



OPEN

Multimorbidity patterns across race/ethnicity as stratified by age and obesity

Manal Alshakhs¹, Bianca Jackson¹, Davina Ikponmwosa¹, Rebecca Reynolds^{1,2,3} & Charisse Madlock-Brown^{1,2,3}✉

The objective of our study is to assess differences in prevalence of multimorbidity by race/ethnicity. We applied the FP-growth algorithm on middle-aged and elderly cohorts stratified by race/ethnicity, 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 stratifying by age and obesity, there are differences in multimorbidity prevalence across races/ethnicities. There are multimorbidity combinations distinct to some racial groups—many of which are understudied. Some multimorbidities are shared by some but not all races/ethnicities. 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.

Multimorbidities become more prevalent as individuals age, and they have been associated with substantial burden and increased mortality¹. The treatment of patients with multimorbidity can be complicated by the involvement of multiple medical specialties². These patients are also at risk of repeated hospitalization, polypharmacy, adverse drug events, and increased care dependence³. 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 patterns⁴. Identifying homogeneous multimorbidity subgroups has been theorized as a way to develop and target interventions more effectively⁵. Studies have also shown that specific racial and ethnic groups are at greater risk of poor health outcomes⁶.

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 combinations⁷. Further, providers should be aware of the most likely multimorbidity combinations to develop and coordinate appropriate care plans when treating patients with these disease groupings⁸. It is crucial to identify specific multimorbidity patterns because the impact of multimorbidity on health-related quality of life varies for different combinations⁹.

Although previous studies have identified factors that may impact multimorbidity, most relevant studies were not conducted in the US, and few studies account for obesity or age^{2,10}. Additionally, there remains a need to identify homogeneous combinations among race/ethnicities. Most previous research identified multimorbidity patterns using counts or cluster analysis and are disease specific¹¹. 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.

¹Health Outcomes and Policy Program, University of Tennessee Health Science Center, Memphis, TN, USA. ²Health Informatics and Information Management Program, University of Tennessee Health Science Center, 66 North Pauline St. rm 221, Memphis, TN 38163, USA. ³Center for Health System Improvement, University of Tennessee Health Science Center, Memphis, TN, USA. ✉email: cmadlock@uthsc.edu

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 interventions¹⁶.

Several studies have identified the prevalence of specific multimorbidity patterns for clusters of two or more diseases. Our previous work identified the most prevalent multimorbidity patterns of two or more diseases for a cohort with obesity¹⁷. Held et al. used association rule analysis mining to find dyads and triads of diseases for a cohort of men of older age in Australia¹⁸. Van den Bussche et al. identified combinations of 3 multimorbidities in a sample of German elderly patients¹⁹. While these studies shed considerable light on how diseases tend to combine, they are limited in that they include only special populations 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 Asians⁶. Race is also associated with faster rates of multimorbidity development. One study demonstrated that African Americans developed multimorbidity about four years earlier than Caucasians in a middle-aged adult population⁶. 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 60–64². Differences in multimorbidity patterns among men and women have also been observed in a Spanish population of adults over 65 years of age^{2,21}. Further, a cross-sectional study conducted in Singapore reported an association between multimorbidity and increasing age, lower socio-economic status, female sex, and mental disorders¹⁰.

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 vulnerability²⁶, 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 population²⁵.

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 multimorbidity¹⁸, 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 conditions²⁸, 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 health-care use and access vary by race²⁹, 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 Institutional Review Board (IRB) approval for access. Because of the de-identified nature of the data, this study is not considered human subject research. The University of Tennessee Health Science Center IRB determined that this research qualifies for 'Not Human Subjects Research' status according to the policies of the National Institutes of Health Office of Human Subjects Research. 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 multimorbidity³⁰. We defined multimorbidity as the presence of two or more chronic conditions within one individual³¹. 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 conditions³³. 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 robust³⁴. 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 whose frequencies are 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 datasets³⁷.

