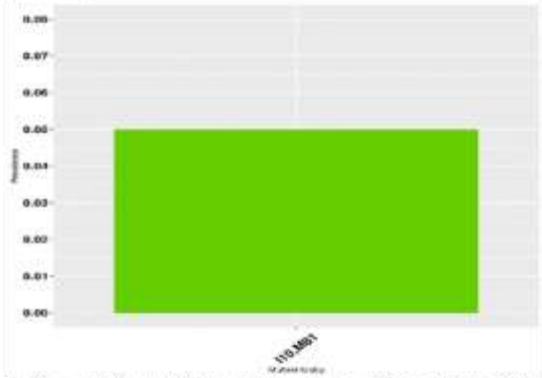


Figure 4b *Distinct Multimorbidities for Asian/Pacific Islander Elderly Cohort Without Obesity*

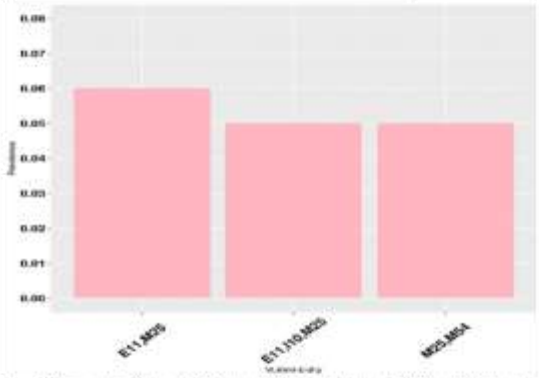


Figure 4c *Distinct Multimorbidities for Biracial Elderly Cohort Without Obesity*

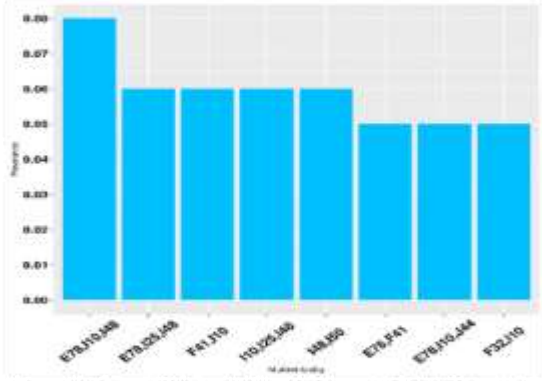


Figure 4d *Distinct Multimorbidities for Caucasian Elderly Cohort Without Obesity*

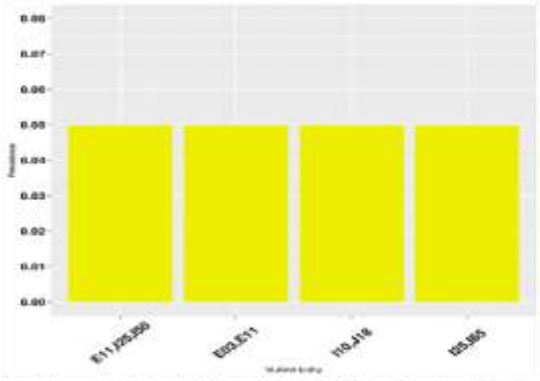


Figure 4e *Distinct Multimorbidities for Native American Cohort Without Obesity*

Figure 4.

Figures 4a – 4e Demonstrate distinct multimorbidities across races in elderly patients without obesity at 5% prevalence threshold. Disease Clinical Categories, D: Blood and immune mechanism disorders, E: Endocrine, nutritional and metabolic, F: Mental, behavior and neurodevelopmental, G: Nervous System, I: Circulatory system, J: Respiratory system, K: Digestive system, N: Genitourinary system, M: musculoskeletal system and connective tissue. The g-test of independence was calculated resulting in all significant results with $p<.05$.

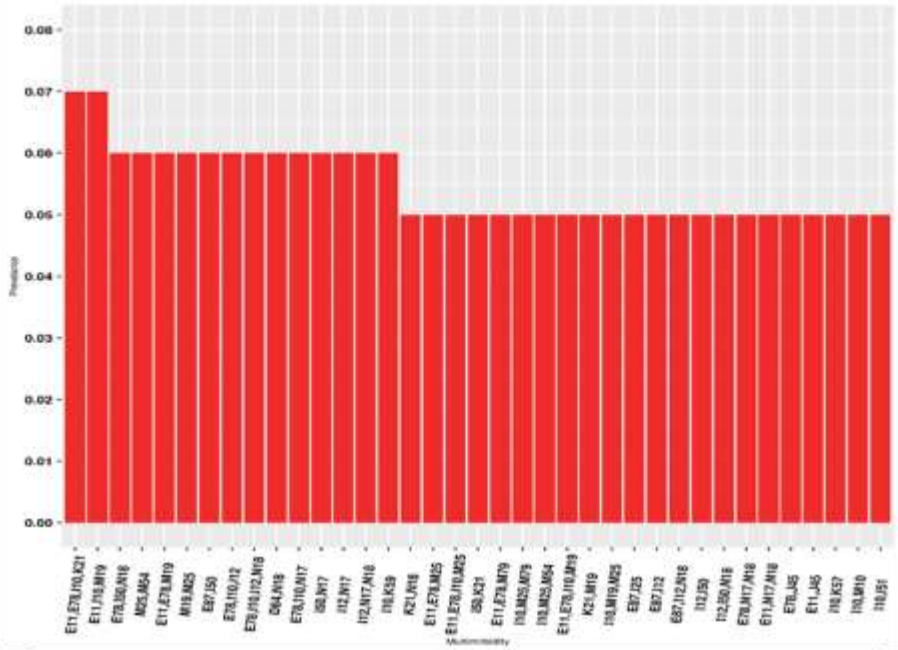


Figure 5a *Distinct Multimorbidities for African American Elderly Cohort with Obesity*

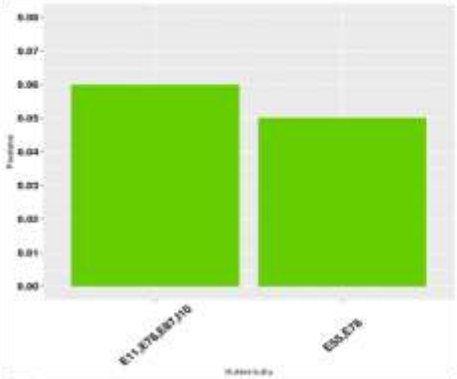


Figure 5b *Distinct Multimorbidities for Asian/Pacific Islander Elderly Cohort with Obesity*

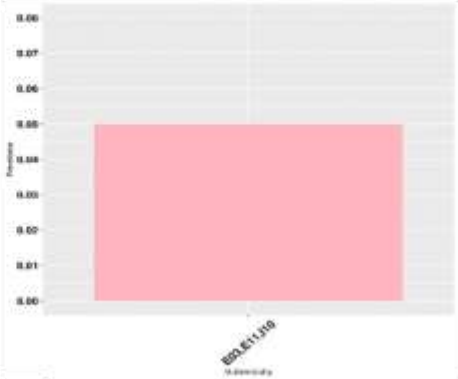


Figure 5c *Distinct Multimorbidities for Biracial Elderly Cohort with Obesity*

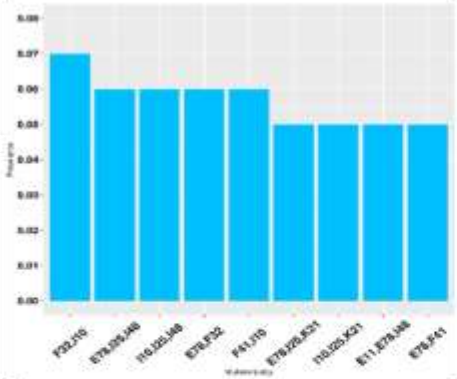


Figure 5d *Distinct Multimorbidities for Caucasian Elderly Cohort with Obesity*

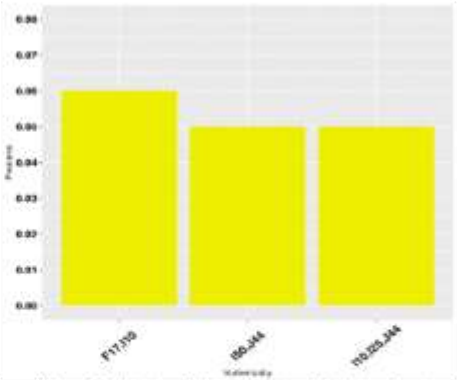


Figure 5e *Distinct Multimorbidities for Native American Elderly Cohort with Obesity*

Figure 5.

Figures 5a – 5e demonstrate distinct multimorbidities across races in elderly patients with obesity at 95% confidence and at 5% prevalence threshold. Disease Clinical Categories; D: Blood and immune mechanism disorders, E: Endocrine, nutritional and metabolic, F: Mental, behavior and neurodevelopmental, G: Nervous System, I: Circulatory system, J: Respiratory system, K: Digestive system, N: Genitourinary system, M: musculoskeletal system and connective tissue. The g-test of independence was calculated resulting in all significant results with $p<.05$.

Table 3. Median BMI across race/ethnicity for each weight class

Race	Middle Aged Without obesity	Middle Aged With obesity	Elderly Without obesity	Elderly With obesity
	Median	Median	Median	Median
African American	25	36	25	34
Asian/Pacific Islander	25	33	24	32
Biracial	26	35	25	33
Caucasian	26	35	25	34
Hispanic	26	34	25	33
Native American	26	35	25	34

Discussion

The present study identified the most prevalent multimorbidity patterns across races/ethnicities, stratified by age and obesity, and compared individual disease prevalence across age cohorts. It also assessed multimorbidity disease prevalence across cohorts. Our findings demonstrate that even after stratifying by age and obesity, there are differences in multimorbidity prevalence across races/ethnicities, and some combinations are distinct to race/ethnic groups. Although we are not the first to use a frequent itemset algorithm to assess multimorbidity18, our study adds to the current body of knowledge by examining the prevalence of *specific* multimorbidity patterns by racial/ethnic category, stratifying by age and obesity status. Many of the identified disease combinations have likely not been studied, as identifying unique patterns among races/ethnicities is unprecedented. Our study showed that the common morbidities present in disease combinations across all races/ethnicities were lipidemia, hypertension, and diabetes regardless of age or obesity level. Multimorbidity increased with age in both with and without obesity groups. Multimorbidity prevalence was the highest among African Americans and lowest among Asian/Pacific Islanders. Even when factoring in age and weight class, the differences remain. The disease composition of multimorbidity also varied by race/ethnicity. African Americans presented with the most distinct multimorbidities at an earlier age compared with other races/ethnicities. Asian/Pacific Islanders, Biracial, and Hispanic groups had no distinct multimorbidities among the middle-aged. By understanding how multimorbidity patterns may clinically present within specific patient populations, health care professionals can implement more structured care plans and provide more appropriate care. Traditional care plans focus on one chronic disease and do not consider the impact of multiple risk factors and multimorbidity. This "single-disease treatment" approach may be inadequate in patients with multimorbidity and result in the involvement of various specialists, as well as an increase in polypharmacy (43). These findings can be used to inform public policy and to develop patient guidelines at various obesity levels and ages.

Our study observed differences in multimorbidity prevalence and composition between the middle-aged and elderly groups and between with and without obesity cohorts. Unique diseases and disease combinations were shared by all race/ethnic groups within the elderly cohorts that weren't shared by all in the middle-aged. For example, we observed that ICD-10-CM diagnosis codes M25: Other joint disorders, M54: Dorsalgia, I25: Heart disease, I12: Hypertensive chronic kidney disease, N18: Chronic kidney disease (CKD), I50: Heart failure, E87: Other disorders of fluid, electrolyte and acid–base balance, E03: Other hypothyroidism, and M79: Other and unspecified soft tissue disorders, not elsewhere classified, in the elderly group but not in the middle-aged group (Figure 1c,d). In addition, these diseases were frequently paired with either I10: Hypertension or E78: Lipidemia, which is consistent with findings from other studies (44). Interestingly, in the elderly cohort with obesity, I25: Heart disease was combined with either E78: Lipidemia, I10: Hypertension, or E11: Diabetes, in three of five multimorbidity triads shared by all race/ethnicities (**Figure 1d**). The triads are similar to those found in other studies. For example, Lim et al. 2018 similarly found that hyperlipidemia, hypertension, and diabetes were the most prevalent in dyad and triad disease combinations (44). However, no other studies have looked for disease combinations of size 4. Therefore, our study is likely to have identified new triads and combinations of more than three diseases.

We found that African Americans presented with the highest number of multimorbidities at an earlier age than patients of other race/ethnicities, consistent with results observed in other studies (6). African Americans are exposed to more traumatic experiences and stressors, such as discrimination and poverty, earlier on in life, which produces additional health risks and contributes to worse health outcomes in later life (45,46). Although mental health disorders are not prevalent, the earlier emergence of multimorbidity could result from psychological distress at an earlier age (47).

We identified a higher number of multimorbidity patterns distinct to African Americans, including many combinations of three or four diseases in the elderly with obesity. Previous studies have primarily identified multimorbidities consisting of up to three diseases. Our findings of multiple combinations of four diseases suggest that patients present with complex disease profiles that have likely never been studied before. The level of social vulnerability of particular groups should also be considered, as a significant correlation between social vulnerability and the total number of chronic conditions was previously demonstrated, with depression/anxiety, obesity, and cardiovascular diseases being the most related (47). Further, these diseases were present in many patterns for various racial groups.

Another combination distinct to elderly, African Americans without obesity (**Figure 3c**) was E87: Other disorders of fluid, electrolyte, and acid–base balance and I10: Hypertension. Alterations in acid–base transporters have been linked to hypertension (48). Converging evidence indicates a pathogenic role of combined high sodium and low potassium levels in the development of hypertension and hypertension-associated cardiovascular complications (49). Both these diagnoses serve as potential precursors of CKD, which is present at a higher prevalence in African Americans than other race/ethnic

groups. As the kidneys play a vital role in regulating body fluids, electrolytes, and acid–base balance, CKD and end-stage renal disease (ESRD) predictably result in multiple complications,
including hyperkalemia, metabolic acidosis, and hyperphosphatemia (49,50). The multimorbidity combination consisting of E11: Diabetes and F17: Nicotine dependence was a unique combination to the elderly without obesity as well as the middle-aged with and without obesity Native Americans. According to the Centers for Disease Control and Prevention, American Indian/Alaska Native youth and adults have the highest prevalence of cigarette smoking in the US compared with other racial groups. Additionally, the risk of developing diabetes is 30–40% higher among smokers compared with nonsmokers (51).

Caucasians had the most multimorbidity combinations in middle-aged with obesity as well as the elderly with and without obesity cohorts that include mental disorders, such as F41: Anxiety and F32: Major depressive disorder. Breslau et al. demonstrated that Hispanics and Non-Hispanic African Americans were at lower risk for common internalizing disorders: depression, generalized anxiety disorder, and social phobia compared to Caucasians (52). F41: Anxiety presented with I10: Hypertension in the list of patterns distinct to Caucasians for the middle-aged and elderly with obesity cohorts. One study reported that having an anxiety disorder was associated with a fourfold increase in the risk of developing hypertension (53). Major depressive disorder coupled with lipidemia was also distinct to Caucasians and was only present in the elderly with obesity. Studies have shown that anxiety and depression may be associated with increased cardiovascular risks and abdominal obesity. One study specifically noted that increased dyslipidemia and obesity risk in patients with severe anxiety disorders and depression might be partly explained by chronic low-grade inflammation and smoking (54). Nicotine dependence was present in any disease combination among Caucasians. Comorbidities associated with mental health related disorders can present significant challenges leading to increased disability, longer hospital stays, and increased mortality (55). Additionally, challenges related to managed care communication between provider and patient is more significant for patients with mental health diagnoses (56).

The biracial group had no distinct multimorbidity combinations in the middle-aged cohort and few in the elderly. Of the multimorbidity combinations presented, hypertension was the most common morbidity present in combination with diseases in the clinical categories of the circulatory system or endocrine, nutritional and metabolic disorders. These findings are similar to previous work, which reported only one connection between disorders of circulatory system and disorders of endocrine, nutritional and metabolic diseases, and immunity amongst the biracial group (20). However, our study is the first to identify multimorbidity prevalence within this group.

Our results demonstrate that Asian/Pacific Islander groups did not have any distinct multimorbidity combinations among the middle-aged and few among the elderly. This is consistent with other studies that find that combined Asian groups are typically healthier than non-Whites (57).

Hispanics had fewer distinct patterns than Caucasians, Native Americans, and African Americans for each cohort except the middle-aged without obesity. Further, there were no multimorbidities distinct to Hispanics among these same cohorts. The comparable rate of multimorbidity patterns could be explained by the Hispanic Paradox, which posits that better health and mortality outcomes in Hispanics compared with non-Hispanics is due to healthier immigrants migrating into the country, while unhealthy people leave (58). Therefore, despite Hispanics being of low socioeconomic status, studies observe Hispanics experience similar or better health outcomes than non-Hispanic Whites (59). Our results could also reflect the underrepresentation of Hispanics within our dataset, as Hispanic is considered an ethnicity rather than race and was not counted twice.

The results of our sensitivity analysis suggests that differences in median BMI distributions by race/ethnicity are not driving our results. While differences within median BMI ranges were statistically significant but typically different by only 1 or 3 points. Differences in multimorbidity patterns and prevalence may be attributed to other factors, such as social determinants of health (SDoH), as socio-cultural and behavioral factors have been shown to influence obesity disparities across races (60). It is suggested that there is a bidirectional relationship between multimorbidity severity and sociodemographic and lifestyle risk factors (61).

We observed that African Americans and Native Americans appeared to share a higher prevalence of certain disease combinations than other races/ethnicities. For example, in the subset of patterns shared by only some of the race/ethnic groups in the middle-aged with obesity. African Americans and Native American groups had similar prevalence for the following combinations: E11: Diabetes + M25: Other joint disorders, E11: Diabetes + M54: Dorsalgia, and M25: Other joint disorders + M54: Dorsalgia as displayed in our results for combinations shared by only some groups as shown in Figure 2b. Further, in the middle-aged without obesity cohort, these two groups shared three patterns with diagnoses in the musculoskeletal disorders (those starting with M) and shared seven in the middle-aged with obesity cohort. The presence of dorsalgia in combination with hypertension and diabetes across these races may be attributed to a deficiency of Vitamin D. One study observed that vitamin D deficiency was common amongst African Americans and Hispanics and may contribute to cardiovascular disease and diabetes (62,63) and can also be a source of leg pain, widespread pain, arthralgia, rib pain, and back pain (64). Musculoskeletal disorders are also commonly seen in patients with Type 1 and Type 2 diabetes (65). These findings suggest that these specific groups have particular needs and certain morbidities that contribute to the emergence of unique multimorbidities, further emphasizing the necessity of understanding multimorbidities by race/ethnicity.

Guidance on how best to manage the health needs of specific patients based on their multimorbidity profile could be beneficial both for patients and the providers providing care. Our findings suggest that multimorbidity profiles should also include information on race/ethnicity and obesity. For example, African Americans were more at risk for multimorbidity earlier, even when without obesity. This suggests that providers should be

vigilant about screening for diseases within this group at earlier ages, regardless of obesity category, compared with other races/ethnicities. This research could further inform larger public health goals by directing initiatives towards what is most prevalent within the patient population and promoting health system design improvement through the provision of patient- and family-centered care approaches (66). Healthcare costs and utilization are also significant implications for patients with multimorbidity. By understanding multimorbidity, identifying the patients who are most likely to be affected by it, and examining common disease patterns, we could identify patients who are most likely to incur healthcare costs (67), thus potentially reducing the economic burden of care that would be incurred after multimorbidity develops.

This study only analyzed patterns that were frequent above the threshold of 5%. There are multimorbidity patterns frequent below that threshold. Many of those combinations could become more prevalent over time. As part of our future work, we plan to analyze disease patterns across time to better understand the complexity of multimorbidity across race/ethnicity.

This study has a few limitations. First, the study sample is limited to patients seeking healthcare during the study period. Therefore, our results may underrepresent diseases that are prevalent among groups that are less likely to seek healthcare. Second, although we can account for certain factors, the cross-sectional design of our study limits our ability to fully understand age and obesity as risk factors for the development of multimorbidity. These results suggest there are differences in multimorbidity prevalence and composition between the middleaged and elderly groups, and between groups with and without obesity. However, due to the study design, we could not assess the progression of multimorbidities in these subgroups over time. Our study did not account for gender due to the complexity of adding an additional stratification. So, we could not understand the impact of gender on multimorbidity. Lastly, the collection and identification of race and ethnicity pose another limitation. The race and ethnicity data is driven by the Uniform Hospital Discharge Data Set (UHDDS) definitions and required reporting for hospitals, and thus may not reflect the accurate representation of these categories (68). The classification of Hispanic ethnicity as a race in this dataset is a further limitation as those patients could be Native American, Black, White, Asian, or a mix of these. However, despite these limitations, our study is unique in that it included biracial and Native American groups that are typically not captured or analyzed alone. Additionally, the study population represents a national sample drawn from across the country and represents the patient population that physicians typically treat in the clinical setting. Although some groups were seemingly under-represented in the study population, we captured trends specific to each group by analyzing each racial group separately.

