ORIGINAL ARTICLE

Epidemiology/Genetics

Associations between cardiometabolic disease severity, social determinants of health (SDoH), and poor COVID-19 outcomes

Carrie R. Howell[1] | **Li Zhang[2]** | **Nengjun Yi[2]** | **Tapan Mehta[3]** |
Andrea L. Cherrington[1] | **W. Timothy Garvey[4]**

[1]Division of Preventive Medicine, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA

[2]Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA

[3]Department of Health Services Administration, School of Health Professions, University of Alabama at Birmingham, Birmingham, Alabama, USA

[4]Department of Nutrition Sciences, School of Health Professions, University of Alabama at Birmingham, Birmingham, Alabama, USA

Correspondence
Carrie R. Howell, Division of Preventive Medicine, Department of Medicine, University of Alabama at Birmingham, Medical Towers 638, 1717 11th Avenue South, Birmingham AL 35205, USA.
Email: chowell@uabmc.edu

Abstract

Objective: This study aimed to determine the ability of retrospective cardiometabolic disease staging (CMDS) and social determinants of health (SDoH) to predict COVID-19 outcomes.

Methods: Individual and neighborhood SDoH and CMDS clinical parameters (BMI, glucose, blood pressure, high-density lipoprotein, triglycerides), collected up to 3 years prior to a positive COVID-19 test, were extracted from the electronic medical record. Bayesian logistic regression was used to model CMDS and SDoH to predict subsequent hospitalization, intensive care unit (ICU) admission, and mortality, and whether adding SDoH to the CMDS model improved prediction was investigated. Models were cross validated, and areas under the curve (AUC) were compared.

Results: A total of 2,873 patients were identified (mean age: 58 years [SD 13.2], 59% were female, 45% were Black). CMDS, insurance status, male sex, and higher glucose values were associated with increased odds of all outcomes; area-level social vulnerability was associated with increased odds of hospitalization (odds ratio: 1.84, 95% CI: 1.38-2.45) and ICU admission (odds ratio 1.98, 95% CI: 1.45-2.85). The AUCs improved when SDoH were added to CMDS ($p < 0.001$): hospitalization (AUC 0.78 vs. 0.82), ICU admission (AUC 0.77 vs. 0.81), and mortality (AUC 0.77 vs. 0.83).

Conclusions: Retrospective clinical markers of cardiometabolic disease and SDoH were independently predictive of COVID-19 outcomes in the population.

INTRODUCTION

End-stage manifestations of cardiometabolic disease (CMD), such as obesity and type 2 diabetes mellitus (T2DM), are associated with poor outcomes in those diagnosed with COVID-19 (1,2), including increased mortality rates (3). CMD begins with insulin resistance, progresses to clinically identifiable high-risk states of metabolic syndrome and prediabetes, and culminates in overt T2DM and cardiovascular disease. Obesity exacerbates insulin resistance and impels progression of this chronic disease process. Not surprisingly, metabolic syndrome (4) and cardiovascular disease (5) have also been

identified as risk factors for poor COVID-19 outcomes. This suggests that poor prognosis in patients with COVID-19 may vary based on an insulin resistance or cardiometabolic continuum, such that patients who have existing suboptimal metabolic and vascular clinical markers progress to worse outcomes.

Of concern, CMD, obesity, and T2DM disproportionately impact racial and ethnic minorities (6) and those with low socioeconomic status (SES) (6), with marked geographic variations in prevalence rates across the United States (7). A recent study of National Health and Nutrition Examination Survey data from 2011 to 2016 estimated diabetes prevalence of 20.4% for non-Hispanic Black adults, 22.1% for

Hispanic adults, and 19.1% for Asian adults compared with 12.1% for non-Hispanic White adults (8). Unfavorable neighborhood factors intersect with race and ethnicity in risk, compounding disparities. Black, Hispanic, and other minority neighborhoods are often characterized by risk factors such as lack of physical activity resources, poor food options, lower SES, and barriers to health care access (9), all associated with higher diabetes and obesity incidence (10,11). Similar disparities have emerged regarding COVID-19 outcomes, including racial (12), social (13), and geographic (13) disparities in incidence and mortality.

Population-based strategies that seek to understand health disparities regarding COVID-19 should consider clinical or biological entities that adversely affect outcomes (i.e., CMD) together with social determinants of health (SDoH), although combined analyses of these factors are rarely performed. Electronic medical records (EMR) of health care systems present an opportunity to integrate SDoH into risk prediction models and stratify populations based on clinical and SDoH parameters. Evaluating how SDoH data can be used in the EMR as a tool for disease prevention is an important step to inform clinical care that can address disparities (14).

To assess contributions of both SDoH and CMD burden on COVID-19 outcomes, we used data retrospectively extracted from the EMR at an academic medical center in the Deep South (i.e., Alabama). To quantify CMD burden, we employed a cardiometabolic disease staging (CMDS) score (15,16). CMDS incorporates presence and severity of metabolic syndrome traits and reflects severity of insulin resistance; it has been validated to predict risks of diabetes and cardiovascular disease (15,17), and it features a Bayesian logistical regression model highly predictive of future diabetes (16). Individual- and neighborhood-level SDoH as well as CMDS data were extracted from medical encounters prior to a positive COVID-19 test. We determined the ability of SDoH and CMDS data to predict subsequent need for hospitalization, intensive care unit (ICU) admission, and mortality once COVID-19 infection occurred, and we investigated the degree to which adding SDoH to the clinical CMDS model improved prediction accuracy.

METHODS

Study design

Retrospective longitudinal patient EMR data on cardiometabolic markers and SDoH were used to predict subsequent COVID-19 outcomes.

Study population and setting

The study population consisted of patients in the University of Alabama at Birmingham (UAB) Health system, located in Jefferson County, Alabama. Approximately 40% of the UAB Hospital's community inpatient discharges per year live in Jefferson County, with an additional 35% residing in 29 surrounding counties. Patients in

Study Importance

What is already known?

► Manifestations of cardiometabolic disease (CMD), such as obesity and type 2 diabetes, are associated with poor outcomes in those diagnosed with COVID-19.

► Poor prognosis in patients with COVID-19 may vary based on an insulin resistance or cardiometabolic continuum, such that patients who have existing suboptimal metabolic and vascular clinical markers progress to worse outcomes.

► Understanding health disparities regarding COVID-19 should consider clinical or biological entities that adversely affect outcomes (i.e., CMD) together with social determinants of health (SDoH), although combined analyses of these factors are rarely performed.

What does this study add?

► CMD severity, assessed up to 3 years in advance of COVID-19 infection, was highly associated with increased hospitalizations, intensive care unit admission, and mortality.

► We have also shown, for the first time, that SDoH (particularly insurance status [none or public] and high area-level social vulnerability) independently affect COVID-19 outcomes in addition to CMD burden.

How might these results change the direction of research or the focus of clinical practice?

