Clinical Research Article

Body Mass Index and Risk of COVID-19 Diagnosis, Hospitalization, and Death: A Cohort Study of 2 524 926 Catalans

Martina Recalde,1,2 Andrea Pistillo,1 Sergio Fernandez-Bertolin,1 Elena Roel,1,2 Maria Aragon,1 Heinz Freisling,3 Daniel Prieto-Alhambra,4,5 Edward Burn,1,4,* and Talita Duarte-Salles1,*

1Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain; 2Universitat Autònoma de Barcelona, Bellaterra, Spain; 3International Agency for Research on Cancer (IARC-WHO), Lyon, France; 4Centre for Statistics in Medicine, NDORMS, University of Oxford, Oxford, UK; and 5Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands

ORCID numbers: 0000-0003-2256-618X (M. Recalde); 0000-0002-4467-0220 (A. Pistillo); 0000-0002-0029-6564 (S. Fernandez-Bertolin); 0000-0002-1964-3546 (E. Roel); 0000-0002-5990-6953 (M. Aragon); 0000-0001-8648-4988 (H. Freisling); 0000-0002-3950-6346 (D. Prieto-Alhambra); 0000-0002-9286-1128 (E. Burn); 0000-0002-8274-0357 (T. Duarte-Salles).

*Joint senior authorship.

Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019; HR, hazard ratio; IQR, interquartile range; MEDEA, Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales; RT-PCR, reverse transcription polymerase chain reaction; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SIDIAP, Information System for Research in Primary Care; SMD, standardized mean difference; WHO, World Health Organization.

Received: 25 March 2021; Editorial Decision: 19 July 2021; First Published Online: 23 July 2021; Corrected and Typeset: 31 August 2021.

Abstract

Context: A comprehensive understanding of the association between body mass index (BMI) and coronavirus disease 2019 (COVID-19) is still lacking.

Objective: To investigate associations between BMI and risk of COVID-19 diagnosis, hospitalization with COVID-19, and death after a COVID-19 diagnosis or hospitalization (subsequent death), accounting for potential effect modification by age and sex.

Design: Population-based cohort study.

Setting: Primary care records covering >80% of the Catalan population, linked to regionwide testing, hospital, and mortality records from March to May 2020.

Participants: Adults (≥18 years) with at least 1 measurement of weight and height.

Main outcome measures: Hazard ratios (HR) for each outcome.

Results: We included 2 524 926 participants. After 67 days of follow-up, 57 443 individuals were diagnosed with COVID-19, 10 862 were hospitalized with COVID-19, and 2467 had a subsequent death. BMI was positively associated with being diagnosed and hospitalized...
with COVID-19. Compared to a BMI of 22 kg/m², the HR (95% CI) of a BMI of 31 kg/m² was 1.22 (1.19-1.24) for diagnosis and 1.88 (1.75-2.03) and 2.01 (1.86-2.18) for hospitalization without and with a prior outpatient diagnosis, respectively. The association between BMI and subsequent death was J-shaped, with a modestly higher risk of death among individuals with BMIs ≤ 19 kg/m² and a more pronounced increasing risk for BMIs ≥ 40 kg/m². The increase in risk for COVID-19 outcomes was particularly pronounced among younger patients.

Conclusions: There is a monotonic association between BMI and COVID-19 diagnosis and hospitalization risks but a J-shaped relationship with mortality. More research is needed to unravel the mechanisms underlying these relationships.

Key Words: obesity, adiposity, SARS-CoV-2, hospitalization, fatality, electronic health records

The coronavirus disease 2019 (COVID-19), the illness caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was declared a global pandemic in March 2020 (1). A high body mass index (BMI) has previously been associated in a linear and nonlinear fashion with an increased risk of multiple health outcomes such as metabolic and cardiovascular conditions, cancer, viral infections, and mortality (2-5). A better understanding of the relation between BMI and the progression of COVID-19 is essential for clinical management of patients and implementation of preventive strategies.

A review and meta-analysis of 75 studies indicated obesity (BMI ≥ 30 kg/m²) as a risk factor for severe COVID-19 and related mortality (6). Additionally, 2 studies with data from a subsample of the UK Biobank and a New York hospital found that BMI was associated in a dose-response manner with an increased risk of testing positive for SARS-CoV-2 and in a J-shaped fashion with the risk of intubation and death, respectively (7,8). These studies have provided relevant insights into this association. However, they have certain limitations that include being restricted to tested or hospitalized populations (increasing the risk of collider bias), having a small sample size, limitedly accounting for potential confounding, or dichotomizing BMI (with/without obesity) (9). Some of these limitations make it difficult to generalize the studies’ conclusions to populations with milder forms of disease or the general population. A study conducted with comprehensive patient-level data containing detailed individuals’ BMI information and capturing incident COVID-19 cases from a large and representative population where outcomes are recorded in diverse healthcare settings could add valuable information to complement the previous evidence in the understanding of the BMI–COVID-19 association.

Catalonia was heavily hit by the first phase (March-May 2020) of the COVID-19 pandemic (10). This region has a universal taxpayer-funded primary care–based health system in which general practitioners have been the first point of contact for care throughout the pandemic. Electronic health records from primary care encompassing demographic, historical lifestyle information, and disease diagnoses linked to SARS-CoV-2 reverse transcription polymerase chain reaction (RT-PCR) test results, hospital records, and regional mortality data offer a unique opportunity to study the role of BMI in the course of COVID-19. We aimed to investigate the associations between BMI and risks of COVID-19 diagnosis, hospitalization with COVID-19, and death after a COVID-19 diagnosis or hospitalization (ie, subsequent death), accounting for potential effect modification by age and sex, using electronic health record data from Catalonia.