Conclusion

To our knowledge, this is the first study to identify *specific* multimorbidity patterns by race/ethnicity, stratifying by age and obesity. This is the first study to identify the prevalence of multimorbidity diseases across these cohorts. We found that multimorbidity was more prevalent in African Americans regardless of age or obesity status, and multimorbidity emerges at an earlier age within this group. Our findings also demonstrate that there are multimorbidity combinations unique to racial/ethnic groups,

particularly amongst the middle-aged cohort with or without obesity. Identifying the most common comorbidity clusters among specific patient populations by weight class provides important clues regarding the underlying mechanisms leading to disease co-occurrence in patients by race/ethnicity and age. This research supports the development of patient-centered care approaches by stratifying patients with obesity-associated multimorbidity and providing care specific to their unique clinical needs.

Declaration of Conflicting Interests
The authors declare no competing interests.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the National Institute of Aging award R15AG067232.

References

1. King, D. E., Xiang, J. & Pilkerton, C. S. Multimorbidity trends in United States Adults, 1988–2014. *J. Am. Board Fam. Med. JABFM.*31(4), 503–513 (2018).

2. Sakib, M. N., Shooshtari, S., St. John, P. & Menec, V. The prevalence of multimorbidity and associations with lifestyle factors among middle-aged Canadians: An analysis of Canadian Longitudinal Study on Aging data. *BMC Public Health* 19(1), 243 (2019).

3. Bao, J., Chua, K.-C., Prina, M. & Prince, M. Multimorbidity and care dependence in older adults: A longitudinal analysis of findings from the 10/66 study. *BMC Public Health* 19(1), 585 (2019).

4. Barnett, K. *et al.* Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross sectional study. *Lancet* 380(9836), 37–43 (2012).

5. Kronick, R., Bella, M., Gilmer, T. & Somers, S. *The Faces of MedicaidII: Recognizing the Care Needs of People with Multiple Chronic Conditions* (Center for Health Care Strategies Inc, 2021).

6. Quinones, A. R. *et al.* Racial/ethnic differences in multimorbidity development and chronic disease accumulation for middle-aged adults. *PLoS One.* 14(6), e0218462 (2019).

7. Muth, C. *et al.* The Ariadne principles: How to handle multimorbidity in primary care consultations. *BMC Med.* 8(12), 223 (2014).

8. Sturgiss, E. A., Elmitt, N., Haelser, E., van Weel, C. & Douglas, K. A. Role of the family doctor in the management of adults with obesity: A scoping review. *BMJ Open* 8(2), e019367 (2018).

9. Wei, M. Y., Kawachi, I., Okereke, O. I. & Mukamal, K. J. Diverse cumulative impact of chronic diseases on physical health-related quality of life: Implications for a measure of multimorbidity. *Am. J. Epidemiol.* 184(5), 357–365 (2016).

10. Low, L. L. *et al.* Epidemiologic characteristics of multimorbidity and sociodemographic factors associated with multimorbidity in a rapidly aging Asian country. *JAMA Netw. Open.* 2(11), e1915245 (2019).

11. Roso-Llorach, A. *et al.* Comparative analysis of methods for identifying multimorbidity patterns: A study of "real-world" data. *BMJ Open* 8(3), e018986 (2018).

12. U.S. Department of Health and Human Services. Multiple chronic conditions—a strategic framework: optimum health and quality of life for individuals with multiple chronic conditions. Wash DC US Dep Health Hum Serv. 2 (2010).

13. Whitson, H. E. *et al.* Identifying patterns of multimorbidity in older americans: Application of latent class analysis. *J. Am. Geriatr. Soc.* 64(8), 1668–1673 (2016).

14. Chong, J. L. & Matchar, D. B. Benefits of population segmentation analysis for developing health policy to promote patient-centered care. *Ann. Acad. Med. Singap.* 46(7), 287 (2017).

15. Olson, J. E., Takahashi, P. Y. & Sauver, J. M. S. Understanding the patterns of multimorbidity. *Mayo Clin Proc.* 93(7), 824–825 (2018).

16. Palmer, K. *et al.* Multimorbidity care model: Recommendations from the consensus meeting of the Joint Action on Chronic Diseases and Promoting Healthy Ageing across the Life Cycle (JA-CHRODIS). *Health Policy Amst. Neth.* 122(1), 4–11 (2018).

17. Madlock-Brown, C. & Reynolds, R. B. Identifying obesity-related multimorbidity combinations in the United States. *Clin. Obes.* 9(6), e12336 (2019).

18. Held, F. P. *et al.* Association rules analysis of comorbidity and multimorbidity: The concord health and aging in men project. *J. Gerontol. A Biol. Sci. Med. Sci.* 71(5), 625–631 (2016).

19. van den Bussche, H. *et al.* Which chronic diseases and disease combinations are specific to multimorbidity in the elderly? Results of a claims data based cross-sectional study in Germany. *BMC Public Health* 14(11), 101 (2011).

20. Kalgotra, P., Sharda, R. & Croff, J. M. Examining multimorbidity differences across racial groups: A network analysis of electronic medical records. *Sci. Rep.* 10(1), 13538 (2020).

21. Guisado-Clavero, M. *et al.* Multimorbidity patterns in the elderly: A prospective cohort study with cluster analysis. *BMC Geriatr.* 18(1), 16 (2018).

22. Canizares, M., Hogg-Johnson, S., Gignac, M. A. M., Glazier, R. H. & Badley, E. M. Increasing trajectories of multimorbidity over time: Birth cohort differences and the role of changes in obesity and income. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 73(7), 1303–1314 (2018).

23. Madlock-Brown, C. R., Reynolds, R. B. & Bailey, J. E. Increases in multimorbidity with weight class in the United States. *Clin. Obes.*11(3), e12436 (2021).

24. Dashputre, A. A., Surbhi, S., Podila, P. S. B., Shuvo, S. A. & Bailey, J. E. Can primary care access reduce health care utilization for patients with obesity-associated chronic conditions in medically underserved areas?. *J. Eval. Clin. Pract.* 26, 1689–1698 (2020).

25. de Carvalho, J. N., de Camargo, C. M. & de Souza, D. L. B. Lifestyle factors and high body mass index are associated with different multimorbidity clusters in the Brazilian population. *PLoS One* 13(11), e0207649 (2018).

26. Nguyen, T. N. *et al.* Social vulnerability in patients with multimorbidity: A cross-sectional analysis. *Int. J. Environ. Res. Public Health.* 16(7), 1244 (2019).

27. Leahy, S., Cassarino, M., O'Connell, M. D., Glynn, L. & Galvin, R. Dynapaenic obesity and its association with health outcomes in older adult populations: Protocol for a systematic review. *BMJ Open* 9(5), e027728 (2019).

28. Shadmi, E. Disparities in multiple chronic conditions within populations. *J. Comorbidity.* 3(2), 45–50 (2013).

29. Manuel, J. I. Racial/ethnic and gender disparities in health care use and access. *Health Serv. Res.* 53(3), 1407–1429 (2018).

30. Almirall, J. & Fortin, M. The coexistence of terms to describe the presence of multiple concurrent diseases. *J. Comorbidity.* 3, 4–9 (2013).

31. Wallace, E. *et al.* Managing patients with multimorbidity in primary care. *BMJ* 20(350), h176 (2015).

32. Goodman, R. A., Posner, S. F., Huang, E. S., Parekh, A. K. & Koh, H. K. Defining and measuring chronic conditions: Imperatives for research, policy, program, and practice. *Prev. Chronic Dis.* 25(10), E66 (2013).

33. Bernell, S. & Howard, S. W. Use your words carefully: What is a chronic disease?. *Front. Public Health.* 2(4), 159 (2016).

34. Simard, M., Rahme, E., Calfat, A. C. & Sirois, C. Multimorbidity measures from health administrative data using ICD system codes: A systematic review. *Pharmacoepidemiol. Drug Saf.* https:// doi. org/ 10. 1002/ pds. 5368 (2021).

35. Heaviest man ever | Guinness World Records. [cited 2021 Oct 13]. https:// www. guinn esswo rldre cords. com/ world- recor ds/ heaviest- man. Accessed 1 October 2021.

36. Terkawi, A. S. *et al.* General anesthesia for the heaviest man in the world. *Saudi J. Anaesth.* 8(Suppl 1), S101–S104 (2014).

37. Zhang, F. *et al.* A distributed frequent itemset mining algorithm using Spark for Big Data analytics. *Clust Comput.* 18(4), 1493–1501 (2015).

38. Qiu, Y., Lan, Y.-J., Xie, Q.-S. An improved algorithm of mining from FP-tree. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat No04EX826)*, vol. 3 1665–1670 (2004).

39. Venkataraman, S., Yang, Z., Liu, D., Liang, E., Falaki, H., Meng, X. *et al*. SparkR: Scaling R programs with spark. In *Proceedings of the 2016 International Conference on Management of Data*. 1099–1104 (ACM, 2016) [cited 2021 Jun 29]. https://doi.org/10.1145/2882903.2903740.

40. Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 26(4), 404–413 (1934).

41. Signorell, A. *et al*. DescTools: Tools for Descriptive Statistics. 2021 [cited 2021 Jun 29]. https://CRAN.R-project.org/package=DescTools. Accessed 1 June 2021.

42. Mair, P. & Wilcox, R. Robust statistical methods in R using the WRS2 package. *Behav. Res. Methods.* 52(2), 464–488 (2020).

43. Salisbury, C., Johnson, L., Purdy, S., Valderas, J. M. & Montgomery, A. A. Epidemiology and impact of multimorbidity in primary care: A retrospective cohort study. *Br. J. Gen. Pract. J. R. Coll. Gen. Pract.* 61(582), e12-21 (2011).

44. Lim, E., Gandhi, K., Davis, J. & Chen, J. J. Prevalence of chronic conditions and multimorbidities in a geographically defined geriatric population with diverse races and ethnicities. *J. Aging Health.* 30(3), 421–444 (2018).

45. Priest, N. *et al.* A systematic review of studies examining the relationship between reported racism and health and wellbeing for children and young people. *Soc. Sci. Med.* 95, 115–127 (2013).

46. Thomas Tobin, C. S. & Moody, M. D. Does early life racial discrimination explain a mental health paradox among black adults?. *J. Aging Health.* 3, 089826432098818 (2021).

47. Lee, H. & Schafer, M. Are positive childhood experiences linked to better cognitive functioning in later life?: Examining the role of life course pathways. *J. Aging Health.* 33(3–4), 217–226 (2021).

48. Boedtkjer, E. & Aalkjaer, C. Disturbed acid-base transport: An emerging cause of hypertension. *Front. Physiol.* 24(4), 388 (2013).

49. Dhondup, T. & Qian, Q. Electrolyte and acid-base disorders in chronic kidney disease and end-stage kidney failure. *Blood Purif.*43(1–3), 179–188 (2017).

50. Rylander, R., Remer, T., Berkemeyer, S. & Vormann, J. Acid–base status affects renal magnesium losses in healthy, elderly persons.*J. Nutr.* 136(9), 2374–2377 (2006).

51. Centers for Disease Control and Prevention (CDC). American Indians/Alaska Natives and Tobacco Use [Internet]. Centers for
Disease Control and Prevention. 2021 [cited 2021 Jun 29].
https://www.cdc.gov/tobacco/disparities/american-indians/index.htm. Accessed 1 June 2021.

52. Breslau, J. *et al.* Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychol. Med.* 36(1),57–68 (2006).

53. Bacon, S. L., Campbell, T. S., Arsenault, A. & Lavoie, K. L. The impact of mood and anxiety disorders on incident hypertension at one year. *Int. J. Hypertens.* 2014, 1–7 (2014).

54. van Reedt Dortland, A. K. B., Giltay, E. J., van Veen, T., Zitman, F. G. & Penninx, B. W. J. H. Longitudinal relationship of depressive and anxiety symptoms with dyslipidemia and abdominal obesity. *Psychosom. Med.* 75(1), 83–89 (2013).

55. Langan, J., Mercer, S. W. & Smith, D. J. Multimorbidity and mental health: Can psychiatry rise to the challenge?. *Br. J. Psychiatry.*202(6), 391–393 (2013).

56. Naylor, C., King's Fund (London E, Centre for Mental Health (London E. *Long-Term Conditions and Mental Health: The Cost of Comorbidities*. (King's Fund, 2012).

57. Zhang, Y., Misra, R. & Sambamoorthi, U. Prevalence of multimorbidity among Asian Indian, Chinese, and non-Hispanic white adults in the United States. *Int. J. Environ. Res. Public Health.* 17(9), 3336 (2020).

58. Teruya, S. A. & Bazargan-Hejazi, S. The Immigrant and hispanic paradoxes: A systematic review of their predictions and effects. *Hisp. J. Behav. Sci.* 35(4), 486–509 (2013).

59. Ruiz, J. M., Steffen, P. & Smith, T. B. Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *Am. J. Public Health.* 103(3), e52-60 (2013).

60. Min, J., Goodale, H., Xue, H., Brey, R. & Wang, Y. Racial-ethnic disparities in obesity and biological, behavioral, and sociocultural influences in the United States: A systematic review. *Adv. Nutr. Bethesda Md.* 12, 1137–1148 (2021).

61. Nguyen, H. *et al.* Factors associated with multimorbidity patterns in older adults in England: Findings from the english longitudinal study of aging (ELSA). *J. Aging Health.* 32(9), 1120–1132 (2020).

62. Harris, S. S. Vitamin D and African Americans. *J. Nutr.* 136(4), 1126–1129 (2006).

63. Jacobs, E. T. *et al.* Vitamin D insufficiency in southern Arizona. *Am. J. Clin. Nutr.* 87(3), 608–613 (2008).

64. Heidari, B., Shirvani, J. S., Firouzjahi, A., Heidari, P. & Hajian-Tilaki, K. O. Association between nonspecific skeletal pain and vitamin D deficiency: Vitamin D deficiency and skeletal pain. *Int. J. Rheum. Dis.* 13(4), 340–346 (2010).

65. Kim, R. P., Edelman, S. V. & Kim, D. D. Musculoskeletal complications of diabetes mellitus. *Clin. Diabetes.* 19(3), 132–135 (2001).

66. Boyd, C. M. & Fortin, M. Future of multimorbidity research: How should understanding of multimorbidity inform health system design?. *Public Health Rev.* 32(2), 451 (2010).

67. Sambamoorthi, U., Tan, X. & Deb, A. Multiple chronic conditions and healthcare costs among adults. *Expert Rev. Pharmacoecon. Outcomes Res.* 15(5), 823–832 (2015).

68. Strmic-Pawl, H. V., Jackson, B. A. & Garner, S. Race counts: Racial and ethnic data on the U.S. census and the implications for tracking inequality. *Sociol. Race Ethn.* 4(1), 1–13 (2018).

Contributions

MA: analysis, methods, data collection, writing. BJ: analysis, writing, literature review. DI: writing, literature review. RR: writing, literature review, critical review. CMB: conception, design, supervision, critical review.

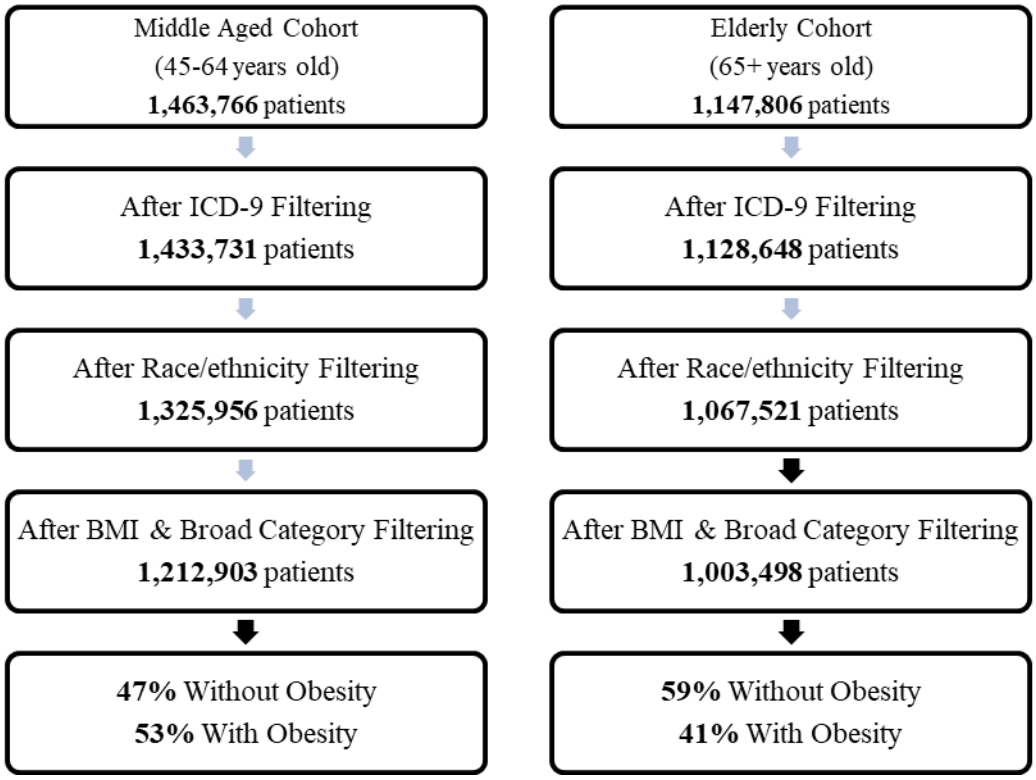
Ethics Declaration

Competing interests

The authors declare no competing interests.

Supplementary Material

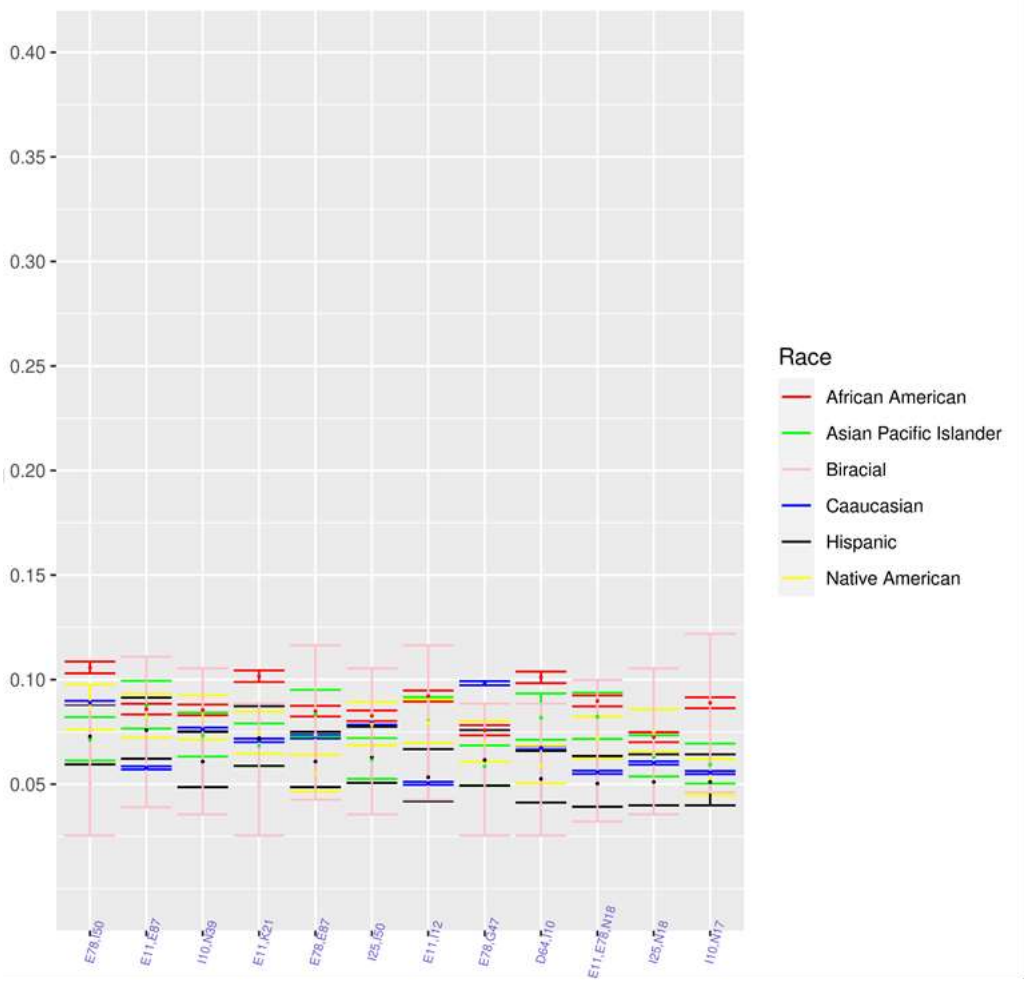
Supplementary Table 1. Patient Population.



