

# Prioritizing Causal Risk Factors for Severe COVID-19: An Exhaustive Mendelian Randomization Study

Yitang Sun

University of Georgia

Jingqi Zhou

University of Georgia

Kaixiong Ye (✉ [kaixiong.ye@uga.edu](mailto:kaixiong.ye@uga.edu))

University of Georgia <https://orcid.org/0000-0003-4658-7292>

---

## Article

**Keywords:** COVID-19, BMI, fat mass, Mendelian randomization, neutrophils, proteins

**Posted Date:** January 25th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-149087/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Communications Medicine on December 9th, 2021. See the published version at <https://doi.org/10.1038/s43856-021-00061-9>.

# Abstract

Identifying causal risk factors for severe coronavirus disease 2019 (COVID-19) is critical for its prevention and treatment. Many associated pre-existing conditions and biomarkers have been reported, but these observational associations suffer from confounding and reverse causation. Here, we perform a large-scale two-sample Mendelian randomization (MR) analysis to evaluate the causal roles of many traits in severe COVID-19. Our results highlight multiple body mass index (BMI)-related traits as risk-increasing: BMI (OR:1.89, 95% CI:1.51–2.37), hip circumference (OR:1.46, 1.15–1.85), and waist circumference (OR:1.82, 1.36–2.43). Our multivariable MR analysis further shows that the BMI-related effect is driven by fat mass (OR:1.63, 1.03–2.58), but not fat-free mass (OR:1.00, 0.61–1.66). Several white blood cell counts are negatively associated with severe COVID-19, including those of neutrophils (OR:0.76, 0.61–0.94), granulocytes (OR:0.75, 0.601–0.93), and myeloid white blood cells (OR:0.77, 0.62–0.96). Furthermore, some circulating proteins are associated with an increased risk of (e.g., zinc-alpha-2-glycoprotein) or protection from severe COVID-19 (e.g., interleukin-3/6 receptor subunit alpha). Our study shows that fat mass and white blood cells underlie the etiology of severe COVID-19. It also identifies risk and protective factors that could serve as drug targets and guide the effective protection of high-risk individuals.

## Introduction

The coronavirus disease 2019 (COVID-19) is a global pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)<sup>1</sup>. As of mid-January 2021, 92 million confirmed cases and two million deaths from COVID-19 have been reported worldwide<sup>2</sup>. Despite substantial public health and medical efforts, COVID-19 continues to cause irreversible damage and death<sup>3–5</sup>. It is essential to identify risk factors and potential drugs for COVID-19 in order to improve primary prevention and to develop treatment strategies.

Many observational studies have reported that age, gender, ethnicity, and pre-existing conditions, such as cardiovascular disease, diabetes, chronic respiratory disease, hypertension, and cancers, are associated with increased COVID-19 susceptibility and severity<sup>5–8</sup>. Moreover, these retrospective observational studies have noted that hospitalized COVID-19 patients, especially those with severe respiratory or systemic conditions, are at increased risks of atrial fibrillation, non-sustained ventricular tachycardia, acute kidney injury, neurologic disorders, and thrombotic complications<sup>9–12</sup>. Vitamin D deficiency, higher body mass index (BMI), and obesity have been associated with an increased risk of COVID-19<sup>13,14</sup>. Some lifestyle factors were also identified as risk-increasing, such as smoking, alcohol consumption, and lack of physical activity<sup>15</sup>. However, it is difficult to infer causal effects from observational studies because they are susceptible to confounding and reverse causation, while data from randomized controlled trials are scarce and inconclusive.

Mendelian randomization (MR) study provides a promising opportunity to validate and prioritize putative risk factors and drug targets. MR studies use randomly allocated genetic variants related to the exposure

as instrumental variables for investigating the effect of the exposure on an outcome<sup>16</sup>. It is expected to be independent of confounding factors and has been demonstrated as an efficient and cost-effective strategy to identify causal effects<sup>17</sup>. Recent MR studies have provided evidence of causality for a range of risk factors on COVID-19 (Supplementary Table 1). For instance, BMI and smoking are associated with an increased COVID-19 risk, while no evidence of causal effects was found for circulating 25-hydroxy-vitamin D (25OHD) levels<sup>18-20</sup>. However, inconsistent results were also reported for some factors, such as Alzheimer's disease, blood lipids, and physical activity<sup>18,19,21-26</sup>. Some of these inconsistencies are likely due to the usage of early genome-wide association studies (GWAS) of COVID-19, which have small sample sizes. Moreover, most studies are limited to a small number of candidate factors, leaving many more to be tested and identified.

In this study, we conducted an unbiased and exhaustive MR analysis to examine the causal effects of an extensive list of exposures on severe COVID-19. All traits with existing GWAS, as compiled by the Integrative Epidemiology Unit (IEU) Open GWAS project, were included<sup>27</sup>. Based on these GWAS, independent genetic variants associated with a trait at the genome-wide significance were selected as instrumental variables for the trait. The associations between genetic instruments and the risk of severe COVID-19 were evaluated based on three GWAS of COVID-19. The COVID-19 Host Genetics Initiative (HGI) study A2 was used in our discovery analysis. HGI A2 compared COVID-19 patients with confirmed severe respiratory symptoms to population controls<sup>28</sup>. The HGI B2 study, comparing hospitalized COVID-19 patients to population controls, was used as one of our two replication datasets. The other replication dataset, labeled as the NEJM study, was drawn from the first published GWAS study of COVID-19 comparing patients with respiratory failure to healthy controls from Italy and Spain<sup>29</sup>. Multiple sensitivity analyses were performed to detect and correct for the presence of pleiotropy. Here, we only report associations that do not have evidence of pleiotropy in genetic instruments and are replicated in at least one of the two replication analyses. As an in-depth investigation into the BMI-related traits, we further conducted a multivariable MR analysis to disentangle the effects of fat mass and fat-free mass. Our findings provide profound insights into the etiology of severe COVID-19 and prioritize candidate causal risk factors for public health intervention and for drug discovery.

## Results

### Study overview

The workflow of our study is summarized in Figure 1. Starting with the 34,519 GWAS compiled by the IEU Open GWAS project, we focused on the 14,422 that are based on European-descent samples, in order to match the major ancestry in the GWAS of COVID-19 and to avoid false positives as results of population discrepancy in genetic effects (Supplementary Tables 2 and 3). Genetic instruments were selected from each GWAS as independent genetic variants at the genome-wide significance, and their effect alleles were harmonized with the outcome GWAS. Three or more genetic instruments are required for statistical tests of pleiotropic effects, and thus exposures with fewer instruments were excluded. For the univariable MR

analysis of each exposure-outcome pair, we first applied the inverse variance-weighted (IVW) method with a multiplicative random-effects model<sup>30</sup>. We then evaluated the possible presence of pleiotropic effects with Cochran's Q test of heterogeneity and the MR-Egger intercept test for directional pleiotropy<sup>31-33</sup>. We excluded all exposures with indications of pleiotropy in their genetic instruments to fulfill the key assumptions underlying MR analysis. We retained 6,442 GWAS for the discovery analysis with the HGI A2 study (Supplementary Table 4), 6,407 GWAS for the replication analysis with the HGI B2 study (Supplementary Table 5), and 6,248 GWAS for the replication analysis with the NEJM study (Supplementary Table 6). The false discovery rate (FDR) approach was utilized to correct for multiple testing of many exposures (Supplementary Tables 7-9). Based on these three sets of analysis, we defined two sets of results: 1) the significant and replicated results, which have a  $q$ -value  $< 0.05$  in the discovery analysis and a nominal  $p$ -value  $< 0.05$  in either one of the replication studies (Supplementary Table 10); and 2) the suggestive and replicated results, which have a nominal  $p$ -value  $< 0.05$  in the discovery analysis and a nominal  $p$ -value  $< 0.05$  in either one of the replication studies (Supplementary Table 11). A total of 49 significant and replicated traits were identified. Among them, 17 were replicated in both replication datasets (Table 1, Supplementary Table 12).

### **BMI-related traits**

In the univariable MR study, eight BMI-related traits are positively associated with severe COVID-19 in our discovery analysis and also in both of our replication analyses (Table 1). Genetically predicted one standard deviation (SD) increase of BMI is associated with a higher risk of severe COVID-19 (OR: 1.89, 95% CI: 1.51–2.37,  $p = 1.78 \times 10^{-6}$ ). Consistent with BMI, genetically instrumented higher hip circumference (OR: 1.46, 95% CI: 1.15–1.85,  $p = 0.0017$ ) and waist circumference (OR: 1.82, 95% CI: 1.36–2.43,  $p = 6.20 \times 10^{-5}$ ) are associated with a higher risk of severe COVID-19. The univariable MR study also provided strong evidence that weight and fat mass in the left arm, right arm, left leg, right leg, trunk, and whole body are positively associated with severe COVID-19 (Supplementary Tables 10 and 11).