Methods

Study Design, Setting, and Data Sources

We conducted a cohort study from March 1 to May 6, 2020. We used prospectively collected primary care records from the Information System for Research in Primary Care (SIDIAP; www.sidiap.org) in Catalonia, Spain. SIDIAP includes data from the Institut Català de la Salut (Catalan Health Institute), the largest public primary healthcare provider of Catalonia (covering 5.8 million people, 80% of the population of Catalonia) since 2006 and is representative of the Catalan population in terms of age, sex, and geographic distribution (11). SIDIAP includes high-quality data on anthropometric measurements, disease diagnoses, laboratory tests, and demographic and lifestyle information. SIDIAP has been linked to COVID-19 RT-PCR test results, hospital records, and regional mortality data and mapped to the Observational Medical Outcomes Partnership Common Data Model (12). The latter allowed the data to be structured in a standardized format and analytical tools to be applied as developed by the open-science Observational Health Data Sciences and Informatics network (13).
Multistate Framework

We addressed our objectives using a multistate framework. Multistate models allow the progression from a time origin until the occurrence of several events to be described, extending on competing risk models by also describing transitions to intermediate events (14). In the context of COVID-19, outpatient diagnoses and hospitalizations with the disease can be considered as intermediate events between not being (identified as) infected on one end to death on the other. We structured our multistate model in 4 states: general population, diagnosed (with COVID-19), hospitalized (with COVID-19), and death (Fig. 1). The following transitions were possible: general population to either diagnosed, hospitalized, or death; diagnosed to either hospitalized or death; or hospitalized to death. This approach is valuable in the context of the first wave of COVID-19 because it provides a more detailed overview of the interaction between individuals and the health system, respective of their BMI. This framework can disentangle the association between BMI and risk of hospitalization with COVID-19, differentiating direct hospitalizations (among the community) from indirect ones (among people already diagnosed with COVID-19 in primary care). Similarly, this approach distinguishes the risk of death related to BMI among individuals who interacted exclusively with primary care (ie, only had an outpatient diagnosis) and those who interacted with secondary care (ie, were hospitalized) before dying. Furthermore, this approach can reduce the risk of collider bias that can be induced by just assessing 1 transition of interest (9).

Participants

We identified all adults (aged ≥ 18 years) registered in the SIDIAP as of March 1, 2020, with a BMI recorded at an age ≥ 18 years. We excluded individuals with less than 1 year of prior history available to have sufficient time to capture participants’ characteristics before study entry and with a previous clinical diagnosis or positive test result for COVID-19. We also excluded those who were hospitalized or living in a nursing home on March 1, 2020, because the transmission dynamics and frequency of testing/diagnosing of these sites differed from the community population, which was the focus of this study (15,16). Finally, individuals without information on smoking and socioeconomic status were also excluded. The flow chart of inclusion and exclusion criteria for this study is presented in Supplementary Figure 1 (17). The descriptive characteristics of the individuals excluded due to living in nursing homes is available in Supplementary Table 1 (17). Individuals’ follow-up period began on March 1, 2020, (index date) and ended for any given transition due to exit from the database (ie, individuals moving out of the catchment area of SIDIAP), the occurrence of the event of interest or a competing event, or the end of the study period.

Variables

The exposure of interest was BMI as a continuous variable (kg/m²). BMI was calculated using the weight and height of patients assessed in a standardized manner by general practitioners or nurses (18). The exposure was assigned as the closest valid BMI (≥15 kg/m² to ≤60 kg/m²) to the index date recorded between January 1, 2006, and February 29, 2020.

The characteristics of interest were sex, age, smoking status, socioeconomic status, and comorbidities. We extracted participants’ sex (female, male), age (in years) at index date, and smoking status (never, former, or current smoker). We assessed socioeconomic status using the

![Figure 1](https://academic.oup.com/jcem/article-lookup/10.1210/clinendo.120.12.112)
Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales (MEDEA) deprivation index, which is calculated at the census tract level in urban areas of Catalonia (19). This measure is categorized into quintiles for anonymization purposes; the first quintile represents the least deprived group of the population, and the fifth, the most deprived one. It also includes a rural category since the MEDEA index is not available for participants living in those areas. We identified the following comorbidities using the individual’s medical history: autoimmune condition, chronic kidney disease, chronic obstructive pulmonary disease, heart disease, hyperlipidemia, hypertension, malignant neoplasm (excluding nonmelanoma skin cancer), and type 2 diabetes. We selected these conditions based on their relevance to the obesity and COVID-19 research fields and their availability in the Observational Medical Outcomes Partnership Common Data Model mapped version of the SIDIAP database and were defined as in previous studies conducted using SIDIAP data (20-22). The definitions are available in a Web application [https://livedataoxford.shinyapps.io/MultiStateCovidCohorts/ (index event breakdown tab)].

The outcomes of interest were an outpatient (primary care) clinical diagnosis of COVID-19, a hospitalization with COVID-19, and death. We defined outpatient COVID-19 diagnoses based on a recorded clinical code for COVID-19 disease (International Classification of Diseases, Tenth Revision, Clinical Modification codes B34.2 “Coronavirus infection, unspecified” and B97.29 “Other coronavirus as the cause of diseases classified elsewhere”). We did not require a positive RT-PCR test result in the definition of clinical diagnoses of COVID-19 due to testing restrictions during the first months of the pandemic (23). At the time of the study, RT-PCR tests were mostly conducted in patients with severe disease who were, or about to be, hospitalized; progressively, specific at-risk populations were also prioritized. Given that the focus of this study was to capture individuals with COVID-19 in the general population, we considered the clinical diagnosis as reported in the SIDIAP. We defined hospitalization with COVID-19 as a hospital admission (hospital stay of at least 1 night) where the individual had a positive RT-PCR test result or a clinical diagnosis of COVID-19 over the 21 days before their admission up to the end of their hospital stay. We defined mortality using regionwide mortality data and so included both deaths during hospitalizations and in the community.

**Statistical Analyses**

We reported the participants’ baseline characteristics by World Health Organization (WHO) categories of BMI [underweight (BMI < 18.5 kg/m²), normal weight (BMI ≥ 18.5 to <25 kg/m²), overweight (BMI ≥ 25 to <30 kg/m²), and obesity (BMI ≥ 30 kg/m²)].

We compared the baseline characteristics of the included individuals to those of the excluded due to unavailability of BMI, smoking status, and/or the MEDEA deprivation index information using standardized mean differences (SMDs). We considered an ISMDI >0.1 indicate meaningful differences in the distribution of a given characteristic between the 2 groups (24).