► Results suggest that using prior patient data, including clinical and SDoH, to identify populations at high risk for severe outcomes has the potential to help guide treatment, intervention, and prevention efforts to improve health and health equity.

the EMR with a positive COVID-19 polymerase chain reaction (PCR) test and appropriate clinical follow-up to define COVID-19 outcome measures were identified. For our analysis, further inclusion criteria were that patients (i) had a previous encounter (within 3 years, from January 2017 to December 2020) in the EMR in which complete clinical data (blood glucose, BMI, blood pressure, high-density lipoprotein [HDL] cholesterol, and triglycerides) were available and (ii) were ≥35 years of age, owing to CMDS being developed and validated in older populations. Data were extracted and transformed through an institutional resource, the COVID-19 Collaborative Outcomes Research Enterprise (CORE) supported by the UAB Center for Clinical and Translational Science (NIH award number UL1TR003096), to facilitate use of institutional data to examine population health outcomes. The study was reviewed and approved

by the institution's institutional review board; informed consent from study participants was not required.

COVID-19 outcomes

Primary outcomes of interest were severe COVID-19 outcomes defined as need for hospitalization, ICU admission, or death during hospitalization. Hospitalization was defined as new hospital admission and positive COVID-19 PCR in the EMR within 14 days of admission. An ICU stay was defined as ICU admission within initial hospitalization. Death during hospitalization was defined using discharge disposition for hospital admission and notification of death record clinical event. Outcomes were dichotomized (Yes/No) and modeled separately.

Cardiometabolic disease

A primary predictor of interest was CMD using risk factors from CMDS. CMDS was originally developed as a discrete staging system using presence and severity of metabolic syndrome traits (15) to predict incident diabetes and cardiovascular disease mortality. Risk factors included in CMDS are BMI, glucose, blood pressure, HDL cholesterol, and triglycerides. Recent CMDS work (16) showed associations between continuous clinical parameters and a robust logistic regression equation for predicting incident diabetes. CMDS scores were generated using the predicted probability equation developed using continuous clinical parameters in previous work (algorithm in Table 1 footnote). Scores range from 0.0 to 0.99, with higher values indicating a higher probability of developing diabetes within 10 years. We also used individual components of CMDS to develop models to predict severe COVID-19 outcomes. To account for multiple encounters, we used mean clinical values.

SDoH

Individual-level SDoH

Guided by the 2015 Institute of Medicine report (18), we considered the following individual measures: educational attainment, individual income, marital status, employment, and insurance status. Educational attainment, individual income, and employment status had either a high level of missingness or they were not captured in discrete fields in the EMR, and thus they not accessible for analysis. Marital status consisted of married, single, or divorced/widowed. Insurance was categorized as no insurance, public insurance, other, and private insurance.

Neighborhood-level social determinants

Patients in the EMR are routinely geocoded and linked to 2010 census tract of residence if address data are available. Neighborhood-level

data were merged by census tract to describe characteristics of the location where a patient resided. The neighborhood-level data included in this analysis are described in the sections that follow.

Social vulnerability index

The social vulnerability index (SVI) (19) is a composite index developed using census tract data indicators on 15 social factors, categorized into 4 main themes: socioeconomic, household composition and disability, minority status and language, and housing and transportation. The index is used to describe social conditions that influence human suffering and financial hardship (social vulnerabilities) for disaster planning and it has been linked to poor health outcomes (20,21). Tracts are assigned a percentile ranking for overall vulnerability. Higher percentile rankings indicate more social vulnerabilities; measures were categorized as low (0.0 to <0.33), moderate (0.33 to <0.66), and high vulnerability (≥0.66) for analysis (19).

Rurality

Rurality was determined using 2010 United States Department of Agriculture (USDA) Rural-Urban Commuting Area codes and categorized as metropolitan area, micropolitan area (10,000-49,999 population), small town (2,500-9,999 population), and rural area.

Health care access

To characterize health care access, we linked census tract with data from the US Health Resources and Services Administration (HRSA) Data Warehouse Primary Care Service Area Data (22). Health professional shortage areas (HPSAs) indicate areas or populations that have a shortage of health care providers. Health care access was determined using census tracts designated as Geographic HPSAs or Population HPSAs, High-Needs Geographic HPSAs, or HPSA population.

Covariates

Age at COVID-19 outcome measurement, race (non-Hispanic Black, non-Hispanic White, and other), and gender were included as covariates in all analyses.

Statistical methods

Descriptive statistics were used to characterize the study population, overall and by presence of each COVID-19 outcome. Characteristics by outcome status were compared using parametric (t tests and χ^2) and nonparametric (Wilcoxon, Fisher exact test)

TABLE 1 Characteristics of the study population, n = 2,873

	Overall	Hospitalization			ICU stay			Mortality		
		Yes (n = 946)	No (n = 1,927)	p value	Yes (n = 392)	No (n = 2,481)	p value	Yes (n = 123)	No (n = 2,750)	p value
Age, mean (SD), years	58.3 (13.2)	62.9 (13.9)	56.1 (12.2)	<0.0001	62.4 (13.7)	57.7 (13.0)	<0.0001	66.5 (14.8)	57.9 (13.0)	<0.0001
Gender, n (%)										
Male	1,175 (40.9)	461 (48.7)	714 (37.1)	<0.0001	227 (57.9)	948 (38.2)	<0.0001	72 (58.5)	1,103 (40.1)	<0.0001
Female	1,698 (59.1)	485 (51.3)	1,213 (62.9)		165 (42.1)	1,533 (61.8)		51 (41.5)	1,647 (59.9)	
Race and ethnicity, n (%)										
Non-Hispanic Black	1,285 (44.7)	494 (52.2)	791 (41.0)	<0.0001	202 (51.5)	1,083 (43.7)	0.014	65 (52.8)	1,220 (44.4)	0.13
Non-Hispanic White	1,460 (50.8)	406 (42.9)	1,054 (54.7)		174 (44.4)	1,286 (51.8)		55 (44.7)	1,405 (51.1)	
Multiple/other	128 (4.5)	46 (4.9)	82 (4.3)		16 (4.1)	112 (4.5)		3 (2.4)	125 (4.5)	
Cardiometabolic markers										
BMI, mean (SD), kg/m^2	32.2 (7.3)	32.0 (7.8)	32.3 (7.1)	0.30	31.8 (7.7)	32.3 (7.3)	0.17	31.9 (7.7)	32.2 (7.3)	0.57
Plasma glucose, mean (SD), mg/dL	122.5 (42.8)	143.1 (46.0)	112.4 (37.2)	<0.0001	147.4 (41.6)	118.6 (41.7)	<0.0001	148.6 (39.9)	121.3 (42.6)	<0.0001
SBP, mean (SD), mm Hg	132.0 (12.2)	132.9 (12.4)	131.6 (12.0)	0.007	131.7 (12.5)	132.0 (12.1)	0.68	131.4 (12.3)	132.0 (12.2)	0.56
DBP, mean (SD), mm Hg	79.9 (6.8)	78.1 (6.6)	80.8 (6.7)	<0.0001	77.5 (6.5)	80.3 (6.8)	<0.0001	77.1 (6.0)	80.1 (6.8)	<0.0001
HDL cholesterol, mean (SD), mg/dL	49.0 (12.9)	45.5 (13.0)	50.7 (12.5)	<0.0001	43.3 (12.8)	49.8 (12.7)	<0.0001	45.0 (14.3)	49.1 (12.8)	0.0005
Triglycerides, mean (SD), mg/dL	139.6 (76.5)	148.1 (81.7)	135.4 (73.5)	<0.0001	154.3 (86.7)	137.3 (74.5)	<0.0001	158.8 (84.0)	138.7 (76.1)	0.004
Mean CMDS score[a]	0.42	0.59	0.33	<0.0001	0.65	0.38	<0.0001	0.66	0.41	<0.0001
Individual SDoH										
Marital status, n (%)										
Married	1,593 (55.4)	428 (45.2)	1,165 (60.5)	<0.0001	197 (50.3)	1,396 (56.3)	0.052	53 (43.1)	1,540 (56.0)	0.0005
Single	689 (24.0)	262 (27.7)	427 (22.2)		99 (25.3)	590 (23.8)		28 (22.8)	661 (24.0)	
Divorced/widowed	591 (20.6)	256 (27.1)	335 (17.4)		96 (24.5)	495 (20.0)		42 (34.1)	549 (20.0)	
Insurance, n (%)										
Private	1,631 (56.8)	265 (28.0)	1,366 (70.9)	<0.0001	96 (24.5)	1,535 (61.9)	<0.0001	16 (13.0)	1,615 (58.7)	0.0005
Public	1,140 (39.7)	629 (66.5)	511 (26.5)		273 (69.6)	867 (34.9)		100 (81.3)	1,040 (37.8)	
None	64 (2.2)	35 (3.7)	29 (1.5)		16 (4.1)	48 (1.9)		6 (4.9)	58 (2.1)	
Other	38 (1.3)	17 (1.8)	21 (1.1)		7 (1.8)	31 (1.2)		1 (0.8)	37 (1.3)	