To pinpoint the different aspects of BMI-related traits, we investigated the roles of fat mass and fat-free mass indices in severe COVID-19 (Supplementary Table 13). In the multivariable MR analysis controlling for fat-free mass, there is strong evidence for direct causal effects of fat mass measured at different body parts, including the whole body, left and right arms, left and right legs, and the trunk. The evidence is consistent across the three GWAS of COVID-19 (Fig. 2, Supplementary Table 13). On the other hand, there is no evidence for direct causal effects of fat-free mass (Fig. 3, Supplementary Table 13). The multivariable MR analysis results indicate that the causal effects of BMI-related traits on severe COVID-19 are mainly driven by fat mass.

### **White blood cell traits**

In the univariable MR analyses, we identified a group of five white blood cell traits to be negatively associated with the risk of severe COVID-19. Specifically, suggestive associations were determined for neutrophil count (OR: 0.76, 95% CI: 0.61–0.94,  $p = 0.013$ ), sum basophil neutrophil counts (OR: 0.71, 95%

CI: 0.57–0.87,  $p = 0.001$ ), sum neutrophil eosinophil counts (OR: 0.76, 95% CI: 0.61–0.95,  $p = 0.015$ ), myeloid white cell count (OR: 0.77, 95% CI: 0.62–0.96,  $p = 0.0197$ ), and granulocyte count (OR: 0.75, 95% CI: 0.601–0.93,  $p = 0.009$ ) (Fig. 4). For all five traits, causal estimates are broadly concordant in weighted median (WM) and weighted mode methods, and consistent directions of effects were also found by the MR-Egger method (Supplementary Table 11). Take neutrophil count as an example, consistent estimates of a protective effect were found with WM (OR: 0.61, 95% CI: 0.42–0.88,  $p = 0.009$ ) and weighted mode (OR: 0.59, 95% CI: 0.39–0.91,  $p = 0.017$ ). Overall, our findings support the causal roles of white blood cells, especially neutrophils, in reducing the risk of developing severe COVID-19.

## Circulating proteins

Our univariable MR analyses revealed evidence of causal effects for some circulating proteins. There are six proteins whose effects on severe COVID-19 are significant in the discovery MR analysis ( $q$ -value  $< 0.05$ ) and also replicated in both replication analyses ( $p$ -value  $< 0.05$ ) (Table 1, Supplementary Table 12). Three of them are negatively associated with the risk of severe COVID-19, including interleukin-3 receptor subunit alpha (OR: 0.87, 95% CI: 0.79–0.94), interleukin-6 receptor subunit alpha (OR: 0.88, 95% CI: 0.83–0.94), and prostate-associated microseminoprotein alpha (OR: 0.71, 95% CI: 0.58–0.86). The other three are risk-increasing, including zinc-alpha-2-glycoprotein (OR: 1.37, 95% CI: 1.14–1.66), C1GALT1-specific chaperone 1 (OR: 1.20, 95% CI: 1.19–1.21), and corneodesmosin (OR: 1.12, 95% CI: 1.09–1.16). There are another six circulating proteins that have significant and replicated effects on severe COVID-19, although they are only replicated in one replication analysis (Supplementary Table 10): inter-alpha-trypsin inhibitor heavy chain H1 (OR: 1.08, 95% CI: 1.04–1.12), alpha-2-macroglobulin receptor-associated protein (OR: 1.14, 95% CI: 1.07–1.23), resistin (OR: 1.09, 95% CI: 1.07–1.11), reticulon-4 receptor (OR: 0.86, 95% CI: 0.79–0.93), C-C motif chemokine 23 (OR: 0.88, 95% CI: 0.83–0.92), and collectin-10 (OR: 0.83, 95% CI: 0.76–0.901). Additionally, our suggestive and replicated results revealed another 14 proteins to be associated with the severe COVID-19 risk (Supplementary Table 11). Overall, our MR analyses prioritized scores of circulating proteins that are likely causal in the development of severe COVID-19.

## Discussion

This exhaustive MR study examined an extensive list of risk factors and prioritized those that are likely to play causal roles in the development of severe COVID-19. It leveraged GWAS of COVID-19 of the largest sample size, and the findings were replicated with one, and for some associations, two additional COVID-19 GWAS. Using univariable MR, we first confirmed that BMI-related traits are causal risk factors for severe COVID-19. Our multivariable MR results further suggested that the effects of BMI-related traits are driven by fat mass but not fat-free mass. Moreover, our findings indicate that white blood cell traits, particularly neutrophils, are inversely associated with the severe COVID-19 risk. We also highlighted scores of circulating proteins that could potentially serve as drug targets.

Our main finding that higher BMI-related traits increase the risk of severe COVID-19 is consistent with several recent MR studies<sup>18,19,23,24</sup>. Furthermore, our multivariable MR analysis further showed that fat

mass is a causal risk factor for severe COVID-19, while fat-free mass is not. These results indicate that the causal effect of BMI on severe COVID-19 is likely driven by fat mass. The causal effects of BMI and fat mass have plausible biological mechanisms. Fat mass has been known to have deleterious effects on lung function, inflammation, and immunity<sup>34-36</sup>. In adipose tissue, high production of circulating proinflammatory cytokines and adipokines may intensify virally induced inflammation and immune dysregulation, and contribute to acute respiratory distress syndrome, which is the leading cause of mortality from COVID-19<sup>37-40</sup>. Notably, other causal risk factors of severe COVID-19 identified in this study are also related to adiposity, including glucosamine, resistin, IL6Ra, prostate-associated microseminoprotein, and zinc-alpha-2-glycoprotein<sup>41-45</sup>. These connections suggest a shared mechanism for their contribution to severe COVID-19. Therefore, further mechanistic understanding of fat mass and other related risk factors will shed light on the etiology of severe COVID-19 and provide multiple targets of intervention for prevention and treatment.

Our study indicates that white blood cell traits, especially neutrophils, reduce the risk of severe COVID-19. In addition to direct evidence from neutrophils, sum basophil neutrophil counts and sum neutrophil eosinophil counts are directly related to neutrophils, and concordant causal effects were obtained using multiple MR methods. We also identified myeloid white blood cell counts and granulocyte counts as being inversely associated with the risk of severe COVID-19, which is consistent with our previous MR findings<sup>46</sup>. In contrast to the negative associations in this MR study, previous observational studies have provided strong evidence that elevated white blood cells and neutrophils but depleted lymphocytes are common in COVID-19 patients<sup>47-50</sup>. This discrepancy highlights the possibility that observed associations are due to confounding and reverse causation. The causal role of neutrophils in preventing the development of severe COVID-19 has biological support. Neutrophils, the integral components of the innate immune system, are the first line of defense against invading pathogens<sup>51</sup>. Moreover, neutrophils participate in elaborate cell signaling networks involving cytokines, chemokines, survival, and growth factors that cause downstream pro-inflammatory effects<sup>52</sup>. On the other hand, neutrophils are involved in the hyperinflammatory responses (e.g., overproduction of neutrophil extracellular traps and cytokine storm) in severe COVID-19 patients. This reflects the reverse causal effect of COVID-19 on neutrophils<sup>53,54</sup>. Overall, our present results support the causal effects of white blood cells, especially neutrophils, on severe COVID-19, likely through an enhanced immune response that suppresses virus infection in the early stage.

To identify potential drug targets, we found that six immune-related proteins are inversely associated with the risk of severe COVID-19. Interleukin-6 receptor subunit alpha is associated with a decreased risk of severe COVID-19, which is consistent with a recent MR finding<sup>55</sup>. Signaling of the interleukin-6 receptor subunit alpha influences many inflammatory molecules and tissue regeneration<sup>56</sup>. Interleukin-3 receptor subunit alpha plays important functions in hemopoietic, vascular, and immune systems<sup>57</sup>. Prostate-associated microseminoprotein may influence inflammation and cancer development<sup>58</sup>. C-C motif chemokine 23 is a chemotactic agent, which probably plays an important role in inflammation and

atherosclerosis<sup>59</sup>. Collectin-10 can act as a cellular chemoattractant in vitro, probably involved in the regulation of cell migration<sup>60</sup>. Reticulon-4 receptor influences the central nervous system and protects motoneurons against apoptosis<sup>61</sup>. The identification of these immune-related circulating proteins highlights the critical role of immune responses in the development of severe COVID-19.