We described the participants’ time at risk at each state and the absolute number of outcomes observed for each transition by WHO categories of BMI. We assessed the relationship between BMI and the risk of transitioning to a subsequent state in the multistate model by estimating cause-specific hazard ratios (HRs) and 95% CIs using Cox proportional hazard regressions. We estimated 3 types of models: (1) with BMI as the sole explanatory variable (unadjusted models); (2) adjusted for age and sex; (3) adjusted for age, sex, smoking status, and the MEDEA deprivation index (fully adjusted models). We used a directed acyclic graph to guide decisions on the control for confounding [Supplementary Figure 2 (17)] (25). We considered nonlinearity in BMI and transitions by fitting models with BMI as a linear term, with a polynomial of degree 2 (ie, quadratic), and with restricted cubic splines (with 3, 4, or 5 knots) (26). We calculated the Bayesian information criterion, and we favored the model with the lowest Bayesian information criterion values. We compared the model where BMI was fitted with a nonlinear term against a linear model using a likelihood ratio test. We fitted age in the adjusted models using the same strategy as for BMI. We checked the proportional hazard assumptions for the variables included in the models by visual inspection of log-log survival curves. We did not model the transition from the general population to death because we were interested in deaths related to COVID-19 (subsequent deaths) which we captured by having gone through the diagnosed or hospitalized states (Fig. 1). However, we considered death among the general population as a competing risk by censoring people at their death.

We assessed effect modification by introducing interaction terms (1 at a time) between BMI and age and sex. We stratified the models in 3 categories of age (18-59, 60-79, and ≥80 years) and sex. As secondary analyses, we re-estimated the models fitting BMI in WHO categories, and we assessed the effect of obesity-related comorbidities (hypertension, type 2 diabetes, and hyperlipidemia) in the studied associations by introducing interaction terms (1 at a time) between BMI and each comorbidity.

For the main analyses, we conducted a complete case analysis (we only included individuals with complete
information on BMI and the covariates of interest). To explore the possibility of selection bias due to excluding those with missing data, in a sensitivity analysis we re-estimated the main models after multiple imputations (using predictive mean matching, with 5 imputations drawn) of missing data on BMI, smoking status, and/or the MEDEA deprivation index. The variables used for the multiple imputations were BMI, sex, age, smoking status, the MEDEA deprivation index, time of follow-up for each transition and outcomes of interest (given that we are using a time-to-event analysis), the Charlson comorbidity index, and a wide range of health conditions (27,28). In a second sensitivity analysis, we considered the impact of exposure misclassification. We replicated the main analyses, first, including only BMI values recorded in the previous 5 years (March 1, 2015-February 29, 2020) and, second, including only BMI values recorded in the previous 2 years (March 1, 2018-February 29, 2020).

We used R version 3.6 for data analysis and visualization. The R packages used for the analyses included numerous tidyverse packages, mstate, survival and rms (29-32). The analytic code we used is available at https://github.com/SIDIAP/MultiStateBmiCovid-19.

This study was approved by the Clinical Research Ethics Committee of the Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol i Gurina (project code: 20/070-PCV).

**Results**

There were 4,765,757 adults from the SIDIAP population registered in the database on March 1, 2020 who were eligible to enter the study. We excluded 104,022 individuals with less than a year of prior clinical history; 306 with a prior COVID-19 clinical diagnosis or positive test; 41,588 who were hospitalized or living in a nursing home on March 1; 1,357,553 who had missing data on smoking status and/or MEDEA deprivation index [Supplementary Figure 1 (17)]. We included 2,524,926 participants, of which 45,382 were living with underweight (2%), 905,898 with normal weight (36%), 952,479 (38%) with overweight, and 621,167 (24%) with obesity (Table 1). The participants’ median BMI [interquartile (IQR) range] was 26 (24-30) kg/m², and age was 52 (39-67) years. People living with underweight or normal weight were younger and more frequently female, current smokers, living in the least-deprived areas of Catalonia and presenting with fewer comorbidities than people living with overweight or obesity (Table 1).

All the analyzed baseline characteristics of the included individuals were meaningfully different (SMDs > 0.1) from those of the excluded individuals due to missing information on BMI, smoking status, and/or the MEDEA deprivation index [Supplementary Table 2 (17)]. Especially, the included participants vs those excluded were older (median age: 52 vs 44 years), more commonly female (55% vs 47%), and more frequently presenting with comorbidities (eg, hypertension prevalence: 20% vs 8%).

After a median follow-up of 67 days of the general population, 57,443 (2.28%) were diagnosed with COVID-19 [median (IQR) BMI: 27 (24-30) kg/m²] and 5,191 (0.21%) were hospitalized without a prior COVID-19 outpatient diagnosis [29 (26-32) kg/m²] [Table 2; also see Supplementary Table 3 (17)]. Among the people diagnosed with COVID-19 in outpatient settings, 5,671 (10.26%) went on to be hospitalized [28 (26-32) kg/m²], and 11,669 (2.43%) died [27 (24-30) kg/m²; median follow-up: 35 days]. Finally, of the people who were hospitalized with COVID-19, 1,301 (19.22%) died [29 (26-32) kg/m²; median follow-up: 37 days]. The time at risk and absolute event rates of the participants by WHO categories of BMI are shown in Table 2, and the descriptive characteristics of people transitioning to each state are available in Supplementary Table 3 (17).

BMI was nonlinearly associated with the risk of COVID-19 diagnosis, hospitalization with COVID-19, and subsequent death for all studied transitions in the fully adjusted models (all P for nonlinearity ≤ 0.001) (Fig. 2). Results for the unadjusted models and the models adjusted for age and sex are shown in Supplementary Figure 3 and Supplementary Table 4 (17); the latter were similarly shaped to the fully adjusted models. There is a modest positive association between BMI and the risk of COVID-19 diagnosis (Fig. 2). Relative to a BMI of 22 kg/m², the estimated HRs were 0.81 (0.79-0.84) for someone with a BMI of 16 kg/m², 1.10 (1.09-1.11) for a BMI of 25 kg/m², 1.22 (1.19-1.24) for a BMI of 31 kg/m², and 1.28 (1.25-1.32) for a BMI of 40 kg/m² (Table 3).

BMI was strongly associated with an increased risk of hospitalization with COVID-19, either with or without a prior outpatient diagnosis (Fig. 2). HRs for hospitalization without and with a prior diagnosis, respectively, relative to a BMI of 22 kg/m², were 0.58 (0.53-0.64) and 0.51 (0.46-0.57) for a BMI of 16 kg/m², 1.27 (1.22-1.31) and 1.37 (1.31-1.43) for a BMI of 25 kg/m², 1.88 (1.75-2.03) and 2.01 (1.86-2.18) for a BMI of 31 kg/m², and 2.85 (2.58-3.13) and 2.66 (2.43-2.91) for a BMI of 40 kg/m² (Table 3).