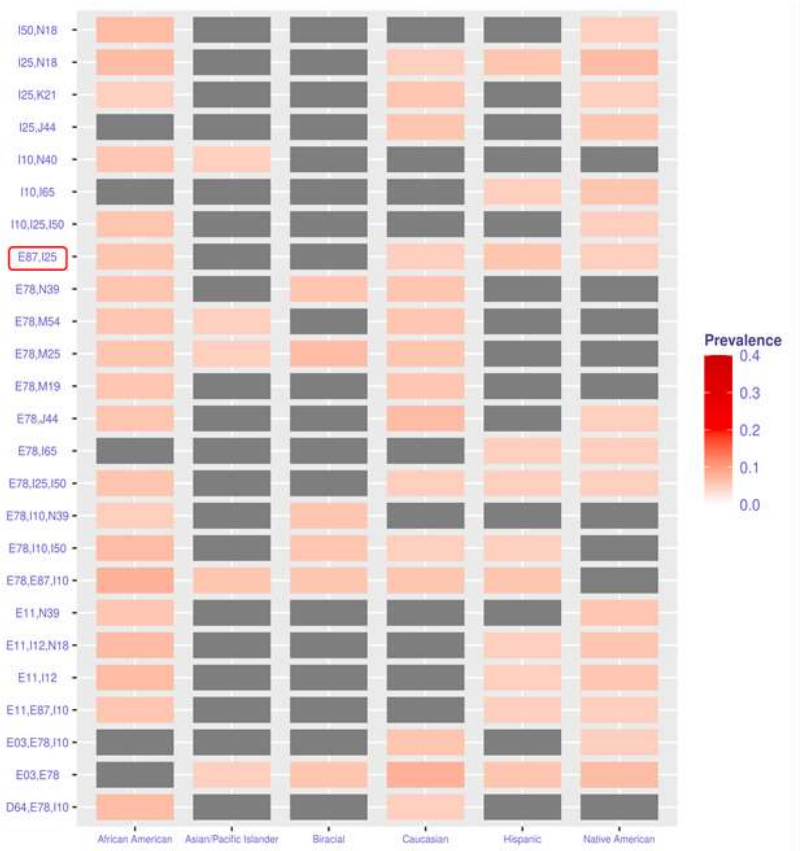
Supplementary Table 2. ICD-10-CM Diagnoses included in the study.

#	Diagnosis
1	**D63:** Anemia in chronic diseases classified elsewhere
2	**D64:** Other anemias
3	**E03:** Other hypothyroidism
4	**E11:** Diabetes
5	**E55:** Vitamin D deficiency
6	**E78:** Lipidemia
7	**E86:**Volume depletion
8	**E87:** Other disorders of fluid, electrolyte and acid-base balance
9	**F03:** Unspecified dementia
10	**F10:** Alcohol related disorders
11	**F17:** Nicotine dependence
12	**F32:** Major depressive disorder, single episode

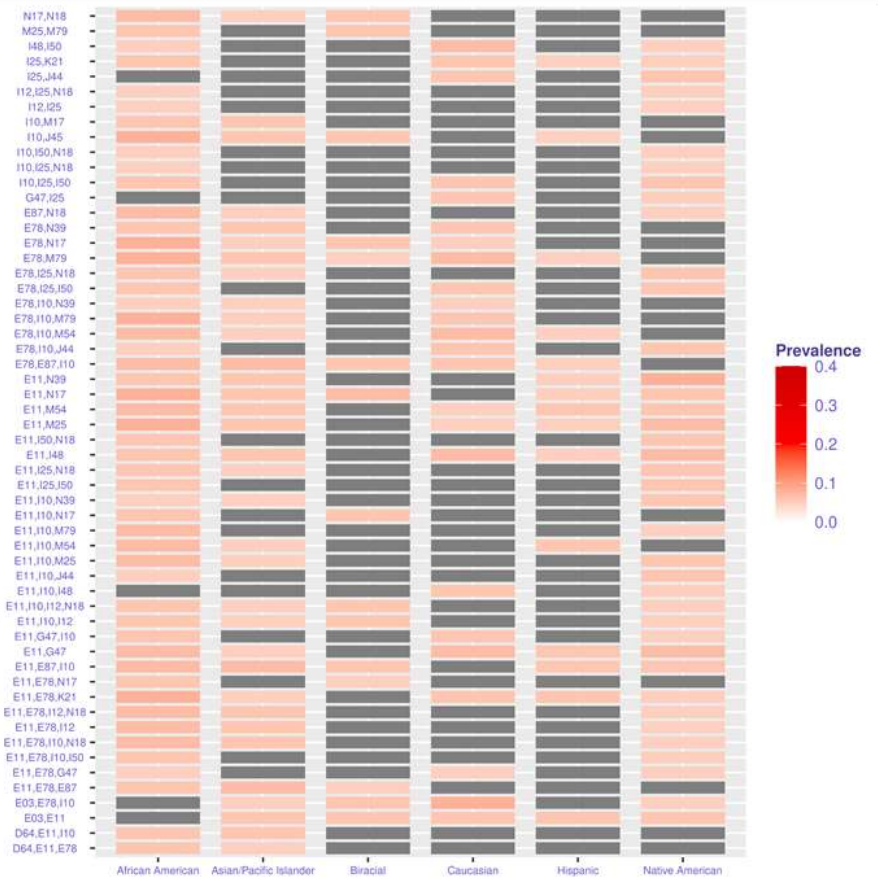
13	**F41:** Anxiety Disorders
14	**G47:** Sleep disorders are represented
15	**G89:** Pain, not elsewhere classified
16	**I10:** Hypertension
17	**I12:** Hypertensive chronic kidney disease
18	**I25:** Heart disease
19	**I48:** Atrial fibrillation and flutter are prevalent
20	**I50:** Heart failure
21	**I65:** Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
22	**I73:** Other peripheral vascular diseases
23	**J44:** Other chronic obstructive pulmonary disease
24	**J45:** Asthma
25	**K21:** GERD
26	**K57:** Diverticular disease of intestine
27	**K59:** Other functional intestinal disorders
28	**M17:** Osteoarthritis of knee
29	**M19:** Other and unspecified osteoarthritis
30	**M25:** Other Joint Disorders
31	**M54:** Dorsalgia
32	**M79:** Other and unspecified soft tissue disorders, not elsewhere classified
33	**M81:** Osteoporosis without current pathological fracture
34	**N17:** Acute kidney failure
35	**N18:** Chronic kidney disease (CKD)
36	**N28:** Other disorders of kidney and ureter, not elsewhere classified
37	**N39:** Other disorders of urinary system
38	**N40:** Benign prostatic hyperplasia



Supplementary Figure 1: *Confidence Interval Overlap for Elderly Patients with Obesity Shared by All Races & Average Prevalence <0.08.*



Supplementary Figure 2: *Multimorbidities Shared by Some Races for Elderly Patients without Obesity & Average Prevalence < 0.065. Red box indicates non-significant g-test.*



Supplementary Figure 3: *Multimorbidities Shared by Some Races for Elderly Patients with Obesity & Average Prevalence <0.065.*

APPENDIX B. CHAPTER 3 ARTICLE

NOTE: Navigation with Adobe Acrobat Reader or Adobe Acrobat Professional: To return to the last viewed page, use key commands Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For "Next view," use Alt/Ctrl+Right Arrow on PC or Command+Right Arrow on Mac. See **Preface** for further details. If needed, use this link to return to **Chapter 3** after navigating within this appendix.

Introduction

This appendix includes the work done for Aim 2: Identify differences in multimorbidity total charges across races adjusted for age and obesity using a comorbidity index. Article reused from prepared manuscript with authors' permission. Alshakhs, M.; Goedecks, P., Bailey, J., Madlock-Brown, C. Racial Differences in Healthcare Expenditures for Prevalent Multimorbidity Combinations in the U.S. (2023). It is included in this dissertation with permission from co-authors.

Article

Racial Differences in Healthcare Expenditures for Prevalent Multimorbidity Combinations in the U.S.

Manal J. Alshakhs, MS[1]
Patricia J. Goedecke, MS[2]
Jim Bailey, MD[3,4]
Charisse Madlock-Brown, PhD, MLS[1,5]

[1]Health Outcomes and Policy Research Program, University of Tennessee Health Science Center, Memphis, Tennessee, USA
[2]Doctoral student, Bredesen Center, University of Tennessee Knoxville
[3]Tennessee Population Health Consortium, University of Tennessee Health Science Center, Memphis, TN, United States
[4]Departments of Medicine and Preventive Medicine, College of Medicine, University of Tennessee Health Science Center, Memphis, TN, United States
[5]Center for Health System Improvement, University of Tennessee Health Science Center, Memphis, Tennessee, USA

Abstract

Background. This research project aims to model total charges for the most prevalent multimorbidity combinations in the U.S. and assess model accuracy across race/ethnicity and multimorbidity composition.

Methods. This cross-sectional study used Cerner HealthFacts data for 2016-2017 to identify most prevalent multimorbidity combinations across races/ethnicities using 38 major diagnostic categories and estimate total charges for each combination. Regression analysis was employed to model total charges for middle-aged (MC) and elderly cohorts (EC) for the most prevalent multimorbidity combinations across races/ethnicities and model accuracy was assessed using residual analysis.

Results. Mean ages were 55 (MC, N=333,094) and 75 (EC, N=327,260) respectively. Actual total charges were highest for African Americans (Means $78,544 [MC], $176,274 [EC]) and lowest for Hispanics (Means $29,597 [MC], $66,911 [EC]) and total charge models demonstrate that African American race was strongly predictive of higher costs ($p < 0.05$ [MC]; $p < 0.05$ [EC]). Although our general total charge models performed well across various races, there was substantial variation by race in model accuracy when multimorbidity grouping was considered. For example, costs were substantially overestimated for elderly Caucasians with multimorbidity combinations including heart disease. Additionally, accuracy varied by age/obesity status. For instance, in the elderly with obesity cohort, model estimates for Hispanic patients were highly underestimated for most multimorbidity combinations compared to other age/obesity status groupings.

Conclusions. This study demonstrates that actual total charges are highest for African Americans, and total charge models showed that African American race are strongly predictive of higher costs. The study shows that total charge models perform well across various races, but there is substantial variation by race in accuracy of models predicting total healthcare costs when multimorbidity grouping is considered. Since models of healthcare expenditures across races/ethnicities are improved through inclusion of prevalent multimorbidity combinations beyond performance seen using a comorbidity index alone, future cost modeling efforts are likely to benefit from inclusion of prevalent multimorbidity combinations.

Key words: multimorbidity, healthcare costs, race/ethnicity stratification

Background

The scarcity of economic models in the field of multimorbidity research represents a serious challenge.[1] A significant limitation of current models is that most do not consider the varying costs of different disease combinations.[2] A better understanding of the burden of multimorbidity through cost assessment for various multimorbidity combinations will assist in targeting highest cost patients for intensive interventions.[1] Most high healthcare utilizers have at least two chronic conditions.[3] Addressing the economic burden of multimorbidity is crucial to developing effective strategies for managing care.

Factoring in multimorbidity has been shown to explain these expenditures better than models based on population characteristics (size and demographics) alone.[4] For example, in the case of diabetes, different comorbidities have a varying impact on cost.[5] Previous studies evaluating the impact of specific multimorbidity combinations on expenditures thus far have focused on a few diseases.[2] The most prevalent multimorbidities in the U.S. represent a broad spectrum of diseases.[6] Effective care planning and resource management requires accurately projecting patient costs for these disease combinations.[2,7]

Modeling the associations of most prevalent multimorbidity combinations with healthcare expenditures is essential to further aging research because the majority of the elderly population have two or more chronic conditions and account for 47% of Medicare spending.[8] Interventions aimed at slowing the aging process need to target patients with multiple diseases to be effective.[9] Mercer *et al.* (2016) found that multimorbidity-focused interventions are cost-effective for this patient population.[10]

Better modeling of expenditures is essential for improving the health of racial and ethnic minorities. Clay *et al.* (2018) found that amongst African American men, comorbidity clusters are associated with poor outcomes, including poor health-related quality of life, disability, and higher mortality rate. As these authors suggest, better modeling of expenditures will be essential for improving the health of racial and ethnic minorities.[11] Multimorbidities exacerbate health inequalities as underserved populations are at greater risk for multimorbidity, increasing their disease burden.[12] Despite a clear need to better understand health disparities, research shows that even robust methods can be susceptible to bias. Predictive models derived from primarily homogenous populations may be poorly generalizable and can exacerbate racial/ethnic disparities.[13] Cost estimates of multimorbidity must address model racial/ethnic bias. To date, no large-scale study of the expenditures associated with common multimorbidity combinations has assessed the accuracy of model predictions across races and obesity status.

This research compares total healthcare expenditures for the most prevalent multimorbidity combinations across racial/ethnic groups. We also aim to determine if multimorbidity expenditure models have similar accuracy across racial/ethnic groups after adjustment for potential confounding factors. In addition, the study aims to assess for possible differences in total charges for middle age versus elderly patients as the incidence of chronic disease rises exponentially with age.[14] This study is among the first to model total charges associated with the most prevalent multimorbidity combinations by race/ethnicity. Our previous work identified the most prevalent multimorbidity combinations by

race/ethnicity, serving as the foundation for this current research.[6] Our primary objectives are to identify the expected total charges associated with the most prevalent multimorbidity combinations by race/ethnicity. Additionally, we sought to assess differences in expenditures for these multimorbidity combinations and assess differences in model accuracy by race/ethnicity.

Methods

Research Design. This cross-sectional study employed de-identified data for 2016-2017 from the Cerner HealthFacts® data warehouse. The dataset includes over 490 million patient encounters for over 70 million patients treated at hospitals and clinics at 792 non-affiliated healthcare systems throughout the United States between 2001 and 2017. Variable categories include encounter type, medical history, diagnoses, labs, prescriptions, patient demographics, clinic type, and procedures. The inclusion criteria for patients were: 1) Age 45+, 2) Body Mass Index (BMI) value present and between 18.5 and 206, 3) assigned race category, 4) assigned gender (i.e., male or female), 5) patient encounters with total charges greater than $0, and 6) an encounter with an International Classification of Diseases-10th Version-Clinical Modification (ICD-10–CM) diagnosis code that is included in one or more of the 38 broad diagnoses that make up the most prevalent multimorbidities in the U.S. as demonstrated by our previous research (See Appendix Table 1).[6] Using a prevalence-based approach for assessing multimorbidity validated through many previous studies,[15-17] we defined multimorbidity as the presence of two or more ICD-10–CM diagnosis codes in an individual during the two-year (i.e., 2016–2017) study period. We aggregated ICD-10-CM sub-classifications of diseases into a broad category for all 38 diagnoses. For example, I11.9 (Hypertensive heart disease without heart failure) would fall under the broader parental category I11 (Hypertensive heart disease). The two-year period was employed to assess multimorbidities to maximize the probability of identifying all major prevalent multimorbidities experienced by individuals during the study period. Since diseases might not be diagnosed at the same visit or within the same year, this longer period allows us to capture more data than would a single year.

Ethical Considerations. The data were de-identified and excluded the 16 identifiable variables that necessitate Internal Review Board (IRB) approval for access. Because the study only employed de-identified data, the study was considered not human subjects' research. Per the National Institutes of Health Office of Human Subjects Research policy, the University of Tennessee Health Science Center (UTHSC) Institutional Review Board (IRB) determined that the research was exempt. We performed this research following all other relevant research requirements.

Independent Variables. Demographic, multimorbidity, and healthcare utilization variables were the primary independent variables. Demographic variables included race, age, gender, BMI, payer information, and rural or urban status. BMI was treated as a dichotomous variable, classifying patients with obesity ($30 \leq BMI < 206$) and without obesity ($18.5 \leq BMI < 30$). The highest recorded BMI value was between 206 and 224, so we considered it valid if the BMI value was 206 or less.[18,19] When assessing the financial burden across races in middle age and older adults, controlling for factors impacting disease severity and socioeconomic issues affecting cost is critical.[20] Therefore, we assessed payer status, rurality, length of stay, and the Elixhauser Comorbidity Index (ECI) score. Hospital information and healthcare usage variables included the number of inpatient visits, number of outpatient visits, number of emergency visits, teaching hospital status, care-type status, and total hospital admission days, if any. Because ethnicity is not a separate variable in the Cerner HealthFacts database, Hispanic is listed as a racial category. Other racial categories included: Caucasian, African American, Biracial, Asian/Pacific Islander, and

Native American. Patients were stratified into two cohorts (i.e., middle-aged [MC; age 45-64] and elderly [EC; age 65+]) according to their age at the beginning of the study. Only patients for whom an assigned gender (i.e., male and female) was listed were included in the study. **Appendix Table 2** clarifies the remaining variables.

Outcome Variable. Our primary outcome variable was the sum of total charges for all encounters over the two-year study period for each patient. Healthcare utilization information included the total charges for each encounter. We categorized patient encounters into one of three categories: inpatient, outpatient, or emergency visit. We chose our outcome variable to be total charges since it is the amount that reflects the expense of the service received before any discounts or negotiations.–Arora *et al.* (2015) described the challenge of answering the question "how much does healthcare cost?" and divided healthcare expenditures into three categories: price or charge, cost, and reimbursement.[21] Price or charge is defined as the amount billed by a provider for a healthcare service. Hospitals in the U.S. use a price list called chargemaster that includes a list of all billable services before any discounts or negotiation to arrive at the price charged, which varies across hospitals.[22,23] The definition of cost varies with perspective. For the provider, the cost is simply the expense incurred to deliver healthcare services to the patient, for the payer, it is the amount that they will pay providers for these services, and for the patient, it is the amount they pay out-of-pocket for healthcare services rendered. Finally, reimbursement is defined as the amount paid a provider by a third party (the payer) for the services rendered to the patient. Due to different agreements and negotiations between hospital providers and payers, cost and reimbursement can vary across patients receiving the same service from the same hospital.[21,24]

Missing Data. Due to their minimal numbers, we deleted hospitals with no census division or rural/urban status information. We imputed hospitals with teaching facility information missing by adding the most prevalent category.[25] We excluded encounters with $0 listed for total charges from the study. According to the Cerner HealthFacts® database data dictionary, total charges of $0 indicate that the administrative staff did not enter the billing information into the database. We compared demographics for sources with missing cost data and those with cost data present; the demographics were not statistically different. For ease of interpretation, the patient record was removed from the study if a patient was treated in two different census divisions or if the patient was treated in a rural and an urban hospital.

Statistical Analysis. We examined the distribution of our outcome variable, total charges over the two years. We checked for skewness and outliers. Having so many variables, we also tested for multicollinearity, a linear relationship between two or more variables.[26] We used a generalized variance inflation factor (GVIF) analysis to identify variables with high multicollinearity, which is appropriate for a mix of categorical and numerical variables.[27,28] We removed the variable with the highest GVIF^(1/2Df) score using the car R package.[29] We repeated this process until no variable had a score above two, a conservative threshold for considering multicollinearity.[28]

We used regression analysis to compare the total charges of the most prevalent multimorbidity combinations by race/ethnicity. A generalized linear model (GLM) with gamma distribution and log link function was applied to estimate the total charges based on the morbidity variables.[30,31] ECI rank was categorized into three categories based on quantile range: low, medium, and high, indicating comorbidity severity. We ran a 3-way ANOVA test on the model residuals to determine whether there was an interaction effect between BMI and race, as a combined effect, and ECI ranks on total charges (the outcome variable).

Results

Demographics. In this study, most patients in both age cohorts were female. **Tables 1 & 2** show the breakdown of demographics by race for the middle-aged and elderly cohorts, respectively. The percentages were calculated relative to the whole patient population. The average age for the middle-aged cohort (333,094 patients) was 55 years and for the elderly cohort (327,260 patients) was 75 years.

Outcomes. The breakdown of visit type, mean emergency room visits, mean ECI score, mean admission days, and mean charges for the middle-aged and elderly cohorts, are shown in **Table 3**. The Cerner HealthFacts® database included data from 1,500,580 middle-aged patients and 1,213,069 elderly patients for the period 2016-2017. We excluded some of the ICD-9-CM diagnosis codes and removed patient encounters with $0 total charges for a total of 647,801 patients remaining in the middle-aged cohort and 534,534 patients remaining in the elderly cohort After excluding patients based on our remaining inclusion/exclusion criteria, 333,094 patients were middle age, and 327,260 remained. A complete breakdown of our exclusion/inclusion criteria on the patient population is displayed in **Appendix Figure 1**.