Our study identified another six circulating proteins to be positively associated with an increased risk of severe COVID-19. Most of them are glycoproteins. The effect of zinc-alpha-2-glycoprotein might be mediated by the depletion of fatty acids from adipose tissues<sup>62</sup>. C1GALT1-specific chaperone 1 might abolish a glycosyltransferase function and disrupt the O-glycan Core 1 synthesis<sup>63,64</sup>. Corneodesmosin is a glycoprotein expressed in the epidermis and the inner root sheath of hair follicles<sup>65</sup>. Inter-alpha-trypsin inhibitor heavy chain H1 is involved in cell adhesion and leukocyte migration in inflammation sites<sup>66</sup>. The alpha-2-macroglobulin receptor-associated protein is responsible for the role of exotoxin A in pseudomonas disease and immunity<sup>67</sup>. Resistin is known as a hormone that potentially links obesity to diabetes through resisting insulin action<sup>45</sup>. More in-depth mechanistic work is needed to better understand the physiological and biological processes through which these druggable proteins contribute to COVID-19 severity. While our study identified scores of circulating proteins, we cannot rule out the possibility that there are more COVID-19-related proteins. The number of genetic instruments is often limited for circulating proteins, precluding many of them from being analyzed. Our findings of circulating proteins not only suggest possible etiological processes but also provide potential druggable targets.

Our study has many strengths. One strength as an MR study is the ability to assess causal effects, avoiding bias from reverse causation and residual confounding. A major feature and strength of our study is an unbiased and exhaustive approach to screen an extensive list of risk factors. To address the issue of multiple testing, we used FDR corrections in the discovery analysis. To ensure robustness and reduce false positives, we only reported results that were replicated in at least one replication analysis. Another strength is that we applied multivariable MR analyses to evaluate the independent causal effects of fat mass and fat-free mass.

Our study also has several weaknesses. Although we applied multiple sensitivity analyses, including the heterogeneity test, MR-Egger, and WM method, we could not fully rule out the possibility that some genetic variants might be pleiotropic. Another limitation is that some GWAS of exposures and the HGI GWAS of COVID-19 have overlapped samples, especially those from the UK Biobank. To mitigate this issue, we also utilized another GWAS of COVID-19, the NEJM study, which does not have overlapping samples. A further weakness is that the statistical power for some analyses was limited, and some null results might be false negatives. Further positive findings may be revealed if more GWAS with larger sample sizes become available. As an effort to reduce complications from population stratification, our study focuses on European ancestry, and thus the findings may not be generalizable to other ethnicities. Lastly, further research is required to decipher the biological pathways underpinning the observed associations with severe COVID-19.

In conclusion, the present study provides evidence that the causal association between BMI-related traits and severe COVID-19 is driven by fat mass, but not by fat-free mass. Our findings suggest that neutrophils, granulocytes, and myeloid white blood cells are inversely associated with the severe COVID-19 risk. Our study also identifies putatively causal associations between 12 circulating proteins and severe COVID-19. These findings provide valuable insights into the etiology of severe COVID-19. These prioritized risk and protective factors could serve as drug targets and guide the effective protection of high-risk populations.

## Methods

### Exposure data sources

To obtain a comprehensive list of traits with existing GWAS, the summary statistics of 34,519 published GWAS were extracted from the latest MRC Integrative Epidemiology Unit (University of Bristol) GWAS database (<https://gwas.mrcieu.ac.uk/>). Details of each GWAS study can be found at <https://gwas.mrcieu.ac.uk/datasets/>. The R package TwoSampleMR (version 0.5.5) was applied to retrieve the IEU GWAS datasets<sup>68</sup>. The univariable MR study was conducted using the same package.

These GWAS were further filtered based on the following criteria: 1) European ancestry; 2) not eQTL studies, those labeled as “eql” from eQTLGen 2019<sup>69</sup>. A total of 14,422 GWAS summary datasets were retained and used in this study. Detailed information on the used data sources is available in Supplementary Table 2, while details for each exposure and its GWAS are available in Supplementary Table 3.

### Outcome data sources

For evaluation of the association with COVID-19 severity, the instrument-outcome effects were retrieved from the recent version of GWAS meta-analysis by the COVID-19 Host Genetics Initiative (HGI, release 4 alpha, accessed on October 9, 2020)<sup>28</sup>. Detailed information has been provided on the COVID-19 HGI website (<https://www.covid19hg.org/results/>). In our primary discovery analysis, we used the summary statistics based on the comparison of 2,972 patients confirmed as “very severe respiratory” COVID-19 with the 284,472 general population samples. This is called “the HGI A2 study” in this study.

To reduce false positives and to ensure the robustness of our discoveries, replication analyses were performed with two additional GWAS of COVID-19. One of them was also from the COVID-19 HGI, comparing 6,492 hospitalized COVID-19 patients with 1,012,809 control participants. We called it “the HGI B2 study”. Only single nucleotide polymorphisms (SNPs) with imputation quality scores > 0.6 were retained. The other GWAS was on 1,610 COVID-19 patients with respiratory failure and 2,180 controls from Italy and Spain, and it was called “the NEJM study”<sup>29</sup>.

### Selection of instrumental variables

For the implementation of MR, SNPs were selected based on the genome-wide significance threshold ( $p < 5 \times 10^{-8}$ ). To ensure SNPs are independent, we pruned the variants by linkage disequilibrium (LD) ( $R^2$  threshold of 0.001 or clumping window within 10,000 kb). When target SNPs were not present in the outcome dataset, proxy SNPs were used instead through LD tagging (minimum LD  $R^2$  threshold of 0.8). The effect alleles of selected genetic variants were harmonized across the exposure and outcome associations.

### Univariable Mendelian randomization

Two-sample Mendelian randomization analysis was undertaken using GWAS summary statistics for each exposure-outcome pair. In order to estimate the causal effect of each trait on severe COVID-19, the IVW method with a multiplicative random-effects model was used as the primary analysis<sup>30,33,70</sup>. Horizontal pleiotropy occurs when SNPs exert a direct effect on severe COVID-19 through other independent biological pathways. To assess the presence of heterogeneity among genetic instruments, Cochran's Q statistic was calculated for heterogeneity for the IVW analyses<sup>31</sup>. An extended version of Cochran's Q statistic (Rücker's Q') can be estimated for the MR-Egger<sup>32</sup>. We used the MR-Egger intercept test to evaluate the extent to which unbalanced horizontal pleiotropy may affect the effect estimate<sup>33</sup>. To account for pleiotropy, additional sensitivity analyses were performed with the MR-Egger<sup>33,70</sup>, weighted median (WM)<sup>71</sup>, and weighted mode methods<sup>72</sup>. The MR-Egger method allows unbalanced horizontal pleiotropic effects even when all SNPs are invalid instruments<sup>33</sup>. The WM method can provide robust causal estimates when at least 50% of SNPs are valid genetic instruments, while the weighted mode method requires that the largest number of instruments that identify the consistent causal effect to be valid instruments<sup>71,72</sup>. The false discovery rate (FDR) approach was utilized to correct for multiple testing, and it was applied to the  $p$  values from the IVW random-effect model<sup>73</sup>. FDR controls the expected proportion of false positives among the hypotheses declared significant if the  $q$ -value is  $< 0.05$ , while the association was deemed to be suggestive if the unadjusted  $p$ -value is  $< 0.05$ .

Two additional exclusion criteria were applied to filter out exposures: 1) the number of genetic instruments was less than three. Three or more are required for statistical tests of pleiotropic effects and for statistical sensitivity analyses to correct for pleiotropy. 2) Exposures with indications of pleiotropy in their genetic instruments. The presence of pleiotropy violates the assumption of MR analysis. For the remaining exposures, FDR correction for multiple testing was applied separately for each analysis with the HGI A2, HGI B2, or NEJM study. To identify potential causal risk factors for severe COVID-19, we used two approaches to consider the evidence strength. First, the significant and replicated results were defined as those with a  $q$ -value  $< 0.05$  in the discovery analysis and a nominal  $p$ -value  $< 0.05$  in either one of the replication studies (Supplementary Table S10). Second, the suggestive and replicated results were defined as those with a nominal  $p$ -value  $< 0.05$  in the discovery analysis and a nominal  $p$ -value  $< 0.05$  in either one of the replication studies (Supplementary Table S11). All MR analyses were conducted in R with the TwoSampleMR package<sup>68</sup>. An analysis flowchart is shown in Fig. 1.

