Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Gastroenterology 2020;159:2221–2225

BRIEF COMMUNICATIONS

Impact of Obesity on Outcomes of Patients With Coronavirus Disease 2019 in the United States: A Multicenter Electronic Health Records Network Study

Shailendra Singh,[1] Mohammad Bilal,[2] Haig Pakhchanian,[3] Rahul Raiker,[4] Gursimran S. Kochhar,[5] and Christopher C. Thompson[6,7]

[1]Division of Gastroenterology and Hepatology, West Virginia University, Morgantown, West Virginia; [2]Division of Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; [3]George Washington University School of Medicine & Health Science, Washington, DC; [4]Department of Medicine, West Virginia University School of Medicine, Morgantown, West Virginia; [5]Division of Gastroenterology, Hepatology & Nutrition, Allegheny Health Network, Pittsburgh, Pennsylvania; [6]Harvard Medical School, Boston, Massachusetts; and [7]Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

Keywords: COVID-19; Coronavirus Disease; Obesity; BMI; Outcomes.

During the 2009 H1N1 influenza A virus pandemic, obesity was significantly associated with increased risk for hospitalization and mortality.[1] In 2020, the coronavirus disease 2019 (COVID-19) pandemic has a higher estimated case fatality rate.[2] It has hit the United States at a time when obesity has also reached epidemic status, with the prevalence of obesity increasing from 30.5% to 42.4% and of severe obesity increasing from 4.7% to 9.2% over the past decade.[3] Comorbidities associated with obesity are widely recognized risk factors for poor COVID-19 outcomes[4]; however, larger population-based data evaluating obesity as an independent risk factor continue to be sparse.

Methods

We performed a retrospective cohort study using TriNetX (Cambridge, MA), a global federated health research network that provided access to electronic medical records of patients from multiple large member health care organizations in the United States. Details of the data source are described in the Supplementary Materials.

A search query was performed to identify all adult patients (≥18 years) with a diagnosis of COVID-19 between January 20, 2020, and May 31, 2020. The search criteria to identify potential patients with COVID-19 were based on specific COVID-19 diagnosis codes (Supplementary Materials) or positive laboratory confirmation of COVID-19. Identified patients with COVID-19 were stratified based on a body mass index (BMI) or a diagnosis code for obesity. Patients with a documented BMI of ≥30 kg/m^2 or a diagnosis of obesity within 1 year before the diagnosis of COVID-19 were included in the obesity group. Patients with a documented BMI of <30 kg/m^2 or with no documented diagnosis of obesity within the last year were included in the control group. We excluded all patients for whom BMI varied between ≥30 kg/m^2 and <30 kg/m^2 in the preceding year before the diagnosis of COVID-19 or for whom diagnosis of obesity was present but BMI was reported as <30 kg/m^2 in the preceding year. Details of patient selection are outlined in Supplementary Figure 1.

The obesity group and control groups were compared after 1:1 propensity score matching (PSM). The primary outcome was a composite of intubation or death up to 30 days after diagnosis of COVID-19. Sensitivity analysis and subgroup analysis based on the obesity class were also performed. Details of the statistical analysis, sensitivity analysis, and limitations are also provided in the Supplementary Materials.

Results

A total of 41,513 adult patients with COVID-19 from 26 health care organizations in the United States were identified. Out of these patients with COVID-19, 8,641 patients with documented BMI of ≥30 kg/m^2 (n = 5,879) or diagnosis of obesity (n = 2,762) were included in the obesity group, and 31,273 patients with BMI of <30 kg/m^2 (n = 6437) or without any reported diagnosis of obesity were included in the control group (Supplementary Figure 1). Sex, racial, and ethnic differences were seen between the groups, and patients in the obesity group had a significantly higher proportion of comorbidities compared to the control group (Table 1). In the crude unadjusted analysis, patients in the obesity group were more likely to have a 30-day composite outcome of death or mechanical ventilation compared to the control group (Risk Ratio [RR] 1.99; 95% confidence interval, 1.84–2.15).

After PSM, a relatively balanced cohort of obese and nonobese patients were obtained (n = 8,112 patients in each group) (Table 1). The risk of composite outcome was higher in the obesity group compared to the control group (RR, 1.56; 95% confidence interval, 1.41–1.73). Kaplan-Meier survival analysis showed that the cumulative probability of being composite event-free up to 30 days remained significantly lower in the obesity group than the control group (87.7% vs 90.5%; *P* log rank < .0001) (Supplementary Figure 2). The risk of mortality, intubation, and hospitalization was higher in the obesity group compared to the control group in the matched cohort (Table 1). In a propensity-matched subgroup

Abbreviations used in this paper: **BMI, body mass index; COVID-19, coronavirus disease 2019; PSM, propensity score matching.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2020.08.028

Table 1. Characteristics and Outcomes of Patients With COVID-19 in the Obesity Group and Control Group Before and After Propensity Score Matching