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 $I = \{P_1, P_2, \dots, P_m\}$ 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 algorithm³⁹. 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 intervals⁴⁰. To compare prevalence rates by race, we used the g -test of

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% (5066)	2% (15,427)	84% (12,982)	16% (2445)
Biracial	< 1% (1220)	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% (3227)	59% (1895)	41% (1332)
Native American	1% (11,578)	38% (4443)	62% (7135)	< 1% (5662)	52% (2963)	43% (2699)

Table 1. Demographic breakdown across age and weight class.

Race/Ethnicity	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

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

independence. These statistical analyses were performed using the R DescTools package⁴¹. 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 test⁴². 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 middle-aged 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 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.

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 (Fig. 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).

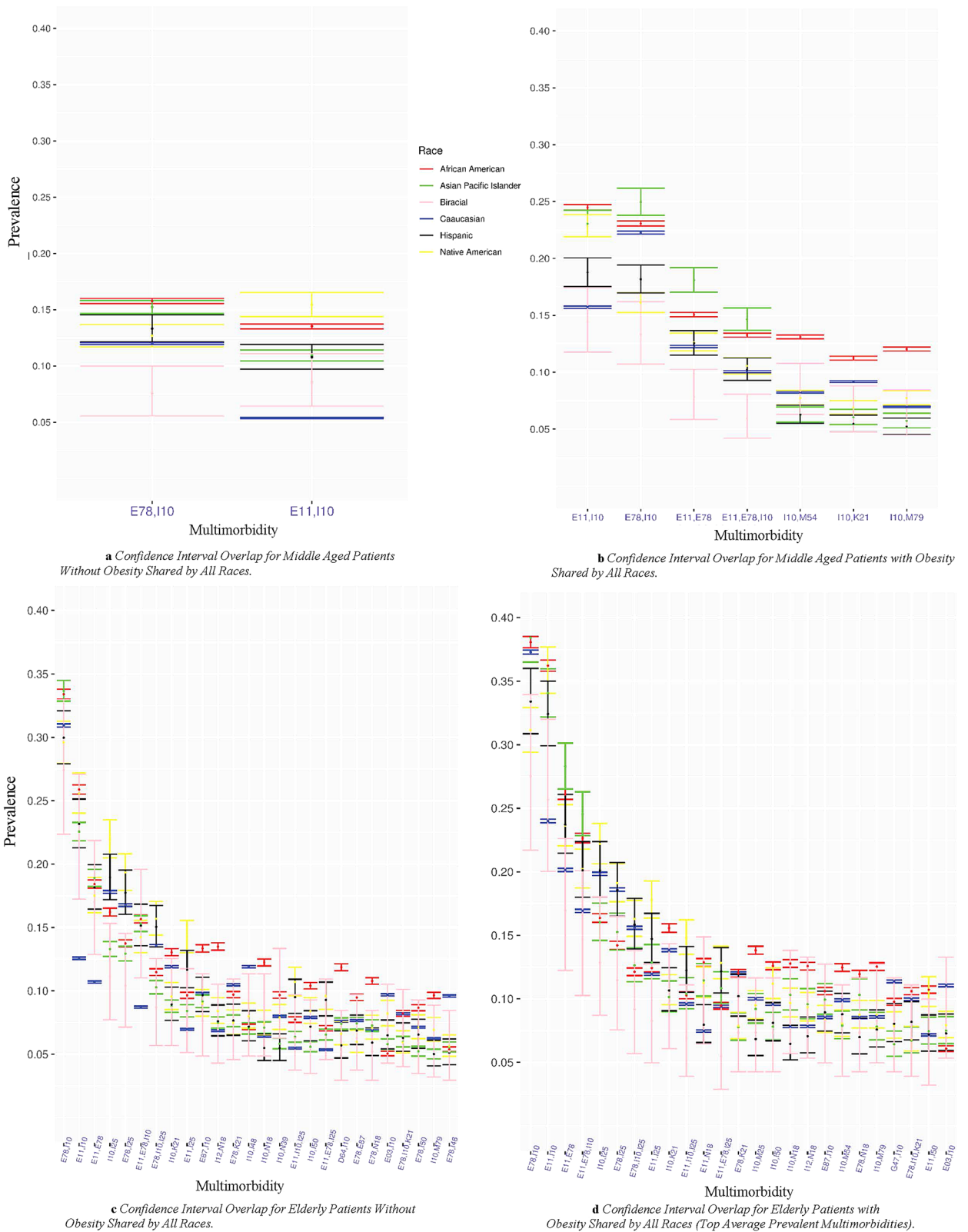


Figure 1. (a,b) Demonstrate the 95% confidence interval overlap for shared multimorbidities among middle aged patients without obesity and patients with obesity, respectively, at 5% prevalence threshold. The g-test of independence was calculated resulting in all significant results with $p < 0.05$. (c,d) 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. (d) Demonstrates shared multimorbidities with average prevalence ≥ 0.08 .

For five of the multimorbidity patterns observed for middle-aged patients with obesity (Fig. 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 Fig. 1c,d and Supplementary Fig. 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 Fig. 1d to only those with prevalence equal to or above 0.08%. In the without obesity cohort (Fig. 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 (Fig. 1d and Supplementary Fig. 1), 25 multimorbidity patterns were present. The following diagnoses appeared in 3 or more multimorbidity patterns: I10: Hypertension (19 patterns), E78: Lipidemia (13 patterns), 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.