## Multivariable Mendelian randomization

As many BMI-related traits are typically correlated with each other, we conducted a two-sample multivariable MR (MVMR) analysis to explore independent causal risk factors for severe COVID-19<sup>74</sup>. SNPs associated with fat mass and fat-free mass were obtained from previous GWAS by MRC IEU and the Neale Lab through the TwoSampleMR package. The effects of genetically predicted fat mass and fat-free mass for each pair of the whole body, left arm, right arm, left leg, right leg, and trunk were estimated using the MVMR package (version 0.2.0) in R.

## Declarations

### Data availability

All analyses were conducted using publicly available data. The exposure data (GWAS summary statistics) used in the analyses described here are freely accessible in the MR-Base platform (<https://www.mrbase.org/>). We downloaded COVID-19 data (GWAS summary statistics) in the COVID-19 Host Genetics Initiative (<https://www.covid19hg.org/>) and COVID GWAS results browser ([https://ikmb.shinyapps.io/COVID-19\\_GWAS\\_Browser/](https://ikmb.shinyapps.io/COVID-19_GWAS_Browser/)).

### Code availability

The codes used in the Mendelian randomization analyses described here are freely accessible in TwoSampleMR R package via GitHub (<https://github.com/MRCIEU/TwoSampleMR/>). Full documentation for the R package is also provided (<https://mrcieu.github.io/TwoSampleMR/>). We implemented the MVMR analysis using the MVMR R package (<https://github.com/WSpiller/MVMR/>).

### Acknowledgments

We thank the investigators of the COVID-19 genome-wide association study, the COVID-19 Host Genetics Initiative, and FinnGen consortium for sharing summary-level data. We would like to express our gratitude to all other Ye lab members for stimulating discussions.

### Contributions

Y.S. and K.Y. conceived the study. Y.S. performed data analysis and prepared visualizations. Y.S., J.Z., and K.Y. interpreted the results. Y.S. and K.Y. wrote the first draft of the manuscript. All authors reviewed and approved the final version.

### Competing interests

The authors declare no competing interests.

## References

1. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).
2. World Health Organization. *Coronavirus disease 2019 (COVID-19) report.*, <<https://www.who.int/emergencies/diseases/novel-coronavirus-2019>> (2021).
3. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506, doi:10.1016/S0140-6736(20)30183-5 (2020).
4. Wenham, C., Smith, J., Morgan, R., Gender & Group, C.-W. COVID-19: the gendered impacts of the outbreak. *Lancet* **395**, 846-848, doi:10.1016/S0140-6736(20)30526-2 (2020).
5. Guzik, T. J. *et al.* COVID-19 and the cardiovascular system: implications for risk assessment, diagnosis, and treatment options. *Cardiovasc Res* **116**, 1666-1687, doi:10.1093/cvr/cvaa106 (2020).
6. Teng, Y., Xu, J., Zhang, Y., Liu, Z. & Zhang, S. Mendelian randomization in COVID-19: Applications for cardiovascular comorbidities and beyond. *EBioMedicine* **57**, 102847, doi:10.1016/j.ebiom.2020.102847 (2020).
7. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*, doi:10.1001/jama.2020.2648 (2020).
8. Rao, S., Lau, A. & So, H. C. Exploring Diseases/Traits and Blood Proteins Causally Related to Expression of ACE2, the Putative Receptor of SARS-CoV-2: A Mendelian Randomization Analysis Highlights Tentative Relevance of Diabetes-Related Traits. *Diabetes Care* **43**, 1416-1426, doi:10.2337/dc20-0643 (2020).
9. Bhatla, A. *et al.* COVID-19 and cardiac arrhythmias. *Heart Rhythm* **17**, 1439-1444, doi:10.1016/j.hrthm.2020.06.016 (2020).
10. Kolhe, N. V., Fluck, R. J., Selby, N. M. & Taal, M. W. Acute kidney injury associated with COVID-19: A retrospective cohort study. *PLoS Med* **17**, e1003406, doi:10.1371/journal.pmed.1003406 (2020).
11. Zhou, J., Liu, C., Sun, Y., Huang, W. & Ye, K. Cognitive disorders associated with hospitalization of COVID-19: Results from an observational cohort study. *Brain Behav Immun* **91**, 383-392, doi:10.1016/j.bbi.2020.10.019 (2021).
12. Klok, F. A. *et al.* Incidence of thrombotic complications in critically ill ICU patients with COVID-19. *Thromb Res* **191**, 145-147, doi:10.1016/j.thromres.2020.04.013 (2020).
13. Meltzer, D. O. *et al.* Association of Vitamin D Status and Other Clinical Characteristics With COVID-19 Test Results. *JAMA Netw Open* **3**, e2019722, doi:10.1001/jamanetworkopen.2020.19722 (2020).
14. Simonnet, A. *et al.* High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Requiring Invasive Mechanical Ventilation. *Obesity (Silver Spring)* **28**, 1195-1199, doi:10.1002/oby.22831 (2020).
15. Hamer, M., Kivimaki, M., Gale, C. R. & Batty, G. D. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: A community-based cohort study of 387,109 adults in UK. *Brain Behav Immun* **87**, 184-187, doi:10.1016/j.bbi.2020.05.059 (2020).

16. Burgess, S., Butterworth, A., Malarstig, A. & Thompson, S. G. Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ* **345**, e7325, doi:10.1136/bmj.e7325 (2012).
17. Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* **32**, 1-22, doi:10.1093/ije/dyg070 (2003).
18. Ponsford, M. J. *et al.* Cardiometabolic Traits, Sepsis, and Severe COVID-19: A Mendelian Randomization Investigation. *Circulation* **142**, 1791-1793, doi:10.1161/CIRCULATIONAHA.120.050753 (2020).
19. Li, S. Modifiable lifestyle factors and severe COVID-19 risk: Evidence from Mendelian randomization analysis. 2020.2010.2019.20215525, doi:10.1101/2020.10.19.20215525 %J medRxiv (2020).
20. Butler-Laporte, G. *et al.* Vitamin D and Covid-19 Susceptibility and Severity: a Mendelian Randomization Study. 2020.2009.2008.20190975, doi:10.1101/2020.09.08.20190975 %J medRxiv (2020).
21. Liu, D. *et al.* Association between Alzheimer's disease and COVID-19: A bidirectional Mendelian randomization. 2020.2007.2027.20163212, doi:10.1101/2020.07.27.20163212 %J medRxiv (2020).
22. Luykx, J. J. & Lin, B. D. Are psychiatric disorders risk factors for COVID-19 susceptibility and severity? a two-sample, bidirectional, univariable and multivariable Mendelian Randomization study. 2020.2011.2029.20240481, doi:10.1101/2020.11.29.20240481 %J medRxiv (2020).
23. Leong, A. *et al.* Cardiometabolic Risk Factors for COVID-19 Susceptibility and Severity: A Mendelian Randomization Analysis. 2020.2008.2026.20182709, doi:10.1101/2020.08.26.20182709 %J medRxiv (2020).
24. Zhang, K. *et al.* Causally Associations of Blood Lipids Levels with COVID-19 Risk: Mendelian Randomization Study. 2020.2007.2007.20147926, doi:10.1101/2020.07.07.20147926 %J medRxiv (2020).
25. Aung, N., Khanji, M. Y., Munroe, P. B. & Petersen, S. E. Causal Inference for Genetic Obesity, Cardiometabolic Profile and COVID-19 Susceptibility: A Mendelian Randomization Study. *Front Genet* **11**, 586308, doi:10.3389/fgene.2020.586308 (2020).
26. Zhang, X. *et al.* Physical activity and COVID-19: an observational and Mendelian randomisation study. *J Glob Health* **10**, 020514, doi:10.7189/jogh-10-020514 (2020).
27. Elsworth BL, M. R., Raistrick CA, Paternoster L, Hemani G, Gaunt TR. MRC IEU UK Biobank GWAS pipeline version 2. doi:10.5523/bris.pnoat8cxo0u52p6ynfaekeigi (2019).
28. Initiative, C.-H. G. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet* **28**, 715-718, doi:10.1038/s41431-020-0636-6 (2020).
29. Severe Covid, G. G. *et al.* Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med* **383**, 1522-1534, doi:10.1056/NEJMoa2020283 (2020).
30. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658-665, doi:10.1002/gepi.21758

