Characteristics and outcomes of 627,044 COVID-19 patients living with and without obesity in the United States, Spain, and the United Kingdom

Martina Recalde1,2,3, Elena Roel1,2,3, Andrea Pistillo1, Anthony G. Sena3,4, Albert Prats-Uribe5, Waheed-Ul-Rahman Ahmed6,7, Heba Alghoul8, Thamir M. Alshammari9, Osaid Alser10, Carlos Areia11, Edward Burn1,5, Paula Casajust12, Dalia Dawoud13, Scott L. DuVall14,15, Thomas Falconer16, Sergio Fernández-Bertolín1, Asieh Golozar17,18, Mengchun Gong19, Lana Yin Hui Lai20, Jennifer C. E. Lane6, Kristine E. Lynch14,15, Michael E. Matheny21,22, Paras P. Mehta23, Daniel R. Morales24, Karthik Natarjan16,25, Fredrik Nyberg26, Jose D. Posada27, Christian G. Reich28, Peter R. Rijnbeek4, Lisa M. Schilling29, Karishma Shah6, Nigam H. Shah27, Vignesh Subbian30, Lin Zhang31,32, Hong Zhu33,34, Patrick Ryan3,16, Daniel Prieto-Alhambra4,5, Kristin Kostka28,35,37 and Talita Duarte-Salles1,37

© The Author(s) 2021

BACKGROUND: A detailed characterization of patients with COVID-19 living with obesity has not yet been undertaken. We aimed to describe and compare the demographics, medical conditions, and outcomes of COVID-19 patients living with obesity (PLWO) to those of patients living without obesity.

METHODS: We conducted a cohort study based on outpatient/inpatient care and claims data from January to June 2020 from Spain, the UK, and the US. We used six databases standardized to the OMOP common data model. We defined two non-mutually exclusive cohorts of patients diagnosed and/or hospitalized with COVID-19; patients were followed from index date to 30 days or death. We report the frequency of demographics, prior medical conditions, and 30-days outcomes (hospitalization, events, and death) by obesity status.

RESULTS: We included 627,044 (Spain: 122,058, UK: 2,336, and US: 502,650) diagnosed and 160,013 (Spain: 18,197, US: 141,816) hospitalized patients with COVID-19. The prevalence of obesity was higher among patients hospitalized (39.9%, 95%CI: 39.8–40.0) than among those diagnosed with COVID-19 (33.1%; 95%CI: 33.0–33.2). In both cohorts, PLWO were more often female. Hospitalized PLWO were younger than patients without obesity. Overall, COVID-19 PLWO were more likely to have prior medical conditions, present with cardiovascular and respiratory events during hospitalization, or require intensive services compared to COVID-19 patients without obesity.

CONCLUSION: We show that PLWO differ from patients without obesity in a wide range of medical conditions and present with more severe forms of COVID-19, with higher hospitalization rates and intensive services requirements. These findings can help guiding preventive strategies of COVID-19 infection and complications and generating hypotheses for causal inference studies.

INTRODUCTION

Obesity is associated with increased mortality and is a well-known risk factor for chronic conditions, such as diabetes, hypertension, cardiovascular disease, and cancer [1, 2]. Due to its proinflammatory state that impairs the immune response, obesity has also been related to an increased risk of viral infections [3]. The novel coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), emerged in December 2019 in Wuhan, China, and rapidly spread around the world [4]. This new virus causes a respiratory tract infection with clinical manifestations ranging from asymptomatic/mild symptoms to severe illness requiring intensive services. Partly due to its similarities with other viral infections such as seasonal influenza or H1N1, people with obesity were soon labeled as “at-risk” individuals [5]. Since obesity is a worldwide public health priority, granular information on patients with COVID-19 and obesity is needed to guide preventive strategies as well as to generate hypotheses for etiological studies [6].

A review and meta-analysis of 75 studies reported that obesity is a risk factor for testing positive for SARS-CoV-2, for severe COVID-19 and for COVID-19 related mortality [7]. While undoubtedly relevant to the field, these studies mainly focused on exploring multiple risk factors related to COVID-19 and thus did not offer a detailed characterization of patients with COVID-19...
living with obesity. For instance, an exhaustive description of the medical conditions and COVID-19 related outcomes, such as thromboembolic events, among these patients is lacking. Other current limitations include the susceptibility to collider bias of studies reporting “risk factors” of COVID-19 infection and progression due to sampling mechanisms (e.g., subsamples of tested or hospitalized populations) [8]. A large characterization study focussing exclusively on patients with COVID-19 living with obesity using real-world data from different health settings and countries could address the limitations of the previous evidence. In this study, we aimed to describe and compare the demographics, medical conditions, and outcomes of COVID-19 patients living with obesity (PLWO) to those of COVID-19 patients living without obesity, in inpatient or outpatient settings.

**METHODS**

**Study design, setting, and data sources**

We conducted a multinational cohort study using routinely collected healthcare data from January to June 2020 from Spain, the United Kingdom (UK), and the United States (US). This study was part of the “Characterizing Health Associated Risks, and Your Baseline Disease In SARS-CoV-2” (CHARYBDIS) study (protocol available at https://www.ohdsi.org/wp-content/uploads/2020/07/Protocol_COVID-19-Charybdis-Characterisation_V5.docx) designed by the Observational Health Data Sciences and Informatics (OHDSI) community. All data were standardized to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) [9]. The OHDSI network maintains the OMOP-CDM, along with a wide range of tools developed by its members to facilitate analyses of mapped data [10]. Data results for this study were extracted on July, 16th, 2020.