The association between BMI and risk of death either after an outpatient diagnosis or a hospitalization with COVID-19 was J-shaped (Fig. 2). Relative to a BMI of 22 kg/m², a BMI of 16 kg/m² was associated with HRs of 1.28 (1.07-1.52) and 1.20 (1.02-1.42) for death after an outpatient diagnosis or a hospitalization with COVID-19.
respectively (Table 3). High BMIs became positively associated with death only at BMIs ≥ 37 kg/m² among those previously hospitalized [HR (95% CI): 1.26 (1.06-1.51)] and 40 kg/m² among those diagnosed in outpatient settings [1.27 (1.03-1.56)].

There was evidence of effect modification by age and sex for 4 out of 5 studied transitions (P for interaction < 0.001) (Fig. 3). The risk of COVID-19 outcomes related to increased BMI was higher for those aged ≤59 years, compared to those in older age groups [Fig. 3; also see Supplementary Table 5A (17)]. Also, the risk of COVID-19 diagnosis for BMIs ≥ 40 kg/m² was higher for the oldest age group [HR (95% CI): 1.52 (1.33-1.74), relative to a BMI of 22 kg/m²] compared to those aged between 60 and 79 [1.10 (1.03-1.18)] or ≤59 years [1.32 (1.28-1.37)] [Fig. 3; also see Supplementary Table 5A (17)]. Associations were similarly shaped for females and males, although males were at a slightly higher risk of being diagnosed or hospitalized with COVID-19 compared to females [Fig. 3; also

### Table 1. Descriptive statistics of the study population by body mass index categories

| BMI categories          | Overall N | Underweight N | Normal weight N | Overweight N | Obesity N |
|-------------------------|-----------|---------------|-----------------|--------------|-----------|
| N                       | 2,524,926 | 45,382        | 905,898         | 952,479      | 621,167   |
| BMI, median (IQR)       | 26 (23.5-29.9) | 17.8 (17.2-18.2) | 22.7 (21.2, 23.9) | 27 (26.1-28.5) | 33 (31.2-35.8) |
| Time elapsed since BMI measurement, median (IQR) | 1.7 (0.6-4.0) | 2.5 (1.0-5.1) | 2.3 (0.9-5.0) | 1.6 (0.5-3.8) | 1.0 (0.3-2.7) |
| Age, median (IQR)       | 52 (39-67) | 35 (26-48) | 45 (34-60) | 56 (43-70) | 58 (45-70) |
| Age, years, n (%)       |           |               |                 |              |           |
| 18-39                   | 633,408 (25.1) | 27,552 (60.7) | 330,538 (36.5) | 175,824 (18.5) | 99,494 (16.0) |
| 40-59                   | 958,492 (38.0) | 11,362 (25.0) | 348,439 (38.5) | 363,878 (38.2) | 234,813 (37.8) |
| 60-69                   | 405,640 (16.1) | 2605 (5.7) | 100,354 (11.1) | 173,258 (18.2) | 129,423 (20.8) |
| 70-79                   | 325,948 (12.9) | 1670 (3.7) | 70,900 (7.8) | 148,561 (15.6) | 104,817 (16.9) |
| ≥80                     | 201,438 (8.0) | 2193 (4.8) | 55,667 (6.1) | 90,958 (9.5) | 52,620 (8.5) |
| Female sex, n (%)       | 1,386,678 (54.9) | 35,139 (77.4) | 549,089 (60.6) | 454,195 (47.7) | 348,255 (56.1) |
| Smoking status, n (%)   |           |               |                 |              |           |
| Never smoker            | 1,343,985 (53.2) | 23,568 (51.9) | 489,677 (54.1) | 503,401 (52.9) | 327,339 (52.7) |
| Former smoker           | 663,383 (26.3) | 6067 (13.4) | 186,830 (20.6) | 274,962 (28.9) | 195,524 (31.5) |
| Current smoker          | 517,558 (20.5) | 15,747 (34.7) | 229,391 (25.3) | 174,116 (18.3) | 98,304 (15.8) |
| MEDEA deprivation index, n (%)   |           |               |                 |              |           |
| Quintile 1 (least deprived) | 394,503 (15.6) | 8706 (19.2) | 168,140 (18.6) | 143,760 (15.1) | 73,897 (11.9) |
| Quintile 2              | 399,883 (15.8) | 7325 (16.1) | 149,848 (16.5) | 151,283 (15.9) | 91,427 (14.7) |
| Quintile 3              | 405,747 (16.1) | 7019 (15.5) | 141,866 (15.7) | 155,003 (16.3) | 101,859 (16.4) |
| Quintile 4              | 410,440 (16.3) | 6735 (14.8) | 135,637 (15.0) | 156,861 (16.5) | 111,207 (17.9) |
| Quintile 5 (most deprived) | 410,231 (16.2) | 7134 (15.7) | 130,020 (14.4) | 153,535 (16.1) | 119,542 (19.2) |
| Rural                   | 504,122 (20.0) | 8463 (18.6) | 180,387 (19.9) | 192,037 (20.2) | 123,235 (19.8) |
| Comorbidities, n (%)    |           |               |                 |              |           |
| Autoimmune condition    | 170,240 (6.7) | 2575 (5.7) | 52,165 (5.8) | 63,801 (6.7) | 51,699 (8.3) |
| Chronic kidney disease  | 141,921 (5.6) | 956 (2.1) | 30,583 (3.4) | 61,692 (6.5) | 48,690 (7.8) |
| COPD                    | 86,723 (3.4) | 1340 (3.0) | 21,163 (2.3) | 35,105 (3.7) | 29,115 (4.7) |
| Heart disease           | 363,012 (14.4) | 2975 (6.6) | 85,868 (9.5) | 153,188 (16.1) | 120,981 (19.5) |
| Hyperlipidemia          | 357,572 (14.2) | 2099 (4.6) | 86,453 (9.5) | 157,070 (16.5) | 111,948 (18.0) |
| Hypertension            | 514,533 (20.4) | 2165 (4.8) | 96,289 (10.6) | 220,109 (23.1) | 195,970 (31.5) |
| Malignant neoplasm      | 197,171 (7.8) | 2109 (4.6) | 56,734 (6.3) | 84,503 (8.9) | 53,825 (8.7) |
| Type 2 diabetes         | 236,253 (9.4) | 656 (1.4) | 34,065 (3.8) | 94,963 (10.0) | 106,569 (17.2) |
| Cause of end of follow-up, n (%) |           |               |                 |              |           |
| End of study            | 2,515,630 (99.6) | 45,131 (99.4) | 902,920 (99.7) | 948,894 (99.6) | 618,685 (99.6) |
| Transferred out of the SIDIAP | 7743 (0.3) | 207 (0.5) | 2347 (0.3) | 3035 (0.3) | 2154 (0.3) |
| Death                   | 1553 (0.1) | 44 (0.1) | 631 (0.1) | 550 (0.1) | 328 (0.1) |

BMI categories: underweight (BMI <18.5 kg/m²), normal weight (BMI ≥18.5 to < 25 kg/m²), overweight (BMI ≥25 to <30 kg/m²), and obesity (BMI ≥30 kg/m²). Malignant neoplasm does not include nonmelanoma skin cancer.