- (2013).
31. Greco, M. F., Minelli, C., Sheehan, N. A. & Thompson, J. R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* **34**, 2926-2940, doi:10.1002/sim.6522 (2015).
  32. Bowden, J. *et al.* Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int J Epidemiol* **47**, 1264-1278, doi:10.1093/ije/dyy101 (2018).
  33. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* **44**, 512-525, doi:10.1093/ije/dyv080 (2015).
  34. Sattar, N., McInnes, I. B. & McMurray, J. J. V. Obesity Is a Risk Factor for Severe COVID-19 Infection: Multiple Potential Mechanisms. *Circulation* **142**, 4-6, doi:10.1161/CIRCULATIONAHA.120.047659 (2020).
  35. Nimptsch, K., Konigorski, S. & Pischon, T. Diagnosis of obesity and use of obesity biomarkers in science and clinical medicine. *Metabolism* **92**, 61-70, doi:10.1016/j.metabol.2018.12.006 (2019).
  36. Tilg, H. & Moschen, A. R. Adipocytokines: mediators linking adipose tissue, inflammation and immunity. *Nat Rev Immunol* **6**, 772-783, doi:10.1038/nri1937 (2006).
  37. Reilly, S. M. & Saltiel, A. R. Adapting to obesity with adipose tissue inflammation. *Nat Rev Endocrinol* **13**, 633-643, doi:10.1038/nrendo.2017.90 (2017).
  38. Lackey, D. E. & Olefsky, J. M. Regulation of metabolism by the innate immune system. *Nat Rev Endocrinol* **12**, 15-28, doi:10.1038/nrendo.2015.189 (2016).
  39. Korakas, E. *et al.* Obesity and COVID-19: immune and metabolic derangement as a possible link to adverse clinical outcomes. *Am J Physiol Endocrinol Metab* **319**, E105-E109, doi:10.1152/ajpendo.00198.2020 (2020).
  40. Ryan, P. M. & Caplice, N. M. Is Adipose Tissue a Reservoir for Viral Spread, Immune Activation, and Cytokine Amplification in Coronavirus Disease 2019? *Obesity (Silver Spring)* **28**, 1191-1194, doi:10.1002/oby.22843 (2020).
  41. Hwang, J. S., Park, J. W., Nam, M. S., Cho, H. & Han, I. O. Glucosamine enhances body weight gain and reduces insulin response in mice fed chow diet but mitigates obesity, insulin resistance and impaired glucose tolerance in mice high-fat diet. *Metabolism* **64**, 368-379, doi:10.1016/j.metabol.2014.11.005 (2015).
  42. Theurich, S. *et al.* IL-6/Stat3-Dependent Induction of a Distinct, Obesity-Associated NK Cell Subpopulation Deteriorates Energy and Glucose Homeostasis. *Cell Metab* **26**, 171-184 e176, doi:10.1016/j.cmet.2017.05.018 (2017).
  43. Smith Byrne, K. *et al.* The role of plasma microseminoprotein-beta in prostate cancer: an observational nested case-control and Mendelian randomization study in the European prospective investigation into cancer and nutrition. *Ann Oncol* **30**, 983-989, doi:10.1093/annonc/mdz121 (2019).

44. Xiao, X. H. *et al.* Zinc alpha2 glycoprotein protects against obesity-induced hepatic steatosis. *Int J Obes (Lond)* **42**, 1418-1430, doi:10.1038/s41366-018-0151-9 (2018).
45. Steppan, C. M. *et al.* The hormone resistin links obesity to diabetes. *Nature* **409**, 307-312, doi:10.1038/35053000 (2001).
46. Sun, Y., Zhou, J. & Ye, K. White blood cells and severe COVID-19: a Mendelian randomization study. 2020.2010.2014.20212993, doi:10.1101/2020.10.14.20212993 %J medRxiv (2020).
47. Behzad, S., Aghaghazvini, L., Radmard, A. R. & Gholamrezanezhad, A. Extrapulmonary manifestations of COVID-19: Radiologic and clinical overview. *Clin Imaging* **66**, 35-41, doi:10.1016/j.clinimag.2020.05.013 (2020).
48. Kuri-Cervantes, L. *et al.* Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci Immunol* **5**, doi:10.1126/sciimmunol.abd7114 (2020).
49. Vabret, N. *et al.* Immunology of COVID-19: Current State of the Science. *Immunity* **52**, 910-941, doi:10.1016/j.immuni.2020.05.002 (2020).
50. Zhou, J., Sun, Y., Huang, W. & Ye, K. Altered blood cell traits underlie a major genetic locus of severe COVID-19. 2020.2009.2009.20191700, doi:10.1101/2020.09.09.20191700 %J medRxiv (2020).
51. Kaplan, M. J. & Radic, M. Neutrophil extracellular traps: double-edged swords of innate immunity. *J Immunol* **189**, 2689-2695, doi:10.4049/jimmunol.1201719 (2012).
52. Cheng, O. Z. & Palaniyar, N. NET balancing: a problem in inflammatory lung diseases. *Front Immunol* **4**, 1, doi:10.3389/fimmu.2013.00001 (2013).
53. Borges, L., Pithon-Curi, T. C., Curi, R. & Hatanaka, E. COVID-19 and Neutrophils: The Relationship between Hyperinflammation and Neutrophil Extracellular Traps. *Mediators Inflamm* **2020**, 8829674, doi:10.1155/2020/8829674 (2020).
54. Mehta, P. *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033-1034, doi:10.1016/S0140-6736(20)30628-0 (2020).
55. Larsson, S. C., Burgess, S. & Gill, D. Genetically proxied interleukin-6 receptor inhibition: opposing associations with COVID-19 and pneumonia. 2003545, doi:10.1183/13993003.03545-2020 %J European Respiratory Journal (2020).
56. Kang, S., Tanaka, T., Narazaki, M. & Kishimoto, T. Targeting Interleukin-6 Signaling in Clinic. *Immunity* **50**, 1007-1023, doi:10.1016/j.immuni.2019.03.026 (2019).
57. Broughton, S. E. *et al.* A dual role for the N-terminal domain of the IL-3 receptor in cell signalling. *Nat Commun* **9**, 386, doi:10.1038/s41467-017-02633-7 (2018).
58. Pei, X. *et al.* PC3-secreted microprotein is a novel chemoattractant protein and functions as a high-affinity ligand for CC chemokine receptor 2. *J Immunol* **192**, 1878-1886, doi:10.4049/jimmunol.1300758 (2014).
59. Kim, J., Kim, Y. S. & Ko, J. CK beta 8/CCL23 induces cell migration via the Gi/Go protein/PLC/PKC delta/NF-kappa B and is involved in inflammatory responses. *Life Sci* **86**, 300-308, doi:10.1016/j.lfs.2009.11.012 (2010).

60. Munye, M. M. *et al.* COLEC10 is mutated in 3MC patients and regulates early craniofacial development. *PLoS Genet* **13**, e1006679, doi:10.1371/journal.pgen.1006679 (2017).
61. Wills, Z. P. *et al.* The nogo receptor family restricts synapse number in the developing hippocampus. *Neuron* **73**, 466-481, doi:10.1016/j.neuron.2011.11.029 (2012).
62. Bao, Y. *et al.* Zinc-alpha2-glycoprotein, a lipid mobilizing factor, is expressed and secreted by human (SGBS) adipocytes. *FEBS Lett* **579**, 41-47, doi:10.1016/j.febslet.2004.11.042 (2005).
63. Kudo, T. *et al.* Molecular cloning and characterization of a novel UDP-Gal:GalNAc(alpha) peptide beta 1,3-galactosyltransferase (C1Gal-T2), an enzyme synthesizing a core 1 structure of O-glycan. *J Biol Chem* **277**, 47724-47731, doi:10.1074/jbc.M205839200 (2002).
64. Zhu, J., Wu, C. & Wu, L. Associations Between Genetically Predicted Protein Levels and COVID-19 Severity. *J Infect Dis* **223**, 19-22, doi:10.1093/infdis/jiaa660 (2021).
65. Levy-Nissenbaum, E. *et al.* Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin. *Nat Genet* **34**, 151-153, doi:10.1038/ng1163 (2003).
66. Bourgeois, R. *et al.* A Comparative Analysis of the Lipoprotein(a) and Low-Density Lipoprotein Proteomic Profiles Combining Mass Spectrometry and Mendelian Randomization. *CJC Open*, doi:<https://doi.org/10.1016/j.cjco.2020.11.019> (2020).
67. Kounnas, M. Z. *et al.* The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds and internalizes Pseudomonas exotoxin A. *J Biol Chem* **267**, 12420-12423 (1992).
68. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, doi:10.7554/eLife.34408 (2018).
69. Vösa, U. *et al.* Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. 447367, doi:10.1101/447367 %J bioRxiv (2018).
70. Burgess, S. *et al.* Guidelines for performing Mendelian randomization investigations [version 2; peer review: 1 approved, 1 approved with reservations]. *Wellcome Open Research* **4**, doi:10.12688/wellcomeopenres.15555.2 (2020).
71. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**, 304-314, doi:10.1002/gepi.21965 (2016).
72. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* **46**, 1985-1998, doi:10.1093/ije/dyx102 (2017).
73. Korthauer, K. *et al.* A practical guide to methods controlling false discoveries in computational biology. *Genome Biol* **20**, 118, doi:10.1186/s13059-019-1716-1 (2019).
74. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* **48**, 713-727, doi:10.1093/ije/dyy262 (2019).