TABLE 1 (Continued)

	Overall	Hospitalization			ICU stay			Mortality		
		Yes (n = 946)	No (n = 1,927)	p value	Yes (n = 392)	No (n = 2,481)	p value	Yes (n = 123)	No (n = 2,750)	p value
Neighborhood SDoH										
Urbanicity, n (%)										
Metropolitan	2,693 (93.7)	871 (92.1)	1,822 (94.6)	0.03	345 (88.0)	2,348 (94.6)	0.0005	107 (87.0)	2,586 (94.0)	0.001
Micropolitan	116 (4.0)	44 (4.7)	72 (3.7)		28 (7.1)	88 (3.5)		7 (5.7)	109 (4.0)	
Rural	27 (0.9)	12 (1.3)	15 (0.8)		7 (1.8)	20 (0.8)		2 (1.6)	25 (0.9)	
Small town	37 (1.3)	19 (2.0)	18 (0.9)		12 (3.1)	25 (1.0)		7 (5.7)	30 (1.1)	
Area-level social vulnerability, n (%)[b]										
Low	1,104 (38.4)	238 (25.2)	866 (44.9)	<0.0001	95 (24.2)	1,009 (40.7)	<0.0001	32 (26.0)	1,072 (39.0)	0.0006
Moderate	751 (26.1)	268 (28.3)	483 (25.1)		99 (25.3)	652 (26.3)		28 (22.8)	723 (26.3)	
High	1,018 (35.4)	440 (46.5)	578 (30.0)		198 (50.5)	820 (33.1)		63 (51.2)	955 (34.7)	
Health care access, n (%)[c]										
Not designated HPSA	1,967 (68.5)	563 (59.5)	1,404 (72.9)	<0.0001	245 (62.5)	1,722 (69.4)	0.007	78 (63.4)	1,889 (68.7)	0.26
Designated HPSA	906 (31.5)	383 (40.5)	523 (27.1)		147 (37.5)	759 (30.6)		45 (36.6)	861 (31.3)	
Comorbidities, n (%)										
Type 2 diabetes	602 (21.0)	295 (31.2)	307 (15.9)	<0.0001	122 (31.1)	480 (19.3)	<0.0001	44 (35.8)	558 (20.3)	<0.0001
Sleep apnea	279 (9.7)	119 (12.6)	160 (8.3)	0.0004	42 (10.7)	237 (9.6)	0.53	16 (13.0)	263 (9.6)	0.27
Chronic obstructive pulmonary disease	430 (15.0)	217 (22.9)	213 (11.1)	<0.0001	91 (23.2)	339 (13.7)	<0.0001	35 (28.5)	395 (14.4)	<0.0001
Chronic kidney disease	918 (32.0)	519 (54.9)	399 (20.7)	<0.0001	225 (57.4)	693 (27.9)	<0.0001	84 (68.3)	834 (30.3)	<0.0001
Coronary artery disease	1,116 (38.8)	528 (55.8)	588 (30.5)	<0.0001	249 (63.5)	867 (34.9)	<0.0001	82 (66.7)	1,034 (37.6)	<0.0001
Cancer	114 (4.0)	78 (8.2)	36 (1.9)	<0.0001	35 (8.9)	79 (3.2)	<0.0001	15 (12.2)	99 (3.6)	<0.0001

Abbreviations: CMDS, cardiometabolic disease staging; DBP, diastolic blood pressure; HDL, high-density lipoprotein; HPSA, health professional shortage area; SBP, systolic blood pressure; SDoH, social determinants of health.

[a]CMDS score calculated using Pr (diabetes) = logit^{-1} (−8.464 − 0.014*Age + 0.053*BMI + 0.006*SBP + 0.003*DBP + 0.062*Blood Glucose − 0.018*HDL + 0.001*Triglycerides − 0.084*Sex − 0.446*Race), in which Pr (diabetes) is the probability of 10-year incident diabetes for any individual; the function, logit^{-1} (x), equals exp(x) / [1 + exp(x)]; Sex equals 1 for male and 0 for female, and Race equals 1 for White and 0 for Black. CMDS score was calculated for White and Black participants only, n = 2,745.

[b]Captured using the social vulnerability index (SVI), a composite index developed by the Centers for Disease Control and Prevention using census data indicators on 15 social factors, categorized into 4 main themes: socioeconomic, household composition and disability, minority status and language, and housing and transportation. The index is used to describe the social conditions that may influence human suffering and financial hardship (social vulnerabilities) for disaster planning. The index is available by census tract where tracts are assigned a percentile ranking for overall vulnerability. Higher percentile rankings indicate more social vulnerabilities; measures were categorized as low (0.0 to <0.33), moderate (0.33 to <0.66), and high (≥0.66) for analysis.

[c]To characterize a patient's neighborhood with respect to health care access, we linked census tract with data from the US Health Resources and Services Administration Data Warehouse Primary Care Service Area data (22). HPSAs indicate areas or populations that have a shortage of health care providers. Health care access was determined using census tracts designated as Geographic HPSAs or Population HPSAs, High-Needs Geographic HPSAs, or HPSA population. Tracts with HPSA status were considered lacking health care access to primary care services; tracts without any HPSA status were considered as having adequate access.

analysis as appropriate. We used Bayesian logistic regression to test associations between CMDS score, SDoH, and COVID-19 outcomes. We then used Bayesian logistic regression to model probability for each binary COVID-19 outcome using individual CMDS components, individual-level and neighborhood-level SDoH, controlling for age, race, and gender. Following Gelman et al. (23), we assigned weakly informative priors (i.e., Cauchy distributions with center 0 and scale 2.5) to the coefficients in the logistic regressions, which have the advantage of providing minimal prior information to constrain co-efficients in a reasonable range, stabilizing the model fitting and improving the model prediction performance. We fit the Bayesian logistic regression models with Cauchy priors by incorporating an approximate expectation-maximization algorithm into the usual it-eratively weighted least squares in classical logistic regression. We first fitted a Bayesian logistic regression model using only CMDS metabolic syndrome traits, then added individual SDoH (marital sta-tus, insurance status) and finally neighborhood SDoH (SVI, rurality, HPSA status). In addition, we ran supplemental models (1) stratified by age owing to the severity of COVID-19 in older-age adults and (2) using an insurance variable that parsed out recipients of Medicare or Medicaid to approximate individual SES in lieu of missing income and education data in the EMR. Results are reported as odds ratios (ORs) with 95% CIs.