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; IQR, interquartile range; MEDEA, Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales.
From diagnosed with COVID-19

Table 2. Time at risk, absolute event rates, and cumulative incidence over time by body mass index categories

| BMI categories | Follow-up in days, median (min, IQR, max) | Events (cumulative incidence at 45 days), n (%). |
|---------------|------------------------------------------|-----------------------------------------------|
| Overall       | 67 (1, 67, 67, 67)                       | 574,443 (2.28)                                |
| Underweight   | 45 (0.21)                                | 5191 (2.18)                                   |
| Normal weight | 99 (0.39)                                | 3276 (2.45)                                   |
| Overweight    | 100 (0.19)                               | 574,443 (2.28)                                |
| Obesity       | 100 (0.19)                               | 574,443 (2.28)                                |

Note: BMI categories: underweight (BMI <18.5 kg/m²), normal weight (BMI ≥18.5 to <25 kg/m²), overweight (BMI ≥25 to <30 kg/m²) and obesity (BMI ≥ 30 kg/m²). 22.2% (12,730/57,443) of individuals diagnosed with COVID-19 also had a positive reverse transcription polymerase chain reaction test result.

Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019; IQR, interquartile range.

Discussion

In this large cohort study that included 2,524,926 participants from the general population in Catalonia, we found a monotonic association between BMI and COVID-19 diagnosis and hospitalization risks and a J-shaped relationship see Supplementary Table 5B (17)]. The risk of death after hospitalization with COVID-19 was stronger for females with BMIs ≥ 43 kg/m² [2.23 (1.66-3.00) relative to a BMI of 22kg/m²] compared to males [1.30 (0.92-1.85)] [Fig. 3; also see Supplementary Table 5B (17)].

The assumption of proportionality was violated for age in the first transition. To account for this, we stratified the main model by calendar month. The risk of COVID-19 diagnosis related to increased BMI was slightly higher for those diagnosed in March compared to April [Supplementary Figure 4 and Supplementary Table 6 (17)].

As a first secondary analysis, we re-estimated the main models with BMI in WHO categories [Supplementary Figure 5 (17)]. Relative to normal weight, overweight and obesity were associated with a higher risk of being diagnosed and hospitalized with COVID-19; no statistically significant associations were observed for the underweight category. No association between categorized BMI and risk of subsequent death was observed. As a second secondary analysis, we assessed the effect of comorbidities in the association between BMI and COVID-19 outcomes. The positive association between BMI and risks of COVID-19 diagnosis and COVID-19 hospitalization (with and without a prior outpatient diagnosis) was higher for individuals without hypertension compared to those with hypertension (P for interaction < 0.01) [Supplementary Figure 6 (17)]. A similar pattern was observed for people without type 2 diabetes for whom the association between BMI and risk of COVID-19 hospitalization (with and without a prior outpatient diagnosis) was higher than for those with type 2 diabetes (P for interaction < 0.01) [Supplementary Figure 7 (17)]. Finally, individuals without hyperlipidemia had a modestly higher risk of COVID-19 diagnosis compared to those with hyperlipidemia [Supplementary Figure 8 (17)].

Our findings were robust to 2 sensitivity analyses. The shape of the studied associations and the estimated effect sizes of our main analyses were similar to those of the analyses in which we did multiple imputations on missing data for BMI and the model's covariates and in which we excluded BMI measurements older than 5 or 2 years [Fig. 2; Table 3; also see Supplementary Figures 9-11 and Supplementary Tables 7-9 (17)], with 2,041,652 and 1,405,484 individuals included in each analysis, respectively. Nevertheless, the association between BMI and death after a COVID-19 diagnosis or a COVID-19 hospitalization was attenuated in the analyses of the multiple imputations.
The associations between BMI and COVID-19 outcomes were stronger for those ≤59 years of age and similarly shaped for females and males, with specific exceptions. The strengths of this study include being a large longitudinal study that investigates the association between BMI and the course of the COVID-19 disease containing mortality. The associations between BMI and COVID-19 outcomes were stronger for those ≤59 years of age and similarly shaped for females and males, with specific exceptions.

![Figure 2](https://academic.oup.com/jcem/article/106/12/e5030/6326782)

**Figure 2.** Association between body mass index (BMI) and the risk of coronavirus disease 2019 outcomes, allowing for nonlinear effects, with 95% CIs. Models are adjusted for age, sex, smoking status, and the Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales deprivation index. Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019.

Table 3. Hazards Ratios of COVID-19 Outcomes Related to Body Mass Index, With 95% CIs

| BMI values (kg/m²) | From general population | From diagnosed with COVID-19 | From hospitalized with COVID-19 | From death |
|-------------------|-------------------------|-----------------------------|-------------------------------|-----------|
|                   | To diagnosed with COVID-19 | To hospitalized with COVID-19 |                               | To death  |
| 16                | 0.81 (0.79-0.84)         | 0.58 (0.53-0.64)            |                               | 1.28 (1.07-1.52) |
| 19                | 0.90 (0.89-0.91)         | 0.77 (0.74-0.81)            |                               | 1.13 (1.04-1.23) |
| 22                | reference                | reference                   |                               | reference |
| 25                | 1.10 (1.09-1.11)         | 1.27 (1.22-1.31)            |                               | 0.90 (0.84-0.97) |
| 28                | 1.17 (1.15-1.19)         | 1.56 (1.47-1.66)            |                               | 0.88 (0.78-0.99) |
| 31                | 1.22 (1.19-1.24)         | 1.88 (1.75-2.03)            |                               | 0.93 (0.82-1.05) |
| 34                | 1.24 (1.22-1.26)         | 2.22 (2.04-2.41)            |                               | 1.02 (0.89-1.17) |
| 37                | 1.26 (1.23-1.29)         | 2.54 (2.33-2.78)            |                               | 1.14 (0.97-1.34) |
| 40                | 1.28 (1.25-1.32)         | 2.85 (2.58-3.13)            |                               | 1.27 (1.03-1.56) |
| 43                | 1.31 (1.26-1.36)         | 3.11 (2.77-3.49)            |                               | 1.42 (1.10-1.83) |
| 47                | 1.34 (1.28-1.40)         | 3.37 (2.87-3.96)            |                               | 1.64 (1.18-2.27) |
| 50                | 1.36 (1.29-1.44)         | 3.48 (2.81-4.31)            |                               | 1.82 (1.24-2.68) |