# Tables

**Table 1. Significant and replicated causal associations with severe COVID-19**

<b>GWAS ID</b>	<b>Exposure</b>	<b>Outcome dataset</b>	<b>SNP#</b>	<b>Odds ratio</b>	<b>95% CI</b>	<b>p-value</b>	<b>q-value</b>
ukb-b-19953	Body mass index (BMI)	HGI A2	454	1.89	1.51, 2.37	3.15E-08	1.78E-06
		HGI B2	453	1.69	1.45, 1.97	1.31E-11	9.34E-10
		NEJM	449	1.53	1.05, 2.25	0.0286	0.2867
ukb-b-15590	Hip circumference	HGI A2	417	1.46	1.15, 1.85	0.0017	0.0319
		HGI B2	417	1.51	1.28, 1.77	5.93E-07	1.88E-05
		NEJM	412	1.66	1.11, 2.50	0.0142	0.1871
ukb-b-9405	Waist circumference	HGI A2	371	1.82	1.36, 2.43	6.20E-05	0.0017
		HGI B2	371	1.92	1.57, 2.34	1.57E-10	9.33E-09
		NEJM	369	1.95	1.21, 3.16	0.0065	0.1108
ukb-b-8338	Arm fat mass (left)	HGI A2	419	1.69	1.33, 2.14	1.33E-05	0.0004
		HGI B2	419	1.61	1.37, 1.90	9.35E-09	3.70E-07
		NEJM	417	1.52	1.02, 2.29	0.0420	0.3645
ukb-b-6704	Arm fat mass (right)	HGI A2	424	1.74	1.36, 2.22	7.64E-06	0.0003
		HGI B2	424	1.68	1.43, 1.96	1.16E-10	7.36E-09
		NEJM	422	1.65	1.12, 2.44	0.0115	0.1613
ukb-a-279	Leg fat mass (left)	HGI A2	287	2.76	2.00, 3.81	8.04E-10	5.55E-08
		HGI B2	287	2.20	1.77, 2.72	5.73E-13	5.66E-11
		NEJM	285	1.82	1.08, 3.07	0.0243	0.2617
ukb-b-	Leg fat mass (left)	HGI A2	423	2.41	1.81,	2.02E-	1.30E-

7212					3.22	09	07
		HGI B2	422	2.14	1.75, 2.61	1.15E- 13	1.36E- 11
		NEJM	418	1.98	1.18, 3.31	0.0094	0.1442
ukb-b- 18096	Leg fat mass (right)	HGI A2	425	2.66	1.98, 3.58	8.66E- 11	7.34E- 09
		HGI B2	424	2.33	1.90, 2.86	5.09E- 16	7.97E- 14
		NEJM	420	2.04	1.23, 3.40	0.0061	0.1059
ukb-b- 18377	Leg fat percentage (left)	HGI A2	381	4.05	2.70, 6.09	1.56E- 11	1.39E- 09
		HGI B2	379	2.76	2.09, 3.63	5.01E- 13	5.25E- 11
		NEJM	375	2.11	1.02, 4.37	0.0453	0.3786
ukb-b- 6591	Age at first sexual intercourse	HGI A2	199	0.39	0.26,0.59	7.25E- 06	0.0003
		HGI B2	198	0.43	0.33, 0.56	7.50E- 10	3.81E- 08
		NEJM	196	0.52	0.27, 0.97	0.0405	0.3566
ukb-b- 4667	Leisure/social activities: Religious group	HGI A2	23	0.00	0.00, 0.12	0.0028	0.0468
		HGI B2	23	0.01	0.00, 0.16	0.0006	0.0108
		NEJM	23	0.02	0.00, 0.84	0.0402	0.3566
ukb-b- 11535	Mineral and other dietary supplements: Glucosamine	HGI A2	5	0.00	0.00, 0.00	1.23E- 11	1.15E- 09
		HGI B2	5	0.00	0.00, 0.04	3.74E- 06	0.0001
		NEJM	5	0.00	0.00, 0.00	3.69E- 06	0.0002
prot-a- 208	Zinc-alpha-2-glycoprotein	HGI A2	3	1.37	1.14, 1.66	0.0009	0.0197
		HGI B2	3	1.24	1.07, 1.45	0.0049	0.0580

		NEJM	3	1.4	1.08, 1.82	0.0121	0.1645
prot-a-294	C1GALT1-specific chaperone 1	HGI A2	3	1.20	1.19, 1.21	< 2E-16	< 2E-16
		HGI B2	3	1.22	1.05, 1.41	0.0074	0.0800
		NEJM	3	1.53	1.26, 1.86	2.07E-05	0.0009
prot-a-500	Corneodesmosin	HGI A2	4	1.12	1.09, 1.16	1.31E-12	1.43E-10
		HGI B2	4	1.06	1.05, 1.08	1.37E-14	1.88E-12
		NEJM	4	1.27	1.14, 1.41	2.40E-05	0.0010
prot-a-1530	Interleukin-3 receptor subunit alpha	HGI A2	3	0.87	0.79, 0.94	0.0011	0.0238
		HGI B2	3	0.84	0.79, 0.90	7.38E-08	2.68E-06
		NEJM	3	0.74	0.58, 0.94	0.0136	0.1804
prot-a-1540	Interleukin-6 receptor subunit alpha	HGI A2	3	0.88	0.83, 0.94	0.0001	0.0036
		HGI B2	3	0.90	0.86, 0.95	2.39E-05	0.0006
		NEJM	3	0.82	0.70, 0.96	0.0160	0.2040
prot-a-1950	Prostate-associated microseminoprotein	HGI A2	3	0.71	0.58, 0.86	0.0005	0.0121
		HGI B2	3	0.81	0.71, 0.93	0.0023	0.0316
		NEJM	3	0.79	0.66, 0.96	0.0176	0.2159

NOTE: Only associations that are replicated in both replication analyses (with the HGI B2 and NEJM datasets) are included here. Odds ratios and 95% confidence intervals were derived using the inverse-variance weighted random-effects model. The full list, including those that are replicated with only one study, is available in Supplementary Table 10. SNP, single nucleotide polymorphism; SNP#, number of SNPs retained for this analysis; CI, confidence interval.

## Supplementary Datasets

Supplementary Table 1. Notable Mendelian randomization studies of COVID-19.

Supplementary Table 2. Details of the datasets used in the present Mendelian randomization study.

Supplementary Table 3. Dataset information of all traits included in our MR analysis.

Supplementary Table 4. All MR results based on the HGI A2 dataset.

Supplementary Table 5. All MR results based on the HGI B2 dataset.

Supplementary Table 6. All MR results based on the NEJM dataset.

Supplementary Table 7. All MR results using FDR, based on the HGI A2 dataset.

Supplementary Table 8. All MR results using FDR, based on the HGI B2 dataset.

Supplementary Table 9. All MR results using FDR, based on the NEJM dataset.

