**Supplementary Methods.**

**MR Sensitivity analyses**

Additionally, the presence of horizontal pleiotropy and heterogeneity within the genetic instruments, which are potential confounders in MR analyses, was tested using MR-Egger regression intercept (1), and IVW and MR-Egger heterogeneity tests (2), respectively. Furthermore, outliers contributing to the heterogeneity (variability in the causal estimates obtained for each SNP), and horizontal pleiotropy (the variant affects the disease indirectly, i.e., outside its effect on the exposure) within the genetic instruments were identified by leave-one-out analyses and the visual assessment of forest and funnel plots. The genetic variants identified as outliers were removed from the analysis and the causal estimates were recalculated to evaluate the potential biases generated by these outliers in the MR results. Similar to other MR studies (3-5) we avoided inference based simply on p value thresholds. The direction and strength of effect for each MR association, together with the corresponding p value, was considered to better reflect the spectrum of evidence related to these results (6).

**Alternative approach to select genetic instruments in the MR analysis to test the effect of COVID-19 outcomes in anthropometric traits**

Within COVID-19 HGI release-5 data, the largest sample is available for the meta-analysis including data from 23andMe in addition to the other cohorts included in the analyses already described in the Data Source section. In this large dataset, COVID-19 was investigated in 5,101 cases and 1,383,241 controls for the A2 phenotype; 9,986 cases and 1,877,672 controls for B2 phenotype; and 38,984 cases and 1,644,784 controls for C2 phenotype. However, due to 23andme data sharing restrictions, genetic association statistics are available only for the top 10,000 variants. Accordingly, we extracted the LD-independent loci among them and used this information to define genetic instruments for the COVID-19 outcomes investigated. These genetic instruments including variants not reaching genome-wide significance might lead to the violation of the MR assumptions. Thus, we considered as primary analysis the estimates obtained from the IVW method and the MR–Robust Adjusted Profile Score (MR-RAPS) approach. The MR-RAPS approach estimates the causal effect under pervasive horizontal pleiotropy and is robust to occasional outliers (7). The reliability of the findings was tested using multiple sensitivity analyses described above.

**References**

1. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512–25
2. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. Epidemiology. 2017;28:30–42
3. Guo, Y., Lu, Y. and Jin, H. (2020) Appraising the role of circulating concentrations of micro-nutrients in epithelial ovarian cancer risk: A Mendelian randomization analysis. Sci Rep, 10, 7356.
4. Fluharty, M. E., Sallis, H. and Munafò, M. R. (2018) Investigating possible causal effects of externalizing behaviors on tobacco initiation: A Mendelian randomization analysis. Drug Alcohol Depend, 191, 338–342.
5. Carvalho, C. M., Wendt, F. R., Stein, D. J., et al. (2020) Investigating Causality Between Blood Metabolites and Emotional and Behavioral Responses to Traumatic Stress: a Mendelian Randomization Study. Mol Neurobiol, 57, 1542–1552.
6. Sterne, J. A. and Davey Smith, G. (2001) Sifting the evidence-what’s wrong with significance tests? BMJ, 322, 226–231.
7. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in twosample summary-data Mendelian randomization using robust adjusted profile score. arXiv. 2019:1801.09652.