To evaluate predictive performance of the fitted model, we used 10-fold cross validation with several measures, including area under the curve (AUC), mean squared error (average squared differ-ence between observed and fitted responses), and misclassification (proportion of wrong predicted). Statistical analysis was performed using R software (version 4.0.3). The model fitting and predictive evaluation were implemented using R function bglm and cv.bh in the BhGLM package (https://github.com/nyiuab/BhGLM).

RESULTS

Descriptive results

A total of 3,989 patients were identified with a positive PCR and COVID-19 outcome in the institution's EMR and with a previous medical encounter in which all clinical data (blood glucose, BMI, blood pressure, HDL cholesterol, and triglycerides) were available (Figure 1). We excluded 548 patients <35 years of age, leaving an eligible sample of 3,441 patients. Of these, 568 (16.5%) were miss-ing one or more SDoH measures, resulting in 2,873 patients with complete data for analysis. Patients missing SDoH did not markedly differ by demographic and clinical parameters from those included in analysis (Supporting Information Table S1). Patients had a mean age of 58 years (SD 13.2) and were mostly female (59%), 45% were non-Hispanic Black, and most were married (55.4%) (Table 1). About 33% had been hospitalized, 13.6% were admitted to the ICU, and 4.2% died during hospitalization. About 40% had public insurance, and 2% had no insurance. In terms of neighborhood SDoH, most were from metropolitan areas (93.7%), 35% were from areas with high social

vulnerability, and 31.5% lived in an HPSA census tract. In bivariate analysis comparing characteristics based on COVID-19 outcomes, patients who were hospitalized, admitted to the ICU, or died were older, were more likely to be male, and had higher CMDS scores. Significantly more patients with no or public insurance or who lived in census tracts with high levels of social vulnerability experienced poor outcomes.

Prediction models

Figure 2 shows OR plots of associations between CMDS score, individual-level and neighborhood-level SDoH, and COVID-19 out-comes. Each 1-SD-unit increase in CMDS score was associated with hospitalization (OR 2.00, 95% CI: 1.83-2.20), ICU admittance (OR 1.88, 95% CI: 1.67-2.11), and death (OR 1.69, 95% CI 1.40-2.04). Patients with no insurance had higher odds of being hospitalized (OR 3.35, 95% CI: 1.88-5.98), ICU admittance (OR 2.99, 95% CI: 1.53-5.56), and mortality (OR 7.27, 95% CI: 2.65-19.94) than those with private insurance. A similar pattern was seen in patients with public insurance. Patients who lived in census tracts with high social vulnerability were more likely to be hospitalized (OR 1.57, 95% CI: 1.21-2.03) or admitted to the ICU (OR 1.66, 95% CI: 1.21-2.28) than those who lived in census tracts with low vulnerability.

Figure 3 shows OR plots of associations between metabolic markers, individual and neighborhood SDoH, and COVID-19 out-comes. Male sex (OR 1.69, 95% CI 1.37-2.09), higher levels of glu-cose (OR 1.79, 95% CI: 1.62-1.97), no insurance (OR 3.56, 95% CI: 2.00-6.34), and public insurance (OR 3.81, 95% CI: 3.00-4.82) were all associated with hospitalization; similar patterns were observed with ICU admittance. Higher levels of glucose (OR 1.43, 95% CI: 1.20-1.70), no insurance (OR 6.33, 95% CI: 2.25-17.79), and public insurance (OR 4.92, 95% CI: 3.70-8.97) were associated with mortal-ity. Patients who lived in census tracts with high social vulnerability were more likely to be hospitalized (OR 1.84, 95% CI: 1.38-2.45) or admitted to the ICU (OR 1.95, 95% CI: 1.35-2.82) than those who lived in census tracts with low vulnerability. Models stratified by age group (Supporting Information Figure S2 and Table S4) and those that used an insurance variable that distinguished between Medicare and Medicaid recipients (Supporting Information Figures S3-S4 and Table S5) produced similar results.

Predictive performance

Table 2 contains the predictive power, mean square error, and mis-classification statistics for each model. The model with CMDS param-eters and hospitalization as the outcome had an AUC of 0.776; adding both individual and neighborhood SDoH increased the AUC to 0.819 (p < 0.05). Similarly, adding SDoH to CMDS parameters increased AUCs for both ICU admittance (AUC 0.765 vs. 0.808, p < 0.05) and mortality (AUC 0.770 vs. 0.827, p < 0.05). Supporting Information Figure S1 shows the receiver operator characteristic (ROC) curves
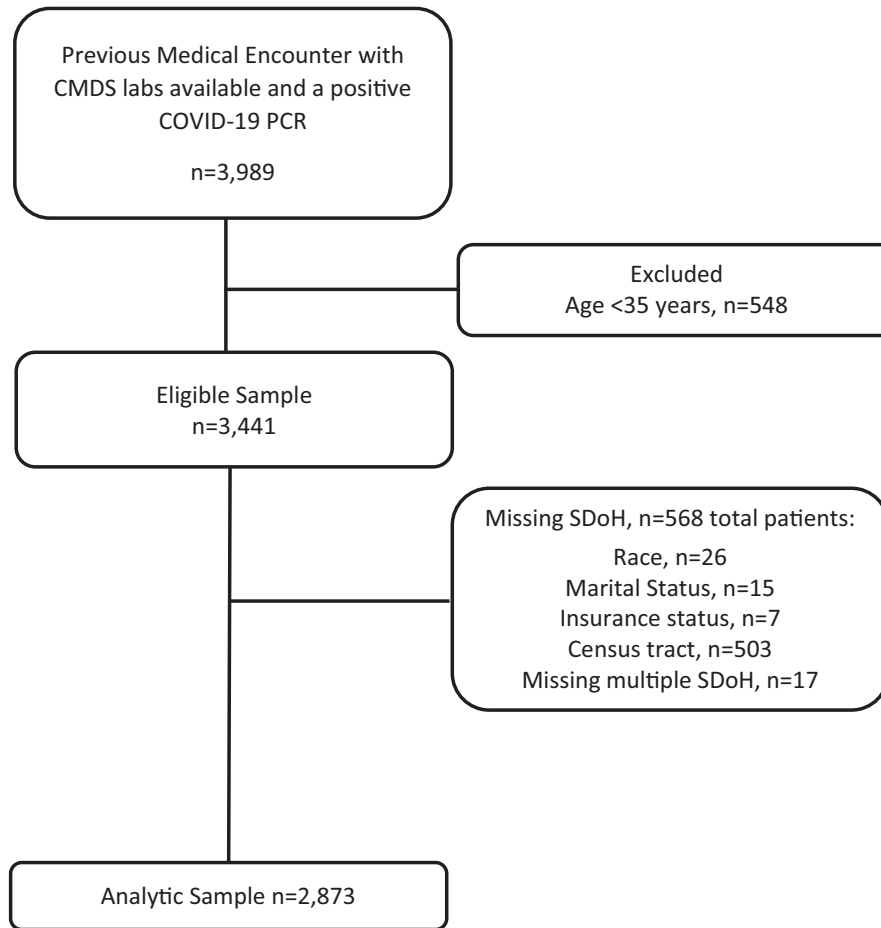
FIGURE 1 Flow of participants extracted from the electronic medical record and selected for analysis. CMDS, cardiometabolic disease staging; SDoH, social determinants of health

for the models shown in Figure 3. Sensitivity analysis (Supporting Information Table S2) of the CMDS-only model in those with complete lab data regardless of presence of SDoH data ($n = 3,441$) produced similar AUCs as the model using the analytical sample ($n = 2,873$).