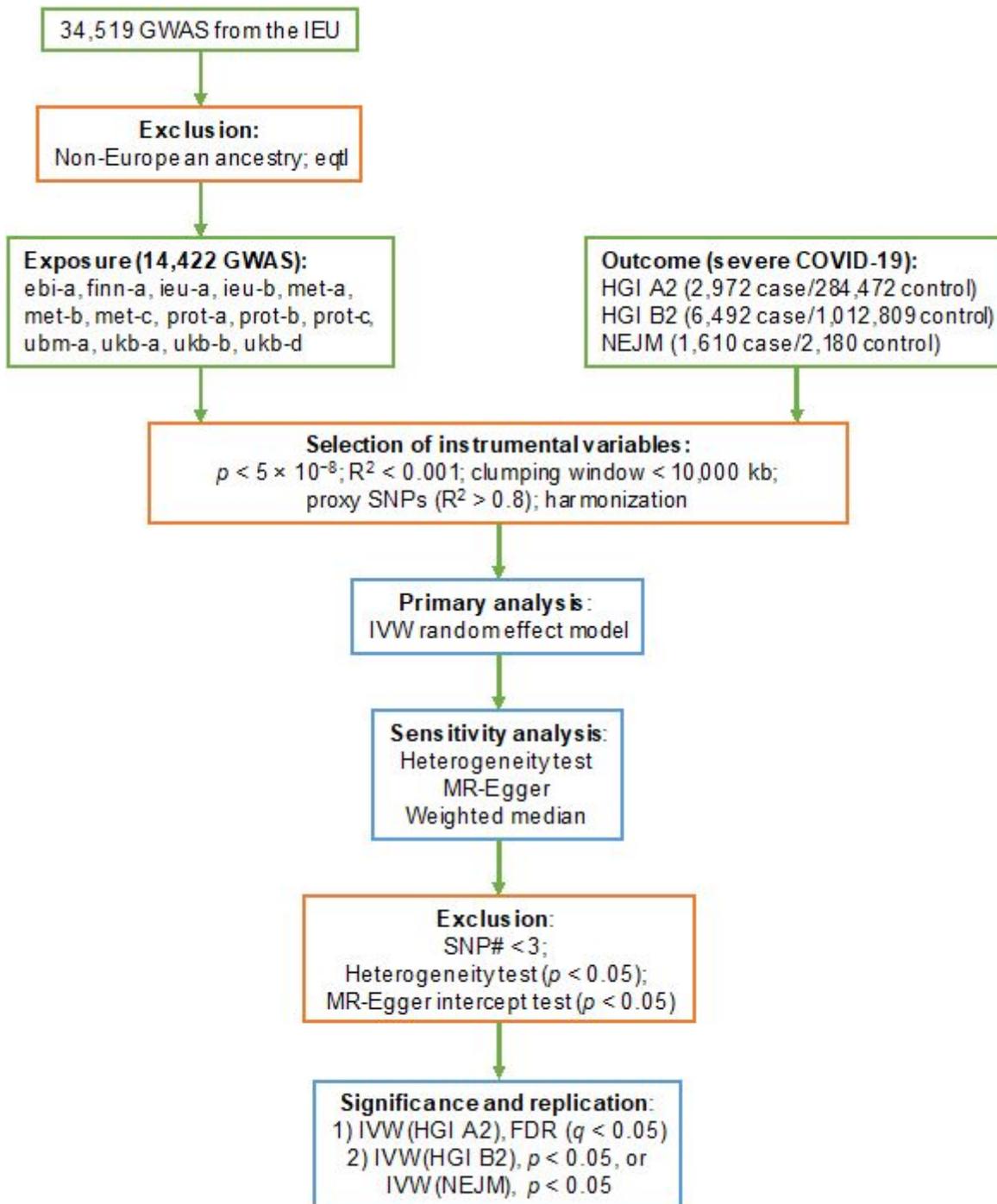
Supplementary Table 10. Significant and replicated results (IVW FDR < 0.05 with HGI A2; and IVW  $p < 0.05$  with HGI B2 OR IVW  $p < 0.05$  with NEJM).

Supplementary Table 11. Suggestive and replicated results (IVW  $p < 0.05$  with HGI A2; and IVW  $p < 0.05$  with HGI B2 OR IVW  $p < 0.05$  with NEJM).

Supplementary Table 12. Significant and replicated results that are replicated with both HGI B2 and NEJM (IVW FDR < 0.05 with HGI A2; and IVW  $p < 0.05$  with HGI B2 AND IVW  $p < 0.05$  with NEJM).

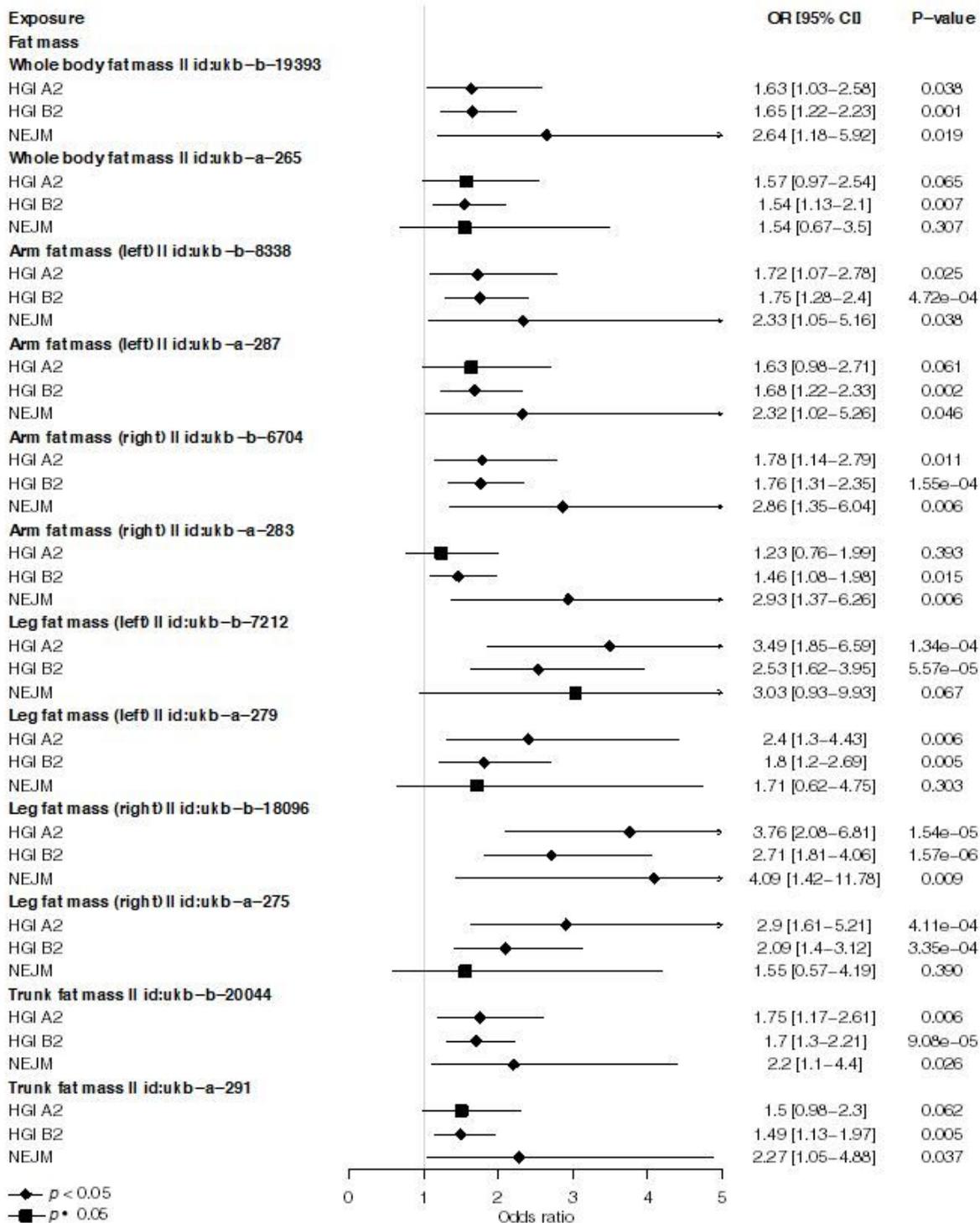
Supplementary Table 13. Multivariable Mendelian randomization of fat mass and fat-free mass indices on severe COVID-19.