We included primary, outpatient and inpatient care data from electronic health records (EHRs) and health insurance claims data from six databases. Data from Spain included the Information System for Research in Primary Care (SIDIAP), which includes primary linked to inpatient care data covering approximately 80% of the population in Catalonia, Spain [11]. The UK data covered the Clinical Research Datalink (CPRD), with patients from over 600 general practices in the UK [12]. Data from the US included: Columbia University Irving Medical Center (CUIMC), covering patients from over 600 general practices in the United Kingdom (UK), and the United States (US). The study was descriptive by nature and, therefore, statistical modeling was out of scope. Differences across the groups compared should not be interpreted as causal effects.

To ensure data privacy at all times, we employed a federated analysis approach [16]. Following a pre-specified analysis plan, a common analytical code for the whole CHARYBDIS study was developed for the OHDSI Methods library, available at https://github.com/ohdsi-studies/Covid19CharacterizationCharybdis, and was run locally in each data-base. Individual-level data remained within host institutions and only aggregate results from each database were provided to the research team and publicly shared. All the results reported in this paper and additional data are available for consultation at a dynamic and interactive website, which changes over time as new databases are added and/or results are updated to CHARYBDIS (https://data.ohdsi.org/Covid19CharacterizationCharybdis). We used R version 3.6 for data visualization. All the data partners obtained Institutional Review Board (IRB) approval or exemption to conduct this descriptive study.

**RESULTS**

**Prevalence of obesity**

We included 627 044 diagnosed and 160 013 hospitalized patients with COVID-19 (Table 1). The diagnosed cohort consisted of 122 058 patients from Spain (SIDIAP), 2336 from the UK (CPRD), and 502 650 from the US (CUIMC: 8519; IQVIA-OpenClaims: 466 191; STARR-OMOP: 3328; VA-OMOP: 24 612). Among diagnosed and hospitalized patients, 207 859 (33.1%, 95% CI 33.0 – 33.2) and 63 866 (39.9%, 95% CI 39.8 – 40.0) had obesity, respectively. In all databases, the prevalence ever-recorded obesity diagnosis (Supplementary Appendix 3) and/or a body mass index (BMI) measurement between 30 and 60 kg/m² and/or a bodyweight measurement between 120 and 200 kg prior or at index date. We included upper cut-off thresholds to discard implausible observations. Patients without obesity were those who did not fulfill the obesity definition.

**Baseline characteristics and outcomes of interest**

Demographics (sex and age) were obtained at the index date. More than 15 000 medical conditions from up to one year prior to the index date were identified based on the Systematized Nomenclature of Medicine (SNOMED) hierarchy, with all descendant codes included [15]. Specific definitions for comorbidities of particular interest were created; the detailed definitions of these variables can be consulted in Supplementary Appendix 3. We reported here a list of key comorbidities based on their prevalence in the cohorts of the participating sites, as well as on their clinical relevance to obesity and the COVID-19 research field [17]. Our 30-day outcomes of interest for the diagnosed cohort were hospitalization and fatality. For the hospitalized cohort, the 30-days outcomes were a requirement of intensive services (IS) (identified by a recorded mechanical ventilation and/or a tracheostomy and/or extra-corporeal membrane oxygenation procedure), respiratory, cardiovascular, thromboembolic, and other events and fatality.

**Data analysis**

We described the number of patients included and the prevalence of obesity in each database as well as the demographics, comorbidities, and outcomes as proportions (calculated by the number of persons within a given category, divided by the total number of persons) with their respective 95% confidence intervals (CIs) for each database, by obesity status. To calculate these proportions in each database, we established a minimum count required (of five individuals), to minimize the risk of re-identification of patients. To compare medical conditions across groups, we calculated standardized mean differences (SMDs) [18], which we summarized in Manhattan-style plots. The SMD can be used to compare the prevalence of a dichotomous variable between two groups and is independent of sample size [19]. A SMD > 0.1 indicates a meaningful difference in the prevalence of a given condition; in the context of this study, a SMD > 0.1 indicates a higher prevalence in PLWO, whereas a SMD < −0.1 indicates a higher prevalence among patients without obesity. This study was descriptive by nature and, therefore, statistical modeling was out of scope. Differences across the groups compared should not be interpreted as causal effects.