Table 1. Demographics of the Middle-Aged Cohort[4,5]

Race	Prevalence n (%)	Gender		Payer Info			Area Status	
		Female n (%)	Male n (%)	Medicaid/Medicare/ Title V n (%)	Other n (%)	Unknown n (%)	Urban n (%)	Rural n (%)
African American	44,595 (13)	26,260 (8)	18,335 (6)	12,145 (4)	28,597 (9)	3,853 (1)	40,203 (12)	4,392 (1)
Asian/Pacific Islander	3,976 (1)	2,341 (1)	1635 (<1)	504 (<1)	2,864(1)	608(<1)	2,925 (1)	1051 (<1)
Biracial	346 (<1)	189 (<1)	157 (<1)	43 (<1)	257 (<1)	46 (<1)	265 (<1)	81 (<1)
Caucasian	278,676 (84)	150,101 (45)	128,575 (39)	55,010 (17)	209,433 (63)	14,233 (4)	227,618 (68)	51,058 (15)
Hispanic	485 (<1)	274 (<1)	211 (<1)	78 (<1)	369 (<1)	38 (<1)	422 (<1)	63 (<1)
Native American	5,016 (2)	2,784 (1)	2,232 (1)	1,393 (<1)	3,524 (1)	99 (<1)	2735 (1)	2,281 (1)
Total	333,094 (100)	181949 (55)	151145 (45)	69173 (21)	245044 (74)	18877 (6)	274168 (82)	58926 (18)

[4] Middle-aged defined as 45 to 64 years of age

[5] Percentages are related to total population

Table 2. Demographics of the Elderly Cohort[6]

Race	Prevalence (n%)	Gender		Payer Info			Area Status	
		Female (n%)	Male (n%)	Medicaid/Medicare/Title V (n%)	Other (n%)	Unknown (n%)	Urban (n%)	Rural (n%)
African American	23,529 (7)	14,381 (4)	9,148 (3)	17,655 (5)	4,004 (1)	1,870 (1)	21,174 (6)	2,355 (1)
Asian/Pacific Islander	3,729 (1)	2,309 (1)	1,420 (<1)	2,426 (1)	635 (<1)	668 (<1)	2,277 (1)	1,452 (<1)
Biracial	1,601 (<1)	84 (<1)	76 (<1)	54 (<1)	61 (<1)	45 (<1)	94 (<1)	66 (<1)
Caucasian	297,299 (91)	164,042 (50)	133,257 (41)	241,109 (74)	47,966 (15)	8,224 (3)	243,273 (74)	54,026 (17)
Hispanic	220 (<1)	111 (<1)	109 (<1)	141 (<1)	68 (<1)	11 (<1)	187 (<1)	33 (<1)
Native American	2,323 (1)	1,352 (<1)	971 (<1)	1,779 (1)	513 (<1)	31 (<1)	1,265 (<1)	1,058 (<1)
Total	327,260 (100)	182,279 (55)	144,981 (45)	263,164 (81)	53,247 (16)	10,849 (3)	268,270 (81)	58,990 (19)

Due to the skewness of the outcome variable (mean total charges), we performed an outlier test and used the interquartile method to eliminate outliers. After testing for collinearity, we removed the *total number of morbidities* variable from the analysis, as it was considered an aliased coefficient in the model, meaning that this particular variable was equivalent to one or more variable(s). We determined the unadjusted and adjusted models' residuals for the middle-ages and elderly cohorts using a generalized linear model with Gamma distribution and log link function (**Appendix Figure 1**) and assessed model performance by inspecting the residuals' quantile-quantile (Q-Q) plots in R. Due to the skewness of the dependent variable, total charges, these models did not fit the data well. To obtain a better-fitting model, we used a log transformation and an exponential transformation and inspected the residuals' Q-Q plots to measure the model performance (**Appendix Figure 2**). We selected the exponential model as optimal for this dataset, since it exhibited the least sum of square error (SSE) in both cohorts.[32]

[6] Elderly defined as 65+ years of age

Table 3. Outcomes of the middle-aged and elderly cohorts

Race	The Middle-Aged Cohort				The Elderly Cohort			
	Mean E.R. Visits	Mean ECI Score	Mean Hospital Admission Days	Mean Charges	Mean E.R. Visits	Mean ECI Score	Mean Hospital Admission Days	Mean Charges
African American	1	2	1	$78,544	1	6	2	$176,274
Asian/Pacific Islander	1	2	0	$54,410	1	4	1	$167,949
Biracial	0	2	0	$47,238	1	4	1	$140,628
Caucasian	1	1	0	$55,704	1	4	1	$146,224
Hispanic	0	1	0	$29,597	0	3	1	$66,911
Native American	1	2	0	$50,496	1	4	1	$111,522

The model. The healthcare total charges model estimates across race in the middle-aged and elderly cohorts with adjusted R-squared values of 0.3906 and 0.4695, respectively, are shown in **Tables 5 and 6** and **Appendix Tables 3 and 4**. **Tables 5 and 6** show the model estimates for key demographic and patient hospital utilization factors. African American was selected as the index race. For the middle-aged cohort (**Appendix Table 3**), all 38 shared morbidities were significant predictors for healthcare charges across races. The Asian/Pacific Islander race, Biracial racial classification, and hospital teaching status were not significant predictors (**Appendix Table 3**). The Caucasian, Hispanic, and Native American races had negative total charges estimates. In the elderly cohort, the Biracial racial classification and living in the South Atlantic region were not significant predictors for healthcare charges (**Table 6** & **Appendix Table 4**). The Asian/Pacific Islander and the Caucasian races had positive total charges estimates, while the Hispanic and the Native American races had negative estimates. **Appendix Tables 3 and 4** include model estimates for the 38 diagnoses that make up the most prevalent multimorbidities across races/ethnicities in the U.S. For each model, all diagnosis estimates were significant. For the middle-age cohort, all diagnoses estimates were positive with the exception of heart failure, vitamin D deficiency, and chronic kidney disease, which were all slightly negative. For the elderly cohort, all were positive except vitamin D deficiency and chronic kidney disease. Estimates for hospital-related variables are also displayed in the **Appendix Tables 3 and 4.**

Table 5. Middle-Aged Cohort Total Charge Model Estimates[7]

Demographics (Variable)	Estimate	P-value	Significance
Age	-0.001	<0.001	***
Race/Asian Pacific Islander	0.006	0.188	-
Race/Caucasian	-0.006	<0.001	***
Race /Native American	-0.015	<0.001	***
Race /Biracial	-0.018	0.175	-
Race/Hispanic	-0.154	<0.001	***
Male	-0.010	<0.001	***
BMI	0.000	<0.001	***
Length of Stay	0.014	<0.001	***
Payer/Unknown	0.035	<0.001	***
Payer/Other	-0.007	<0.001	***
ECI	0.008	<0.001	***
Emergency visits	0.028	<0.001	***
Outpatient visits	0.003	<0.001	***

Table 6. Elderly Cohort Total Charge Model Estimates

Demographics (Variable)	Estimate	P-value	Significance
Age	0.001	<0.001	***
Race/Asian Pacific Islander	0.123	<0.001	***
Race/Caucasian	0.015	<0.001	***
Race /Biracial	-0.028	0.550	-
Race /Native American	-0.046	0.001	***
Race/Hispanic	-0.369	<0.001	***
Male	-0.007	0.005	**
BMI	-0.001	<0.001	***
Length of Stay	0.025	<0.001	***
Payer/Unknown	0.042	<0.001	***
Payer/Other	-0.075	<0.001	***
ECI	0.018	<0.001	***
Emergency visits	0.024	<0.001	***
Outpatient visits	-0.008	<0.001	***

The overall mean of the absolute value of the model residuals and the standard deviations for the middle-aged and elderly cohorts are shown in **Table 7**. The model best predicted the total charges for the Hispanic race and was least accurate for the African American race. This table also displays extreme standard deviation values for the model's residuals.

[7] Indexed categorical variables for both cohorts were: African American race, Female gender, Medicaid/Medicare/Title V payer type, and Inpatient visit type

All standard deviations were greater than the mean and some races exhibited remarkably high standard deviations.

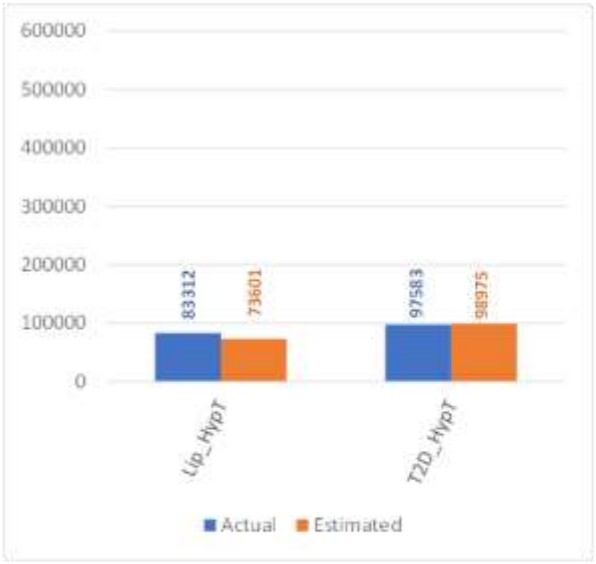
Table 7. Mean Model Residuals for the Middle-Aged and Elderly Cohorts

Race	Middle-Aged Cohort		Elderly Cohort	
	Mean Residuals	Standard Deviation	Mean Residuals	Standard Deviation
African American	175,514	19,636,829	157,118	646,311
Asian/Pacific Islander	45,864	108,516	120,527	271,469
Biracial	40,425	77,686	109,464	180,946
Caucasian	47,112	515,279	205,310	20,528,300
Hispanic	27,569	65,028	56,234	134,076
Native American	70,938	1,673,968	83,095	161,816

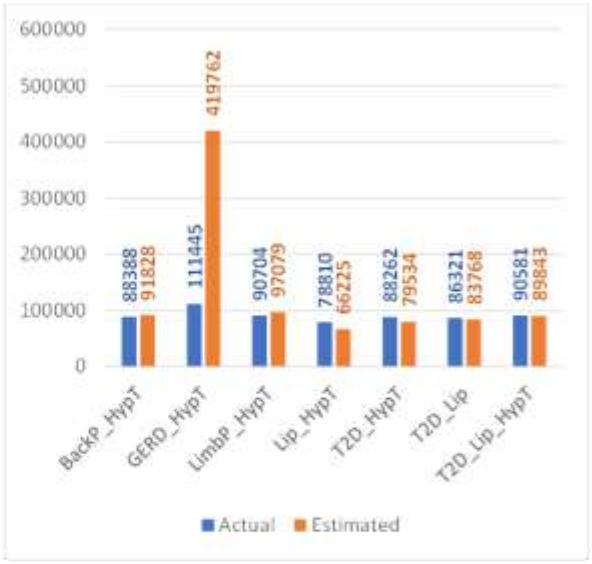
The actual vs. the estimated mean total charges over the study period (2016-2017) for the most prevalent multimorbidity combinations by race in the middle-aged and elderly cohorts, respectively, are shown in **Figure 1**. The variance between the actual and estimated mean total charges for the *hypertension + GERD* multimorbidity combination in the middle-aged cohort with obesity was almost double. In the elderly cohort with obesity, *hypertension + heart disease*, *lipidemia + hypertension + heart disease*, and *lipidemia + heart disease* multimorbidities exhibited the highest variance between the actual and estimated mean total charges with values that were also almost double. In general, the mean total charges and the variance between actual and estimated mean total charges were higher in the elderly cohorts than in the middle-aged cohorts.

The mean residuals by race for each of the most prevalent multimorbidity combinations in both cohorts with and without obesity are shown in **Figure 2**. Due to many prevalent multimorbidity combinations in the elderly cohorts, we examined the residuals by race for only the top seven most prevalent multimorbidity combinations. The mean model residuals for the shared multimorbidity patterns by race in the middle-aged cohort without obesity are shown in **Figure 2a**. The model overestimated the total charges for both shared multimorbidities for the African American race and one shared multimorbidity for the Hispanic race and it underestimated the total charges for the Caucasian race. The best estimates were for the Native American race. In contrast, the mean model residuals for the shared multimorbidity patterns by race in the middle-aged cohort with obesity (**Figure 2b**) indicated that the model highly overestimated the total charges for the GERD + hypertension multimorbidity pattern in the African American race. The model underestimated all of the total charges for the multimorbidity patterns for the Hispanic race, while the estimates for the Native American race fluctuated between over- and underestimation.

The mean model residuals for the shared multimorbidity patterns by race in the elderly cohort without obesity are shown in **Figure 2c**. The total charge estimates for the African American race were overestimated or highly overestimated, while those for almost all of the Asian/Pacific Islander race were highly underestimated. The model also overestimated two patterns for the Hispanic race and highly underestimated the remainder. In contrast, the mean model residuals for the shared multimorbidity patterns by race in the elderly cohort with obesity (**Figure 2d**) indicated that the model underestimated all of the total charge estimates for the Asian/Pacific Islander race. The Hispanic race exhibited the most patterns that were highly underestimated, while those for the Caucasian race had one triad pattern that was significantly overestimated (-316995). The model estimated two multimorbidity patterns for the African American race better than others, but overestimated the remaining patterns. The remaining patterns were either highly overestimated or underestimated.

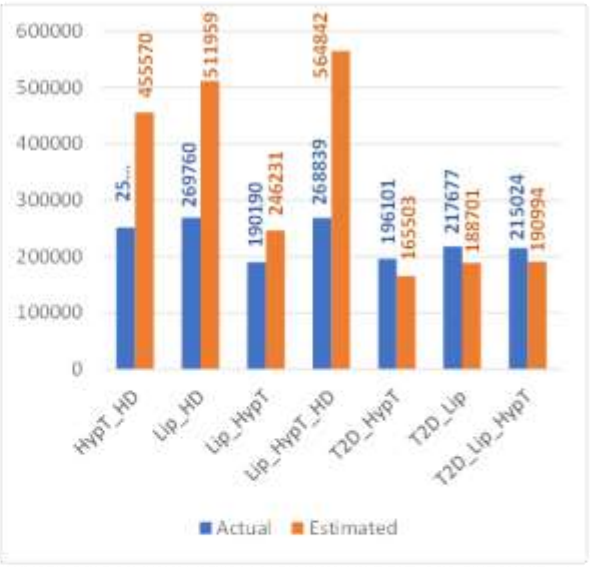


a. *Actual vs. Estimated Mean Total Charges for the Most Prevalent Multimorbidity Combinations by Race in the Middle-Aged Cohort Without Obesity*



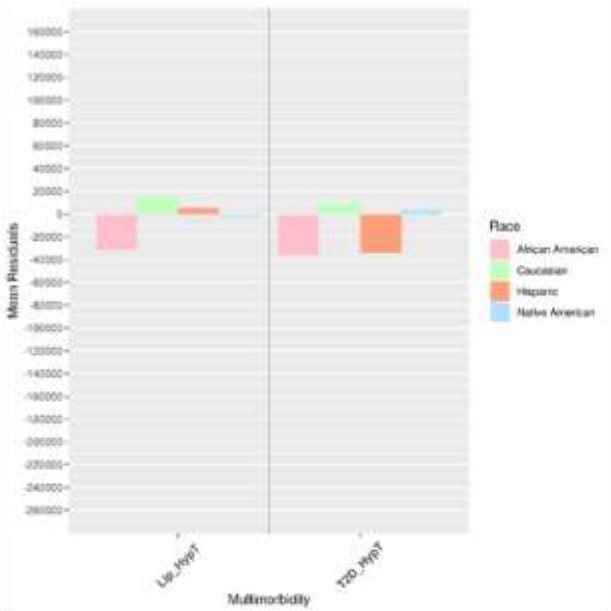
b. *Actual vs. Estimated Mean Total Charges for the Most Prevalent Multimorbidity Combinations by Race in the Middle-Aged Cohort, With Obesity*

c. *Actual vs. Estimated Mean Total Charges for the Most Prevalent Multimorbidity Combinations by Race in the Elderly Cohort Without Obesity*

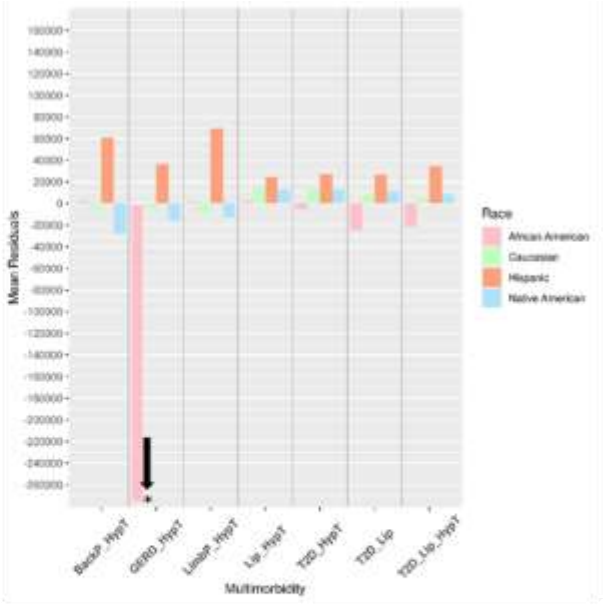


d. *Actual vs. Estimated Mean Total Charges for the Most Prevalent Multimorbidity Combinations by Race in the Elderly Cohort with Obesity*

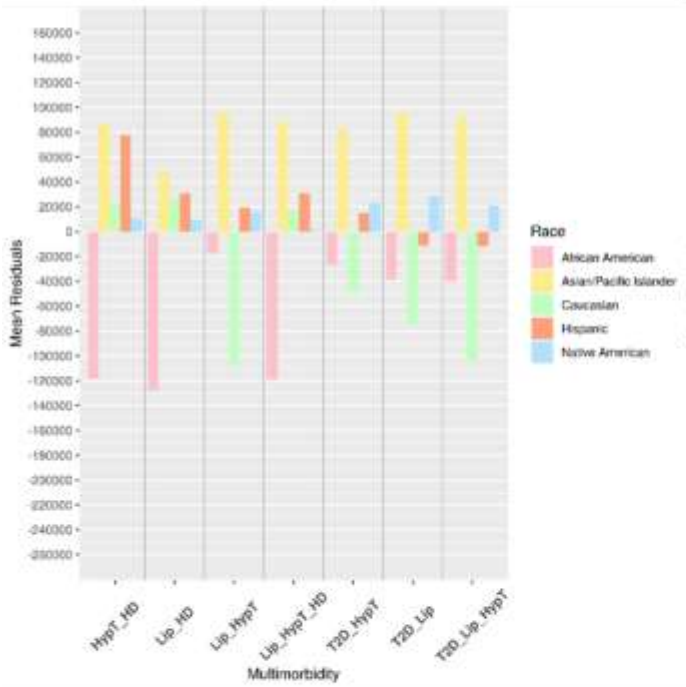
Figure 1. Actual and Estimated Mean Total Charges for the Most Prevalent Multimorbidity Combinations by Race in the Middle-Aged and Elderly Cohorts, With and Without Obesity. Abbreviations: BackP, Severe Back Pain; **GERD**, Gastroesophageal reflux disease; **HD**, Heart Disease; **HypT**, Hypertension; **Lip**, Lipidemia; **LimbP**, Pain in limb, hand, foot, fingers, and toes; **T2D**, Type 2 Diabetes Mellitus.



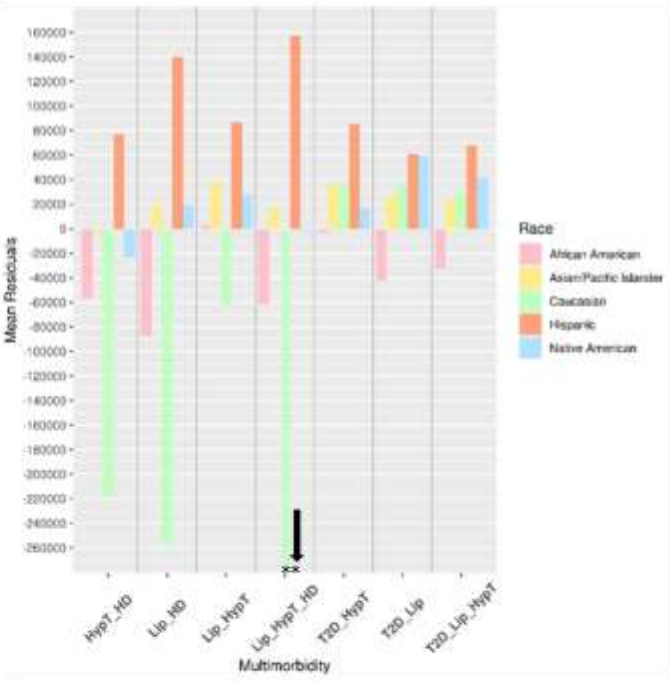
a. *Mean Model Residuals for Shared Multimorbidity Patterns by Race in the Middle-Aged Cohort Without Obesity.*



b. *Mean Model Residuals for Shared Multimorbidity Patterns by Race in the Middle-Aged Cohort With Obesity*



c. *Mean Model Residuals for Shared Multimorbidity Patterns by Race in the Elderly Cohort without Obesity (top 7 patterns with significant model results)*



d. *Mean Model Residuals for Shared Multimorbidity Patterns by Race in the Elderly Cohort with Obesity (top 7 patterns with significant model results)*

Figure 2. Mean Model Residuals for Shared Multimorbidity Patterns by Race in the elderly and middle-aged cohorts with and without obesity. * = -1749059 residual value , ** = -316995 residual value (these values were too large to display in the figure). **Abbreviations: BackP,** Severe back pain; **CKD,** Chronic kidney disease; **GERD,** Gastroesophageal reflux disease; **HD,** Heart disease; **HypT,** Hypertension; **Lip,** Lipidemia; **LimbP,** Pain in limb, hand, foot, fingers, and toes; **OJD,** Other joint disorder; **T2D,** Type 2 diabetes mellitus.