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 Fig. 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 (Fig. 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 (Fig. 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 prevalences of the F17: Nicotine dependence–I10: Hypertension combination than the Caucasian cohorts in both the with and without obesity middle-aged samples (Fig. 2a–d). In the middle-aged patients with obesity cohort (Fig. 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 (Fig. 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 (Fig. 2b). The only two combinations present for the Asian/Pacific Islander without obesity (Fig. 2a) include both diabetes E11: Diabetes and E78: Lipidemia.

Results for the elderly cohorts appear in Fig. 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 Fig. 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 Fig. 2, and those for the with obesity appear in Supplementary Fig. 3. No new clinical categories are represented. Combinations in the cohort without obesity (Fig. 2c and Supplementary Fig. 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 (Fig. 2c and Supplementary Fig. 2), there are 11 patterns comprised of three morbidities and eight comprised of four morbidities. The patients in the elderly with obesity cohort (Fig. 2d and Supplementary Fig. 3) exhibited 46 multimorbidity combinations comprised of three morbidities and five comprised of four morbidities. In the cohort without obesity, the three most prevalent patterns comprised three diagnoses (Fig. 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 (Fig. 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 Fig. 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 (Fig. 2c) and for all except Asian/Pacific Islander in the elderly with obesity patients (Fig. 2d). F17: Nicotine dependence appeared in the African American and Native American elderly cohort without obesity (Fig. 2c), yet it did not appear in any shared pattern in elderly patients with obesity (Fig. 2d).

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 (Fig. 3a–d). Gray sectors indicate the



Figure 2. (a,b) Show shared multimorbidities across two or more races in middle aged patients without and with obesity, respectively, at 5% prevalence threshold. 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. (c,d) 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.