To ensure data privacy at all times, we employed a federated analysis approach [16]. Following a pre-specified analysis plan, a common analytical code for the whole CHARYBDIS study was developed for the OHDSI Methods library, available at https://github.com/ohdsi-studies/Covid19CharacterizationCharybdis, and was run locally in each data-base. Individual-level data remained within host institutions and only aggregate results from each database were provided to the research team and publicly shared. All the results reported in this paper and additional data are available for consultation at a dynamic and interactive website, which changes over time as new databases are added and/or results are updated to CHARYBDIS (https://data.ohdsi.org/Covid19CharacterizationCharybdis). We used R version 3.6 for data visualization. All the data partners obtained Institutional Review Board (IRB) approval or exemption to conduct this descriptive study.
|                     | SIDIAP (Spain) | CPRD (UK) | CUIMC (US) | IQVIA-Open Claims (US) | STARR-OMOP (US) | VA-OMOP (US) |
|---------------------|----------------|-----------|------------|------------------------|----------------|--------------|
| Diagnosed with COVID-19 |                |           |            |                        |                |              |
| All, n              | 36 409         | 85 649    | 976        | 1360                   | 3446           | 5073         |
| All, %* (95%CI)     | 29.8 (29.5–30.1) | 70.2 (69.9–70.5) | 41.8 (39.8–43.8) | 58.2 (56.2–60.2) | 40.5 (39.5–41.5) | 33.1 (33.0–33.2) |
| Laboratory confirmed, % (95%CI) | 39.2 (38.9–39.5) | 27.2 (27–27.4) | 51.5 (49.5–53.5) | 50.3 (48.3–52.3) | 66.7 (65.7–67.7) | – – |
| Female sex, % (95%CI) | 62.5 (62.2–62.8) | 55.9 (55.6–56.2) | 56.2 (54.2–58.2) | 57.2 (55.2–59.2) | 60.9 (59.9–61.9) | 64.8 (55.3–57.5) |
| Age, % (95%CI)      |                |           |            |                        |                |              |
| <18 years           | 1.2 (1.1–1.3)  | 4.7 (4.6–4.8) | 1.2 (1–1.4) | 3.5 (3.1–3.9)          | 1.0 (0.97–1.03) | 3.9 (3.8–4) |
| 18–64 years         | 58.9 (58.6–59.2) | 75.6 (75.4–75.8) | 55.4 (53.4–57.4) | 52.3 (50.3–54.3) | 65.2 (64.2–66.2) | 65.8 (64.8–66.8) |
| >65 years           | 40.0 (39.7–40.3) | 19.6 (19.4–19.8) | 44.7 (42.7–46.7) | 47.7 (45.7–49.7) | 33.6 (32.6–34.6) | 30.6 (29.6–31.6) |
| Hospitalized with COVID-19 |                |           |            |                        |                |              |
| All, n              | 8403           | 9794      | – –        | 1408                   | 1192           | 50863        |
| All, % (95%CI)      | 46.2 (45.5–46.9) | 53.8 (53.1–54.5) | – –        | 54.2 (52.3–56.1) | 45.8 (43.9–47.7) | 38.2 (37.9–38.5) |
| Laboratory confirmed, % (95%CI) | 76.0 (75.4–76.6) | 72.9 (72.3–73.5) | – –        | 88.5 (87.3–89.7) | 91.6 (90.5–92.7) | 10.6 (8.2–13) |
| Female sex, % (95%CI) | 51.2 (50.5–51.9) | 39.9 (39.2–40.6) | – –        | 54.7 (52.8–56.6) | 39.8 (37.9–41.7) | 55.0 (54.7–55.3) |
| Age, % (95%CI)      |                |           |            |                        |                |              |
| <18 years           | – –            | – –       | – –        | – –                    | 0.5 (0.5–0.5)  | 2.0 (1.9–2.1) |
| 18–64 years         | 37.9 (37.2–38.6) | 50.6 (49.9–51.3) | – –        | 53.3 (51.4–55.2) | 34.5 (32.7–36.3) | 35.4 (32.7–36.3) |
| >65 years           | 62.1 (61.4–62.8) | 49.4 (48.7–50.1) | – –        | 46.7 (46.1–48.6) | 65.6 (63.8–67.4) | 47.8 (47.5–48.1) |

Notes. * Proportion of patients with and without obesity among all patients (row percentage); – data not available or below the minimum cell count required (five individuals)
Abbreviations: CI confidence interval, COVID-19 coronavirus disease 2019, CPRD Clinical Practice Research Datalink, CUIMC Columbia University Irving Medical Center, SIDIAP Information System for Research in Primary Care, STARR-OMOP Stanford Medicine Research Data Repository, UK United Kingdom, US United States, VA-OMOP United States Department of Veterans Affairs.
of obesity was lower among diagnosed patients than among those hospitalized, with differences ranging from 5 (IQVIA-OpenClaims) to 16% (SIDIAP).

Baseline demographics
The sex distribution (proportions and 95% CIs) of the patients are reported in Table 1. Aside from VA-OMOP, in the diagnosed cohort, patients with and without obesity were mostly female. The proportion of females was higher among PLWO compared to patients without obesity in SIDIAP (63% vs 56%), CUIMC (61% vs 56%), and IQVIA-OpenClaims (61% vs 52%), while in VA-OMOP the opposite was observed (13% vs 19%). No differences were observed in CPRD and STARR-OMOP. In the hospitalized cohort, patients without obesity were predominantly male (female ranged from 40 to 49%, VA-OMOP: 4%) but PLWO still were more commonly female in all databases aside from VA-OMOP (range: 51–55%, VA-OMOP: 7%). Differences in the proportion of females between PLWO and patients without obesity ranged from 3 (VA-OMOP) to 15% (CUIMC).