Characteristics	Before Propensity Matching			After Propensity Matching		
	COVID-19 with obesity (n = 8,641)	COVID-19 without obesity (n = 31,273)	*P* value	COVID-19 with obesity (n = 8,112)	COVID-19 without obesity (n = 8,112)	*P* value
Age, *y*, mean ± SD	49.68 ± 15.84	49.87 ± 19.27	.395	49.47 ± 16.07	50.68 ± 16.93	<.001
Age, *y*, n (%)						
<40	2,563 (29.66)	11,087 (35.45)	<.001	2,496 (30.77)	2,388 (29.44)	.065
40–60	3,596 (41.62)	10,409 (33.28)	<.001	3,260 (40.19)	3,237 (39.9)	.712
60–80	2,216 (25.65)	6,986 (22.34)	<.001	2,090 (25.76)	2,217 (27.33)	.024
>80	266 (3.08)	2,791 (8.93)	<.001	266 (3.28)	270 (3.33)	.861
Female, n (%)	5,374 (62.19)	16,469 (52.66)	<.001	4,963 (61.18)	4,929 (60.76)	.584
Race, n (%)						
White	3,901 (45.15)	14,302 (45.73)	.332	3,719 (45.85)	3,698 (45.59)	.741
Black or African American	3,114 (36.04)	7,534 (24.09)	<.001	2,818 (34.74)	2,886 (35.58)	.264
Asian	96 (1.11)	1,050 (3.36)	<.001	96 (1.18)	110 (1.36)	.326
Unknown race	1,482 (17.15)	8,202 (26.23)	<.001	1,434 (17.68)	1,369 (16.88)	.177
Ethnicity: Hispanic or Latino, n (%)	1,308 (15.14)	4,047 (12.94)	<.001	1,235 (15.22)	1,246 (15.36)	.81
Hypertensive disease, n (%)	4,661 (53.94)	8,091 (25.87)	<.001	4,138 (51.01)	4,202 (51.80)	.315
Disorders of lipoprotein metabolism and other lipidemias, n (%)	3,453 (39.96)	5,823 (18.62)	<.001	2,998 (36.96)	3,037 (37.44)	.526
Diabetes mellitus, n (%)	2,816 (32.59)	3,945 (12.62)	<.001	2,359 (29.08)	2,379 (29.33)	.73
Chronic lower respiratory diseases, n (%)	2,606 (30.16)	4,355 (13.93)	<.001	2,224 (27.42)	2,279 (28.09)	.335
Ischemic heart diseases, n (%)	1,252 (14.49)	2,488 (7.96)	<.001	1,092 (13.46)	1,095 (13.5)	.945
Heart failure, n (%)	1,016 (11.76)	1,618 (5.17)	<.001	844 (10.4)	793 (9.78)	.184
Pulmonary heart diseases, n (%)	395 (4.57)	496 (1.59)	<.001	290 (3.58)	261 (3.22)	.209
Cerebrovascular diseases, n (%)	669 (7.74)	1,768 (5.65)	<.001	603 (7.43)	629 (7.75)	.441
Chronic kidney disease, n (%)	1,051 (12.16)	2,043 (6.53)	<.001	913 (11.26)	933 (11.5)	.621
Fatty liver disease, n (%)	599 (6.93)	511 (1.63)	<.001	409 (5.04)	386 (4.76)	.403
Cirrhosis of liver, n (%)	130 (1.5)	270 (0.86)	<.001	106 (1.31)	105 (1.29)	.945
Malignant neoplasm of breast, n (%)	143 (1.66)	317 (1.01)	<.001	130 (1.6)	138 (1.7)	.622
Malignant neoplasms of lymphoid, hematopoietic and related tissue, n (%)	140 (1.62)	308 (0.99)	<.001	115 (1.42)	125 (1.54)	.515

Table 1. Continued

Characteristics	Before Propensity Matching			After Propensity Matching		
	COVID-19 with obesity (n = 8,641)	COVID-19 without obesity (n = 31,273)	*P* value	COVID-19 with obesity (n = 8,112)	COVID-19 without obesity (n = 8,112)	*P* value
Malignant neoplasms of digestive organs, n (%)	112 (1.3)	278 (0.89)	<.001	95 (1.17)	94 (1.16)	.942
Malignant neoplasm of prostate, n (%)	79 (0.91)	296 (0.95)	.7832	75 (0.93)	80 (0.99)	.687
Nicotine dependence, n (%)	931 (10.77)	1,891 (6.05)	<.001	817 (10.07)	865 (10.66)	.216
Anthropometric parameters (within last 1 year)[a]						
Body height, *inches*, mean ± SD	66.08 ± 4.54 (n[b] = 7,407)	66.32 ± 4.28 (n[b] = 14,376)	.0001	66.13 ± 4.49 (n[b] = 6,880)	66.12 ± 4.29 (n[b] = 4,920)	.891
Body weight, *lb*, mean ± SD	224.72 ± 62.38 (n[b] = 6,739)	171.59 ± 46.95 (n[b] = 14,892)	<.001	223.64 ± 61.98 (n[b] = 6,275)	175.88 ± 48.73 (n[b] = 4,811)	<.001
BMI, kg/m^2, mean ± SD	37.07 ± 6.74 (n[b] = 5,879)	24.62 ± 3.23 (n[b] = 6,437)	<.001	36.89 ± 6.58 (n[b] = 5,508)	24.70 ± 3.16 (n[b] = 1,995)	<.001
Body surface area, m^2	2.18 ± 0.71 (n[b] = 994)	1.83 ± 0.27 (n[b] = 2,132)	<.001	2.17 ± 0.36 (n[b] = 922)	1.86 ± 0.27 (n[b] = 750)	<.001