The model estimated the mean total charges for the Hypertension + GERD multimorbidity pattern as much higher than the actual charges. The mean model residual for the same cohort was highly overestimated for the African American race. The model also estimated

the mean total charges for the Lipidemia + hypertension + heart disease triad multimorbidity pattern as much higher than the actual charges. The mean model residual for the same cohort was highly overestimated for the Caucasian race. The multimorbidities with heart disease also showed extreme differences in actual vs. estimated mean total charges and an extreme over or underestimation for the mean model residuals of certain races.

The variability of mean model residuals for the African American race increased with obesity in the middle-aged cohort, yet variability decreased with obesity in the elderly cohort. The model best estimated the Lipidemia + hypertension multimorbidity pattern for this race across all cohorts. The Asian/Pacific Islander race model residuals were more extreme for patients in the elderly cohort without obesity than those with obesity. Although the Asian/Pacific Islander race exhibited negative model estimates compared to the African American race, when the mean model residuals were categorized as a function of multimorbidity, some combinations were better estimated than others. For the Caucasian race, the variability of mean model residuals was comparable by weight class in the middle-aged cohort, but it was more accurate for patients without obesity in the elderly cohort. The most accurate overall model estimates were for the Native American race, although variability in mean model residuals increased with obesity and aging. The mean model residuals also increased substantially by age group for the African American, Hispanic, and Native American races.

Since BMI and race as a combined effect were not significant (p-value = 0.870) when analyzed by 3-way ANOVA, we removed this interaction so that only ECI rank was significant (p-value = 0.0353) for the middle-aged cohort. For the elderly cohort, BMI and race as a combined effect were not significant (p-value = 1.000,) and only ECI rank was significant (p-value < 0.001).

Discussion

Our well-fit cost models contributes to understanding the relationship between cost and multimorbidity, which is largely missing from the literature.[33,34] The models show that the accuracy of estimating cost varies across race and by multimorbidity, age group, and obesity status. However, it exhibited varying patterns of over- or underestimating total charges for specific racial groups, suggesting that more robust methods will be necessary to ensure accurate cost capture, particularly for vulnerable populations. Capturing such a complex interplay is challenging. While this type of modeling has some limitations, it can help to identify the costs associated with multimorbidities to help project future patient costs. This study also showed that aging does not have a straightforward relationship with cost estimates for certain races. For example, African Americans were the index race in both models, and the Caucasian race had a negative total charges estimate for the middle-aged cohort and a positive total charges estimate for the elderly cohort.

While previous literature notes that levels and most prevalent categories of multimorbidity vary by race,[35] our research shows that the relationship between cost and multimorbidity is inconsistent for each racial group. Specific multimorbidity patterns were more inaccurate for some groups. Additionally, our study demonstrated that some racial groups could be

driving the overall inaccuracy of cost estimates for specific multimorbidity combinations. For example, the average estimated total charges for hypertension + GERD significantly deviated from the actual total charges. Residual analysis indicated that these estimates were significantly overestimated for the African American population in particular. As multimorbidity is associated with higher outpatient and inpatient utilization of healthcare services,[36] the importance of accurately modeling cost cannot be overstated. Given the seriousness of the inequalities in healthcare access and outcomes by race,[37] it is crucial that we generate models that are accurate across racial groups. Our findings provide necessary information on understanding the complexity of the relationship between cost and multimorbidity. Researchers modeling multimorbidity and cost must analyze estimates for specific patterns stratified by race to know how much specific estimates can be trusted.

Our results indicate that the pattern of model accuracy across the obesity category varies by race. A good understanding of the situations in which models are inaccurate for specific groups can help the research community identify areas for improved modeling to estimate costs for patient populations better. Except for one multimorbidity combination (Hypertension + GERD), multimorbidities in the patient population with obesity exhibited less extreme average residuals in the African American group in both the middle-aged and elderly cohorts. We observed a similar relationship in the elderly Asian population. On the other hand, the Hispanic population exhibited more extreme residuals in the absence of obesity. We observed a similar trend in the elderly Native American and Caucasian populations. In some instances, this could be attributed to differences in the type of multimorbidities in distinct groups, but we also observed this trend in cases where the multimorbidity is the same (e.g., Lipidemia + heart disease in the elderly cohort). Our results demonstrate the importance of stratification by weight category for improved model accuracy.

Limitations. The cross-sectional design of our study restricts our comprehension of multimorbidity, race, age, BMI, and ECI as risk factors that impacting patients' mean total charges. The results could not produce a single model consistent in predicting total charges across races in the same weight and age groups. The Cerner HealthFacts database contains patient records with $0 charges due to information not being transferred to the data warehouse. Consequently, we excluded these records from our study. However, as noted in the Methods section, this data is likely missing at random. The Uniform Hospital Discharge Data Set (UHDDS) definitions and regulations drive hospital reporting requirements for race and ethnicity data, which may not accurately reflect these categories.[38] The Cerner HealthFacts database categorized the Hispanic ethnicity as a race, yet these patients could identify as a member of the Native American, Black, White, or Asian races, or could be Biracial. Our exclusion of patients with unknown race, gender, BMI, or age data substantially reduced our sample, which could impact specific groups disproportionally. For the middle-aged cohort's model, the Asian/Pacific Islander and Biracial races were not statistically significant, nor was the Biracial racial category significant for the elderly cohort's model, which is most likely because we had small samples for these two races.

Despite these limitations, our study is unique because it included the Biracial and the

Native American groups, which are often not studied. Also, the study population reflects a nationwide sample selected from all corners of the nation and is representative of the patient group that doctors generally treat in a clinical setting.

If building a model for the most prevalent multimorbidity combinations by race is so challenging, how accurate will expenditure models be for multimorbidities that are not shared by all racial groups and how can we evaluate them? Although the model we developed was a good fit for the data we accessed, its variability in predicting total charges by race demonstrated that we need more robust models that accurately predict total healthcare charges for all racial groups. In particular, multimorbidity and race need to be studied more comprehensively in this regard.

Conclusions

To our knowledge, this is the first study to assess the accuracy of model predictions across races for total charges of the most prevalent multimorbidity combinations. Our study is the first to identify total charge trends for the most prevalent multimorbidity combinations. Our research showed that members of the African American race had the highest mean total charges and members of the Hispanic race had the lowest in both cohorts. We also demonstrated that our model was inconsistent in its ability to predict total charges by race based on multimorbidity patterns**.** In general, the total charges were either over- or underestimated across multimorbidity patterns, and in some cases, the estimates were extreme. This highlights the difficulty in modeling total charge estimates for diseases that may interact in a multimorbidity, since they do not have a simple additive effect. This demonstrates the need to develop more robust models to ensure the healthcare system can better serve all populations. Improved modeling of underserved populations is necessary and multimorbidity and race need to be studied more comprehensively.

List of abbreviations

BackP, Severe back pain; BMI, Body mass index; **CKD,** Chronic kidney disease; **ECI**, Elixhauser Comorbidity Index; **GERD,** Gastroesophageal reflux disease; **HD,** Heart disease; **HypT,** Hypertension; ICD-10–CM, International classification of diseases-10th version-clinical modification; **Lip,** Lipidemia; **LimbP,** Pain in limb, hand, foot, fingers, and toes; **OJD,** Other joint disorder; **T2D,** Type 2 diabetes mellitus; UHDDS, Uniform hospital discharge data set.

Declarations

Ethics approval and consent to participate. Because the study only employed de-identified data, the University of Tennessee Health Science Center (UTHSC) Institutional Review Board (IRB) determined that the research was exempt.

Consent for publication. Not applicable

Availability of data and materials. The datasets generated and/or analyzed during the current study are not publicly available due as they must be licensed by Oracle Cerner (https://www.cerner.com/ap/en/solutions/data-research).

Competing interests. The authors declare that they have no competing interests.

Funding. The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the National Institute of Aging award R15AG067232.

Authors' contributions. M.A.: conception, design, analysis, methods, data collection, writing. T.G. writing, design, and review. J.B. writing, review. C.M.B.: conception, supervision, design, review. All authors read and approved the final manuscript.

Acknowledgements. The authors also wish to thank Dr. Kyle Johnson Moore in the UTHSC Office of Scientific Writing for editing the manuscript.

Appendix

Appendix Table 1. Most Prevalent Morbidities at the 5% Threshold

#	Diagnosis
1	**D63:** Anemia in chronic diseases classified elsewhere
2	**D64:** Other anemias
3	**E03:** Other hypothyroidism
4	**E11:** Diabetes
5	**E55:** Vitamin D deficiency
6	**E78:** Lipidemia
7	**E86:** Volume depletion
8	**E87:** Other disorders of fluid, electrolyte, and acid-base balance
9	**F03:** Unspecified dementia
10	**F10:** Alcohol related disorders
11	**F17:** Nicotine dependence
12	**F32:** Major depressive disorder, single episode
13	**F41:** Anxiety Disorders
14	**G47:** Sleep disorders are represented
15	**G89:** Pain, not elsewhere classified
16	**I10:** Hypertension
17	**I12:** Hypertensive chronic kidney disease
18	**I25:** Heart disease
19	**I48:** Atrial fibrillation and flutter are prevalent
20	**I50:** Heart failure
21	**I65:** Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
22	**I73:** Other peripheral vascular diseases
23	**J44:** Other chronic obstructive pulmonary disease
24	**J45:** Asthma
25	**K21:** GERD
26	**K57:** Diverticular disease of intestine
27	**K59:** Other functional intestinal disorders
28	**M17:** Osteoarthritis of knee
29	**M19:** Other and unspecified osteoarthritis
30	**M25:** Other Joint Disorders
31	**M54:** Dorsalgia
32	**M79:** Other and unspecified soft tissue disorders, not elsewhere classified
33	**M81:** Osteoporosis without current pathological fracture
34	**N17:** Acute kidney failure
35	**N18:** Chronic kidney disease (CKD)
36	**N28:** Other disorders of kidney and ureter, not elsewhere classified
37	**N39:** Other disorders of urinary system
38	**N40:** Benign prostatic hyperplasia

Appendix Table 2. Study Variables.

Variable	Notes
Comorbidity	38 comorbidities in Table 1 with ICD-10-CM broad categories
Age	45+
Gender	Male and Female
Race	African American, Asian/Pacific Islander, Biracial, Caucasian, Hispanic, and Native American
BMI	With obesity ($30 \leq BMI < 206$) and without obesity ($18.5 \leq BMI < 30$). 18.5 is the minimum BMI for normal weight, and 206 is the maximum BMI recorded for a human when this study was first conducted
Total number of morbidities	Added the total number of morbidities for each patient during the two years.
Elixhauser Comorbidity Index (ECI)	Assess disease burden.
Total Charges	Total charges greater than $0 were included, and the average charges of all encounters for the two years were assigned for each patient. The records in the Cerner HealthFacts database have the total charges of each patient encounter. An encounter is an interaction between a patient and healthcare provider(s) to provide healthcare service(s) or assess a patient's health status of a patient.[39] Hospital charges are more accurate in reflecting the patient population receiving a hospital service.[40]
Payer Type	Medicare/Medicaid/Title V, Other, and Unknown
Hospital Division	Based on the U.S. Census Bureau, nine hospital divisions were in the data set. Some studies have projected that geographical location impact cost.[41]
Length of Stay	If the patient was hospitalized, their length of stay was calculated by subtracting the discharge date from the admission date.
Number of Emergency Visits	Adding the number of emergency room visits during the two years
Number of Outpatient Visits	Adding the number of outpatient visits during the two years
Urban Status of Area	The urban or rural status of a hospital impacts cost[42]
Teaching Status of Hospital	Indicating the teaching status of the hospital
Acute Care Status	Hospitals providing acute care services are more profitable than other services and therefore are projected to impact cost.[43]

Appendix Table 3. Morbidity and Hospital Information Model Estimates for the Middle-Aged Cohort

Morbidity (Variable)	Estimate	P-value	Significance
K57: Diverticular disease of intestine	0.134	<0.001	***
F10: Alcohol related disorders	0.115	<0.001	***
K21: GERD	0.096	<0.001	***
F17: Nicotine dependence	0.095	<0.001	***
I25: Heart disease	0.092	<0.001	***
N17: Acute kidney failure	0.086	<0.001	***
I12: Hypertensive chronic kidney disease	0.084	<0.001	***
N28: Other disorders of kidney and ureter, not elsewhere classified	0.083	<0.001	***
E86: Volume depletion	0.082	<0.001	***
D64: Other anemias	0.081	<0.001	***
M17: Osteoarthritis of knee	0.081	<0.001	***
I65: Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction	0.079	<0.001	***
E87: Other disorders of fluid, electrolyte, and acid-base balance	0.060	<0.001	***
F32: Major depressive disorder, single episode	0.059	<0.001	***
I10: Hypertension	0.056	<0.001	***
M19: Other and unspecified osteoarthritis	0.054	<0.001	***
N40: Benign prostatic hyperplasia	0.054	<0.001	***
K59: Other functional intestinal disorders	0.051	<0.001	***
E11: Diabetes	0.042	<0.001	***
G89: Pain, not elsewhere classified	0.041	<0.001	***
N39: Other disorders of urinary system	0.040	<0.001	***
J45: Asthma	0.038	<0.001	***
I48: Atrial fibrillation and flutter are prevalent	0.034	<0.001	***
G47: Sleep disorders are represented	0.031	<0.001	***
M81: Osteoporosis without current pathological fracture	0.029	<0.001	***
F03: Unspecified dementia	0.028	0.006	**
E03: Other hypothyroidism	0.027	<0.001	***
F41: Anxiety Disorders	0.025	<0.001	***
J44: Other chronic obstructive pulmonary disease	0.024	<0.001	***
D63: Anemia in chronic diseases classified elsewhere	0.023	0.002	**
M54: Dorsalgia	0.022	<0.001	***
E78: Lipidemia	0.020	<0.001	***
M25: Other Joint Disorders	0.015	<0.001	***
M79: Other and unspecified soft tissue disorders, not elsewhere classified	0.013	<0.001	***
I73: Other peripheral vascular diseases	0.011	0.011	*
I50: Heart failure	-0.008	0.012	*
E55: Vitamin D deficiency	-0.010	<0.001	***
N18: Chronic kidney disease (CKD)	-0.054	<0.001	***

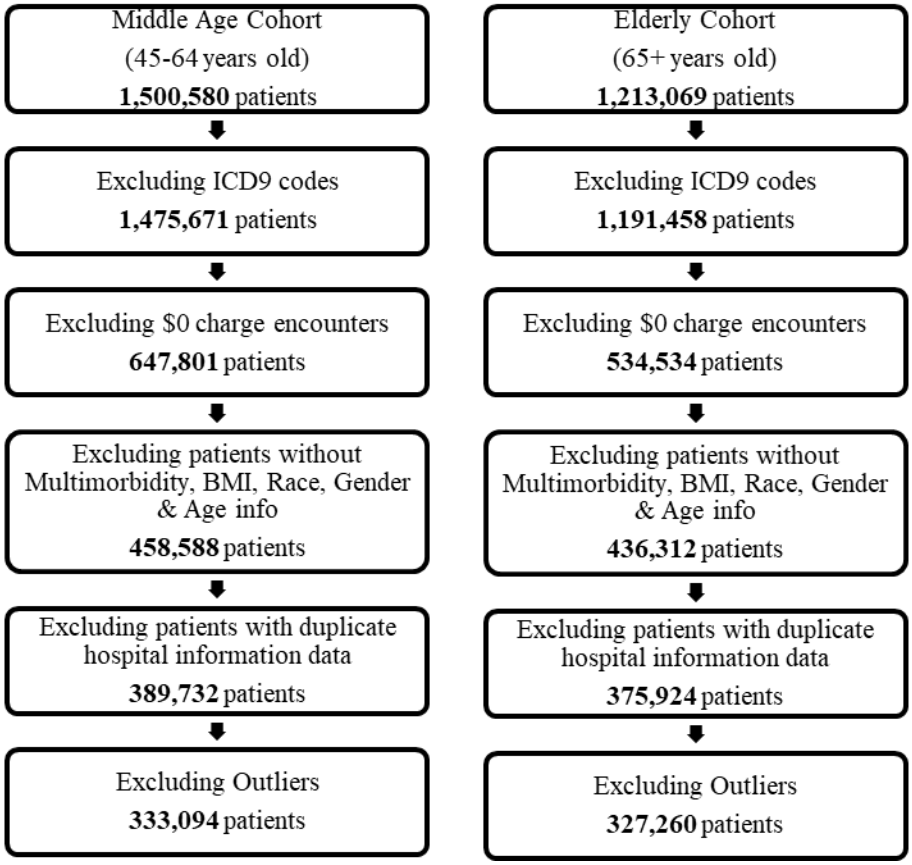
Hospital Information (Variable)	Estimate	P-value	Significance
Urban hospital	0.009	<0.001	***
Acute care hospital	0.288	<0.001	***
Teaching hospital	0.000	0.731	-
Census Division 5: East South Central (A.L., KY, MS, TN)	-0.180	0.016	*
Census Division 6: South Atlantic (D.E., DC, FL, GA, MD, NC, SC, VA, WV)	-0.185	0.013	*
Census Division 3: West North Central (I.A., KS, MN, MO, ND, SD)	-0.260	0.001	***
Census Division 9: Pacific (A.K., CA, HI, OR, WA)	-0.272	<0.001	***
Census Division 7: West South Central (A.R., LA, OK, TX)	-0.279	<0.001	***
Census Division 4: East North Central (I.L., IN, MI, OH, WI)	-0.368	<0.001	***
Census Division 8: Mountain (AZ, CO, ID, MT)	-0.375	<0.001	***
Census Division 2: Middle Atlantic (N.J., NY, PA)	-0.390	<0.001	***

Appendix Table 4. Morbidity and Hospital Information Model Estimates for the Elderly Cohort

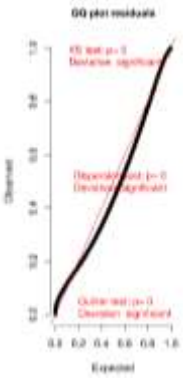
Morbidity (Variable)	Estimate	P-value	Significance
M17: Osteoarthritis of knee	0.263	<0.001	***
K57: Diverticular disease of intestine	0.230	<0.001	***
F10: Alcohol related disorders	0.229	<0.001	***
I12: Hypertensive chronic kidney disease	0.219	<0.001	***
N17: Acute kidney failure	0.217	<0.001	***
K21: GERD	0.208	<0.001	***
I25: Heart disease	0.203	<0.001	***
I65: Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction	0.197	<0.001	***
F17: Nicotine dependence	0.194	<0.001	***
F03: Unspecified dementia	0.169	<0.001	***
I10: Hypertension	0.169	<0.001	***
E86: Volume depletion	0.166	<0.001	***
E87: Other disorders of fluid, electrolyte, and acid-base balance	0.160	<0.001	***
N40: Benign prostatic hyperplasia	0.155	<0.001	***
G89: Pain, not elsewhere classified	0.152	<0.001	***
F32: Major depressive disorder, single episode	0.151	<0.001	***
D64: Other anemias	0.132	<0.001	***
N28: Other disorders of kidney and ureter, not elsewhere classified	0.125	<0.001	***
N39: Other disorders of urinary system	0.111	<0.001	***
J45: Asthma	0.109	<0.001	***
M19: Other and unspecified osteoarthritis	0.107	<0.001	***
I48: Atrial fibrillation and flutter are prevalent	0.104	<0.001	***
J44: Other chronic obstructive pulmonary disease	0.102	<0.001	***
E78: Lipidemia	0.097	<0.001	***
D63: Anemia in chronic diseases classified elsewhere	0.097	<0.001	***
K59: Other functional intestinal disorders	0.097	<0.001	***
E11: Diabetes	0.096	<0.001	***
E03: Other hypothyroidism	0.092	<0.001	***
G47: Sleep disorders are represented	0.091	<0.001	***
F41: Anxiety Disorders	0.079	<0.001	***
M81: Osteoporosis without current pathological fracture	0.076	<0.001	***
I73: Other peripheral vascular diseases	0.065	<0.001	***
M54: Dorsalgia	0.056	<0.001	***
M25: Other Joint Disorders	0.041	<0.001	***
M79: Other and unspecified soft tissue disorders, not elsewhere classified	0.023	<0.001	***
I50: Heart failure	0.020	<0.001	***
E55: Vitamin D deficiency	-0.074	<0.001	***
N18: Chronic kidney disease (CKD)	-0.133	<0.001	***
Hospital Information (Variable)	**Estimate**	**P-value**	**Significance**
Urban hospital	0.023	<0.001	***
Acute care hospital	0.776	<0.001	***
Teaching hospital	0.012	<0.001	***