The age distribution in each database is summarized in Table 1 with proportions and their respective 95% CIs and in Fig. 1 with histograms. In the diagnosed cohort, PLWO were slightly older than those without obesity (i.e., the age distribution for PLWO was slightly skewed to the left compared to patients without obesity). This was particularly marked in SIDIAP, where 40% of the PLWO were aged above 65 years and only 20% were so without obesity. Hospitalized patients were older than those diagnosed. In the hospitalized cohorts, PLWO were fairly consistently younger than those without obesity (except for SIDIAP). The proportion of patients aged above 65 ranged from 36 to 63% for PLWO and from 43 to 73% for those without obesity.

Baseline medical conditions
We compared baseline medical conditions of PLWO to those of patients without obesity in the diagnosed and hospitalized cohorts using SMDs, which are summarized in Fig. 2. We depicted the SMDs of 485 (CPRD) to 5050 (VA-OMOP) medical conditions in the diagnosed cohort, and 529 (STARR-OMOP) to 5240 (IQVIA-OpenClaims) in the hospitalized cohort. In both cohorts, medical conditions were largely more frequent among PLWO than patients without obesity.

The distribution of the selected key comorbidities is shown in Fig. 3, and the proportions with their respective 95% CIs and SMDs between PLWO and patients without obesity are available in Supplementary Appendices 4 and 5. In the diagnosed cohorts, PLWO consistently had a higher prevalence of comorbidities compared to those without obesity; these differences were meaningful (i.e., with a SMD > 0.1, which indicates a meaningfully higher prevalence among PLWO) for the majority of comorbidities across databases. For example, while the prevalence of hypertension for PLWO ranged from 30 to 32% in Europe (SIDIAP and CPRD) and from 55 to 81% in the US, in those without obesity it ranged from 12 to 16% and from 26 to 53%, respectively. The SMD for hypertension was above 0.1 in all databases. As in the diagnosed cohort, PLWO hospitalized with COVID-19 had a higher prevalence of comorbidities than those without obesity, and these differences were meaningful for the majority of comorbidities. However, the differences between groups were less obvious. For example, heart disease differed by 20% among those diagnosed in VA-OMOP (PLWO: 60%, without obesity: 40%) and by 9% among those hospitalized (PLWO: 74%, without obesity: 65%); although the SMD was still above 0.1 in all databases.

30-day outcomes of interest
The distribution of 30-days outcomes is shown in Fig. 4, the proportions with their respective 95% CI and SMDs between PLWO and patients without obesity are available in Table 2. In the diagnosed cohorts, hospitalization rates were higher among PLWO than those without obesity in all databases. For example, in SIDIAP the proportion of patients hospitalized was 20% for PLWO and 10% for patients without obesity. However, these differences were meaningful (SMD > 0.1) only in three databases: SIDIAP, CUIMC, and STARR-OMOP. In PLWO, mortality ranged from 5 to 12% and was higher than in patients without obesity in SIDIAP and CUIMC (7% vs 3% and 8% vs 5%, respectively), while in CPRD and VA-OMOP it was similar in both groups. SIDIAP was the only database with a meaningful difference in the proportion of mortality.

Overall, in the hospitalized cohort, PLWO more frequently had adverse events occurring in the 30 days after the index date than patients without obesity. For example, PLWO required IS and presented with ARDS more frequently than patients without obesity in the largest databases: IQVIA-OpenClaims (IS: 13% vs 10%; ARDS: 35% vs 31%) and VA-OMOP (IS: 22% vs 15%; 46% vs 41%), whereas in CUIMC and STARR-OMOP percentages were similar. VA-OMOP was the only database with a meaningful
difference in the proportion of IS. Similarly, heart failure was also more frequent among PLWO than among patients without obesity in CUIMC: 7% vs 3%, IQVIA-OpenClaims: 7% vs 5%, STARR-OMOP: 16% vs 9%, and VA-OMOP: 23% vs 17%), these differences were meaningful in CUIMC and STARR-OMOP. Sepsis, cardiac arrhythmia, and cardiovascular disease events were slightly more frequent among PLWO, although SMDs were below 0.1 in all databases. Acute kidney injury was the only outcome that was more frequent among patients without obesity; however, this difference was not meaningful in any database.
As for fatality, there were no consistent nor meaningful differences between PLWO and patients without obesity in the hospitalized cohort: while it was higher for PLWO in SIDIAP (14% vs 11%), there were no differences in CUIMC (20% vs 21%) nor in VA-OMOP (16% vs 18%).

**DISCUSSION**

In this large cohort study including 627,044 COVID-19 patients from Spain, the UK, and the US, we found that the prevalence of obesity was higher among COVID-19 patients hospitalized (40%) compared to those diagnosed (31%). PLWO diagnosed and hospitalized with COVID-19 were more commonly female, and those hospitalized were younger than patients without obesity. The extraction of more than 15,000 medical conditions revealed PLWO were not only more prone to have obesity-related comorbidities, such as hypertension, heart disease, and type 2 diabetes but also to more than a thousand different health conditions. After 30-days of follow-up, PLWO presented with higher hospitalization rates and intensive services requirements, although these differences were only meaningful in some databases.