Outcomes Before PSM

Outcome	COVID-19 with obesity, % (n/total) (n = 8,641)	Control group, % (n/total) (n = 31,273)	Risk ratio (95% CI)	*P* value
Composite outcome (intubation or death)	10.32 (892/8,641)	5.19 (1,623/31,273)	1.99 (1.84–2.15)	<.001
Mortality	4.57 (395/8,641)	3.32 (1,037/31,273)	1.38 (1.23–1.54)	<.001
Intubation	8.46 (731/8,641)	3.32 (1,014/31,273)	2.61 (2.38–2.86)	<.001
Hospitalization	25.66 (2,217/8,641)	14.27 (4,463/31,273)	1.80 (1.72–1.88)	<.001

Outcomes After PSM				
Outcome	COVID-19 with obesity, % (n/total) (n = 8,112)	Control group, % (n/total) (n = 8,112)	Risk ratio (95% CI)	*P* value
Composite outcome (intubation or death)	10.15 (823/8,112)	6.50 (527/8,112)	1.56 (1.41–1.73)	<.001
Mortality	4.56 (368/8,112)	3.87 (314/8,112)	1.17 (1.01–1.36)	.035
Intubation	8.32 (675/8,112)	4.55 (369/8,112)	1.83 (1.62–2.07)	<.001
Hospitalization	25.33 (2,055/8,112)	18.11 (1,469/8,112)	1.40 (1.32–1.49)	<.001

	COVID-19 with stage 2 obesity (n = 2,568)	Control group (n = 2,568)	Risk ratio (95% CI)	*P* value
Composite outcome (intubation or death)	10.83 (278/2,568)	7.20 (185/2,568)	1.50 (1.26–1.80)	<.001
Mortality	5.45 (140/2,568)	4.17 (107/2,568)	1.31 (1.02–1.67)	.031
Intubation	8.18 (210/2,568)	5.30 (136/2,568)	1.54 (1.25–1.90)	<.001
Hospitalization	25.94 (666/2,568)	19.08 (490/2,568)	1.36 (1.23–1.51)	<.001

Outcome	COVID-19 with stage 3 obesity (n = 2,538)	Control group (n = 2,538)	Risk ratio (95% CI)	*P* value
Composite outcome (intubation or death)	13.6 (345/2,538)	7.39 (187/2,538)	1.85 (1.56–2.19)	<.001
Mortality	6.22 (158/2,538)	4.61 (117/2,538)	1.35 (1.07–1.76)	.011
Intubation	10.95 (278/2,538)	4.89 (124/2,538)	2.24 (1.83–2.75)	<.001
Hospitalization	29.55 (750/2,538)	20.57 (522/2,538)	1.44 (1.30–1.58)	<.001

[a]Not included in PSM.
[b]Number of patients with available data.

analysis based on obesity class, the risk of composite outcome and other poor outcomes was highest in patients with obesity class 3 (Table 1). The results of the sensitivity analysis confirmed the robustness of our main findings (Supplementary Materials).

Discussion

Our study using a large nationally representative database showed that patients with COVID-19 with any degree of obesity had a significantly higher risk of hospitalization and intubation or death compared to patients without obesity. A substantial incremental risk of intubation or death in the obesity cohort persisted even after meticulous PSM to adjust for confounding comorbidities. Patients with severe obesity were at highest risk of these poor outcomes.

The COVID-19 pandemic has exposed the delivery of health care in the United States and has provoked a reckoning regarding our health care model moving forward. The US obesity epidemic has continued to grow for decades without any signs of abating. Obesity and its associated comorbidities are now a significant determinant of COVID-19 outcomes[5] in a population where more than 90 million adults have obesity and are highly susceptible. The disproportionate prevalence of obesity and associated comorbidities probably also have played a significant role in the racial and ethnic disparities seen during the COVID-19 pandemic. The obesity cohort derived from our data source showed a higher proportion of African Americans and Hispanics in the obesity group. Obesity increases the risk of poor outcomes in this vulnerable population with limited access to health care. Advanced age and male sex are major risk factors for worse prognosis and higher mortality in patients with COVID-19.[6] However, a larger proportion of patients with obesity in our cohort were female, and the impact of this can be dramatic enough to shift severe COVID-19 outcomes toward female patients. Similarly, a large number of younger patients with obesity are also affected by severe COVID-19 with poor outcomes. In the United States where obesity is an epidemic, its impact is not only limited to clinical outcomes. Along with the psychosocial impacts of social distancing and quarantining that are applicable to the entire society, persons with obesity must contend with "weight stigma." Derogation of persons with obesity is not uncommon and, unfortunately, more socially acceptable than other marginalized groups.[7] These biases and behaviors are not limited to the general public, and studies have shown that many health care workers can also have negative attitudes and stereotypes about persons with obesity.[8]