Census Division 6: South Atlantic (D.E., DC, FL, GA, MD, NC, SC, VA, WV)	-0.222	0.182	-
Census Division 9: Pacific (A.K., CA, HI, OR, WA)	-0.329	0.048	*
Census Division 7: West South Central (A.R., LA, OK, TX)	-0.400	0.016	*
Census Division 5: East South Central (A.L., KY, MS, TN)	-0.412	0.013	*
Census Division 3: West North Central (I.A., KS, MN, MO, ND, SD)	-0.447	0.007	**
Census Division 8: Mountain (AZ, CO, ID, MT)	-0.671	<0.001	***
Census Division 2: Middle Atlantic (N.J., NY, PA)	-0.681	<0.001	***
Census Division 4: East North Central (I.L., IN, MI, OH, WI)	-0.686	<0.001	***

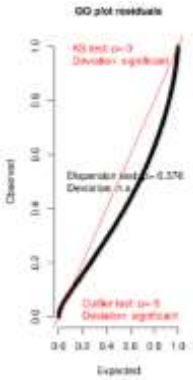
Appendix Figure 1. Patient Population



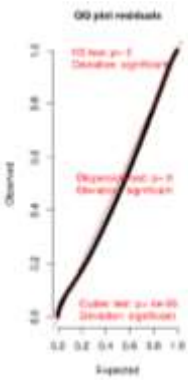
Appendix Figure 2. Adjusted and Unadjusted Model Residuals for the Middle-Aged and Elderly Cohorts



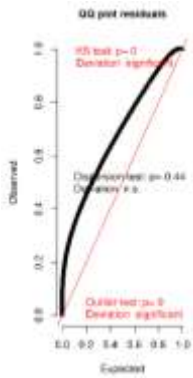
A. *Unadjusted Model Residuals for the Middle-*



B. *Adjusted Model Residuals for the Middle-Aged Cohort*

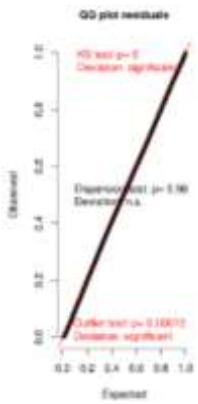


C. *Unadjusted Model Residuals for the Elderly*

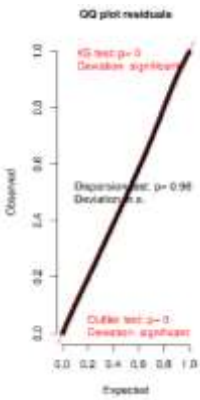


D. *Adjusted Model Residuals for the Elderly*

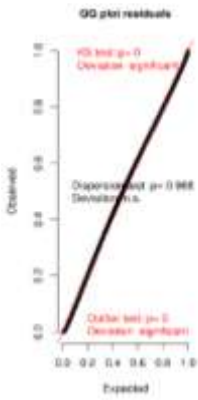
Appendix Figure 3. Log and Exponential Adjusted Model Residuals for the Middle-Aged and Elderly Cohorts



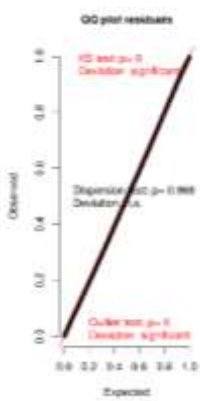
A. *Log Model Residuals for the Middle-Aged Cohort*



B. *Exponential Model Residuals for the Middle-Aged Cohort*



C. *Log Model Residuals for the Elderly Cohort*



D. *Exponential Model Residuals for the Elderly Cohort*

References

1. McPhail SM. Multimorbidity in chronic disease: impact on health care resources and costs. *Risk Manag Healthc Policy*. 2016;9:143-56. https://doi.org/10.2147/rmhp.S97248

2. Sambamoorthi U, Tan X, Deb A. Multiple chronic conditions and healthcare costs among adults. *Expert Rev Pharmacoecon Outcomes Res*. 2015;15(5):823-32. https://doi.org/10.1586/14737167.2015.1091730

3. Harris LJ, Graetz I, Podila PS, Wan J, Waters TM, Bailey JE. Characteristics of Hospital and Emergency Care Super-utilizers with Multiple Chronic Conditions. *J Emerg Med*. Apr 2016;50(4):e203-14. https://doi.org/10.1016/j.jemermed.2015.09.002

4. Caballer-Tarazona V, Guadalajara-Olmeda N, Vivas-Consuelo D. Predicting healthcare expenditure by multimorbidity groups. *Health Policy*. 2019/04/01/ 2019;123(4):427-434. https://doi.org/10.1016/j.healthpol.2019.02.002

5. Sancho-Mestre C, Vivas-Consuelo D, Alvis-Estrada L, Romero M, Usó-Talamantes R, Caballer-Tarazona V. Pharmaceutical cost and multimorbidity with type 2 diabetes mellitus using electronic health record data. *BMC Health Serv Res*. Aug 17 2016;16(1):394. https://doi.org/10.1186/s12913-016-1649-2

6. Alshakhs M, Jackson B, Ikponmwosa D, Reynolds R, Madlock-Brown C. Multimorbidity patterns across race/ethnicity as stratified by age and obesity. *Sci Rep*. Jun 11 2022;12(1):9716. https://doi.org/10.1038/s41598-022-13733-w

7. Schoenberg NE, Kim H, Edwards W, Fleming ST. Burden of common multiple-morbidity constellations on out-of-pocket medical expenditures among older adults. *Gerontologist*. Aug 2007;47(4):423-37. https://doi.org/10.1093/geront/47.4.423

8. Lochner KA, Cox CS. Prevalence of multiple chronic conditions among Medicare beneficiaries, United States, 2010. *Prev Chronic Dis*. Apr 25 2013;10:E61. https://doi.org/10.5888/pcd10.120137

9. Espeland MA, Crimmins EM, Grossardt BR, et al. Clinical Trials Targeting Aging and Age-Related Multimorbidity. *J Gerontol A Biol Sci Med Sci*. Mar 1 2017;72(3):355-361. https://doi.org/10.1093/gerona/glw220

10. Mercer SW, Fitzpatrick B, Guthrie B, et al. The CARE Plus study - a whole-system intervention to improve quality of life of primary care patients with multimorbidity in areas of high socioeconomic deprivation: exploratory cluster randomised controlled trial and cost-utility analysis. *BMC Med*. Jun 22 2016;14(1):88. https://doi.org/10.1186/s12916-016-0634-2

11. Clay OJ, Perkins M, Wallace G, Crowe M, Sawyer P, Brown CJ. Associations of Multimorbid Medical Conditions and Health-Related Quality of Life Among Older African American Men. *J Gerontol B Psychol Sci Soc Sci*. Jan 11 2018;73(2):258-266. https://doi.org/10.1093/geronb/gbx090

12. Lynch CP, Gebregziabher M, Axon RN, Hunt KE, Payne E, Egede LE. Geographic and racial/ethnic variations in patterns of multimorbidity burden in patients with type 2 diabetes. *J Gen Intern Med*. Jan 2015;30(1):25-32. https://doi.org/10.1007/s11606-014-2990-y

13. Noseworthy PA, Attia ZI, Brewer LC, et al. Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis. *Circ Arrhythm Electrophysiol*. Mar 2020;13(3):e007988. https://doi.org/10.1161/circep.119.007988

14. St Sauver JL, Boyd CM, Grossardt BR, et al. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open*. Feb 3 2015;5(2):e006413. https://doi.org/10.1136/bmjopen-2014-006413

15. Almirall J, Fortin M. The coexistence of terms to describe the presence of multiple concurrent diseases. *J Comorb*. 2013;3:4-9. https://doi.org/10.15256/joc.2013.3.22

16. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Managing patients with multimorbidity: systematic review of interventions in primary care and community settings. *BMJ : British Medical Journal*. 2012;345:e5205. https://doi.org/10.1136/bmj.e5205

17. Simard M, Rahme E, Calfat AC, Sirois C. Multimorbidity measures from health administrative data using ICD system codes: A systematic review. *Pharmacoepidemiol Drug Saf*. Jan 2022;31(1):1-12. https://doi.org/10.1002/pds.5368

18. Guinness World Records. Heaviest man ever Guinness World Records, Ltd. Accessed June 22, 2022, 2022. https://www.guinnessworldrecords.com/world-records/heaviest-man

19. Terkawi AS, Rafiq M, Algadaan R, et al. General anesthesia for the heaviest man in the world. *Saudi J Anaesth*. Nov 2014;8(Suppl 1):S101-4. https://doi.org/10.4103/1658-354x.144087

20. Kuo RN, Lai MS. The influence of socio-economic status and multimorbidity patterns on healthcare costs: a six-year follow-up under a universal healthcare system. *Int J Equity Health*. Aug 20 2013;12:69. https://doi.org/10.1186/1475-9276-12-69

21. Arora V, Moriates C, Shah N. The Challenge of Understanding Health Care Costs and Charges. *AMA J Ethics*. Nov 1 2015;17(11):1046-52. https://doi.org/10.1001/journalofethics.2015.17.11.stas1-1511

22. Batty M, Ippolito B. Mystery Of The Chargemaster: Examining The Role Of Hospital List Prices In What Patients Actually Pay. *Health Aff (Millwood)*. Apr 1 2017;36(4):689-696. https://doi.org/10.1377/hlthaff.2016.0986

23. Linde S, Egede LE. Hospital Price Transparency in the United States: An Examination of Chargemaster, Cash, and Negotiated, Price Variation for 14 Common Procedures. *Medical Care*. 2022;60(10)

24. Brookings Health System. Charge, Cost & Price. Brookings Health System. https://www.brookingshealth.org/why-brookings-health/health-care-value/understanding-medical-prices/charge-cost-price

25. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj*. Jun 29 2009;338:b2393. https://doi.org/10.1136/bmj.b2393

26. Alin A. Multicollinearity. *WIREs Computational Statistics*. 2010/05/01 2010;2(3):370-374. https://doi.org/10.1002/wics.84

27. Allison P. When Can You Safely Ignore Multicollinearity? . *Statistical Horizons* blog. September 10, 2012, 2012. Accessed November 14, 2022. https://statisticalhorizons.com/multicollinearity/

28. Fox J, Monette G. Generalized Collinearity Diagnostics. *J Am Stat Assoc*. 1992/03/01 1992;87(417):178-183. https://doi.org/10.1080/01621459.1992.10475190

29. Fox J, Weisberg S. *An R Companion to Applied Regression* 3rd ed. SAGE Publications, Inc.; 2018.

30. Turner H. Introduction to Generalized Linear Models. April 22-24, 2008 2008; ESRC National Centre for Research Methods, UK. http://statmath.wu.ac.at/courses/heather_turner/glmCourse_001.pdf

31. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ*. Jul 2001;20(4):461-94. https://doi.org/10.1016/s0167-6296(01)00086-8

32. Hessing T. Sum of Squares (S.S.). Six Sigma Study Guide. Accessed July 10, 2022, 2022. https://sixsigmastudyguide.com/?s=Sum+of+Squares

33. Koné Pefoyo AJ, Bronskill SE, Gruneir A, et al. The increasing burden and complexity of multimorbidity. *BMC Public Health*. Apr 23 2015;15:415. https://doi.org/10.1186/s12889-015-1733-2

34. Hajat C, Siegal Y, Adler-Waxman A. Clustering and Healthcare Costs With Multiple Chronic Conditions in a US Study. *Front Public Health*. 2020;8:607528. https://doi.org/10.3389/fpubh.2020.607528

35. Rocca WA, Boyd CM, Grossardt BR, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. *Mayo Clin Proc*. Oct 2014;89(10):1336-49. https://doi.org/10.1016/j.mayocp.2014.07.010

36. Zulman DM, Pal Chee C, Wagner TH, et al. Multimorbidity and healthcare utilisation among high-cost patients in the US Veterans Affairs Health Care System. *BMJ Open*. Apr 16 2015;5(4):e007771. https://doi.org/10.1136/bmjopen-2015-007771

37. Dickman SL, Himmelstein DU, Woolhandler S. Inequality and the health-care system in the USA. *Lancet*. Apr 8 2017;389(10077):1431-1441. https://doi.org/10.1016/s0140-6736(17)30398-7

38. Strmic-Pawl HV, Jackson BA, Garner S. Race Counts: Racial and Ethnic Data on the U.S. Census and the Implications for Tracking Inequality. *Sociology of Race and Ethnicity*. 2018/01/01 2017;4(1):1-13. https://doi.org/10.1177/2332649217742869

39. Health Level 7 (HL7) International Fast Healthcare Interoperability Resources (FHIR) Foundation. FHIR Release 4. Ann Arbor, MI: HL7 FHIR Foundation; 2019.

40. Minnesota Department of Health. *Finance in rural and urban hospitals*. 2017:7. April 2017. www.health.state.mn.us/facilities/ruralhealth/pubs/docs/2017finances.pdf

41. Johnson WC, Biniek JF. Sources of Geographic Variation in Health Care Spending Among Individuals With Employer Sponsored Insurance. *Med Care Res Rev*. Oct 2021;78(5):548-560. https://doi.org/10.1177/1077558720926095

42. Wengle E, Blumberg LJ, Holahan J. *Are Marketplace Premiums Higher in Rural Than in Urban Areas?* Vol. 2022. 2018. November 15, 2018. https://www.rwjf.org/en/library/research/2018/11/are-marketplace-premiums-higher-in-rural-than-in-urban-areas.html

43. Gee E. *The High Price of Hospital Care*. 2019. June 2019. https://www.americanprogress.org/article/high-price-hospital-care/

APPENDIX C. CHAPTER 4 ARTICLE

NOTE: Navigation with Adobe Acrobat Reader or Adobe Acrobat Professional: To return to the last viewed page, use key commands Alt/Ctrl+Left Arrow on PC or Command+Left Arrow on Mac. For the next page, use Alt/Ctrl or Command+Right Arrow. See the **Preface** for further details. If needed, use this link to return to **Chapter 4** after navigating within this appendix.

Introduction

This appendix includes the work done for Aim 3: "Assess the association of BMI with multimorbidities and healthcare burden across races." Article reused from prepared manuscript with authors' permission. Alshakhs, M., Goedecks, P., Chinthala, L., Weiskopf, N., Madlock-Brown, C. Assessing the Relationship Between Healthcare Burden and BMI as Stratified by Race in the US. (2023). It is included in this dissertation with permission from co-authors.

Article

Assessing the Relationship Between Healthcare Burden and BMI as Stratified by Race in the US

Manal J.Alshakhs, MS[1]
Patricia J Goedecke, MS[2]
Lokesh K Chinthala, MS[1]
Nicole G. Weiskopf, PhD[3]
Charisse Madlock-Brown, PhD, MLS[1,4]

[1]Health Outcomes and Policy Research Program, University of Tennessee Health Science Center, Memphis, Tennessee, USA
[2]Doctoral student, Bredesen Center, University of Tennessee Knoxville
[3]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University
[4]Center for Health System Improvement, University of Tennessee Health Science Center, Memphis, Tennessee, USA

Abstract

Background. This research study aims to assess the relationship between BMI and healthcare burden as stratified by race and healthcare utilization among middle-aged patients in the US.

Methods. Using regression analysis, we compared the area under the curve (AUC) for the receiver operating characteristic (ROC) curve across race and healthcare utilization for 2016-2017 data from the Cerner HealthFacts database. Additionally, we evaluated the relationship between body mass index (BMI) and healthcare burden for middle-aged patients in the US.

Results. Based on the coefficient values, BMI significantly impacted the CCI (Charlson comorbidity index), the healthcare burden measure, of the Asian/Pacific Islander and Caucasian races across all utilization cohorts. In the healthcare utilizer cohort, BMI coefficient values almost doubled for these two races relative to those for the African American and Native American races. The model best predicted the CCI for the Caucasian cohort and predicted for the African American cohort the least well among the healthcare utilizer cohort.

Conclusions. These results demonstrated that the relationship between BMI and CCI varied across race and within the same healthcare care utilization category. More work needs to be done to understand how multimorbidity, BMI, and healthcare burden associate across race.

Key words: multimorbidity, race stratification, healthcare utilization, healthcare burden

Background

Body mass index (BMI) is a widely used measure in healthcare to assess an individual's weight status and determine their risk for developing health problems related to obesity. However, it is essential to note that BMI does not distinguish between muscle and fat, so it may overestimate body fat in athletes and underestimate body fat in older adults. Additionally, BMI does not consider individual differences in body shape and composition. The World Health Organization (WHO) has recommended rescaling the BMI for Asian individuals,[1] based on studies performed in China that have produced new population-based BMI scales for obesity. [2] Indonesia, Turkey, Singapore, Japan, India, and other countries have done the same.[3] These BMI cutoffs vary depending on the data and resources available. Some measures used to set the new BMI cutoffs include adiposity, all-cause mortality, waist circumference, waist-to-hip ratio, and incidence rates of certain prevalent diseases. [4–6] A 2019 study also recommended new BMI cutoffs for Black, Hispanic, and White males and females, based on the risk of developing one of three metabolic diseases, namely dyslipidemia, diabetes mellitus, and hypertension. These authors found differences in BMI cutoffs across race and gender for these three diseases. [7] The WHO recommends that population-based obesity BMI cutoffs be founded on morbidity and mortality.[6]

Assessing the relationship between BMI and healthcare burden across race in the United States (U.S.) is crucial in determining a person's overall health status and informing the development of a tailored treatment plan. This relationship between BMI and healthcare utilization is generally positive, meaning that as BMI increases, so does healthcare utilization.[8] Higher BMIs are often associated with a greater risk of obesity-related health conditions such as type 2 diabetes mellitus, cardiovascular disease, and certain cancers.[9] These health conditions can lead to increased hospitalizations, medication use, and doctor visits, thereby straining the healthcare system. While individuals with lower BMIs are less likely to require as many healthcare resources, regular healthcare utilization results in better healthcare outcomes regardless of BMI status.[10,11]

The Charlson Comorbidity Index (CCI) is a widely used tool for evaluating comorbidity in patients with chronic diseases and is often used to assess health outcomes.[12] This study aims to assess the relationship between BMI and CCI as stratified by race. We hypothesize that BMI does not operate equally across race. This will help address health disparities and design target interventions to improve health outcomes.

Methods

Research Design. This cross-sectional study used data from the Cerner HealthFacts® data warehouse for 2016-2017 that did not contain any identifying information. The dataset comprises the encounter data from more than 490 million patient encounters with over 70 million patients who received treatment at hospitals and clinics from 792 non-affiliated healthcare systems across the U.S. from 2001 to 2017. The data includes information about the type of encounter, patient medical history, diagnoses, laboratory results, prescriptions, patient demographics, clinic type, and procedures performed. The Cerner HealthFacts® database did not capture waist circumference, which has strongly correlated with multimorbidity in previous research. However, this database did capture insurance and visit types. We will use these two variables to investigate their impact on the burden of multimorbidity across races.

For this study, the patient inclusion criteria were 1) age 45-64 years, 2) BMI value present and between 18.5 and 75, 3) assigned race category, and 4) assigned gender (i.e., male or female). The exclusion criteria for patients were 1) patients with a cancer diagnosis, 2) pregnant patients, and 3) patients with International Classification of Diseases-9th Version-Clinical Modification (ICD-9-CM) diagnostic codes. We selected a two-year study period to enhance our analytical sensitivity in capturing healthcare utilization patterns and gather a more comprehensive dataset.

We stratified the patients by race and divided them into three subgroups based on patterns of healthcare utilization, which we defined as a minimum of two outpatient visits within the two-year study period. Patients with no outpatient visits were classified as "non-utilizers," while those with only one such visit were classified as "low utilizers." It should be noted that all utilization subgroups may have had inpatient and/or emergency department visits. Stratifying by utilization refers to grouping individuals or data based on their use of healthcare services. This can include the frequency of doctor visits, use of prescription medications, and hospitalization rates. By stratifying the data in this way, we can examine how different levels of healthcare utilization may be associated with various health outcomes. It is also crucial to stratify the data by race because some racial groups may have different levels of access to healthcare that influences health-seeking behavior. By stratifying the data, we can examine how these differences in health-seeking behavior may impact health outcomes for different racial groups.

Ethical Considerations. This research utilized only de-identified data that excluded the 16 identifiable variables. Per the National Institutes of Health policy, the University of Tennessee Health Science Center (UTHSC) Institutional Review Board (IRB) classified the research as exempt. This study was conducted in compliance with all other necessary research requirements.