Our study has several strengths, such as its large amount of data. By bringing together harmonized data using a federated approach, we have conducted a large-scale study while respecting the confidentiality of patient records. The international approach of this study is a strong asset given that we are investigating the intersection of two major global threats, namely the obesity epidemic and the COVID-19 pandemic. The former, together with the diverse healthcare settings and populations described in this study, increase the generalizability of our findings. Further, we provide a wide overview of the characteristics and outcomes of patients with and without obesity, using data visualization tools to summarize large amounts of medical data. This exhaustive characterization goes far beyond prior studies reporting few comorbidities and supports the generation of new hypotheses that can be tested in future studies. In addition, for the sake of transparency and reproducibility, we have made methods, tools, and all results publicly available. As CHARITYDIS is an ongoing study, results (including longer follow-up time) will be updated and new studies focussing on obesity could be conducted. All of the above has been accomplished through the coordinated efforts of the OHDSI community to provide a rapid response to the COVID-19 pandemic.

Our study also has limitations. First, we cannot exclude a selection bias of COVID-19 cases due to underreporting in the context of testing restrictions and asymptomatic or paucisymptomatic cases. Additionally, testing policies have varied across countries and time depending on the course of the pandemic. Nevertheless, the inclusion of patients clinically diagnosed (not tested) in different settings likely provided consistency to our data, although it might have incurred in false positives. Second, we did not have information on BMI as a continuous variable, which prevented us from investigating the impact of different categories of obesity in COVID-19 outcomes. This might explain the higher proportion of comorbidities and outcomes observed in the US databases, as PLWO from the US might have higher BMIs than those from Europe. In addition, our definition of obesity included diagnoses and measurements recorded at any time prior to or at the index date, and therefore some individuals might have been misclassified due to changes in BMI since the most recently recorded status. However, previous evidence shows that BMI trajectories in adults are relatively stable, with a tendency to increase with age. Finally, this study was underpinned by routinely collected data which can raise concerns about the quality of the data. Some databases are prone to oversampling certain groups of people as a result of how these
|                  | SIDIAP (Spain) | CPRD (UK) | CUIMC (US) | IQVIA-Open Claims (US) | STARR-OMOP (US) | VA-OMOP (US) |
|------------------|---------------|-----------|------------|------------------------|----------------|-------------|
|                  | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD | With obesity  | Without obesity | SMD |
| Patients diagnosed, n | 36 409         | 854 649    | NA          | 976         | 1360             | NA          | 154 325       | 311 866            | NA          | 1157         | 2171              | NA          | 11 546       | 13 066             | NA          |
| Hospitalization  | 20.3 (19.9 20.7) | 10.3 (10.2 10.4) | 0.20   | –            | –                | 0.26   | 32.1 (31.9 32.3) | 25.8 (25.6 26.0) | 0.10   | 22.1 (19.7 24.5) | 15.8 (14.3 17.3) | 0.11   | 16.6 (15.9 17.3) | 12.8 (12.2 13.4) | 0.08   |
| Death            | 71.6 (68.7 74) | 3.3 (3.3 3.3) | 0.12   | 12.3 (10.2 14.4) | 11.3 (9.6 13.0) | 0.02   | 8.0 (7.1 8.9) | 5.0 (4.4 5.6) | 0.09   | –            | –                | –            | –            | –            | –            | –            |
|                  |               |           |          |            |                  |          |                |                    |        |               |                  |        |               |                  |        |               |                  |        |
| Patients hospitalized, n | 8403         | 9794      | NA          | –            | –                | NA          | 50 863         | 8228              | NA        | 274          | 341              | NA         | 2918         | 2592            | NA         |
| Intensive service requirement | –            | –        | –        | –            | –                | –            | 2.3 (1.5 3.1) | 2.0 (1.2 2.8) | 0.01   | 13.0 (12.7 13.3) | 9.7 (9.1 10.3) | 0.06   | 9.1 (5.7 12.3) | 6.2 (3.6 8.8) | 0.08   |
| Death            | 13.9 (13.2 14.6) | 10.9 (10.3 11.5) | 0.06   | –            | –                | –            | 19.5 (17.4 21.6) | 20.6 (18.3 22.9) | –0.02  | –            | –                | –            | –            | –            | –            | –            |
| Cardiovascular events | –            | –        | –        | –            | –                | –            | 2.6 (1.8 3.4) | 2.5 (1.6 3.4) | 0.00   | 17.1 (16.8 17.6) | 19.1 (18.1 20.2) | –0.01  | 2.4 (1.8 3.8) | 4.0 (3.2 4.8) | –0.06  |
|                   |               |          |          |            |                  |          |                |                    |        |               |                  |        |               |                  |        |               |                  |        |
| Total cardiovascular disease events | –            | –        | –        | –            | –                | –            | 7.3 (5.9 8.7) | 5.9 (4.6 7.2) | 0.04   | 82.0 (80.0 84.0) | 7.6 (7.0 8.2) | 0.02   | 11.7 (7.9 15.5) | 8.8 (5.8 11.8) | 0.07   |
| Thromboembolic events | –            | –        | –        | –            | –                | –            | 2.1 (1.4 2.8) | 1.8 (1.0 2.6) | 0.02   | 2.2 (2.1 2.3) | 1.7 (1.4 2.0) | 0.03   | –            | –            | –            |
|                   |               |          |          |            |                  |          |                |                    |        |               |                  |        |               |                  |        |               |                  |        |
| Acute kidney injury* | –            | –        | –        | –            | –                | –            | 18.1 (16.1 20.1) | 25.4 (22.9 27.9) | –0.01  | 11.1 (10.8 11.4) | 10.4 (9.7 11.1) | 0.04   | 8.4 (5.1 11.7) | 9.1 (6.0 12.2) | 0.08   |
| Acute pancreatitis | –            | –        | –        | –            | –                | –            | 17.9 (15.9 19.9) | 15.2 (13.2 17.2) | 0.05   | 34.9 (34.5 35.3) | 30.5 (29.5 31.5) | 0.07   | 14.6 (10.4 18.8) | 10.0 (6.8 13.2) | 0.10   |
| ARDS*            | –            | –        | –        | –            | –                | –            | 0.02         | 0.00          | –        | –            | –                | –        | –            | –            | –        | –            | –0.06  |
The prevalence of obesity was greater in COVID-19 patients living with obesity. A SMD > 0.1 indicates a meaningful difference in the prevalence of a given condition. The consistency of our findings across databases that differ by setting and country lends credence to the generalizability of our findings.