Our findings highlight the need for a vast improvement in the care of patients with obesity during this pandemic and moving forward. Physicians should manage patients with COVID-19 with obesity aggressively because outcomes can be significantly worse than in the general population. In the long term, to prepare for future pandemics or if COVID-19 becomes seasonal, there is also a serious need to develop and implement weight-loss strategies. There is a necessity for more health care professionals, including gastroenterologists, to play a central role in caring for patients with obesity.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2020.08.028.

References

1. Morgan OW, et al. PLoS One 2010;5(3):e9694.
2. Faust JS, del Rio C. JAMA Intern Med 2020;180:1045–1046.
3. Hales C, et al. NCHS Data Brief 2020;(360):1–8.
4. Zhou F, et al. Lancet 2020;395:1054–1062.
5. Hajifathalian K, Kumar S, Newberry C, et al. Obesity (Silver Spring) 2020;28:1606–1612.
6. Huang C, et al. Lancet 2020;395:497–506.
7. Pearl RL. Obesity (Silver Spring) 2020;28:1180–1181.
8. Phelan SM, et al. Obes Rev 2015;16:319–326.

Received July 16, 2020. Accepted August 18, 2020.

Correspondence
Address correspondence to: Christopher C. Thompson, MSc, MD, FASGE, FACG, AGAF, Harvard Medical School, Brigham and Women's Hospital, Division of Gastroenterology, Hepatology, and Endoscopy, 75 Francis Street, Thorn 1404, Boston, Massachusetts 02115. e-mail: cthompson@hms.harvard.edu; fax: (617) 525-8740; or Shailendra Singh, MD, Division of Gastroenterology and Hepatology, West Virginia University, 5th Floor Suite 5500, Morgantown, WV 26506 e-mail: Shail121@gmail.com.

Acknowledgments
The authors acknowledge West Virginia Clinical and Translational Science Institute for providing us with access to and training on the TriNETX global health care network. The authors also acknowledge the TriNETX (Cambridge, MA) health care network for design assistance to complete this project. The authors acknowledge these additional contributors: Ahmad Khan, Monica Chowdhry, Sergio A. Sánchez-Luna, Arka Chatterjee, and Diogo Turiani Hourneaux de Moura.

CRediT Authorship Contributions
Shailendra Singh, MD (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing.);
Mohammad Bilal, MD (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing);
Haig Pakhchanian, BS (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing);
Rahul Raiker, BS (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing);
Gursimran Kochhar, MD (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing);
Christopher C. Thompson, MSc, MD, FASGE, FACG, AGAF (Conceptualization, investigation, methodology, manuscript writing – original draft, review & editing).

Conflicts of interest
These authors disclose the following: Christopher C. Thompson has served as a consultant/provided research support for Apollo Endosurgery (consulting fees/institutional research grants), provided research support for Aspire Bariatrics (institutional research grant), served as general partner for BlueFlame Healthcare Venture Fund, served as a consultant for Boston Scientific (consulting fees), served as a consultant for Covidien/Medtronic (consulting fees), served as board member for EnVision Endoscopy, served as a consultant/advisory board member for Fractyl (consulting fees), served as a consultant/provided research support for GI Dynamics (consulting fees/institutional research grant), holds ownership interest in GI Windows, served as a consultant for/provided research support for Olympus/Spiration (consulting fees/equipment loans), provided research support for Spatz (institutional research grant), and served as a consultant/advisory board member and provided research support for USGI Medical (consulting fees/research grant). The remaining authors disclose no conflicts.

Supplementary Methods

Data Source

TriNetX (Cambridge, MA) is a global federated health research network providing access to the electronic health records (EHRs) of patients from 34 large member health care organizations (HCOs) in United States. COVID-19 data were incorporated in TriNetX by using specific diagnosis and terminology following the World Health Organization and Centers for Disease Control and Prevention (CDC) COVID-19 criteria. Real-time access to Health Insurance Portability and Accountability Act–compliant, deidentified, longitudinal clinical data to member HCOs is provided on a cloud-based platform. A typical HCO is a large academic health center with data coming from the majority of its affiliates. In addition to EHR data available in a structured fashion (eg, demographics, diagnoses, procedures, medications, laboratory test results, and vital signs), TriNetX can also extract facts of interest from the narrative text of clinical documents using natural language processing. Data are mapped to a standard and controlled set of clinical terminologies and transformed into a proprietary data schema. This transformation process includes an extensive data quality assessment to reject records that do not meet quality standards.