DISCUSSION

In this study using EMR data from an academic medical institution in the Deep South during the COVID-19 pandemic, we found that the CMDS score, assessed up to 3 years in advance of COVID-19 infection, was highly associated with increased hospitalization, ICU admission, and mortality. We also showed, for the first time, that SDoH, particularly insurance status (none or public) and high area-level social vulnerability, independently affect COVID-19 outcomes in addition to CMD burden.

CMDS quantitatively reflects the burden of CMD as assessed by the presence and severity of metabolic syndrome traits. Insulin resistance is central to the pathophysiology of CMD. Our data support the hypothesis that the inflammation, oxidative stress, and endothelial dysfunction that accompany the insulin-resistant state are responsible for poor COVID-19 outcomes perhaps by contributing

to hyperimmune responses, tissue injury, and a clotting diathesis (24). Because obesity can accelerate the progression of CMD and its end-stage manifestations include diabetes and hypertension, insulin resistance could constitute a common mechanism explaining the associations of these diseases with poor COVID-19 outcomes.

We originally considered that SDoH could affect the biology of COVID-19 infection via adverse effects on the severity of CMD. In this instance, once the contribution of CMD was accounted for, SDoH would no longer be found to independently worsen the course of COVID-19 infection. However, we observed consistent associations of insurance status (none or public) and high area-level social vulnerability with severe outcomes independent of metabolic parameters. Thus, adding individual-level and neighborhood-level SDoH to COVID-19 outcome models significantly improved predictability beyond the pathobiological contribution of CMD. It is possible SDoH reflect the embodiment of chronic social related stress via dysregulation of physiological systems, epigenetic modification, and immune and inflammation responses (25), although further research in this area is needed.

Our results using a CMDS score reflecting CMD risk are consistent with robust associations linking diabetes and obesity with COVID-19 morbidity and mortality (26) and are similar to a study
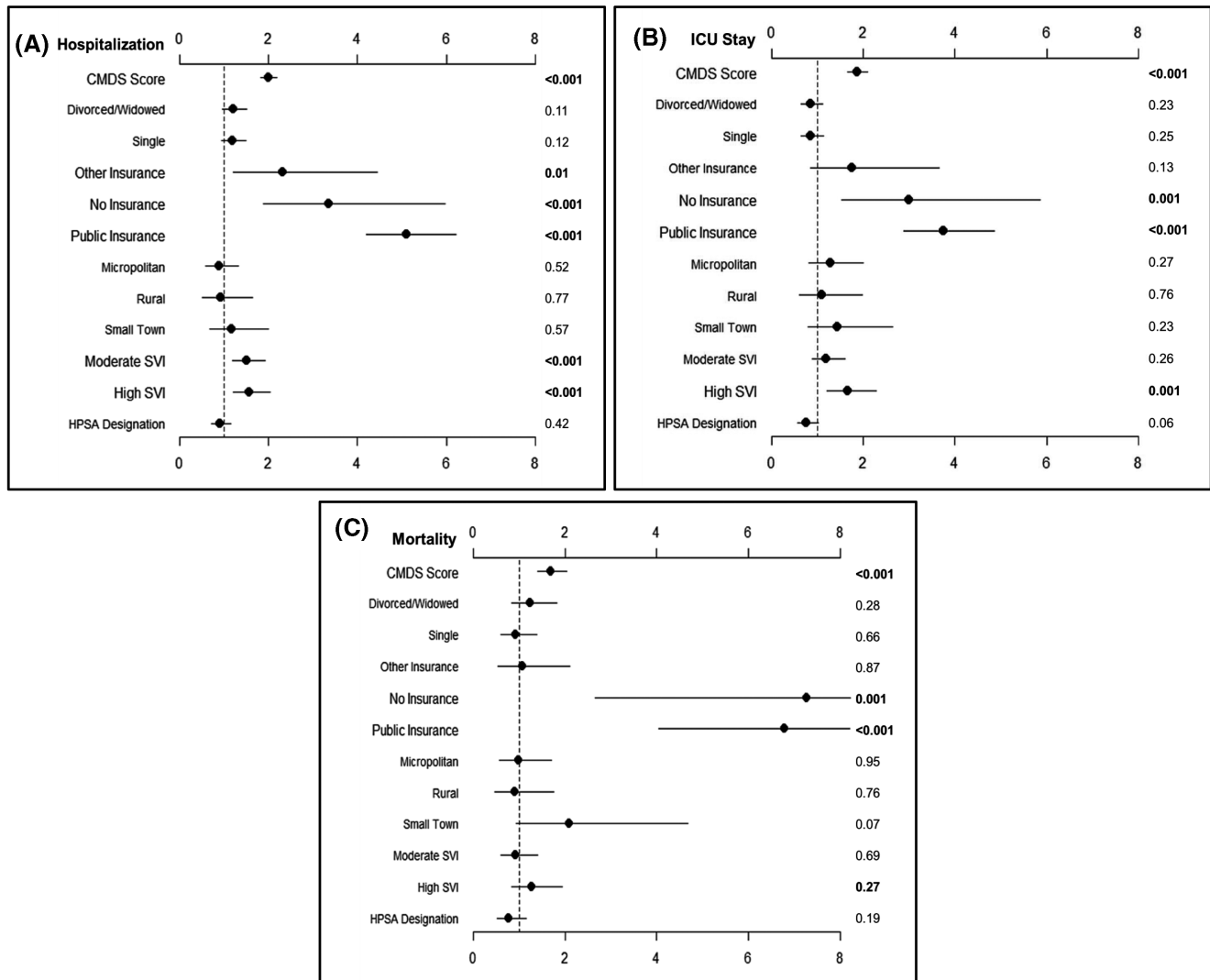
FIGURE 2 Odds ratio plots of calculated CMDS score, individual-level SDoH (marital status, insurance status), and neighborhood-level SDoH (rurality, SVI, HPSA status) in $n = 2,745$ White and Black participants. The points and lines present the estimated values and 95% CIs, respectively, and the values at the right side are p values. CMDS calculated using Pr (diabetes) = logit^{-1} (−8.464 − 0.014*Age + 0.053*BMI + 0.006*SBP + 0.003*DBP + 0.062*Blood Glucose − 0.018*HDL + 0.001*Triglycerides − 0.084*Sex − 0.446*Race), in which Pr (diabetes) is the probability of 10-year incident diabetes for any individual; the function, logit^{-1} (x), equals exp(x) / [1 + exp(x)]; Sex equals 1 for male and 0 for female, and Race equals 1 for White and 0 for Black. CMDS, cardiometabolic disease staging; HPSA, health professional shortage area; ICU, intensive care unit; SDoH, social determinants of health; SVI, social vulnerability index

that found a 3-fold increase in mortality and a 4-fold increase in ICU stays among patients hospitalized with COVID-19 who had metabolic syndrome (27). A recent study employed longitudinal medical records to predict risk of mortality among patients with COVID-19 (28) using factors deduced with machine learning methods. While a prior diagnosis of diabetes was associated with a 3-fold increase in mortality, metabolic correlates associated with diabetes and CMD (e.g., glucose, HDL) were not investigated; no SDoH were considered. While few studies have capitalized on the rich longitudinal medical records accessible via EMR, there is advantage in using these data to perform predictive analysis particularly in new disease states for which classical epidemiological cohorts do not exist (29).