Models are adjusted for age, sex, smoking status, and the Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales deprivation index. Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019.
individual detailed BMI information and incident COVID-19 outcomes recorded in diverse healthcare settings from a large and representative population. Also, the possibility to investigate COVID-19 trajectories in a single and sufficiently powered data set, including systematic investigation of nonlinearity and effect modification, is a major strength. Further, our results were robust when we explored the violation of the models’ assumptions, the possibility of selection bias, and exposure misclassification.

This study also has weaknesses. The exposure was captured using a 14-year window, which for certain individuals relied on the assumption that BMI measurements were constant for a long period. However, we observed that the median of time elapsed since the BMI measurement was 1.7 years (IQR: 0.6-4.0) for the included participants. Moreover, in the sensitivity analyses where we used BMI measurements that were no older than 5 or 2 years, the obtained results were very similar to those of the main analysis. We defined COVID-19 cases as individuals who had a clinical diagnosis of the disease. Although this could have resulted in false positives, we decided not to require a confirmation of an RT-PCR positive test because testing was mainly restricted to severe cases of COVID-19 and specific at-risk populations during the first wave of the pandemic. This decision resulted in including only COVID-19 diagnoses of individuals who interacted with the health system, missing asymptomatic patients or individuals who did not seek medical care. However, Catalonia has a tax-funded almost universal healthcare system. Further, the results of this study are not generalizable to people living in nursing homes.

Figure 3. Effect modification by age and sex in the association between body mass index and the risk of COVID-19 outcomes, allowing for non-linear effects, with 95% CIs. Models are adjusted for age, sex, smoking status, and the Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales deprivation index. *P*-values for interaction were obtained by comparing the fully adjusted model, which included an interaction term (left side of the figure with age, right side with sex) against the fully adjusted model using a likelihood ratio test. Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019.
since we decided to exclude this subgroup of the population. We did not have the cause of death (only death after being diagnosed/hospitalized with COVID-19), which prevented us from attributing deaths to the disease. However, subsequent deaths were more frequent and happened more quickly than the deaths among the general population. The cumulative incidence of death was 0.2% in the general population, compared to 2.4% and 19.2% in those diagnosed and hospitalized with COVID-19, respectively (Table 2). The median time to death after a COVID-19 diagnosis or hospitalization was much shorter (35 and 37 days, respectively) than for those in the general population (67 days) [Supplementary Figure 12 (17)], which suggests subsequent deaths were COVID-19 related. Additionally, we missed individuals who died with COVID-19 but who were not identified as having been diagnosed or hospitalized with the disease. The likelihood of this outcome misclassification was probably reduced with the exclusion of nursing homes’ residents. We did not have data on hospital visits that did not lead to an overnight stay nor admission to intensive services units; these data can be useful to further study the progression of COVID-19 in detail. We did not have information on individual socioeconomic status nor the type of occupation of the participants; we tried to minimize this limitation by including the MEDEA deprivation index. Finally, the use of routinely collected data for research can raise concerns about data quality; however, BMI and COVID-19 data from the SIDIAP have successfully been repurposed for research (22,23,33,34).

The mechanisms by which higher BMI can increase COVID-19 severity include physical mechanisms (eg, altered ventilation due to reduced diaphragm excursion), chronic inflammation and impaired immune function (6). Higher BMI is also a risk factor for several medical conditions that could mediate the association between adiposity and the risk of COVID-19 severity such as type 2 diabetes or hypertension, which were also common in this study among patients with obesity (6,21). Our findings support the latter hypothesis: the positive association between BMI and the risk of being hospitalized with COVID-19 was attenuated among people with hypertension or type 2 diabetes (compared to those without). This suggests that shared biological mechanisms between obesity, hypertension, and type 2 diabetes might partially explain the higher susceptibility to COVID-19 hospitalization among individuals living with these conditions. Other proposed explanations include delayed seek for medical care among individuals with obesity due to fear of stigmatization (eg, 26% and 39% of those diagnosed and hospitalized without an outpatient diagnosis of COVID-19, respectively, had obesity) and the difficulty of care in hospital settings for supportive therapies (35,36).

Obesity has been associated with the risk of SARS-CoV-2 infection and COVID-19 diagnosis (6). Our dose-response analysis revealed that the risk of COVID-19 diagnosis increased linearly with higher BMI values, which is in line with a study of UK Biobank participants (8). Our findings are also aligned with a Mendelian randomization analysis, which reported that genetically increased BMI was causally associated with COVID-19 positivity (37). These results highlight the importance of avoiding extremely high BMI cutoffs to determine vulnerable groups to the COVID-19 (eg, the National Health Service only considers BMIs > 40 kg/m\(^2\) as risk groups) (20).

Our findings revealed a much stronger association between BMI and COVID-19 diagnosis among those aged >80 years and a modestly higher risk among males. While our findings are congruent with another study of the UK Biobank regarding sex differences in risk, no effect modification by age group (younger vs older than 70 years) was reported there (38). The underlying age distribution of those participants could explain this discrepancy; unfortunately, this information was unavailable.

Our findings of a strong positive association between BMI and risk of COVID-19 hospitalization are in line with a large meta-analysis and a population-based study conducted in another Spanish region (Navarra) (6,39). Our results also suggest the necessity to lower BMI cutoffs to establish risk groups for disease severity.

The risk of hospitalization with COVID-19 was systematically higher for those ≤59 years of age, which is congruent with 2 hospital-based studies from the United States. One reported a negative correlation between BMI and age among COVID-19 patients in 6 hospitals and another a positive association only among patients <60 years of age compared to older adults (40,41). Further, these results are congruent with the previously mentioned study from Navarra, where it was reported that the association between severe obesity and risk of hospitalization was much higher for those aged between 25 and 49 years (39).