TriNetX data have been granted a waiver from the Western institutional review board because TriNetX is a federated network, and only aggregate counts and statistical summaries of the deidentified information without any protected health information were received from participating HCOs. Both the patients and HCOs as data sources remain anonymous.

Data Quality Checks

The software checks the basic formatting to ensure, for example, that dates are appropriately represented. It enforces a list of fields that are required (eg, patient identifier) and rejects those records where the required information is missing. Referential integrity checking is done to ensure that data spanning multiple database tables can be successfully joined together. As the data are refreshed, the software monitors changes in volumes of data over time to ensure data validity. TriNetX requires at least 1 nondemographic fact for a patient to be counted in our data set. Patient records with only demographic information are not included in data sets.

Coding Systems to Present Data

Demographics are coded to HL7, version 3, administrative standards. Diagnoses are represented by International Classification of Diseases, 10th Revision–Clinical Modification (ICD-10-CM) codes. If an HCO provides data in the International Classification of Diseases, Ninth Revision–Clinical Modification (ICD-9-CM), the data source uses a 9–to–10-CM mapping based on general equivalence mappings plus custom algorithms and curation to transform data from ICD-9-CM to ICD-10-CM. Diagnoses data are enriched with the Chronic Condition Indicator. Depending on the coding system used by an HCO, procedure data are coded in ICD-10 Procedure Coding System or Current Procedural Terminology (CPT). For many procedures, both ICD-10 Procedure Coding System and CPT codes are added to a query to define a cohort. Medications are represented at the level of ingredients, coded to RxNorm, and organized by National Drug File - Reference Terminology therapeutic classes. Laboratory test results, vitals, and findings are coded to Logical Observation Identifiers, Names, and Codes (LOINC). To ease finding and using common laboratory tests, LOINC codes are combined up to clinically significant levels for the most frequent laboratory tests and coded as TNX: LAB.

Selection of Patients With Coronavirus Disease 2019

The search was conducted following the CDC's COVID-19 coding guidance. These codes included ICD-9-CM and ICD-10-CM codes U07.1 (COVID-19, virus identified), B34.2 (Coronavirus infection, unspecified), B97.29 (Other coronavirus as the cause of diseases classified elsewhere), and J12.81 (Pneumonia due to SARS [severe acute respiratory syndrome]-associated coronavirus). Patients identified with diagnosis code 079.89 (Other specified viral infection) were excluded. Only patients diagnosed with the codes, as mentioned, between January 20, 2020, (the first confirmed case in the United States) and May 31, 2020, were included. The B97.29 code was specifically included based on the recommendation from the general guidance of the ICD-10-CM Official Coding Guidelines released by the CDC on February 20, 2020. Similarly, U07.1 is the new specific code for a confirmed diagnosis of COVID-19 with a positive COVID-19 test result starting April 1, 2020, as per the new CDC guidelines. The codes B34.2 and J12.81 were used more often before the CDC guidelines. Patients with ICD-9 code 079.89 (mapped to ICD-10 codes B34.2 and B97.2) were excluded to reduce any false positive COVID-19 patients because this ICD-9 code can still be used occasionally as catch-all code for more than 50 viral infections.

In addition to the ICD codes, the following LOINC codes with positive laboratory test results were also used to identify patients with COVID-19: 94533-7 SARS coronavirus 2 N gene [Presence] in Respiratory specimen by nucleic acid amplification with probe detection OR 94534-5 SARS coronavirus 2 RdRp gene [Presence] in Respiratory specimen by NAA with probe detection OR 94505-5 SARS coronavirus 2 IgG Ab [Units/volume] in Serum or Plasma by Immunoassay OR 41458-1 SARS coronavirus RNA [Presence] in Unspecified specimen by NAA with probe detection OR 94309-2 SARS coronavirus 2 RNA [Presence] in Unspecified specimen by NAA with probe detection OR 94531-1 SARS Coronavirus 2 RNA panel—Respiratory specimen by NAA with probe detection OR 94506-3 SARS coronavirus 2 IgM Ab [Units/volume] in Serum or Plasma by Immunoassay OR 94500-6 SARS coronavirus 2 RNA [Presence] in Respiratory specimen by NAA with probe detection OR 94315-9 SARS coronavirus 2 E gene [Presence] in Unspecified specimen by NAA with probe detection.

Variables and Outcomes

TriNetX has the capability of analyzing data based on a temporal relationship to the index event. The index event in

our study was defined as the diagnosis of COVID-19. Baseline characteristics were estimated from any time before the index event. Presenting laboratory values and medications were recorded from the time of the index event up to 2 weeks before the index event. Outcomes were assessed from the index event up to 30 days after the index event. The risk for intubation (mechanical ventilation), hospitalization, and mortality after diagnosis of COVID-19 was recorded. The primary outcome was a composite of intubation or death.