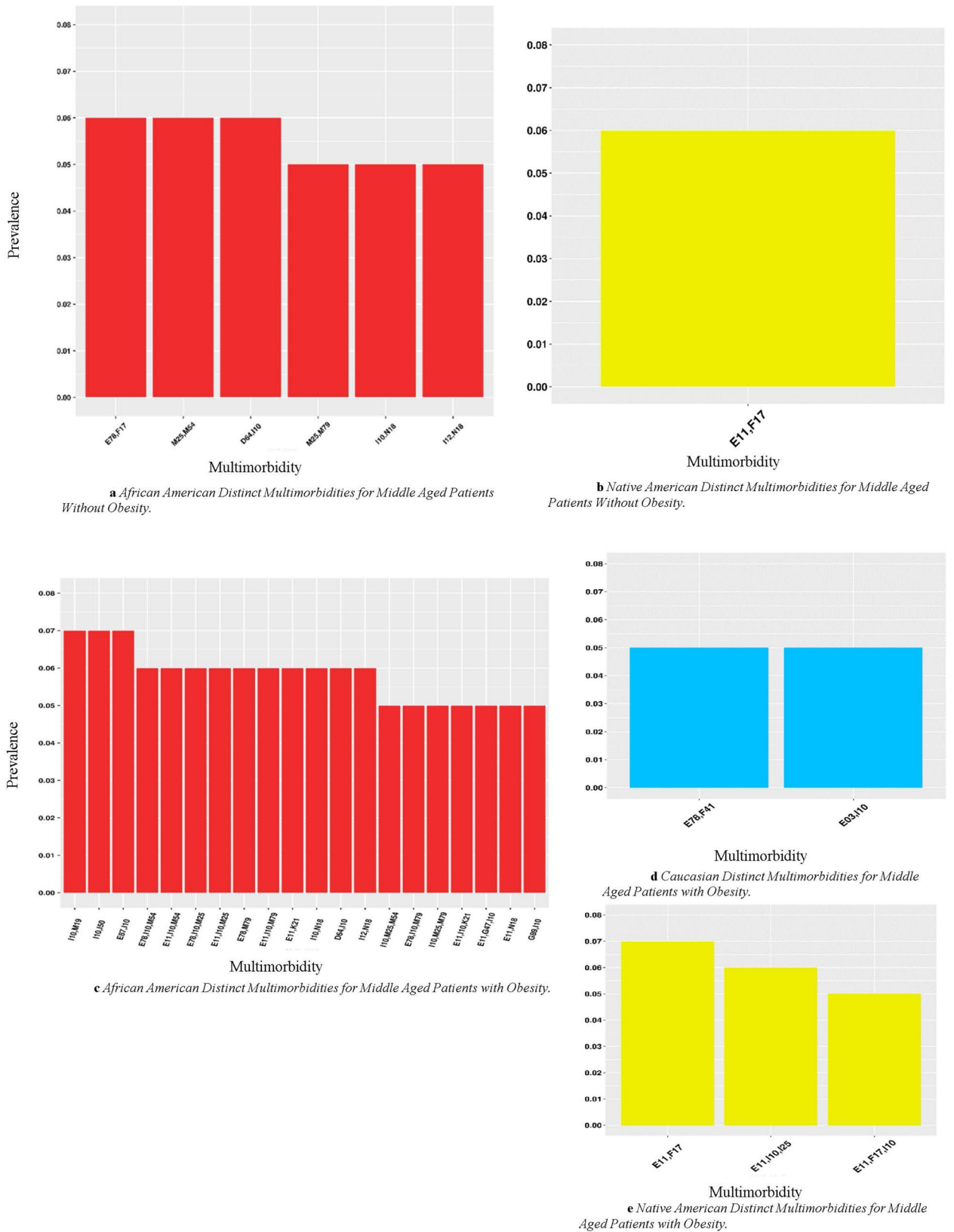
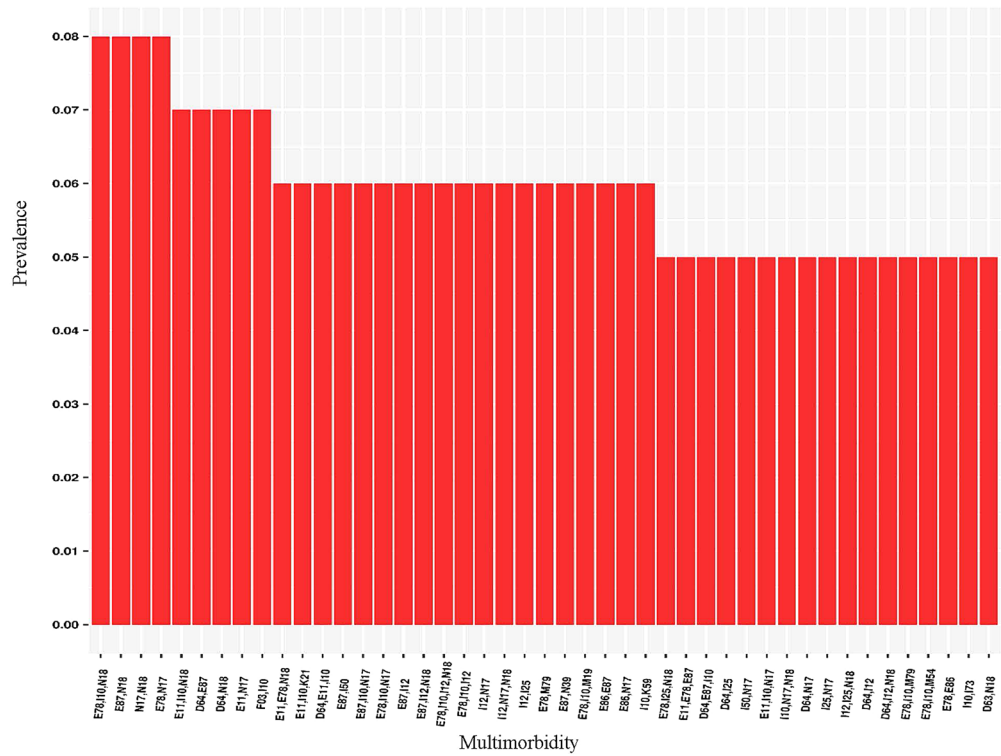
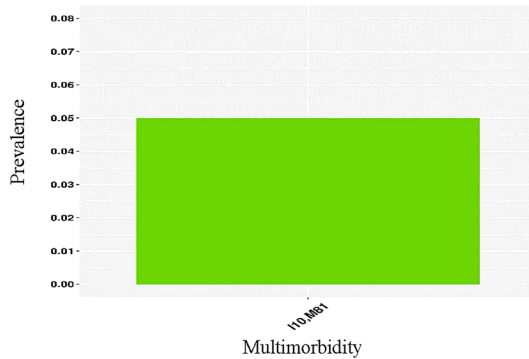