Table 2 continued

| Obesity | SIDIAP (Spain) | CPRD (US) | VA-OMOP (US) | CUIMC (US) | STARR-OMOP (US) | IQVIA-Open Claims (US) |
|---------|----------------|-----------|--------------|------------|----------------|----------------------|
| With    | SMD            | With      | SMD          | With       | SMD            | With                 |
| Obesity | With           | Without    | SMD          | With       | SMD            | Without              |
|         | 1.2 (1.1)      | 1.6 (1.1)  | 1.6 (1.0)    | 1.3 (0.9)  | 1.3 (1.2)      |                     |
| 1.3 (1.1)| 1.9 (1.0)      | 1.7 (1.0)  | 1.2 (1.1)    | 1.2 (1.0)  | 1.0 (0.4)      |                     |
| 1.4 (1.2)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 1.5 (1.2)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 1.6 (1.1)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 1.7 (1.0)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 1.8 (1.1)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 1.9 (1.1)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.0 (1.2)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.1 (1.3)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.2 (1.4)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.3 (1.5)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.4 (1.6)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.5 (1.7)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.6 (1.8)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.7 (1.9)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.8 (2.0)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 2.9 (2.1)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |
| 3.0 (2.2)| 1.9 (1.0)      | 1.2 (1.0)  | 1.2 (1.0)    | 1.3 (1.1)  | 1.3 (1.2)      |                     |

Notes: - data not available or below the minimum cell count required five individual events marked with an * were recorded during hospitalization. SMD < 0 means the prevalence was greater in COVID-19 patients living without obesity. SMD > 0 means the prevalence was greater in COVID-19 patients living with obesity. A SMD > |0.1| indicates a meaningful difference in the prevalence of a given condition. ARDS acute respiratory distress syndrome, CI confidence interval, COVID-19 coronavirus disease 2019, CPRD Clinical Practice Research Datalink, CUIMC Columbia University Irving Medical Center, VA United States Department of Veterans Affairs, OMOP Observational Medical Outcomes Partnership.
studies aiming to disentangle the effect of obesity in COVID-19 outcomes should have this information present and consider data-driven techniques to account for confounding, such as propensity score estimation and its adjustment methods [18].

Finally, PLWO experienced adverse events more frequently than those without obesity, particularly hospitalization and the requirement of intensive services. Certainly, our results must be interpreted carefully considering the differences in demographics and comorbidities between these groups. Interestingly, in patients hospitalized, we did not observe clear differences in fatality between patients with and without obesity. While two meta-analyses reported that obesity is associated with a higher risk of COVID-19 related mortality; other large observational studies from the US and the UK using finer categories of BMI only found an association with mortality for morbid obesity (BMIs ≥ 35 kg/m² or ≥40 kg/m²) [7, 25, 28, 29, 34]. Given the scarcity of evidence regarding the frequency of specific adverse events during hospitalization among PLWO, our findings are of special interest to the field and should be addressed in upcoming etiological studies.

In this large international cohort, we showed that among COVID-19 cases, PLWO were more likely to be female, have more comorbidities, and worse outcomes than patients without obesity. The prevalence of obesity was higher among hospitalized patients with COVID-19 compared to patients diagnosed with COVID-19. Our results may be useful in guiding clinical practice and aid future preventative strategies for patients living with obesity, as well as providing useful data to support subsequent etiological studies focussed on obesity and COVID-19.