Statistical Analysis

All statistical analyses were performed in real-time by using TriNetX. The means, standard deviations, and proportions were used to describe and compare patient characteristics. Categorical variables were compared using the Pearson chi-square test and continuous variables using an independent-samples t test. We performed a 1:1 PSM to reduce the effects of confounding. Covariates in the propensity score model included age, race, ethnicity, dyslipidemia, diabetes mellitus, chronic lower respiratory diseases (chronic obstructive pulmonary disease and asthma), ischemic heart diseases, heart failure, pulmonary heart diseases, cerebrovascular diseases, chronic kidney disease, fatty liver, cirrhosis of liver, malignant neoplasm, and nicotine use (Table 1). Logistic regression on these input matrices was used to obtain propensity scores for each patient in both cohorts. Logistic regression was performed in Python 3.6.5 (Python Software Foundation) using standard libraries numpy and sklearn. The same analyses were also performed in R 3.4.4 software (R Foundation for Statistical Computing, Vienna, Austria) to ensure the matching of outputs. After the calculation of propensity scores, matching was performed by using a greedy nearest-neighbor matching algorithm with a caliper of 0.1 pooled standard deviations. The order of the rows in the covariate matrix can affect the nearest neighbor matching; therefore, the order of the rows in the matrix was randomized to eliminate this bias. For each outcome, the risk ratio with a 95% confidence interval was calculated to compare the association of obesity with the outcome. Kaplan-Meier survival analyses were used to estimate the survival probability of composite outcome at the end of 30 days after the index event. Patients were censored when the time window ended or on the day after the last fact in their record. Hypothesis testing for Kaplan-Meier survival curves was conducted by using the log-rank test. An a priori defined 2-sided alpha of $<.05$ was used for statistical significance.

Sensitivity Analysis

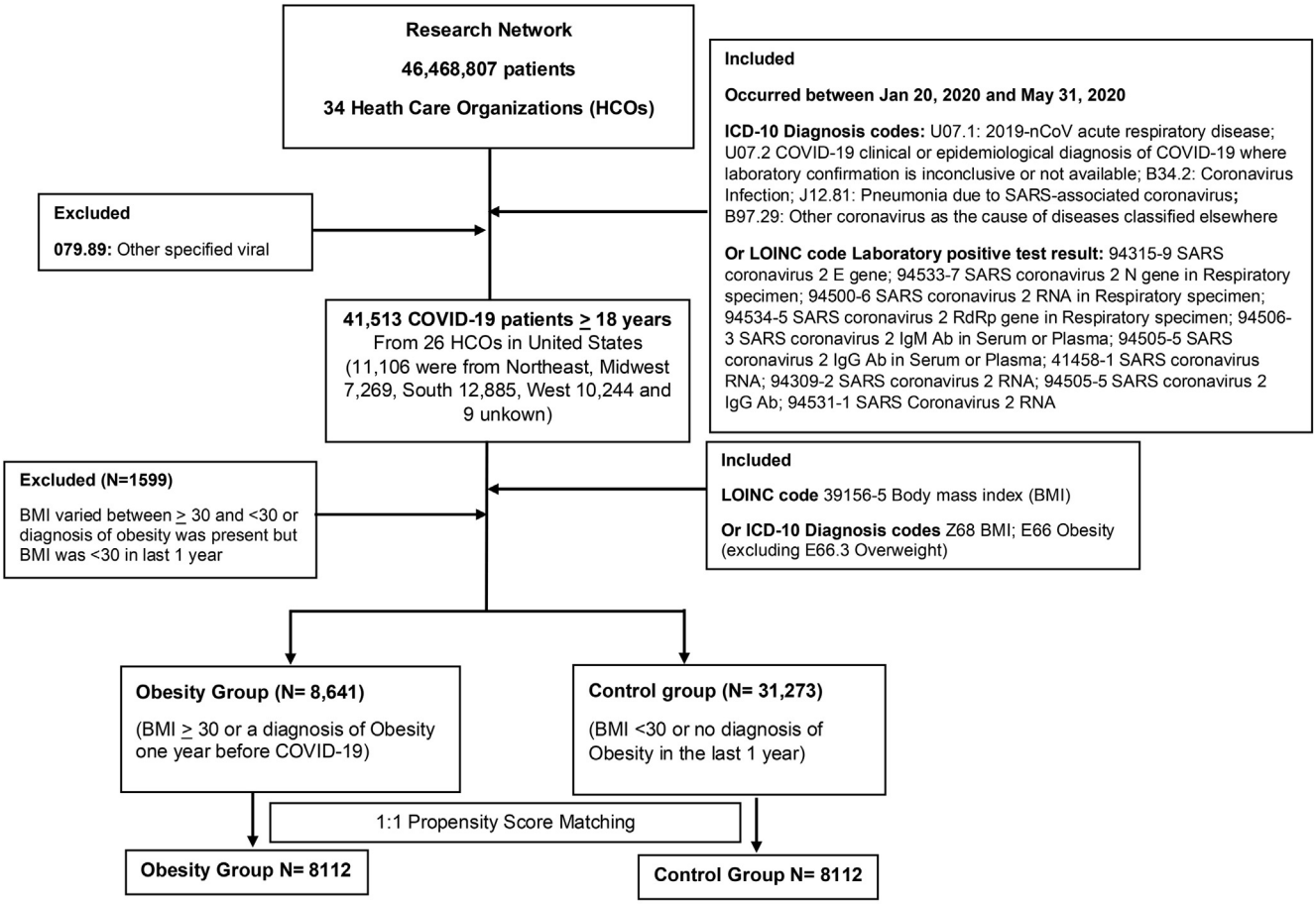
Selection bias in the obesity group and the control group was possible. Therefore, we performed a sensitivity analysis by varying the inclusion criteria. We first included all patients with a diagnosis of obesity in their health records at any time before COVID-19 diagnosis and compared them to a cohort of patients with no record of obesity. Second, we compared patients with a diagnosis of obesity in the last 3 months and 1 month to a cohort of patients with no reported obesity. Additional sensitivity analyses included the same set of main analyses but also adjusting for medications (angiotensin-converting enzyme inhibitors or angiotensin receptor blockers) and presenting laboratory values (ferritin, C-reactive protein, and lactic acid dehydrogenase). Finally, given the possibility that poor outcomes in patients with obesity might be higher at presentation or related to late presentation and access to health care, we performed an analysis excluding the composite outcomes in the first 2 days after diagnosis.