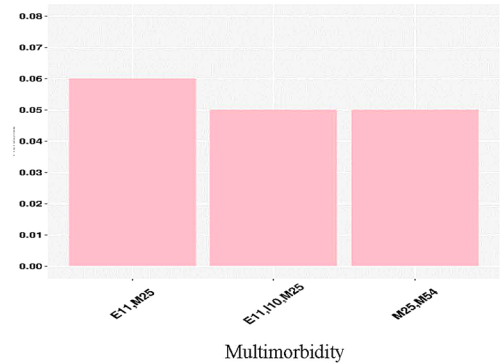
Figure 3. (a,b) Demonstrates distinct multimorbidities across races in middle-aged without obesity at 5% prevalence threshold. Gray sectors indicate the combination was not prevalent above the 5% threshold for a given group. (c–e) Demonstrate the distinct multimorbidities for each race in middle aged 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 $p < 0.05$.



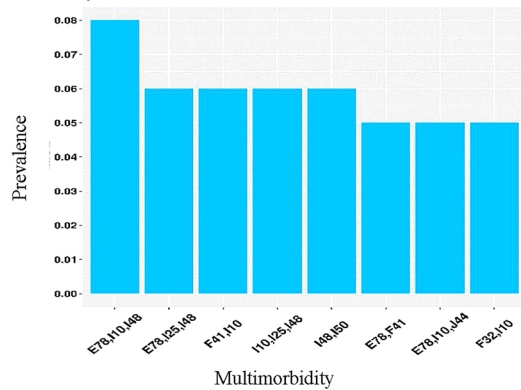
a Distinct Multimorbidities for African American Elderly Cohort Without Obesity.