Independent Variables. Demographics, payer type, smoking, alcohol use, urbanism, and 2014-2015 healthcare utilization were the primary independent variables. Demographic variables were race, age, gender, marital status, and body mass index (BMI), while racial categories were African American, Asian/Pacific Islander, Caucasian, and Native American. Given the limitations in sample size, we chose to include only these four racial categories due to their adequate representation within the Cerner HealthFacts dataset, despite the availability of additional racial categories. We categorized marital status as those patients currently with a partner (i.e., married or life partner), having an ex-partner (i.e., divorced, widowed, or legally separated), single, and unknown. We categorized payer types as those patients with public insurance, private insurance, self-pay, or unknown, based on the information available in the dataset. We limited BMI values to the range $18.5 \leq BMI < 75$, based on the distribution of this variable within our sample population. We defined 2014-2015 healthcare utilization as a patient having at least two outpatient visits during that period. The variables of age and BMI were standardized to fall within a range of 0 to 1. Patients with unknown urban status was assigned to the urban category, since it had the most prevalence in the dataset.

Outcome Variable. Our primary outcome variable was the healthcare burden measure, CCI, which we calculated from the International Classification of Diseases-10th Version-

Clinical Modification (ICD-10-CM) diagnosis codes. The CCI weighted score was further ranked 0 for patients who did not have one of 17 listed comorbidities and 1 for those who had at least one of these comorbidities. [13] Thus it is important to note that a patient might have morbidities or comorbidities that were not on this list. We chose CCI as our outcome variable since it is a reliable tool for evaluating health outcomes by quantifying the degree of illness in an individual patient[12] and provides a standardized measure of disease that allows patient comparisons over time.[14]

Statistical Analysis. In this study, we applied logistic regression with a binomial family distribution to the training dataset (70% of the final population) to model the relationship between the predictor variables and the binary outcome variable, CCI.[15] The Kolmogorov-Smirnov (KS) test was performed to compare the distribution of the outcome variable among the different cohorts.[16] The results indicated no statistically significant deviation from normality for all the cohorts except for the Caucasian utilizer cohort, which the KS test showed to be statistically significant. We checked for multicollinearity, which refers to linear relationships between multiple variables[17] using a generalized variance inflation factor (GVIF) analysis suitable for a combination of numerical and categorical variables.[18,19] We used the car R package to remove variables with a high GVIF^(1/2Df) score (i.e., alcohol use)[20] and repeated this process until no variable had a score above the conservative threshold of two.[19]

We examined how BMI impacted the CCI score by comparing the area under the curve (AUC) of the receiver operating characteristic (ROC) curve for the adjusted model with and without inclusion of the BMI. We determined the best BMI cutoff for each group by analyzing the ROC curve and finding the point with the highest AUC within the range of 20 to 38 BMI for each group. We had 19 ROC curves for each cohort and considered only the BMI cutoff with the highest AUC.

Results

Demographics. In this study, most patients in all cohorts were female and the average age was 55 years. **Table 1** shows the breakdown of demographics by race. The percentages were determined in relation to the entire patient population. **Table 2** shows the breakdown of the number of patients in different utilization cohorts by race.

Table 1. Patient Demographics by Race

Race	Gender		Prevalence
	Female n (%)	Male n (%)	n (%)
African American	140,209 (11)	104,743 (8.24)	244,952 (19.3)
Asian/Pacific Islander	13,424 (1.06)	9,727 (<1)	23,151 (1.82)
Caucasian	544,570 (42.8)	447,680 (35.2)	992,250 (78)
Native American	6,196 (<1)	5,148 (<1)	11,344 (~1)
Total	704,399 (55)	567,298 (45)	1271697 (100)

Table 2. Number of Patients in All Utilization Cohorts by Race

Race/Utilization	Utilizers n (%)	Non-Utilizers n (%)	Low-Utilizers n (%)	Total n (%)
African American	111,161 (9)	96,436 (8)	37,355 (3)	244,952 (19)
Asian/Pacific Islander	13,665 (1)	5,152 (<1)	4,334 (<1)	23,151 (2)
Caucasian	526,115 (41)	274,813 (22)	191,322 (15)	992,250 (78)
Native American	4,642 (<1)	4,446 (<1)	2,256 (<1)	11,344 (1)
Total	655,583 (51.5)	380,847 (29.9)	235,267 (18.5)	1,271,697 (100)

Outcomes. The Cerner HealthFacts® database included data from 1,500,636 middle-aged patients for 2016-2017. The CONSORT diagram used to recruit the final patient population is shown in **Appendix Table 1**. After data cleaning procedures, the exclusion of patients with ICD-9-CM diagnostic codes, patients with a diagnosis of cancer or pregnancy, and patients who received treatment in multiple urban locations, the remaining included patient population was comprised of 1,271,697 individuals. The mean age, BMI, and weighted CCI scores for each racial cohort is shown in **Table 3**, while the distribution of BMI across races is shown in **Figure 1.**

Table 3. Mean Age, BMI, and Weighted CCI Score by Race for Healthcare Utilizer Cohorts

Race/Variable	Mean Age	Mean BMI	Mean Weighted CCI
African American	55	33	1
Asian/Pacific Islander	54	27	0
Caucasian	55	31	0
Native American	54	32	1

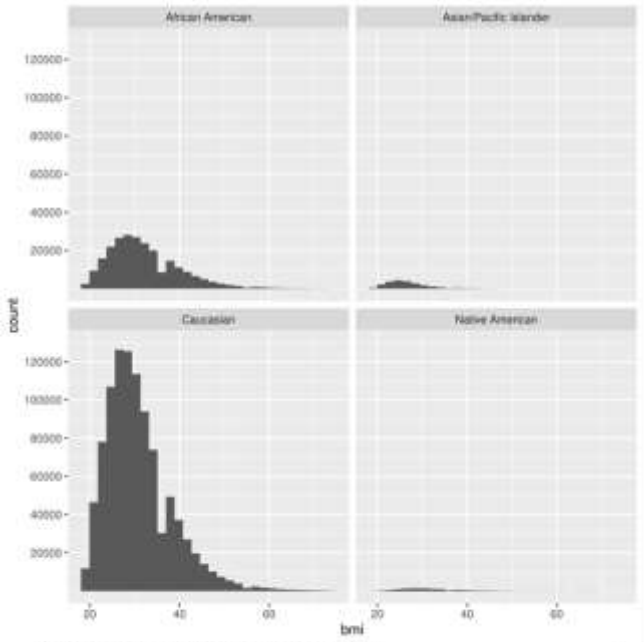


Figure 1. BMI Distribution by Race

Model. In our logistic regression model, the outcome variable CCI was modeled as a function of several predictor variables: BMI, age, gender, health insurance type, smoking, marital status, 2014-2015 healthcare utilization, and urbanism. The following are the indexed variables: female gender, self-pay health insurance, with a partner marital status, non-smoking, in a rural area, and having less than 2 outpatient visits in 2014-2015. **Table 4** and **Figure 2** show the odds ratio across race and healthcare utilization. Based on the odds ratio, BMI significantly impacted the healthcare burden measure, CCI, in the Asian/Pacific Islander and Caucasian races across all utilization cohorts. In the healthcare utilizer cohort, BMI coefficient values almost doubled for those two races compared to the African American and the Native American races. BMI was the most crucial predictor of CCI across race and utilization. The top three variables predicting CCI varied by healthcare utilization. For example, BMI, age, and smoking were the top three variables predicting CCI in the healthcare utilizer cohorts for all races. Yet, BMI, age, and public insurance were the top three variables predicting CCI in the healthcare non-utilizer cohorts for the African American and Native American races. At the same time, the rank remained the same for the remaining races. For the healthcare low-utilizers, BMI and smoking were the top two variables positively predicting CCI in the African American race. The third was unknown marital status which negatively impacted CCI prediction. Age, smoking, and BMI remained the top three predictors for the remaining races in the low-utilizer cohorts.
We assessed the effect of BMI on the CCI score by displaying the AUC of the ROC curve for the adjusted model with and without BMI in the health utilizer cohort alongside the highest AUC of the adjusted model with BMI in the 20-38 range, as shown in **Figure 3**.

The assessment of the remaining utilization cohorts is shown in **Appendix Figures 1** and **2.** Considering all variables in the utilization cohorts, the model best predicted CCI for the Caucasian racial group, followed by that for the Asian/Pacific Islander racial group, with the CCI for the African American racial group being predicted least well by the model. The average BMI for the largest AUC in the healthcare utilizer cohorts was 34 for the African American cohort, 27 for the Asian/Pacific Islander cohort, 32 for the Caucasian cohort, and 35 for the Native American cohort.

Table 4. Odds Ratio Across Race and Healthcare Utilization.

Utilizers								
Variable/Race	**African American / 95% CI**		**Asian/Pacific Islander / 95% CI**		**Caucasian / 95% CI**		**Native American / 95% CI**	
BMI	5.04*	(4.54,5.58)	38.18*	(24.14, 60.62)	24.30*	(22.98,25.69)	3.82*	(2.28,6.42)
Age	1.82*	(1.72,1.91)	3.07*	(2.62,3.61)	2.01*	(1.95,2.06)	2.11*	(1.64,2.72)
Male Gender	1.55*	(1.50,1.60)	1.65*	(1.49,1.81)	1.60*	(1.57,1.62)	1.60*	(1.38,1.86)
Private Insurance	0.82*	(0.78,0.86)	0.58*	(0.49,0.67)	0.74*	(0.72,0.76)	0.77*	(0.62,0.96)
Public Insurance	1.20*	(1.14,1.26)	0.99	(0.83,1.20)	1.29*	(1.26,1.33)	1.21	(0.97,1.52)
Unknown Insurance	0.94*	(0.88,1.0)	0.86	(0.72,1.02)	0.94*	(0.91,0.97)	0.69*	(0.50,0.96)
Smoking	1.57*	(1.51,1.64)	1.70*	(1.33,2.18)	2.22*	(2.17,2.27)	1.74*	(1.43,2.12)
Single	1.06*	(1.03,1.10)	1.02	(0.89,1.17)	1.21*	(1.18,1.23)	0.99	(0.83,1.18)
Unknown Marital Status	0.71*	(0.63,0.81)	0.60*	(0.46,0.78)	0.63*	(0.60,0.67)	0.67	(0.42,1.05)
X-Partner	1.10*	(1.05,1.15)	1.17	(0.98,1.38)	1.28*	(1.25,1.30)	1.07	(0.86,1.32)
2014-2015 Utilization	1.49*	(1.44,1.54)	1.25*	(1.13,1.38)	1.24*	(1.22,1.26)	1.56*	(1.34,1.81)
Urban Hospital	1.27*	(1.23,1.31)	0.81*	(0.73,0.89)	0.93*	(0.91,0.95)	1.03	(0.88,1.21)
Non-Utilizers								
Variable/Race	**African American / 95% CI**		**Asian/Pacific Islander / 95% CI**		**Caucasian / 95% CI**		**Native American / 95% CI**	
BMI	9.59*	(8.48,10.86)	11.25*	(5.54,22.78)	23.66*	(21.76,25.73)	6.53*	(3.68,11.61)
Age	2.44*	(2.29,2.59)	3.00*	(2.26,4.01)	2.54*	(2.44,2.65)	2.50*	(1.89,3.31)
Male Gender	1.11*	(1.07,1.15)	1.42*	(1.19,1.69)	1.38*	(1.34,1.41)	1.00	(0.85,1.18)
Private Insurance	1.20*	(1.14,1.27)	0.81	(0.61,1.09)	0.94*	(0.90,0.97)	1.74*	(1.39,2.19)
Public Insurance	2.16*	(2.05,2.28)	1.53*	(1.14,2.08)	1.62*	(1.56,1.68)	2.31*	(1.85,2.91)
Unknown Insurance	0.97*	(0.95,1.05)	0.87	(0.63,1.21)	0.82*	(0.77,0.87)	0.99	(0.54,1.74)
Smoking	2.04*	(1.95,2.14)	2.13*	(1.58,2.85)	2.33*	(2.26,2.40)	1.85*	(1.52,2.24)
Single	0.93*	(0.89,0.97)	0.98	(0.78,1.22)	0.97*	(0.94,1.00)	1.00	(0.83,1.20)
Unknown Marital Status	0.76*	(0.65,0.88)	1.69	(0.95,2.88)	0.77*	(0.70,0.84)	0.87	(0.46,1.58)
X-Partner	1.06*	(1.00,1.12)	1.04	(0.77,1.39)	1.14*	(1.10,1.18)	0.96	(0.76,1.21)
2014-2015 Utilization	1.28*	(1.19,1.37)	1.62*	(1.12,2.32)	1.38*	(1.32,1.45)	1.16	(0.82,1.62)
Urban Hospital	1.35*	(1.29,1.41)	0.75*	(0.63,0.89)	1.03	(1.00,1.06)	1.22*	(1.02,1.46)
Low-Utilizers								
Variable/Race	**African American / 95% CI**		**Asian/Pacific Islander / 95% CI**		**Caucasian / 95% CI**		**Native American / 95% CI**	

BMI	6.29*	(5.18,7.63)	56.47*	(21.32,150.38)	23.96*	(21.45,26.76)	4.17*	(1.93,8.97)
Age	1.83*	(1.66,2.01)	3.25*	(2.32,4.56)	2.23*	(2.11,2.35)	2.10*	(1.41,3.13)
Male Gender	1.44*	(1.36,1.53)	1.52*	(1.23,1.87)	1.52*	(1.47,1.57)	1.16	(0.92,1.47)
Private Insurance	0.80*	(0.73,0.88)	0.66*	(0.47,0.95)	0.75*	(0.71,0.80)	0.99	(0.69,1.42)
Public Insurance	1.34*	(1.22,1.48)	1.49*	(1.01,2.22)	1.40*	(1.32,1.49)	1.69*	(1.19,2.43)
Unknown Insurance	0.76*	(0.67,0.86)	0.94	(0.64,1.39)	0.62*	(0.58,0.68)	0.58	(0.27,1.16)
Smoking	2.12*	(1.95,2.30)	2.91*	(1.82,4.6)	3.31*	(3.16,3.56)	2.51*	(1.85,3.39)
Single	0.97	(0.90,1.03)	1.39*	(1.05,1.83)	1.11*	(1.06,1.16)	0.92	(0.70,1.21)
Unknown Marital Status	0.50*	(0.40,0.61)	0.52*	(0.31,0.83)	0.57*	(0.51,0.63)	0.56	(0.24,1.17)
X-Partner	1.09	(1.00,1.18)	1.12	(0.74,1.65)	1.24*	(1.19,1.30)	1.13	(0.82,1.56)
2014-2015 Utilization	0.99	(0.93,1.06)	0.85	(0.65,1.12)	0.90*	(0.87,0.94)	0.96	(0.73,1.27)
Urban Hospital	1.48*	(1.39,1.58)	0.87	(0.71,1.07)	1.16*	(1.11,1.21)	1.12	(0.87,1.45)

***** Indicates significant P-value results

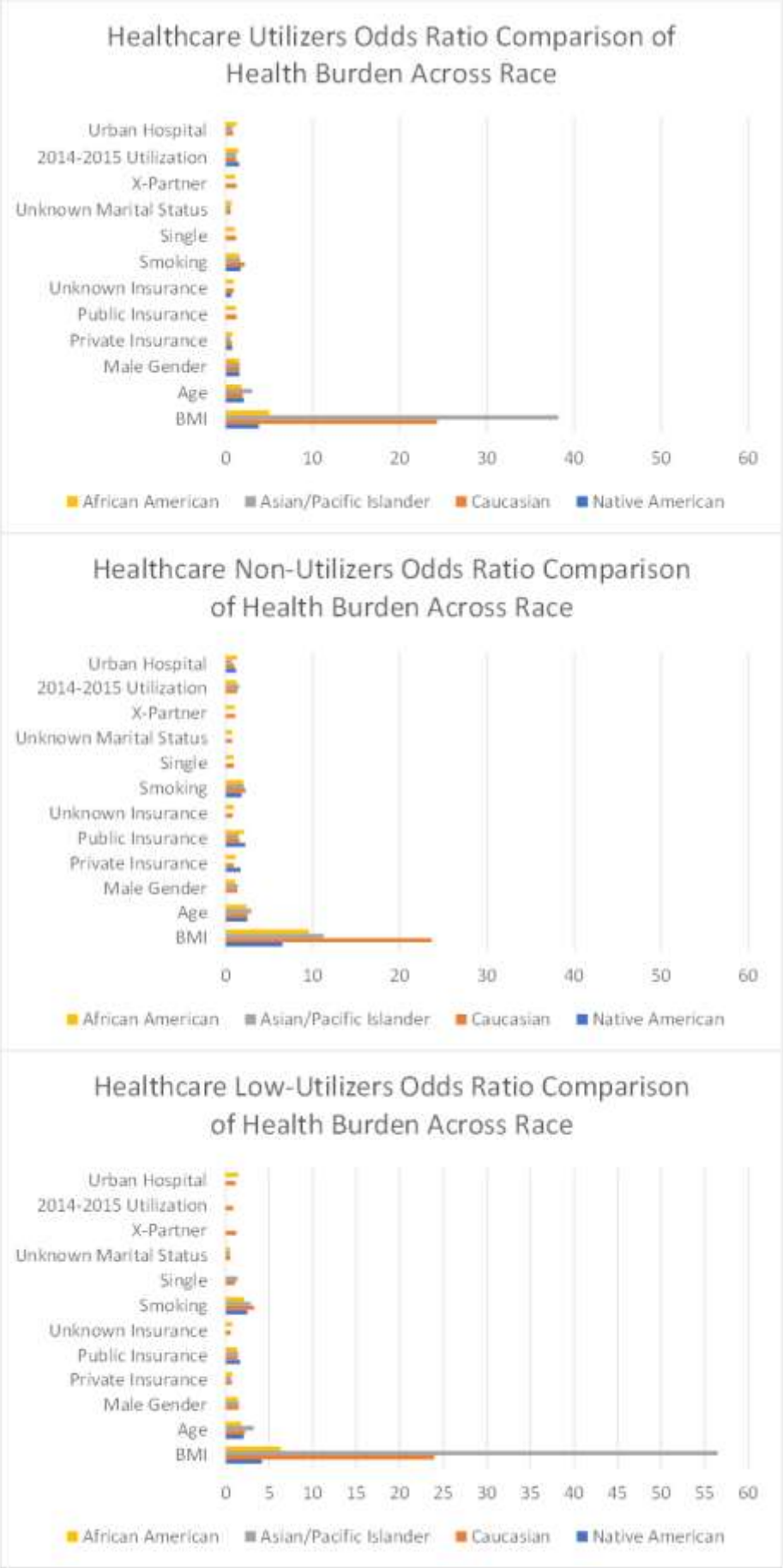


Figure 2. Relationship Between Healthcare Burden and Model Variables

Missing bars indicates non-significant P-value

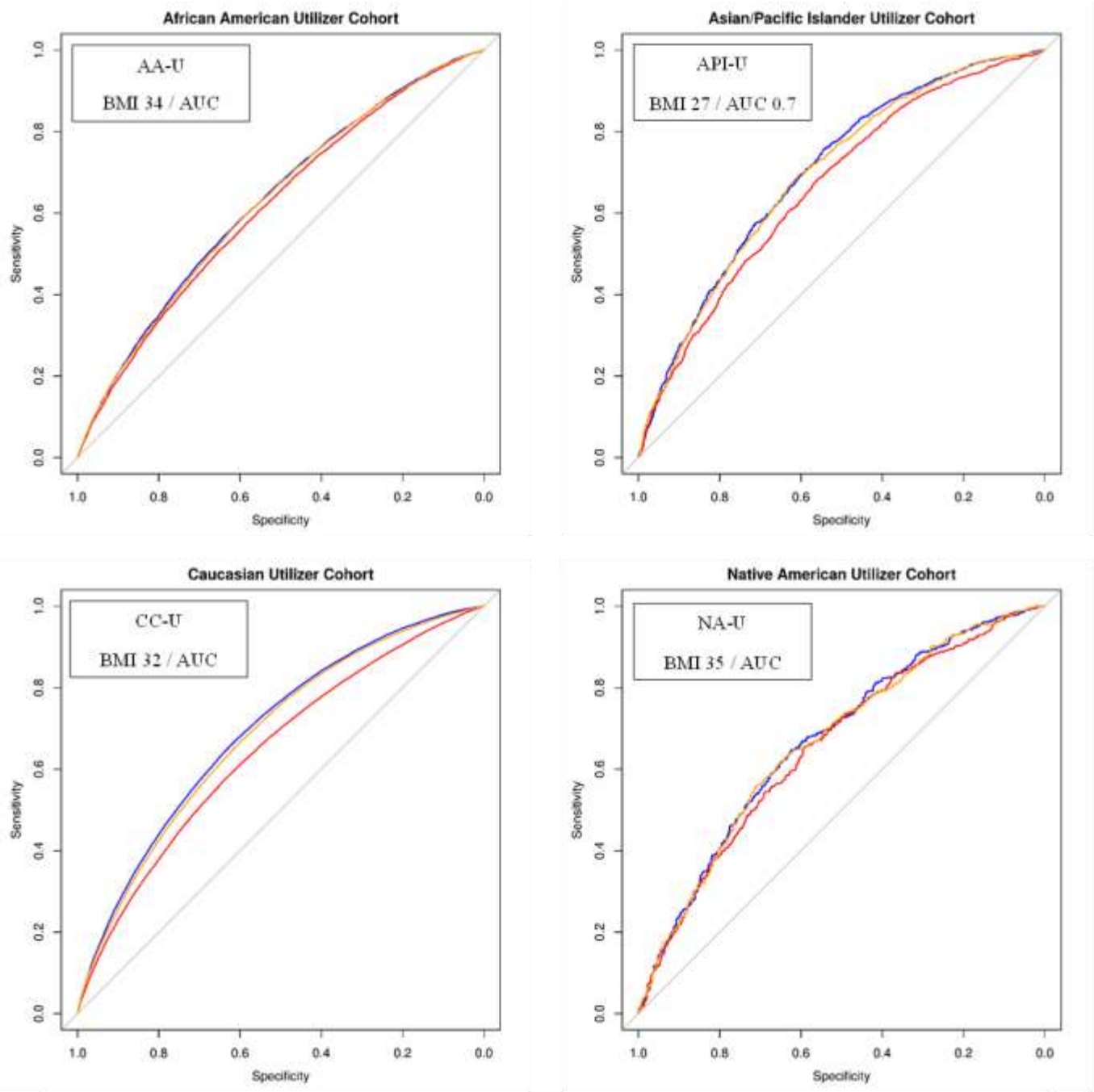


Figure 3. Comparing Area Under the Curve (AUC) for Different Models to Evaluate BMI Impact for the Healthcare Utilizer Cohort. The red curve represents the adjusted model without BMI, the blue curve represents the adjusted model with BMI, and the orange curve represents the highest AUC for the specified BMI cutoff.