Results of Sensitivity Analysis

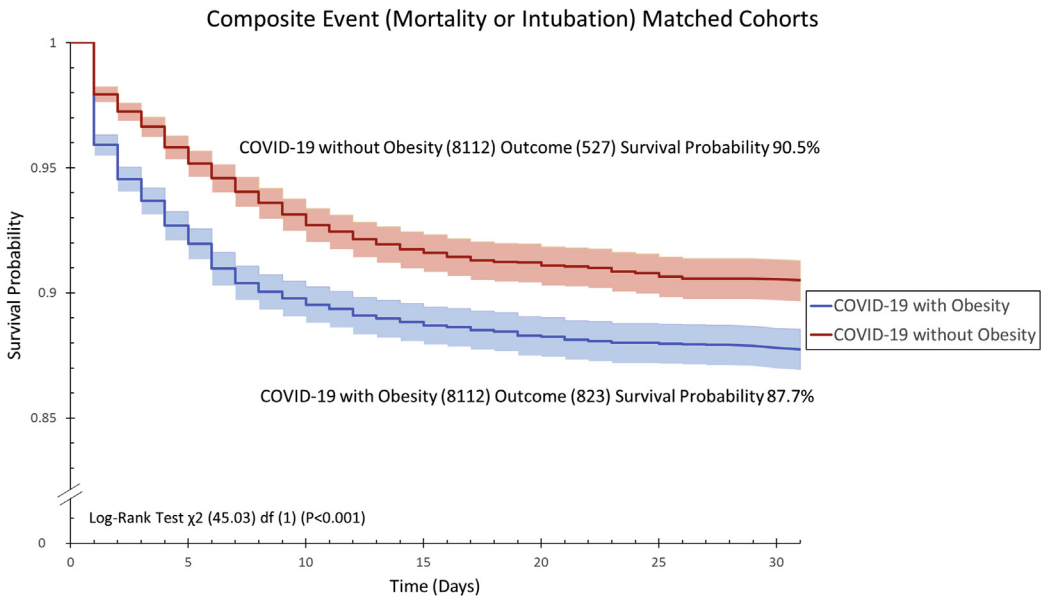
An analysis of a larger group of selected patients using diagnostic criteria of obesity as any time before the index event (after PSM: n = 9,769) showed a higher risk for composite outcomes in the obesity group (RR, 1.32; 95% confidence interval [CI], 1.20–1.46; $P < .0001$) compared to control individuals. Similarly, using an obesity diagnosis period of 3 months (after PSM: n = 6,780) also yielded a higher risk of composite outcome in obesity group (RR, 1.81; 95% CI, 1.62–2.02; $P < .0001$). Likewise, using an obesity diagnosis period of 1 month (after PSM: n = 5,825) also showed a higher risk of composite outcomes in the obesity group (RR, 2.18; 95% CI, 1.93–2.45; $P < .0001$). Composite outcome after adjustment for medications, angiotensin-converting enzyme inhibitors, and angiotensin receptor blockers (RR, 1.53; 95% CI, 1.38–1.70; $P < .0001$) or laboratory values (RR, 1.32; 95% CI, 1.19–1.46; $P < .001$) were similar. After excluding patients with outcomes in the first 2 days after diagnosis, the risk for composite outcomes was still higher in patients with obesity (RR, 1.54; 95% CI, 1.37–1.73; $P < .0001$).