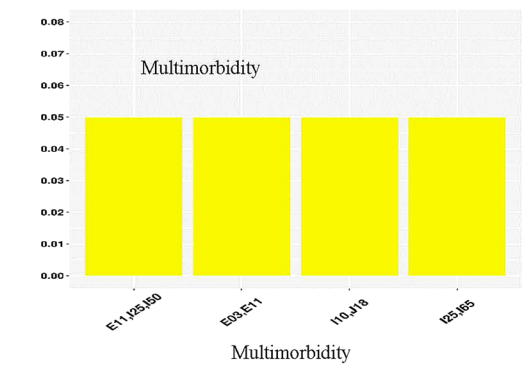
b Distinct Multimorbidities for Asian/Pacific Islander Elderly Cohort Without Obesity.



c Distinct Multimorbidities for Biracial Elderly Cohort Without Obesity.



d Distinct Multimorbidities for Caucasian Elderly Cohort Without Obesity.



e Distinct Multimorbidities for Native American Cohort Without Obesity.

Figure 4. (a–e) 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 $p < 0.05$.

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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 3e). For the middle-aged cohort, the Caucasian group had two distinct patterns among the patients with obesity (Fig. 3d). Both include a diagnosis in the E: Endocrine, nutritional and metabolic. One includes F41: Other anxiety disorders.

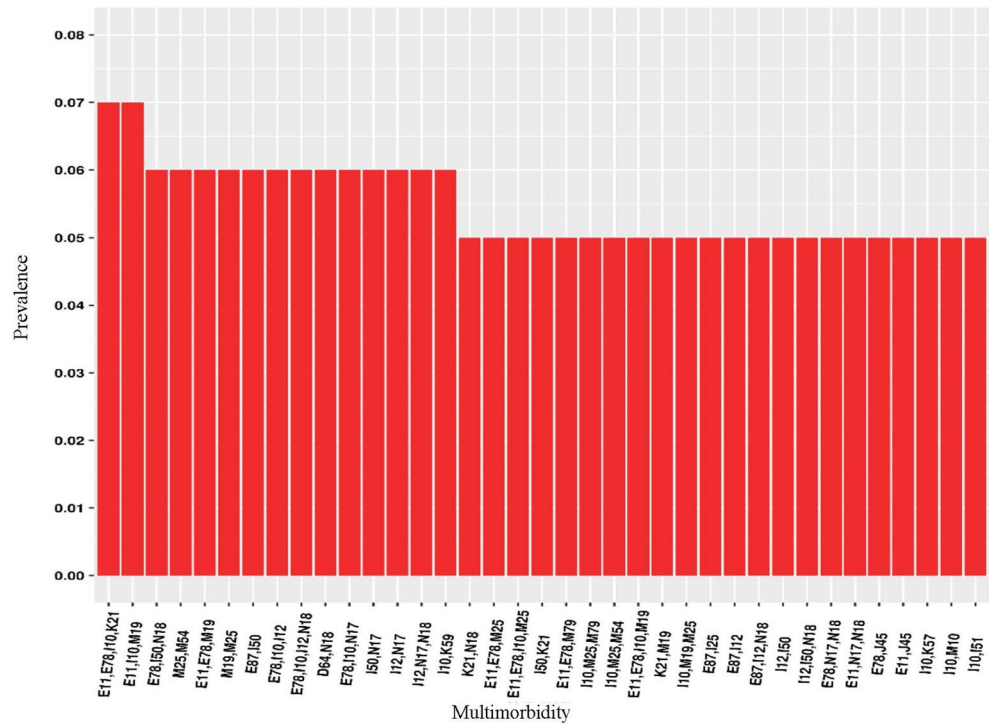
In the elderly, the African American cohort without obesity (Fig. 3c) had many combinations that combined with I10: Hypertension, E78: Lipidemia, or E11: Diabetes, including N18: Chronic kidney disease (CKD), I12: Hypertensive chronic kidney disease, E87: Other disorders of fluid, electrolyte and acid–base balance, F17: Nicotine dependence, K21: GERD, D64: Other anemias, and I50: Heart failure. In the cohort without obesity (Fig. 4a), there are 21 multimorbidities comprised of 3 or 4 diseases. In the cohort with obesity (Fig. 4a), there are 21 multimorbidities comprised of three or four diseases and 19 in the cohort with obesity (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 (Fig. 4d), there are four multimorbidities comprised of three or four diseases as five in the cohort with obesity (Fig. 5a). For the Asian/Pacific Islander, without obesity group (Fig. 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 (Fig. 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 (Fig. 4c), all patterns include musculoskeletal disorders (those starting with M). The only distinct pattern for the Biracial cohort with obesity (Fig. 5c) includes E03: Other hypothyroidism. In the Native American without obesity cohort (Fig. 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 (Fig. 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 (Fig. 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 (Table 3).