While the CMDS score was strongly associated with all outcomes, the models involving individual metabolic components showed varying associations. Elevated glucose levels were associated with severe outcomes, consistent with multiple studies showing strong associations between diabetes (1,30) and poor glucose control (31,32). Higher levels of HDL were protective against hospitalization and ICU stays, similar to previous work (Dierckx et al. medRxiv, doi:10.1101/2020.11.09.20228221, unpublished data) that investigated whether blood metabolites predicted COVID-19 severity progression as well as Zhu et al. (32), who found HDL to be protective (OR 0.64, 95% CI: 0.50-0.83) from severe symptoms. While the associations found in our study were not surprising, they further reaffirm the ability of the CMDS score in predicting poor COVID-19 outcomes. Unique to our study, we assessed preexisting CMD clinical parameters collected up to 3 years prior to COVID-19 diagnosis; therefore, these clinical parameters are not acutely affected by the infection per se.
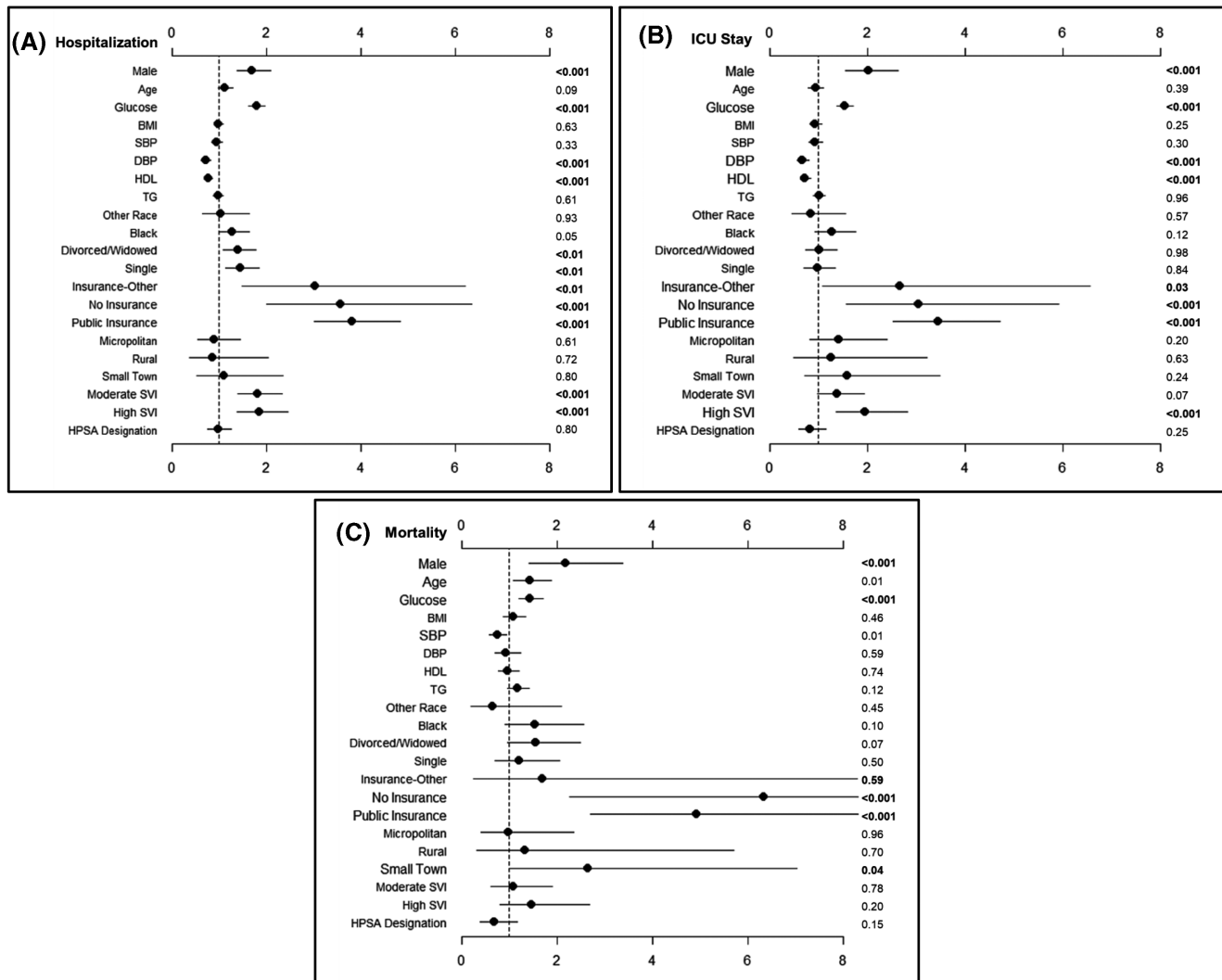
FIGURE 3 Odds ratio plots of models using cardiometabolic disease staging components, individual-level and neighborhood-level SDoH for each outcome, $n = 2{,}873$. The points and lines present the estimated values and 95% CIs, respectively, and the values at the right side are p values. DBP, diastolic blood pressure; HDL, high-density lipoprotein cholesterol; HPSA, health professional shortage area; ICU, intensive care unit; SBP, systolic blood pressure; SDoH, social determinants of health; SVI, social vulnerability index; TG, triglycerides

Interestingly, our work did not find independent associations between BMI and severe COVID-19 outcomes, which is inconsistent with previous reports associating higher BMI with poor outcomes (33,34). We believe this is explained in part by the fact that obesity worsens insulin resistance and cardiometabolic risk, and once severity of CMD is considered, obesity loses its significance as an independent risk factor (24). Previous studies have not examined the full impact of CMD in studies of obesity and COVID-19 outcomes. Furthermore, BMI collected in the clinical setting and captured in the EMR was shown to be subject to measurement and data entry errors (35), whereas biomarkers such as glucose and lipids are not as susceptible to such biases. Multiple investigations have shown that younger adult patients have stronger associations between increased BMI and poor COVID-19 outcomes (33,36). Our analysis was limited to individuals aged ≥35 years with a mean age of 58 years, which likely produced differing results for BMI. Indeed, when we looked at bivariate associations in those aged 35 to 59 in our sample versus those 60 years of age or older (Supporting Information Table S3), we found higher BMI and poor COVID-19 outcomes among the younger patients but not the older patients. Lastly, methodological issues using BMI, such as reverse causation, collider bias, and using BMI as a proxy for adiposity, may contribute to our findings (37).