REFERENCES

1. Di Angelantonio E, Bhupathiraju SN, Wormser D, Gao P, Kaptoge S, de Gonzalez AB, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet. 2016;388:776–86.
2. Hu FB. Obesity Epidemiology. New York: Oxford University Press, 2009.
3. Mozaffarian D, Flegal K, Dietz W, Flood P, Sacks F. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001–2004. JAMA. 2007;298:1884–92.
4. WHO. WHO Director-General’s opening remarks at the media briefing on COVID-19—11 March 2020. WHO. 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11–march–2020.
5. Louie JK, Acosta M, Samuel MC, Schechter R, Vigia DJ, Harriman K, et al. A novel risk factor for a novel virus: obesity and 2009 pandemic influenza a (H1N1). Clin Infect Dis. 2011;52:301–12.
6. GOV.UK. New obesity strategy unveiled as country urged to lose weight to beat coronavirus and protect the NHS. GOV.UK. 2020. https://www.gov.uk/government/news/new-obesity-strategy-unveiled-as-country-urged-to-lose-weight-to-beat-coronavirus-covid-19-and-protect-the-nhs.
7. Popkin BM, Du S, Green WD, Beck MA, Algaith T, Herbst CH, et al. Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. Obes Rev. 2020;21:e13128.
8. Kobusingye J, Hering-Smith S, Morris JS, Jenkinson P, Xiang Z, Simmonds A, et al. 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;13:3–9.
9. WHO. WHO Director-General’s opening remarks at the media briefing on COVID-19—11 March 2020. WHO. 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11–march–2020.
10. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics. The Book of OHDSI (Independent published). 2019. https://ohdsi.github.io/TheBookOfOhdsi/.
11. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. JAMA. 2020;323:2052–9.
12. Austin PC. An introduction to propensity score methods for reducing the bias of observational studies. Multivariate Behav Res. 2011;46:399–424.
13. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. Commun Stat Simul Comput. 2009;38:1228–34.
14. M. Recalde et al. [Preprint]. 2020 [cited 2020 Aug 20]. Available from: https://arxiv.org/abs/2003.10534v1
15. Burn E, Tebe C, Fernandez-Bertolini S, Aragon M, Recalde M, Roel E, et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. Nat Commun. 2021;12:1–12.
16. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O’Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966.
17. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ. 2020;369:m1985.
18. Gupta S, Hayek SS, Wang W, Chan L, Mathews KS, Melamed ML, et al. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. JAMA Intern Med. 2020;180:1436–47.
19. Anderson MR, Geleris J, Andersen DR, Zucker J, Nobel YR, Freedberg D, et al. Body mass index and risk for intubation or death in SARS-CoV-2 infection: a retrospective cohort study. Ann Intern Med. 2020;173:773–81.
20. Du Y, Lv Y, Zha W, Zhou N, Hong X. Association of body mass index (BMI) with mortality for morbid obesity (BMIs ≥ 35 kg/m² or ≥40 kg/m²) trajectories across adulthood and smoking in relation to prostate cancer risks: the NIH-AARP diet and health study. Int J Epidemiol. 2019;48:644–73.
21. Sattar N, Ho FK, Gill JM, Ghouri N, Gray SR, Cellis-Morales CA, 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 Clin Res Rev. 2020;14:1149–51.
22. Tartof SY, Qian L, Hong Y, Wei R, Nadjaif RF, Fischer H, et al. Obesity and mortality among patients diagnosed with COVID-19: results from an integrated health care organization. Ann Intern Med. 2020;173:773–81.
23. Popkin BM, Du S, Green WD, Beck MA, Algaith T, Herbst CH, et al. Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. Obes Rev. 2020;21:e13128.
24. Griffith GJ, Morris TT, Tuddall MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. Nat Commun. 2020;11:1–12.
25. Voss EA, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. J Am Med Informatics Assoc. 2015;22:533–64.
26. Hripceusk D, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI) opportunities for observational researchers. Stud Health Technol Inform. 2015;216:574–8.
27. Bolibar B, Fina Avilés F, Morros R, Del Mar García-Gil M, Hermosilla E, Ramos R, et al. Base de datos SIDAP: La historia clínica informatizada de Atención Primaria como fuente de información para la investigación epidemiológica. Med Clin. 2012;138:167–21.
28. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, Staa Tvan, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44:827–36.
29. Datta S, Posada J, Olson G, Li W, O’Reilly C, Balraj D, et al. A new paradigm for accelerating clinical data science at Stanford Medicine. arXiv:2003.10534v1

ACKNOWLEDGEMENTS

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.

AUTHOR CONTRIBUTIONS
MR, ER, APu, PR, DPA, KK, and TDS conceived and designed the study. SDL, TF, KEL, MEM, KN, JDP, CGR, NHS, FR, KK, and TDS coordinated data contributions at their respective sites. AP, AGS, TF, SBE, JDP, KK, and TDS analyzed the data; MR, ER, and AP produced the figures and tables. MR, ER, EB, DRM, FN, PR, LMS, DPA, KK, and TDS interpreted the data. MR, ER, and TDS searched the literature and wrote the first draft with insightful contributions from EB, LYHL, JCEL, DRM, FN, PR, LMS, DPA, and KK. All authors contributed to the revision of the first draft, reviewed, and approved the final version of the paper.

FUNDING
The European Health Data & Evidence Network has 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 program and EFPIA. This research received partial support from the National Institute for Health Research (NIHR) Oxford Biomedical Research Center (BRC), US National Institutes of Health, US Department of Veterans Affairs, Janssen Research & Development, and IQVIA. The University of Oxford received funding related to this work from the Bill & Melinda Gates Foundation (Investment ID INV-016201 and INV-019257). APu has received funding from the Medical Research Council (MRC) [MR/K501256/1, MR/N013468/1] and Fundación Alfonso Martín Escudero (FAME) (APu). Vinc1 [VA HSR RES 13-457] (SDL, MEM, KEL). JCEL has received funding from the Medical Research Council (MRC) [MR/K501256/1] and Versus Arthritis (17170). MR is funded by the Research Council (MRC) [MR/K501256/1, MR/N013468/1] and Aurora Healthcare (Aurora Healthcare, Aurora, Alaska, USA). The authors contributed to the revision of the first draft, reviewed, and approved the final version of the paper.