Limitations

We acknowledge the limitations due to the retrospective nature of the study. The data derived from an EHR-based database is susceptible to errors in coding when patient information is translated into codes. However, extensive data quality assessment that includes data cleaning and quality checks minimizes the risk of data collection errors at the investigator's end. Adjustments for missing data are not currently possible on the TriNetX platform. Cases of COVID-19 could have been misdiagnosed as other cases of pneumonia or viral infections due to diagnosis or coding errors, especially early in the pandemic. We likely missed patients who were asymptomatic or had mild disease and did not seek medical attention; therefore, our cohort may represent the more the severe spectrum of COVID-19. Data on exposure history, incubation time, and dynamic changes in patients' clinical conditions could not be estimated from the EHR database. Socioeconomic and structural determinants, psychological elements, geographical factors, and health care delivery during COVID-19 could have affected the care of patients with obesity but were beyond the scope of our study. Despite these limitations, our study uses a large national database to evaluate the impact of obesity in patients with COVID-19. Given that our study population is representative of multiple centers across the United States, the results are more generalizable than single-center or regional experiences. In addition, even though our study was not randomized, we performed a robust statistical analysis using PSM.



Supplementary Figure 1. Flow diagram showing patient selection in the obesity group and control group. SARS, severe acute respiratory syndrome.



Supplementary Figure 2. Kaplan-Meier survival curve showing the probability of being composite event (intubation or death)–free up at the end of 30 days after COVID-19 diagnosis.

Supplementary Table 1. Codes Used

Variable	Coding System and Codes
Codes Used for Patient Characteristics Included in the PSM	
Race	HL7 version 3
White	2106-3
Black or African American	2054-5
Unknown race	2131-1
Asian	2028-9
Ethnicity	HL7 version 3
Hispanic or Latino	2135-2
Hypertensive diseases	ICD-10 I10-I16
Disorders of lipoprotein metabolism and other lipidemias	ICD-10 E78
Diabetes mellitus	ICD-10 E08-E13
Chronic lower respiratory diseases	ICD-10 J40-J47
Chronic kidney disease	ICD-10 N18
Ischemic heart diseases	ICD-10 I20-I25
Heart failure	ICD-10 I50
Nicotine dependence	ICD-10 F17
Cerebrovascular diseases	ICD-10 I60-I69
Fatty (change of) liver	ICD-10 K76.0
Cirrhosis of liver	ICD-10 K74.6
Pulmonary heart diseases	ICD-10 I27
Malignant neoplasm of breast	ICD-10 C50
Malignant neoplasms of lymphoid, hematopoietic and related tissue	ICD-10 C81-C96
Malignant neoplasm of prostate	ICD-10 C61
Malignant neoplasms of digestive organs	ICD-10 C15-C26
Codes Used for Anthropometric Parameters and Obesity Diagnosis	
Body height	TNX 9077 (Included LOINC codes 8307-1 Body height –preoperative, 8306-3 Body height –lying, 8302-2 Height, 8301-4 Body height Estimated, 3138-5 Body height Stated, 8308-9 Body height –standing, 8305-5 Body height –postpartum, 3137-7 Body height Measured)
Body weight	TNX 9081 (Included LOINC codes 8335-2 Body weight Estimated, 3142-7 Body weight Stated, 3141-9 Weight, 29463-7 Body weight)
Body surface area	TNX 9087 (Included LOINC codes 8277-6 Body surface area, 3139-3 Body surface area Measured, 3140-1 Body surface area)
BMI	LOINC code 39156-5 Body mass index ICD-10 Z68
Obesity	ICD-10 E66 (excluding ICD-10 E66.3 overweight)

Supplementary Table 1. Continued

Variable	Coding System and Codes
Codes Used to Define Outcomes of the Study	
Variable	Codes and
Mortality	"Deceased" (Known deceased documented)
Mechanical Ventilation	"31500" (CPT: Intubation, endotracheal, emergency procedure) OR "1015098" (CPT: Ventilator management) OR "5A1935Z" (ICD-10: Respiratory Ventilation, Less than 24 Consecutive hours) OR "5A1945Z" (ICD-10: Respiratory Ventilation, 24–96 Consecutive hours) OR "5A1955Z" (ICD-10: Respiratory Ventilation, Greater than 96 Consecutive hours) OR "0BH17EZ" (ICD-10: Insertion of Endotracheal Airway into Trachea, Via Natural or Artificial Opening) OR 0BH18EZ (ICD-10: Insertion of Endotracheal Airway into Trachea, Via Natural or Artificial Opening Endoscopic) OR 0BH13EZ (ICD-10: Insertion of Endotracheal Airway into Trachea, Percutaneous Approach) OR 1022227 (CPT: Extracorporeal membrane oxygenation [ECMO]/ extracorporeal life support [ECLS] provided by physician) OR 39.65 (ICD9: Extracorporeal membrane oxygenation [ECMO])
Hospitalization:	"1013659" (CPT: Hospital Inpatient Services) OR "1013609" (CPT: Initial Inpatient Consultation) OR "1013729" (CPT: Critical Care Services) OR "Visit: Inpatient Acute" OR "Visit: Inpatient Encounter" OR "Visit: Inpatient Non-acute" OR "Visit: Short Stay"