While numerous investigations have explored various SDoH and their association with COVID-19 incidence and outcomes (38,39), this is the first investigation to use retrospective metabolic parameters and both individual-level and neighborhood-level SDoH in the EMR to predict subsequent COVID-19 outcomes. We found striking associations between insurance status (none and public) and all outcomes, regardless of how we modeled metabolic parameters (e.g., CMDS score or individual components). We saw a 3- to 4-fold increased risk of hospitalization and ICU stays among individuals with no or public insurance, with nearly a 6- to 8-fold increase in risk of mortality, even after controlling for metabolic markers and other SDoH. This is consistent with previous

TABLE 2 Predictive power and validation using CMDS metabolic syndrome traits, individual-level and neighborhood-level SDoH

	AUC	MSE	Misclassification
Hospitalized			
CMDS	0.776[a]	0.176	0.262
CMDS+ individual SDoH	0.815[b]	0.161	0.238
CMDS+ neighborhood and individual SDoH	0.819[c]	0.159	0.239
ICU			
CMDS	0.765[a]	0.106	0.143
CMDS+ individual SDoH	0.803[b]	0.103	0.144
CMDS+ neighborhood and individual SDoH	0.808[b]	0.099	0.133
Death			
CMDS	0.770[a]	0.040	0.044
CMDS+ individual SDoH	0.823[b]	0.039	0.043
CMDS+ neighborhood and individual SDoH	0.827[b]	0.038	0.044

Abbreviations: AUC, area under the curve; CMDS, cardiometabolic disease staging; MSE, mean square error; SDoH, social determinants of health.
[a,b,c]Groups with the different superscript letters are significantly different from each other, DeLong's test for 2 correlated ROC curves.

findings showing increases in risk of poor COVID-19 outcomes for both no and public-based insurance (40,41). Area-level indicators were linked to COVID-19 (39), and several studies have shown that higher SVI is associated with higher COVID-19 mortality (42) and incidence (42,43). We found an almost 2-fold increased risk of hospital admission and ICU stays among patients who lived in census tracts with moderate and high social vulnerability, a result not surprising among our local population. The Deep South is a geographic and cultural region in the United States that includes the states of Alabama, Georgia, Louisiana, Mississippi, and South Carolina. Alabama is at the core of this geographic region, with a higher prevalence of diabetes than the United States overall (14.0% vs. 10.8%) and persistent disparity in SES factors such as education and household income within the state. These socio-economic and social vulnerability disparities contribute to regional disparities in diabetes prevalence, with a recent study linking higher SVI with increased COVID-19 incidence rates in Alabama (43). Interestingly, in age-stratified models, the addition of SDoH to the model provided more information (increases in AUC) in the younger age models (35-59 years of age) than the 60+ models. While a recent study found that different SDoH were associated with COVID-19 incidence across different age strata (44), our findings warrant further exploration.

Our study has several limitations. Our data were derived from an EMR from a local population, which may not be representative of the entire hospital catchment population or national demographics. Generalizing results to other populations should be done with caution. Furthermore, it is possible patients with a positive test were hospitalized later in another local health system, introducing bias in outcome

ascertainment. In addition, we had a high rate of missing income and education in our discrete EMR fields, which we replaced with area-level data such as SVI to capture SES. Although these measures may not capture the individual-level effect of SES on outcomes and are prone to the ecological fallacy, such measures can give health care systems a "rough" estimate of SES in the absence of complete data. It is also important to note that insurance status has been used as a proxy for individual SES in other studies, particularly when using EMR data in which income data are incomplete or nonexistent (45). Although it would have been informative to model insurance based on Medicare versus Medicaid status to infer some degree of income, because of a lack of Medicaid expansion in Alabama, only 3% of our sample had Medicaid, leaving a narrow category for analysis. Although we provide supplementary analyses distinguishing between these 2 types of insurance, results should be interpreted with caution. Furthermore, while census tract was provided in the EMR in a discrete field, roughly 15% of patients were missing these data, and accuracy of geocoding was not available in our data set. Geocoding is not foolproof; the inability to limit our data to accurately matched census tracts may have introduced some measurement error. Moving forward, ensuring the accurate and complete capture of SDoH data is imperative to promote health equity and not inadvertently exacerbate disparities.

CONCLUSION

In summary, we found that cardiometabolic markers collected within 3 years prior to a positive COVID-19 diagnosis among hospitalized patients were predictive of subsequent poor outcomes. SDoH improved model predictions and independently contributed to poor outcomes, suggesting effects beyond the pathobiological mechanisms of CMD. Results suggest that using prior patient data, including clinical and SDoH, to identify populations at high risk of severe outcomes has the potential to help guide treatment, intervention, and prevention efforts to improve health and health equity.**O**

CONFLICT OF INTEREST
The authors declared no conflict of interest.

AUTHOR CONTRIBUTIONS
Carrie R. Howell, W. Timothy Garvey, and Andrea L. Cherrington contributed to the study conception and design. Material preparation and data collection were performed by Carrie R. Howell, Li Zhang, and Nengjun Yi. Data analysis and interpretation were performed by Li Zhang, Nengjun Yi, Tapan Mehta, and Carrie R. Howell.

The first draft of the manuscript was written by Carrie R. Howell, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ORCID

Carrie R. Howell https://orcid.org/0000-0002-6554-6237
W. Timothy Garvey https://orcid.org/0000-0003-0822-0860