ETHICAL APPROVAL
All the data partners received Institutional Review Board (IRB) approval or exemption. STARR-MOP had approval from IRB Panel #B (RB-53248) registered to Leland Stanford Junior University under the Stanford Human Research Protection Program (HRPP). The use of VA data was reviewed by the Department of Veterans Affairs Central IRB, was determined to meet the criteria for exemption under Exemption Category 4 (3), and approved for Waiver of HIPAA Authorization. The research was approved by the Columbia University Institutional Review Board as an OHDSI network study. The use of SIDIAp was approved by the Clinical Research Ethics Committee of the IDIAPJGol (project code: 20/070-PCV). The use of CPRD was approved by the Independent Scientific Advisory Committee (ISAC) (protocol number 20_059RA2). The use of IQVIA-OpenClaims was exempted from IRB approval.

COMPETING INTERESTS
All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: Mr. Sena reports personal fees from Janssen Research & Development, outside the submitted work; Dr. DuVall reports grants from Anolinx, LLC, grants from AstraZeneca Pharmaceuticals LP, grants from Boehringer Ingelheim International GmbH, grants from Celgene Corporation, grants from Eli Lilly and Company, grants from Genentech Inc., grants from Genomic Health, Inc., grants from Gilead Sciences Inc., grants from GlaxoSmithKline PLC, grants from Innocin Pharmaceuticals Inc., grants from Janssen Pharmaceuticals, Inc., grants from Kantar Health, grants from Myriad Genetic Laboratories, Inc., grants from Novartis International AG, grants from Parexel International Corporation through the University of Utah or Western Institute for Biomedical Research, outside the submitted work; Mr. Ahmed reports funding from the NIH Oxford Biomedical Research Center (BRC), AstaZeneca Foundation, and the Royal College Surgeons of England; Dr. Golozar reports personal fees from Regeneron Pharmaceuticals, outside the submitted work. She is a full-time employee at Regeneron Pharmaceuticals. This work was not conducted at Regeneron Pharmaceuticals. Miss Lane is supported by a Medical Research Council Doctoral Research Fellowship (MR/K501256/1) and a Versus Arthritis Clinical Research Fellowship (21605). Dr. Morales is supported by a Wellcome Trust Clinical Research Development Fellowship (Grant 214588/Z/18/Z) and reports grants from Chief Scientist Office (CSO), grants from Health Data Research UK (HDr-UK), grants from National Institute of Health Research (NIHR), outside the submitted work; Dr. Nyberg reports other from AstraZeneca, outside the submitted work; Dr. Subban reports grants from National Science Foundation, grants from State of Arizona; Arizona Board of Regents, grants from Agency for Healthcare Research and Quality, outside the submitted work; Dr Prieto-Alhambra reports grants and other from AMGEN; grants, non-financial 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 consortia, and Synapse Management Partners have supported training programs organized by DPA’s department and open for external participants. Ms. Kostka and Dr. Reich report being employees of IQVIA Inc at the time the analysis was conducted. Dr. Rijnbeek reports grants from Innovative Medicines Initiative, grants from Janssen Research and Development, during the conduct of the study. Dr. Ryan is an employee of Janssen Research and Development and shareholder of Johnson & Johnson. The views expressed are those of the authors and do not necessarily represent the views or policy of the Department of Veterans Affairs or the United States Government. No other relationships or activities that could appear to have influenced the submitted work.

DATA SHARING STATEMENT
Analyses were performed locally in compliance with all applicable data privacy laws. Although the underlying data is not readily available to be shared, authors contributing to this paper have direct access to the data sources used in this study. All results (e.g., aggregate statistics, not presented at a patient-level with redactions for minimum cell count) are available for public inquiry. These results are inclusive of site-identifiers by contributing data sources to enable interrogation of each contributing site. All analytic code and result sets are made available at: https://github.com/ohdsi-studies/Covid19CharacterizationCharybdis.

ADDITIONAL INFORMATION
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41366-021-00893-4.
Correspondence and requests for materials should be addressed to T.D-S.
Reprints and permission information is available at http://www.nature.com/reprints
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1Faciudat Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol i Gurina (IDIAPIJGol), Barcelona, Spain.
2Universitat Autònoma de Barcelona, Bellaterra, Spain.
3Janssen Research & Development, Titusville, NJ, USA.
4Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.
5Centre for Statistics in Medicine, NDORMS, University of Oxford, Oxford, UK.
6Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Botnar Research Centre, Oxford, UK.
7College of Medicine and Health, University of Exeter, St Luke’s Campus, Exeter, UK.
8Faculty of Medicine, Islamic University of Gaza, Gaza, Palestine.
9College of Pharmacy, Riyadh Elm University, Riyadh, Saudi Arabia.
10Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.
11Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK.
12Real-World Evidence, Trial Form Support, Barcelona, Spain.
13Cairo University, Faculty of Pharmacy, Cairo, Egypt.
14VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT, USA.
15Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA.
16Pharmacoeconomics, Regeneron Pharmaceuticals, Tarrytown, NY, USA.
17DHCon10 Limited, Beijing, China.
18Division of Cancer Sciences, School of Medical Sciences, University of Manchester, Manchester, UK.
19College of Medicine, University of Tennessee Health Science Center, Memphis, TN, USA.
20Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA.
21College of Medicine, The University of Arizona.
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the