Discussion

This study found that the relationship between BMI and CCI varied across race. The BMI values from the utilizer cohort with the highest AUC value represent a potentially alarming health risk. The variables impacting the prediction of CCI varied depending on the degree of healthcare utilization being analyzed. BMI and age were the two most important variables impacting the prediction of the CCI, regardless of healthcare utilization and race. These findings highlight the complexity of the relationship between BMI, healthcare burden, and healthcare utilization across race.

The estimated relationship between BMI and healthcare burden varied depending on the type of healthcare utilization being analyzed. The strength and direction of this may differ depending on the type and number of hospital visits. The relationship may also vary among different racial groups. Clearly it will be important to consider multiple factors when analyzing the complex relationship between BMI and healthcare burden.

This research can help evaluate the healthcare services provided to different utilization groups, identify disparities in access, quality, or outcome, and better understand how patients with various healthcare utilization patterns interact with the healthcare system. Identifying patient subgroups that may have different healthcare utilization patterns and different levels of access to healthcare can aid in treatment plans and resource allocation, identify factors associated with increased or decreased utilization, evaluate the effectiveness of interventions targeted at specific utilization subgroups, and identify potential areas for cost savings. Overall, this data can help researchers and healthcare providers to understand and address health disparities.[21–24]

Limitations. A cross-sectional design may not fully capture the relationship between BMI, healthcare burden, and healthcare utilization complexities and may not control for other important factors influencing healthcare utilization, such as genetics, lifestyle, and overall health. Many patients were excluded from the analysis due to missing data on race, sex, BMI, or age. This exclusion of data could have a disproportionate impact on specific groups, potentially leading to unintended bias in the results. We also limited our study to middle-aged patients, which means that our conclusions cannot be generalized to the whole population. Cerner HealthFacts had a limited number of variables that we were able to control for. Future work that includes more variables such as education, physical activity, stress levels, and diet and includes more younger and older age groups over a longer period of time could provide a more comprehensive approach for understanding the BMI-CCI relationship.

Conclusion

To our knowledge, this is the first study to assess the relationship between BMI and healthcare burden as stratified by race and healthcare utilization. This research demonstrated that the relationship between BMI and CCI varied across race within same healthcare care utilization cohorts. Some of this variation could be driven by access to healthcare resources. Most of the research regarding multimorbidity focused on a specific point in time and for a few multimorbidities at the same time. Understanding how multimorbidity accumulates over time across populations is not addressed. More work needs to be done to understand how multimorbidity, BMI, and healthcare burden associate across race.

References

1. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *The Lancet*. 2004;363(9403):157-163. https://doi.org/10/fm3fgw

2. He W, Li Q, Yang M, et al. Lower BMI cutoffs to define overweight and obesity in China: Chinese BMI Cutoffs. *Obesity*. 2015;23(3):684-691. https://doi.org/10/f3pg2s

3. Deurenberg-Yap M, Deurenberg P. Is a Re-evaluation of Who Body Mass Index Cut-off Values Needed? the Case of Asians in Singapore. *Nutrition Reviews*. 2003;61(suppl_5):S80-S87. https://doi.org/10/c4gmb5

4. Misra A. Ethnic-Specific Criteria for Classification of Body Mass Index: A Perspective for Asian Indians and American Diabetes Association Position Statement. *Diabetes Technol Ther*. 2015;17(9):667-671. https://doi.org/10.1089/dia.2015.0007

5. Wildman RP, Gu D, Reynolds K, Duan X, He J. Appropriate body mass index and waist circumference cutoffs for categorization of overweight and central adiposity among Chinese adults. *The American Journal of Clinical Nutrition*. 2004;80(5):1129-1136. https://doi.org/10.1093/ajcn/80.5.1129

6. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *The Lancet*. 2004;363(9403):157-163. https://doi.org/10.1016/S0140-6736(03)15268-3

7. Stanford FC, Lee M, Hur C. Race, Ethnicity, Sex, and Obesity: Is It Time to Personalize the Scale? *Mayo Clinic Proceedings*. 2019;94(2):362-363. https://doi.org/10.1016/j.mayocp.2018.10.014

8. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of Obesity and Trends in Body Mass Index Among US Children and Adolescents, 1999-2010. *JAMA*. 2012;307(5):483-490. https://doi.org/10.1001/jama.2012.40

9. Obesity and overweight. Accessed February 9, 2023. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

10. Shiroma EJ, Lee IM. Physical activity and cardiovascular health: lessons learned from epidemiological studies across age, gender, and race/ethnicity. *Circulation*. 2010;122(7):743-752. https://doi.org/10.1161/CIRCULATIONAHA.109.914721

11. Hunger JM, Smith JP, Tomiyama AJ. An Evidence-Based Rationale for Adopting Weight-Inclusive Health Policy. *Social Issues and Policy Review*. 2020;14(1):73-107. https://doi.org/10.1111/sipr.12062

12. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: A Critical Review of Clinimetric Properties. *PPS*. 2022;91(1):8-35. https://doi.org/10.1159/000521288

13. Concept: Charlson Comorbidity Index. Accessed November 28, 2022. http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?printer=Y&conceptID=1098

14. Jassal SV, Schaubel DE, Fenton SSA. Baseline Comorbidity in Kidney Transplant Recipients: A Comparison of Comorbidity Indices. *American Journal of Kidney Diseases*. 2005;46(1):136-142. https://doi.org/10.1053/j.ajkd.2005.03.006

15. Logistic Regression. Accessed August 21, 2021. https://www.sheffield.ac.uk/polopoly_fs/1.233565!/file/logistic_regression_using_SPSS_level1_MASH.pdf

16. 1.3.5.16. Kolmogorov-Smirnov Goodness-of-Fit Test. Accessed February 2, 2023. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm

17. Alin A. Multicollinearity. *WIREs Computational Statistics*. 2010;2(3):370-374. https://doi.org/10.1002/wics.84

18. Allison P. When Can You Safely Ignore Multicollinearity? Statistical Horizons. Published September 10, 2012. Accessed November 14, 2022. https://statisticalhorizons.com/multicollinearity/

19. Fox J, Monette G. Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*. 1992;87(417):178-183. https://doi.org/10.2307/2290467

20. Fox J, Weisberg S. *An R Companion to Applied Regression*. SAGE Publications; 2018.

21. National Academies of Sciences E, Division H and M, Services B on HC, Disabilities C on HCU and A with. *Factors That Affect Health-Care Utilization*. National Academies Press (US); 2018. Accessed January 21, 2023. https://www.ncbi.nlm.nih.gov/books/NBK500097/

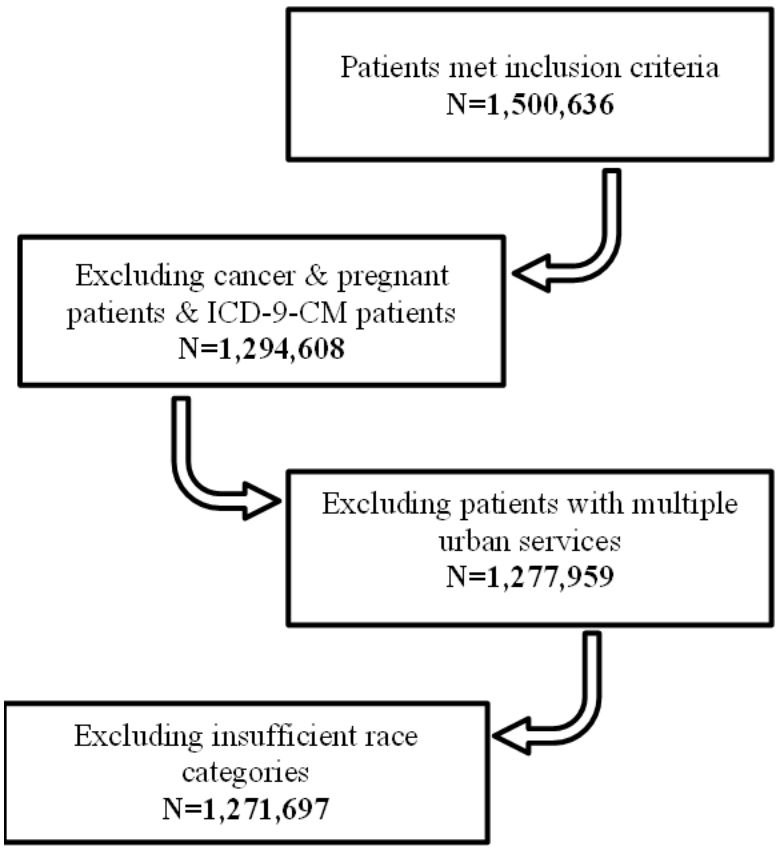
22. Dashputre AA, Surbhi S, Podila PSB, Shuvo SA, Bailey JE. Can primary care access reduce health care utilization for patients with obesity-associated chronic conditions in medically underserved areas? *J Eval Clin Pract*. Published online February 20, 2020. https://doi.org/10.1111/jep.13360

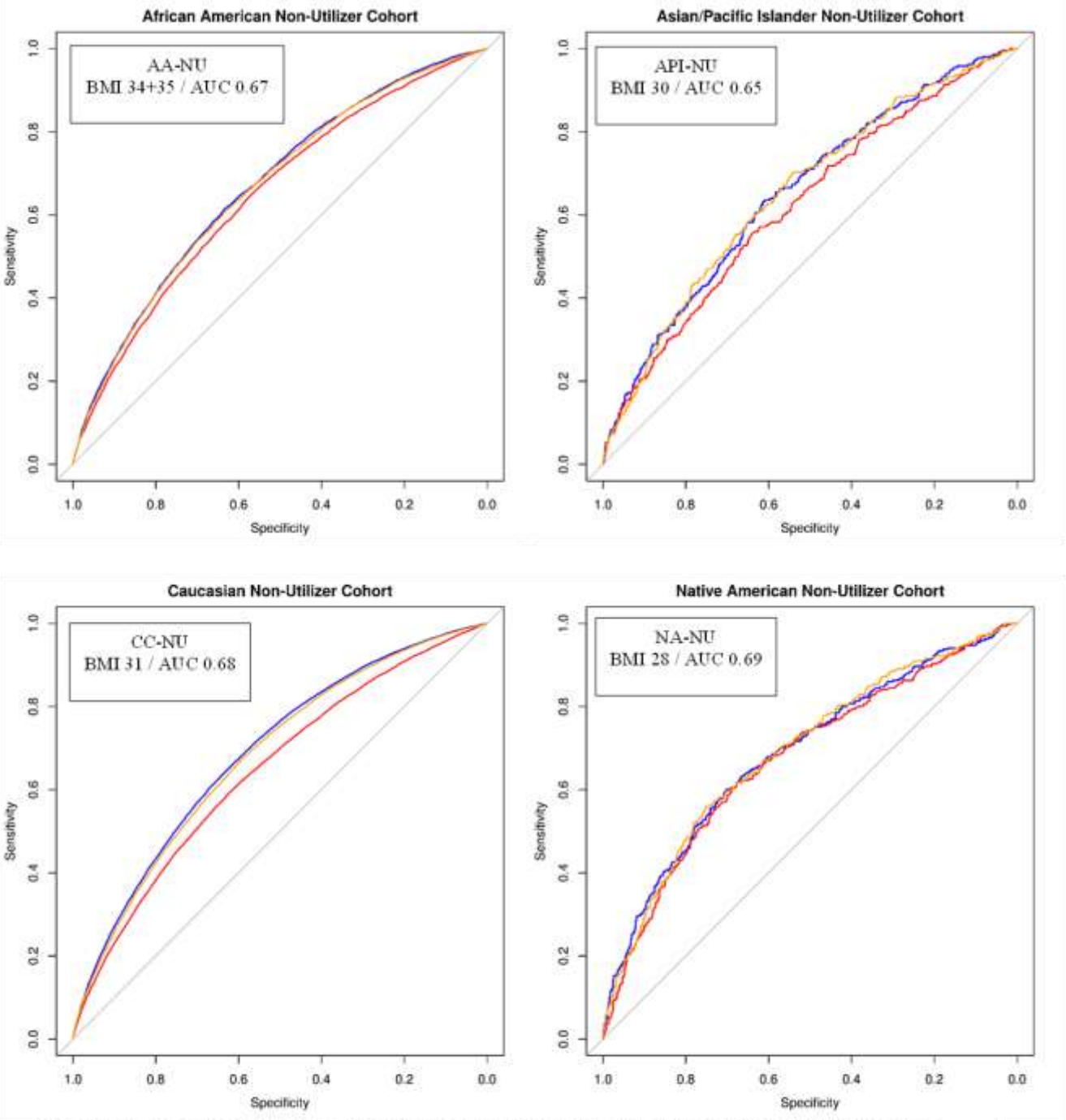
23. McCormack LA, Jones SG, Coulter SL. Demographic factors influencing nonurgent emergency department utilization among a Medicaid population. *Health Care Manag Sci*. 2017;20(3):395-402. https://doi.org/10.1007/s10729-016-9360-8

24. Glynn LG, Valderas JM, Healy P, et al. The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Family Practice*. 2011;28(5):516-523. https://doi.org/10.1093/fampra/cmr013

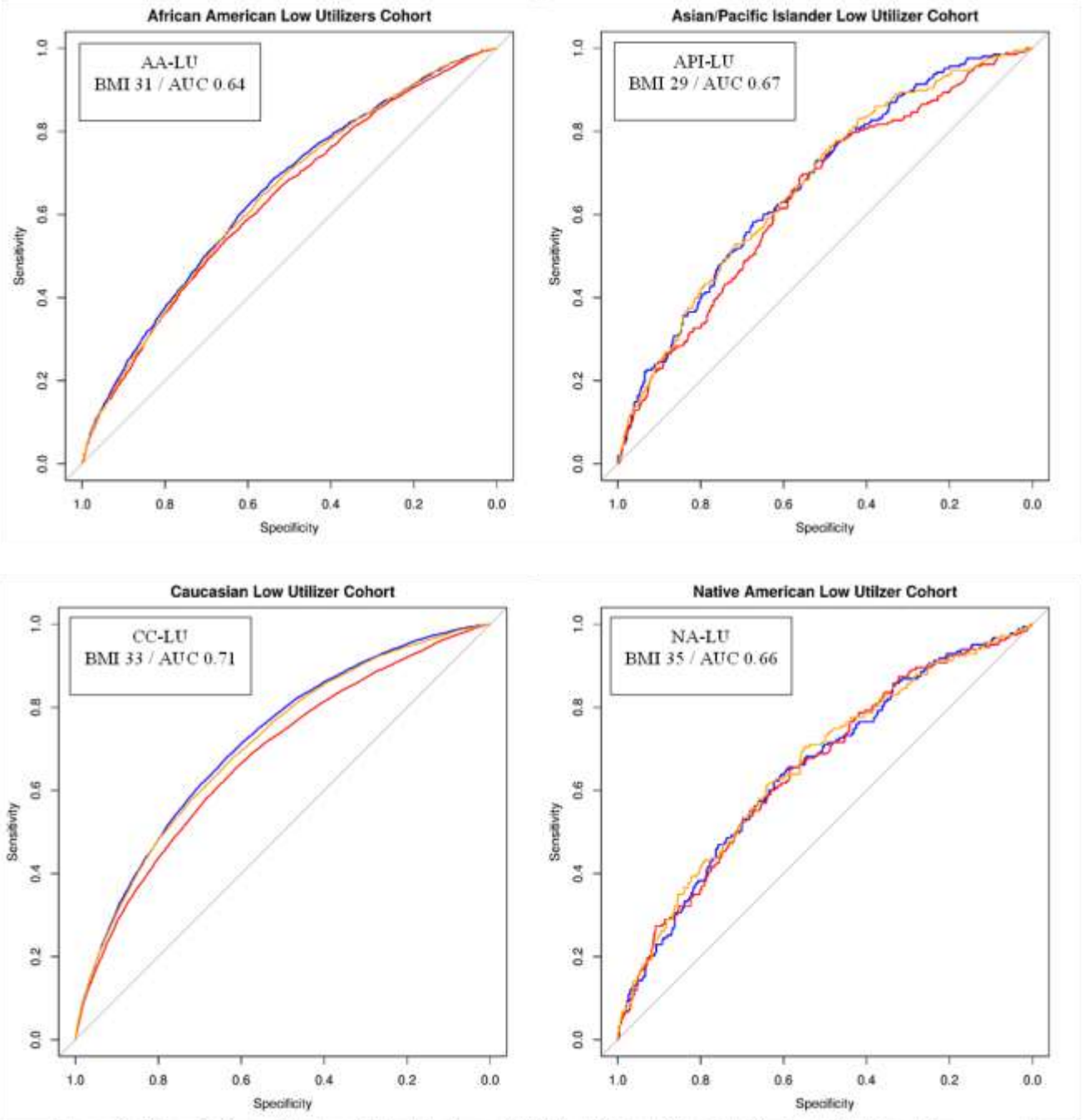
Appendix

Appendix Table 1. Patient





Appendix Figure 1. Comparing Area Under the Curve (AUC) for Different Models to Evaluate BMI Impact for Healthcare Non-Utilizer Cohorts. The red curve represents the adjusted model without BMI, the blue curve represents the adjusted model with BMI, and the orange curve represents the highest AUC for the specified BMI cutoff.



Appendix Figure 2. Comparing Area Under the Curve (AUC) for Different Models to Evaluate BMI Impact for Healthcare Low-Utilizer Cohorts. The red curve represents the adjusted model without BMI, the blue curve represents the adjusted model with BMI, and the orange curve represents the highest AUC for the specified BMI cutoff.

VITA

Manal Jawad Alshakhs was born in Alkohbar, Saudi Arabia, in 1971. She received her BS in Mathematics in 1995 from Syracuse University, Syracuse, New York. Her MS was in Information Resources Management which she got from the same university in 1997. In 1999, she started her teaching job as a lecturer at Imam Abdulrahman Bin Faisal University in Dammam, Saudi Arabia. She taught in many colleges there, from first-year students to graduate students. With her social skills, initiative, and passion for serving her community, her work was not limited to teaching. She held various administrative and academic positions, including College Coordinator and Department Coordinator. She also initiated and co-chaired one of the largest events in the University's history, The Career Forum. She always participated in many university-level events as a lead organizer throughout the years. In 2017, she got accepted at The University of Tennessee Health Science Center (UTHSC) to pursue a lifelong dream. She started the Health Informatics Ph.D. track at the Health Outcomes and Policy Research program. Manal's research interest was understanding the relationship between obesity and multimorbidity across races. Her research was mentored by Dr. Charisse Madlock Brown, along with committee members Dr. Jim Bailey, Dr. Simonne Nouer, Dr. Rebecca Reynolds, and Dr. Shelley White-Means. During her studies at UTHSC, she published one paper and is working on the final touches to publish two more where she is the primary investigator. She also presented her work at The American Medical Informatics Association (AMIA), Graduate Research Day, 3M Thesis, and several other virtual events. While at UTHSC, she founded the first Saudi Students Association on campus in 2019 and was the Treasure and Vice President of the International Students Association as well 2020-2022. She received The Imhotep Society award for her contribution to student life on campus and The Outstanding Leadership and Dedication award for her role as co-chair in organizing the 2019 UTHSC Biomedical Symposium. Manal is expected to get her Doctorate of Philosophy degree in April, 2023.