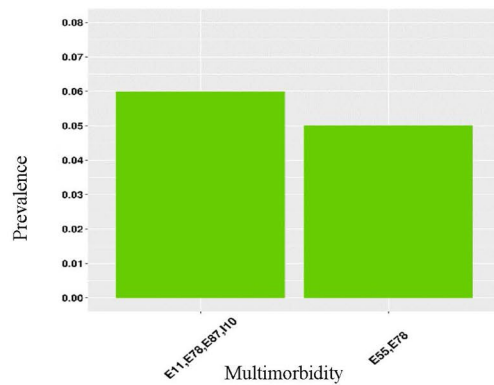
Discussion

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 polypharmacy⁴³. These findings can be used to inform public policy and to develop patient guidelines at various obesity levels and ages.

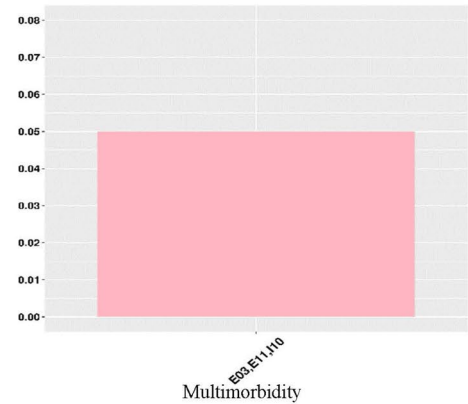
Our study observed differences in multimorbidity prevalence and composition between the middle-aged and elderly groups and between with and without obesity cohorts. Unique diseases and disease combinations were shared by all race/ethnic groups within the elderly cohorts that weren't shared by all in the middle-aged. For



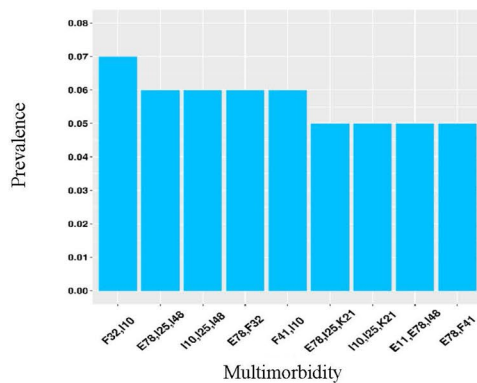
a Distinct Multimorbidities for African American Elderly Cohort with Obesity.



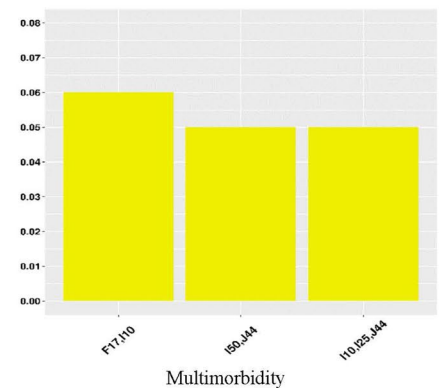
b Distinct Multimorbidities for Asian/Pacific Islander Elderly Cohort with Obesity.



c Distinct Multimorbidities for Biracial Elderly Cohort with Obesity.



d Distinct Multimorbidities for Caucasian Elderly Cohort with Obesity.



e Distinct Multimorbidities for Native American Elderly Cohort with Obesity.

Figure 5. (a–e) Demonstrate distinct multimorbidities across races in elderly patients with 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 $p < 0.05$.