## Figures



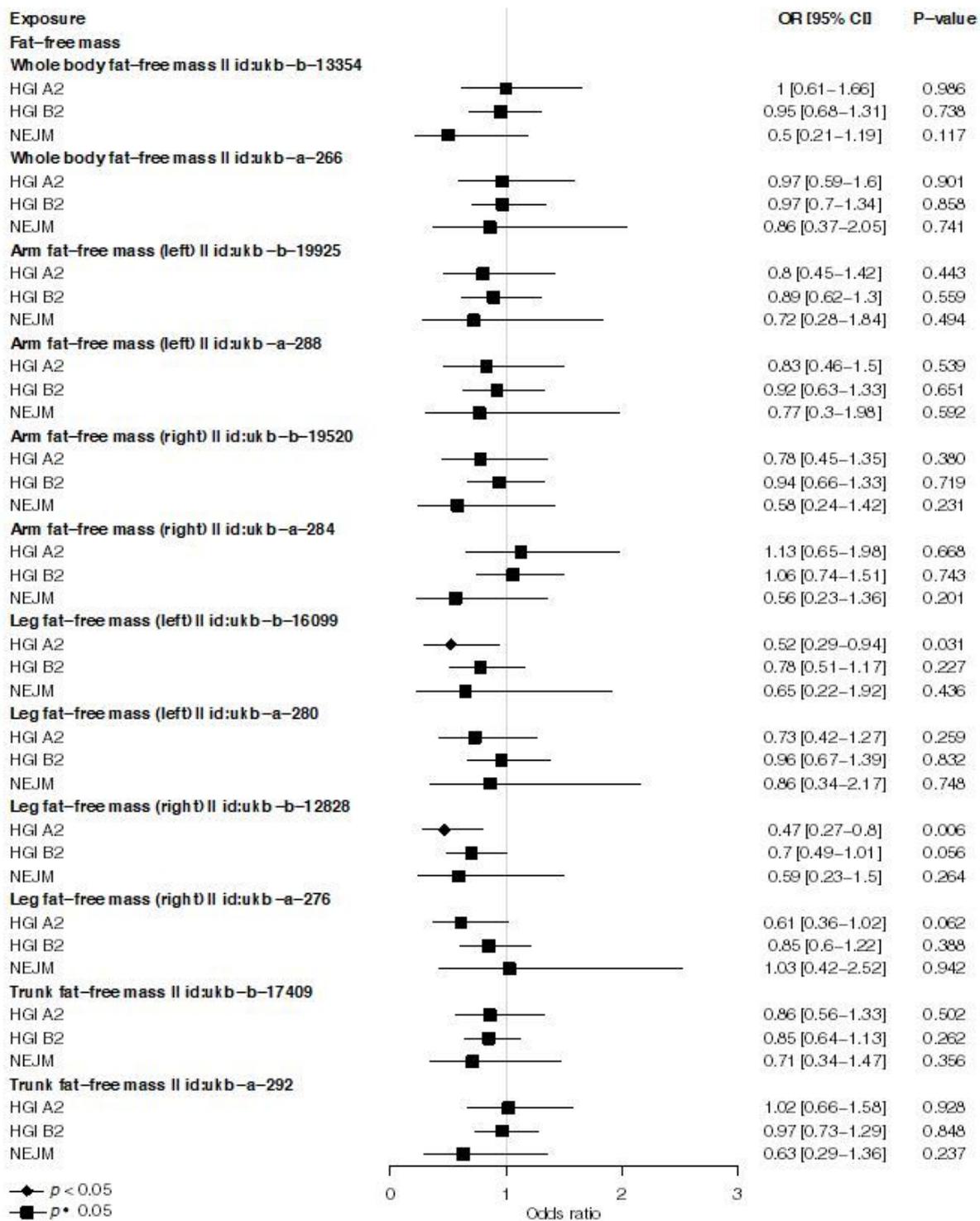
**Figure 1**

The workflow of our exhaustive MR study for severe COVID-19. SNP, single nucleotide polymorphism; SNP#, number of SNPs retained for this analysis.



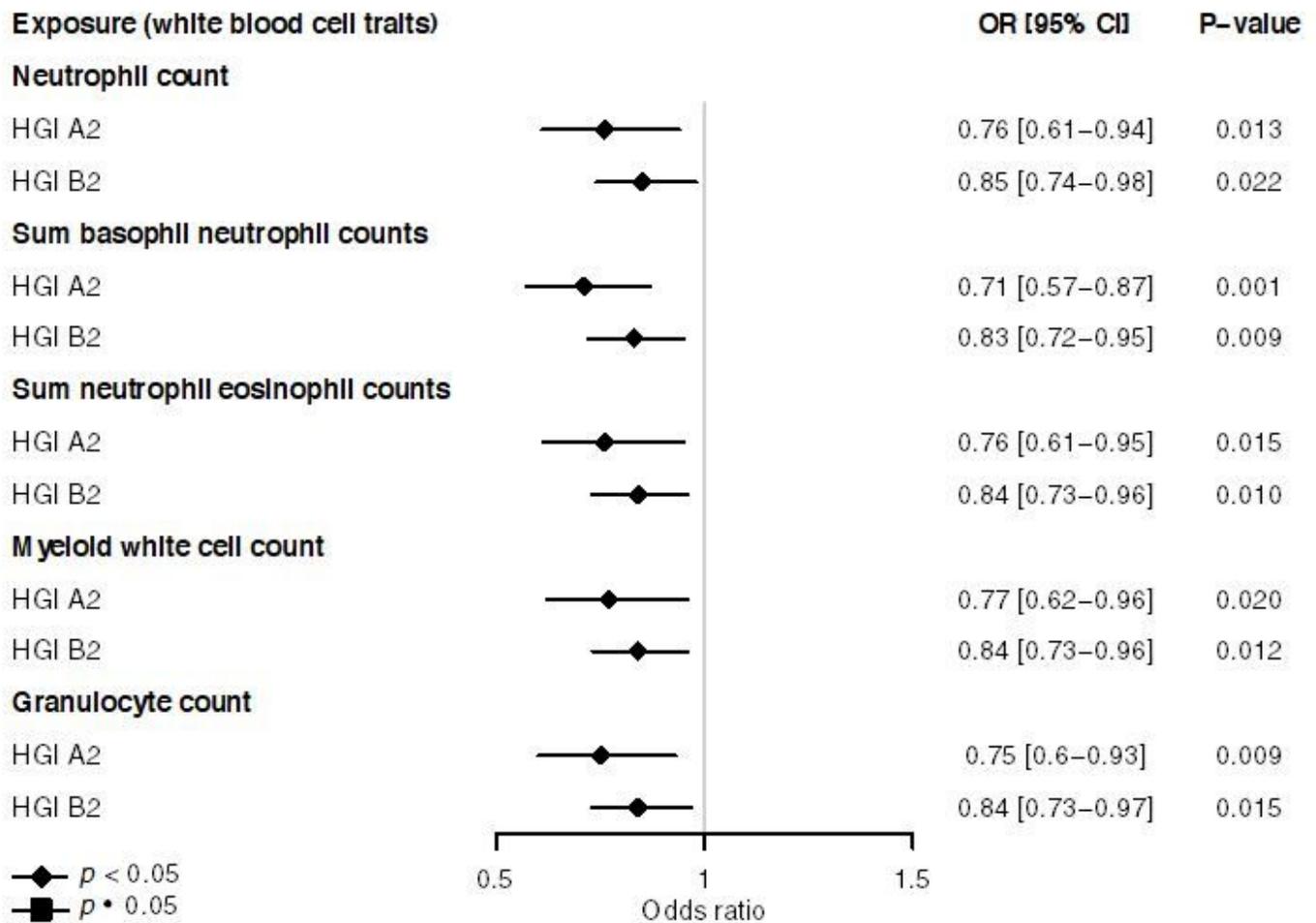
**Figure 2**

Multivariable analysis to determine independent causal roles of fat mass in severe COVID-19. Odds ratios are scaled to a genetically predicted 1 SD increase in fat mass. Associations with p-value < 0.05 were indicated with diamonds, while others with squares. Detailed summary statistics could be found in Supplementary Table 13.



**Figure 3**

Multivariable analysis to determine independent causal roles of fat-free mass in severe COVID-19. Odds ratios are scaled to a genetically predicted 1 SD increase in fat-free mass. Associations with p-value < 0.05 were indicated with diamonds, while others with squares. Detailed summary statistics could be found in Supplementary Table 13.



**Figure 4**

Mendelian randomization estimates of white blood cell traits on severe COVID-19 risk. Odds ratios are scaled to a genetically predicted 1 SD increase in white blood cell trait. Associations with p-value < 0.05 were indicated with diamonds, while others with squares. Detailed summary statistics could be found in Supplementary Table 11.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTableS1.docx](#)
- [SupplementaryTables213.xlsx](#)