REFERENCES

1. Singh AK, Gupta R, Ghosh A, Misra A. Diabetes in COVID-19: prevalence, pathophysiology, prognosis and practical considerations. *Diabetes Metab Syndr*. 2020;14:303-310.
2. Kalligeros M, Shehadeh F, Mylona EK, et al. Association of obesity with disease severity among patients with coronavirus disease 2019. *Obesity (Silver Spring)*. 2020;28:1200-1204.
3. Hussain A, Mahawar K, Xia Z, Yang W, El-Hasani S. Obesity and mortality of COVID-19: meta-analysis. *Obes Res Clin Pract*. 2020;14:295-300.
4. Marhl M, Grubelnik V, Magdic M, Markovic R. Diabetes and metabolic syndrome as risk factors for COVID-19. *Diabetes Metab Syndr*. 2020;14:671-677.
5. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med*. 2020;8:e21. doi:10.1016/S2213-2600(20)30116-8
6. Beckles GL, Chou CF. Disparities in the prevalence of diagnosed diabetes - United States, 1999–2002 and 2011–2014. *MMWR Morb Mortal Wkly Rep*. 2016;65:1265-1269.
7. Thornton PL, Kumanyika SK, Gregg EW, et al. New research directions on disparities in obesity and type 2 diabetes. *Ann N Y Acad Sci*. 2020;1461:5-24.
8. Cheng YJ, Kanaya AM, Araneta MRG, et al. Prevalence of diabetes by race and ethnicity in the United States, 2011–2016. *JAMA*. 2019;322:2389-2398.
9. Gaskin DJ, Thorpe RJ, McGinty EE, et al. Disparities in diabetes: the nexus of race, poverty, and place. *Am J Public Health*. 2014;104:2147–2155.
10. Christine PJ, Auchincloss AH, Bertoni AG, et al. Longitudinal Associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA). *JAMA Intern Med*. 2015;175:1311-1320.
11. Fisher-Hoch SP, Vatcheva KP, Rahbar MH, McCormick JB. Undiagnosed diabetes and pre-diabetes in health disparities. *PLoS One*. 2015;10:e0133135. doi:10.1371/journal.pone.0133135
12. Yancy CW. COVID-19 and African Americans. *JAMA*. 2020;323: 1891-1892.
13. Ramirez IJ, Lee J. COVID-19 emergence and social and health determinants in Colorado: a rapid spatial analysis. *Int J Environ Res Public Health*. 2020;17:3856. doi:10.3390/ijerph17113856
14. Bilal U, Auchincloss AH, Diez-Roux AV. Neighborhood environments and diabetes risk and control. *Curr Diab Rep*. 2018;18:62. doi:10.1007/s11892-018-1032-2
15. Guo F, Garvey WT. Development of a weighted cardiometabolic disease staging (CMDS) system for the prediction of future diabetes. *J Clin Endocrinol Metab*. 2015;100:3871-3877.
16. Wilkinson L, Yi N, Mehta T, Judd S, Garvey WT. Development and validation of a model for predicting incident type 2 diabetes using quantitative clinical data and a Bayesian logistic model: a nationwide cohort and modeling study. *PLoS Med*. 2020;17:e1003232. doi:10.1371/journal.pmed.1003232
17. Deemer S, Garvey WT. CMDS is a practical clinical estimate of insulin resistance in adults with overweight/obesity. [poster abstract #092]. *Obesity (Silver Spring)*. 2020;28(suppl S2):68-69. doi:10.1002/oby.23063
18. Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records; Board on Population Health and Public Health Practice; Institute of Medicine. *Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2*. National Academies Press (US); 2015.
19. Flanagan BE, Hallisey EJ, Adams E, Lavery A. Measuring community vulnerability to natural and anthropogenic hazards: the centers for disease control and prevention's social vulnerability index. *J Environ Health*. 2018;80:34-36.
20. Carmichael H, Moore A, Steward L, Velopulos CG. Using the social vulnerability index to examine local disparities in emergent and elective cholecystectomy. *J Surg Res*. 2019;243:160-164.
21. Yee CW, Cunningham SD, Ickovics JR. Application of the social vulnerability index for identifying teen pregnancy intervention need in the United States. *Matern Child Health J*. 2019;23:1516-1524.
22. Health Resources and Services Administration. HPSA data downloads. Accessed April 30, 2019. https://data.hrsa.gov/data/download
23. Gelman A, Jakulin A, Pittau MG, Su Y-S. A weakly informative default prior distribution for logistic and other regression models. *Ann Appl Stat*. 2008;2:1360-1383.
24. Nadolsky KZ, Hurley DL, Garvey WT. COVID-19 & obesity: beyond BMI. *Endocr Pract*. 2020;26:923-925.
25. Palmer RC, Ismond D, Rodriquez EJ, Kaufman JS. Social determinants of health: future directions for health disparities research. *Am J Public Health*. 2019;109(S1):S70-S71.
26. Zhang JY, Shang T, Ahn D, et al. How to best protect people with diabetes from the impact of SARS-CoV-2: report of the international COVID-19 and diabetes summit. *J Diabetes Sci Technol*. 2021;15:478-514.
27. Xie J, Zu Y, Alkhatib A, et al. Metabolic syndrome and COVID-19 mortality among adult black patients in New Orleans. *Diabetes Care*. 2020;44:188-193.
28. Estiri H, Strasser ZH, Klann JG, Naseri P, Wagholikar KB, Murphy SN. Predicting COVID-19 mortality with electronic medical records. *NPJ Digit Med*. 2021;4:15. doi:10.1038/s41746-021-00383-x
29. Estiri H, Strasser ZH, Murphy SN. Individualized prediction of COVID-19 adverse outcomes with MLHO. *Sci Rep*. 2021;11:5322. doi:10.1038/s41598-021-84781-x
30. Guo WN, Li MY, Dong YL, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res*. 2020;36:e3319. doi:10.1002/dmrr.3319
31. Wang S, Ma P, Zhang S, et al. Fasting blood glucose at admission is an independent predictor for 28-day mortality in patients with COVID-19 without previous diagnosis of diabetes: a multi-centre retrospective study. *Diabetologia*. 2020;63:2102-2111.
32. Zhu B, Jin S, Wu L, et al. J-shaped association between fasting blood glucose levels and COVID-19 severity in patients without diabetes. *Diabetes Res Clin Pract*. 2020;168:108381. doi:10.1016/j.diabres.2020.108381
33. Gao M, Piernas C, Astbury NM, et al. Associations between body-mass index and COVID-19 severity in 6.9 million people in England: a prospective, community-based, cohort study. *Lancet Diabetes Endocrinol*. 2021;9:350-359.
34. Hendren NS, de Lemos JA, Ayers C, et al. Association of body mass index and age with morbidity and mortality in patients hospitalized with COVID-19: results from the American Heart Association COVID-19 Cardiovascular Disease Registry. *Circulation*. 2021;143:135-144.
35. Goodloe R, Farber-Eger E, Boston J, Crawford DC, Bush WS. Reducing clinical noise for body mass index measures due to unit and transcription errors in the electronic health record. *AMIA Jt Summits Transl Sci Proc*. 2017;2017:102-111.
36. Kass DA, Duggal P, Cingolani O. Obesity could shift severe COVID-19 disease to younger ages. *Lancet*. 2020;395:1544-1545.
37. Banack HR, Stokes A. The 'obesity paradox' may not be a paradox at all. *Int J Obes (Lond)*. 2017;41:1162-1163.

38. Upshaw TL, Brown C, Smith R, Perri M, Ziegler C, Pinto AD. Social determinants of COVID-19 incidence and outcomes: a rapid review. *PLoS One*. 2021;16:e0248336. doi:10.1371/journal.pone.0248336

39. Su C, Zhang Y, Flory JH, et al. Clinical subphenotypes in COVID-19: derivation, validation, prediction, temporal patterns, and interaction with social determinants of health. *NPJ Digit Med*. 2021;4:110. doi:10.1038/s41746-021-00481-w

40. Himmelstein DU, Woolhandler S. Health insurance status and risk factors for poor outcomes with COVID-19 among US health care workers: a cross-sectional study. *Ann Intern Med*. 2020;173:410-412.

41. Gregory JM, Slaughter JC, Duffus SH, et al. COVID-19 severity is tripled in the diabetes community: a prospective analysis of the pandemic's impact in type 1 and type 2 diabetes. *Diabetes Care*. 2021;44:526-532.

42. Karmakar M, Lantz PM, Tipirneni R. Association of social and demographic factors with COVID-19 incidence and death rates in the US. *JAMA Netw Open*. 2021;4:e2036462. doi:10.1001/jamanetworkopen.2020.36462

43. Oates GR, Juarez LD, Horswell R, et al. The association between neighborhood social vulnerability and COVID-19 testing, positivity, and incidence in Alabama and Louisiana. *J Community Health*. 2021;46:1115-1123.

44. Bauer C, Zhang K, Lee M, et al. Census tract patterns and contextual social determinants of health associated with COVID-19 in a Hispanic population from South Texas: a spatiotemporal perspective. *JMIR Public Health Surveill*. 2021;7:e29205. doi:10.2196/29205

45. Hirsch AG, Durden TE, Nordberg C, Berger A, Schwartz BS. Associations of four community factors with longitudinal change in hemoglobin A1c levels in patients with type 2 diabetes. *Diabetes Care*. 2017;41:461-468.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.