Two meta-analyses reported that obesity is associated with a higher risk of COVID-19 mortality (6,42). However, nonlinear associations cannot be ignored in BMI-related research, especially concerning mortality (4,5). Large observational studies from the United States, the United Kingdom, and Spain using multiple categories of BMI only found an association between morbid obesity (BMIs > 35 or 40 kg/m\(^2\)) and COVID-19 mortality (39,43-46). Our results for high BMIs are consistent with the latter studies and revealed BMI was associated in a J-shaped fashion with the risk of subsequent death. Only BMIs above 37 kg/m\(^2\) and 40 kg/m\(^2\) were linked with a higher risk of death after a COVID-19 hospitalization and after a COVID-19 outpatient diagnosis, respectively. The J-shaped association between BMI and risk of COVID-19 related death has also
been reported in a study conducted in a New York hospital and England using a large primary care database (7,47). Interestingly, we observed a lower risk of death for individuals living with overweight among people diagnosed with COVID-19 in outpatient settings. These results are congruent with the UK study, which also reported HRs below 1 for the risk of confirmed or suspected deaths due to COVID-19 in BMIs in the overweight range (47). Our results provide important insights on the higher risk of subsequent death for low BMIs (≤19 kg/m²); while other studies also found this trend, these were not significant, likely due to their smaller sample sizes (7,45,46).

We also found that mortality risk related to an increased BMI was higher among individuals ≤59 years of age compared to older adults. Four previous studies are much in line with our findings, while a meta-analysis reported the opposite (7,38,42,45,46). The risk of death after a hospitalization with COVID-19 associated with BMI was higher among females, which is congruent with a UK Biobank study (38). However, a study performed in a New York hospital found a higher risk among males, and others found opposite or null differences by sex (7,42,45,46).

We provided a comprehensive analysis of the association between BMI and the course of COVID-19 during the first wave of the pandemic in Catalonia. Our analyses revealed that BMI is positively associated with being diagnosed and hospitalized with COVID-19 and in a J-shaped fashion with the risk of death following a COVID-19 diagnosis or hospitalization; the associations were particularly pronounced among younger patients. These findings highlight the necessity to consider individuals with both overweight and obesity as vulnerable groups to COVID-19 and its severity. Defining this high-risk group is especially important for the prioritization of individuals in preventive strategies such as vaccination campaigns. More broadly, our results reinforce the need for public health strategies focusing on the reduction of overweight and obesity, which can not only help prevent COVID-19 outcomes but also other well-established obesity-related diseases such as cardiometabolic conditions and certain cancer types.

**Acknowledgments**

We would like to acknowledge the patients who suffered from or died of this devastating disease and their families and carers. We would also like to thank the healthcare professionals involved in the management of COVID-19 during these challenging times, from primary care to intensive care units in the Catalan healthcare system.

**Authors’ Contributions:** S.F.B., M.A., and T.D.S. mapped source data to the Observational Medical Outcomes Partnership Common Data Model. M.R., A.P., E.B., and T.D.S. led the data analysis. M.R. and H.F. performed a literature review. All authors were involved in the study conception and design, interpretation of the results, and the preparation of the manuscript.

**Funding:** This project is funded by the Health Department from the Generalitat de Catalunya with a grant for research projects on SARS-CoV-2 and COVID-19 disease organized by the Direcció General de Recerca i Innovació en Salut. This project has received support from the European Health Data and Evidence Network (EHDEN) project. EHDEN received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement no. 806968. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and EFPIA. The University of Oxford received a grant related to this work from the Bill & Melinda Gates Foundation (Investment ID INV-016201) and partial support from the UK National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. D.P.-A. is funded through a National Institute for Health Research (NIHR) Senior Research Fellowship (grant number SRF-2018-11-ST2-004). MR is funded by Wereld Kanker Onderzoek Fonds (WKOF), as part of the World Cancer Research Fund International grant program (grant number: 2017/1630). The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

**Disclaimer:** Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

**Additional Information**

**Correspondence:** Daniel Prieto-Alhambra, MD, MSc, PhD, Botnar Research Centre, Windmill Road, OX37LD, Oxford, UK. Email: daniel.prietoalhambra@ndorms.ox.ac.uk; or Talita Duarte-Salles, MPH, PhD, Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Gran Via Corts Catalanes, 587 atic, 08007, Barcelona, Spain. Email: tduarte@idiapjgol.org.

**Disclosures:** All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; D.P.-A. reports grants and other from AMGEN; grants, nonfinancial support, and other from UCB Biopharma; grants from Les Laboratoires Servier, outside the submitted work; and Janssen, on behalf of IMI-funded EHDEN and EMIF consortiums, and Synapse Management Partners have supported training programs organized by D.P.-A.’s department and open for external participants. No other relationships or activities that could appear to have influenced the submitted work.

**Data Availability:** In accordance with current European and national law, the data used in this study is only available for the researchers participating in this study. Thus, we are not allowed to distribute or make publicly available the data to other parties. However, researchers from public institutions can request data from SIDIAP if they comply with certain requirements. Further information is available online (https://www.sidiap.org/index.php/menu-solicitudesen/application-procedure) or by contacting Anna Moleras (amoleras@idiapjgol.org).
References

1. World Health Organization. WHO Director-General’s opening remarks at the media briefing on COVID-19. Published March 11, 2020. 2020. Accessed September 21, 2020. https://www.who.int/ dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

2. Hu F. Obesity Epidemiology. Oxford University Press; 2008.

3. Moser JS, Galindo-Fraga A, Ortiz-Hernández AA, et al; La Red ILI 002 Study Group. Underweight, overweight, and obesity as independent risk factors for hospitalization in adults and children from influenza and other respiratory viruses. Influenza Other Respir Viruses. 2019;13(1):3-9.

4. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet*. 2014;384(9945):755-765.

5. Bhaskaran K, Dos-Santos-Silva I, Leon DA, Douglas IJ, Smeeth L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. *Lancet Diabetes Endocrinol*. 2018;6(12):944-953.

6. Popkin BM, Du S, Green WD, et al. Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. *Obes Rev*. 2020;21(11):e13128.