Race/Ethnicity	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

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

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 (Fig. 1c,d). In addition, these diseases were frequently paired with either I10: Hypertension or E78: Lipidemia, which is consistent with findings from other studies⁴⁴. 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 (Fig. 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 combinations⁴⁴. 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 studies⁶. 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 age⁴⁷.

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 related⁴⁷. Further, these diseases were present in many patterns for various racial groups.

Another combination distinct to elderly, African Americans without obesity (Fig. 3c) was E87: Other disorders of fluid, electrolyte, and acid–base balance and I10: Hypertension. Alterations in acid–base transporters have been linked to hypertension⁴⁸. Converging evidence indicates a pathogenic role of combined high sodium and low potassium levels in the development of hypertension and hypertension-associated cardiovascular complications⁴⁹. 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 nonsmokers⁵¹.

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 Caucasians⁵². 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 hypertension⁵³. 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 smoking⁵⁴. 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 mortality⁵⁵. Additionally, challenges related to managed care communication between provider and patient is more significant for patients with mental health diagnoses⁵⁶.

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 group²⁰. 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-Whites⁵⁷.

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 leave⁵⁸. Therefore, despite Hispanics being of low socioeconomic status, studies observe Hispanics experience similar or better health outcomes than non-Hispanic Whites⁵⁹. 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 races⁶⁰. It is suggested that there is a bidirectional relationship between multimorbidity severity and sociodemographic and lifestyle risk factors⁶¹.

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 Fig. 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 pain⁶⁴. Musculoskeletal disorders are also commonly seen in patients with Type 1 and Type 2 diabetes⁶⁵. 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 approaches⁶⁶. 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 costs⁶⁷, 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 middle-aged 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 categories⁶⁸. 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.

Received: 19 October 2021; Accepted: 12 May 2022

Published online: 11 June 2022

References

- 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).
- 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).
- 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).
- 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).
- Kronick, R., Bella, M., Gilmer, T. & Somers, S. *The Faces of Medicaid III: Recognizing the Care Needs of People with Multiple Chronic Conditions* (Center for Health Care Strategies Inc, 2021).
- Quiñones, A. R. *et al.* Racial/ethnic differences in multimorbidity development and chronic disease accumulation for middle-aged adults. *PLoS One*. **14**(6), e0218462 (2019).
- Muth, C. *et al.* The Ariadne principles: How to handle multimorbidity in primary care consultations. *BMC Med.* **8**(12), 223 (2014).
- Sturgiss, E. A., Elmitt, N., Haelsler, 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).
- 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).
- 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).
- 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).
- 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).
- 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).
- Chong, J. L. & Matchar, D. B. Benefits of population segmentation analysis for developing health policy to promote patient-centred care. *Ann. Acad. Med. Singap.* **46**(7), 287 (2017).
- Olson, J. E., Takahashi, P. Y. & Sauver, J. M. S. Understanding the patterns of multimorbidity. *Mayo Clin Proc.* **93**(7), 824–825 (2018).
- 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).
- Madlock-Brown, C. & Reynolds, R. B. Identifying obesity-related multimorbidity combinations in the United States. *Clin. Obes.* **9**(6), e12336 (2019).
- 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).
- 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).
- 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).
- Guisado-Clavero, M. *et al.* Multimorbidity patterns in the elderly: A prospective cohort study with cluster analysis. *BMC Geriatr.* **18**(1), 16 (2018).
- 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).
- 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).
- 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).
- 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).
- 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).
- Leahy, S., Cassarino, M., O’Connell, M. D., Glynn, L. & Galvin, R. Dynapenic 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.guinnessworldrecords.com/world-records/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. Boedtker, 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. *Hispanic 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).

Author contributions

M.A.: analysis, methods, data collection, writing. B.J.: analysis, writing, literature review. D.I.: writing, literature review. R.R.: writing, literature review, critical review. C.M.B.: conception, design, supervision, critical review.

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.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-13733-w>.

Correspondence and requests for materials should be addressed to C.M.-B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022