7. Anderson MR, Geleris J, Anderson DR, et al. Body mass index and risk for intubation or death in SARS-CoV-2 infection: a retrospective cohort study. *Ann Intern Med*. 2020;173(10):782-790.

8. Yates T, Raziieh C, Zaccardi F, Davies MJ, Khunti K. Obesity and risk of COVID-19: analysis of UK biobank. *Prim Care Diabetes*. 2020;14(5):566-567.

9. Griffith GJ, Morris TT, Tidball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun*. 2020;11(1):5749.

10. Ministerio de Ciencia e Innovación. *Gobierno de España. Panel COVID-19. Evolución pandemia*. CCAD; 2020.

11. Bolíbar B, Fina Avilés F, Morros R, et al; Grupo SIDIAP. [SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research]. *Med Clin (Barc)*. 2012;138(14):e671-621.

12. Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc*. 2015;22(3):553-564.

13. Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform*. 2015;216:574-578.

14. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007;26(11):2389-2430.

15. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al; ENE-COVID Study Group. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet*. 2020;396(10250):535-544.

16. Arons MM, Hatfield KM, Reddy SC, et al; Public Health–Seattle and King County and CDC COVID-19 Investigation Team. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N Engl J Med*. 2020;382(22):2081-2090.

17. Recalde M, Puntillo A, Fernandez-Bertolin S, et al. Supplementary material for: Body mass index and risk of COVID-19 diagnosis, hospitalisation, and death: a cohort study of 2 524 926 Catalans. *Figshare*. Posted July 12, 2021. doi:10.6084/m9.figshare.14959758.v1

18. Lecube A, Monereo S, Rubio MA, et al. Prevención, diagnóstico y tratamiento de la obesidad. Posicionamiento de la Sociedad Española para el Estudio de la Obesidad de 2016. *Endocrinol Diabetes y Nutr*. 2017;64:15-22.

19. Domínguez-Berjón MF, Borrell C, Cano-Serral G, et al. Constructing a deprivation index based on census data in large Spanish cities [the MEDEA project]. *Gac Sanit*. 2008;22(3):179-187.

20. National Health Service. Who’s at higher risk from coronavirus (COVID-19). 2020. https://www.nhs.uk/conditions/coronavirus-covid-19/people-at-higher-risk/whos-at-higher-risk-from-coronavirus/. Accessed August 20, 2020.

21. Centers for Disease Control and Prevention. Certain medical conditions and risk for severe COVID-19 illness. 2020. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html. Accessed August 20, 2020.

22. Burn E, You SC, Sena AG, et al. Deep phenotyping of 34 128 adult patients hospitalised with COVID-19 in an international network study. *Nat Commun*. 2020;11(1):5009.

23. Burn E, Tebe C, Fernandez-Bertolin S, et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. *Nat Commun*. 2021;12(1):777.

24. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424.

25. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37-48.

26. Harrell FE Jr. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer; 2001.

27. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol*. 2017;9:157-166.

28. Moons KG, Donders RA, Stijnen T, Harrell FE Jr. Using the out-come for imputation of missing predictor values was preferred. *J Clin Epidemiol*. 2006;59(10):1092-1101.

29. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse Tidyverse package. *J Open Source Softw*. 2019;4:1-6.

30. de Wreede LC, Fiocco M, Putter H. mstate: an R package for the analysis of competing risks and multi-state models. *J Stat Softw*. 2011;38(7). doi:10.18637/jss.v038.i07

31. Therneau TM. A package for survival analysis in R. 2020. https://m9.figshare.14959758.v1

32. Harrell FE Jr. A regression modeling strategies. R package version 5.1-0. 2017. https://cran.r-project.org/package=survival

33. Prieto-Alhambra D, Balló E, Coma E, et al. Filling the gaps in the characterization of the clinical management of COVID-19: 30-day hospital admission and fatality rates in a cohort of 118 150 cases diagnosed in outpatient settings in Spain. *Int J Epidemiol*. 2021;49(6):1930-1939.

34. Recalde M, Davila-Batista V, Diaz Y, et al. Body mass index and waist circumference in relation to the risk of 26 types of cancer: a prospective cohort study of 3.5 million adults in Spain. *BMC Med*. 2021;19(1):10.
35. Czeisler Mé, Marynak K, Clarke KEN, et al. Delay or avoidance of medical care because of COVID-19-related concerns—United States, June 2020. MMWR Morb Mortal Wkly Rep. 2020;69(36):1250-1257.

36. Sattar N, McInnes IB, McMurray JJV. Obesity is a risk factor for severe COVID-19 infection: multiple potential mechanisms. Circulation. 2020;142(1):4-6.

37. Leong A, Cole JB, Brenner LN, Meigs JB, Florez JC, Mercader JM. Cardiometabolic risk factors for COVID-19 susceptibility and severity: a mendelian randomization analysis. PLoS Med. 2021;18(3):e1003553.

38. Sattar N, Ho FK, Gill JM, et al. BMI and future risk for COVID-19 infection and death across sex, age and ethnicity: preliminary findings from UK biobank. Diabetes Metab Syndr. 2020;14(5):1149-1151.

39. Fressán U, Guevara M, Elía F, Albéniz E, Burgui C, Castilla J; Working Group for the Study of COVID-19 in Navarra. Independent role of severe obesity as a risk factor for COVID-19 hospitalization: a Spanish population-based cohort study. Obesity (Silver Spring). 2021;29(1):29-37.

40. Lighter J, Phillips M, Hochman S, et al. Obesity in patients younger than 60 years is a risk factor for COVID-19 hospital admission. Clin Infect Dis. 2020;71(15):896-897.

41. Kass DA, Duggal P, Cingolani O. Obesity could shift severe COVID-19 disease to younger ages. Lancet. 2020;395(10236):1544-1545.

42. Du Y, Lv Y, Zha W, Zhou N, Hong X. Association of body mass index (BMI) with critical COVID-19 and in-hospital mortality: a dose-response meta-analysis. Metabolism. 2020;154373.

43. Gupta S, Hayek SS, Wang W, et al; STOP-COVID Investigators. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. JAMA Intern Med. 2020;180(11):1436-1447.

44. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020;584(7821):430-436.

45. Tartof SY, Qian L, Hong V, et al. Obesity and mortality among patients diagnosed with COVID-19: results from an integrated health care organization. Ann Intern Med. 2020;173(10):773-781.

46. Hendren NS, de Lemos JA, Colby A, 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(2):135-144.

47. 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(6):350-359.