Vitamin D and Covid-19 Susceptibility and Severity: a Mendelian Randomization Study

Authors:
*Guillaume Butler-Laporte[1,2], *Tomoko Nakanishi[1,3,4,5], Vincent Mooser[3,6], David R. Morrison[1], Tala Abdullah[1], Olumide Adeleye[1], Noor Mamlouk[1], Nofar Kimchi[1,7], Zaman Afrasiabi[1], Nardin Rezk[1], Annarita Giliberti[8], Alessandra Renieri[8,9], Yiheng Chen[1], Sirui Zhou[1,2], Vincenzo Forgetta[1], J Brent Richards[1,2,3,10]

*These authors made equal contributions to this manuscript.

Affiliations:

1) Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Québec, Canada
2) Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, Canada
3) Department of Human Genetics, McGill University Montréal, Québec, Canada
4) Kyoto-McGill International Collaborative School in Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan
5) Research Fellow, Japan Society for the Promotion of Science
6) Canada Excellence Research Chair in Genomic Medicine, McGill University, Montréal, Québec, Canada
7) Ruth and Bruce Rappaport Faculty of Medicine, Technion Israel Institute of Technology, Israel
8) Medical Genetics, University of Siena, Siena, Italy
9) Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy
10) Department of Twin Research, King's College London, London, United Kingdom

Corresponding author:
Brent Richards, Professor of Medicine
McGill University
Senior Lecturer, King's College London (Honorary)
Contact:
Pavilion H-413, Jewish General Hospital
3755 Côte-Ste-Catherine
Montréal, Québec, Canada, H3T 1E2
T: +1 514 340 8222 x24362 F: +1 514 340 7529
E: brent.richards@mcgill.ca
www.mcgill.ca/genepi

Funding:
The Richards research group is supported by the Canadian Institutes of Health Research (CIHR: 365825; 409511), the Lady Davis Institute of the Jewish General Hospital, the Canadian Foundation for Innovation, the NIH Foundation, Cancer Research UK, Genome Québec, the Public Health Agency of Canada and the Fonds de Recherche Québec Santé (FRQS). GBL is supported by a joint scholarship from the FRQS and Québec's Ministry of Health and Social Services. TN is supported by Research Fellowships of Japan Society for the Promotion of Science (JSPS) for Young Scientists and JSPS Overseas Challenge Program for Young Researchers. JBR is supported by a FRQS Clinical Research Scholarship. Support from Calcul Québec and Compute Canada is acknowledged. TwinsUK is funded by the Welcome Trust, Medical

Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. These funding agencies had no role in the design, implementation or interpretation of this study.

Disclosures:
JBR has served as an advisor to GlaxoSmithKline and Deerfield Capital. These agencies had no role in the design, implementation or interpretation of this study.

Article type: Research
Word count: 3315 (Introduction, Methods, Results, and Discussion)
Key words: Vitamin D, Covid-19, Sars-CoV-2, Mendelian Randomization, Outcome

Abstract

Introduction

Increased vitamin D levels, as reflected by 25OHD measurements, has been proposed to protect against Covid-19 disease based on *in-vitro*, observational, and ecological studies. However, vitamin D levels are associated with many confounding variables and thus associations described to date may not be causal. Vitamin D MR studies have provided results that are concordant with large-scale vitamin D randomized trials. Here, we used two-sample MR to assess the effect of circulating 25OHD levels on Covid-19 susceptibility.

Methods

Genetic variants strongly associated with 25OHD levels in a 443,734-participant genome-wide association study (GWAS) were used as instrumental variables. GWASs of Covid-19 susceptibility and severity from the Covid-19 Host Genetics Initiative were used to test the effect of 25OHD levels on these outcomes. Cohorts from the Covid-19 Host Genetics Initiative GWAS included up to 966,395 individuals of European ancestry.

Results

Genetically increased 25OHD levels by one standard deviation on the logarithmic scale had no clear effect on susceptibility but tended to increase the odds ratio of hospitalization (OR = 2.34; 95% CI: 1.33, 4.11) and severe disease (OR = 2.21; 95% CI: 0.87, 5.55). Extensive sensitivity analyses probing the assumptions of MR provided consistent estimates.

Conclusion

These findings do not support a protective role of increased 25OHD levels on Covid-19 outcomes and may suggest harm. At present, individuals should not use vitamin D supplements to protect against Covid-19 outcomes, and on-going supplementation trials should closely monitor for signals of harm.

Introduction

SARS-CoV-2 infection has killed hundreds of thousands of individuals and has led to the largest economic contraction since the Great Depression[1]. Vaccines are many months away and their resultant immunity is unknown[2]. Therefore, therapies are required to treat severe Covid-19 disease and to prevent its complications. Therapeutic development, in turn, requires well-validated drug targets to lessen Covid-19 severity.

Recently, vitamin D status, as reflected by 25-hydroxy-vitamin D (25OHD) level has been identified potentially actionable drug target in the prevention and treatment of Covid-19[3]. As the pre-hormone to the biologically active calcitriol, 25OHD has been epidemiologically linked to many health outcomes[4,5]. Given calcitriol's recognized *in-vitro* immunomodulatory role[6], as well as observational and ecological studies on measured 25OHD blood levels and Covid-19[7,8], the vitamin D pathway might be a biologically plausible target in Covid-19. This could be of public health importance, given that the prevalence of vitamin D insufficiency is high in most countries, and that more than 37% of elderly adults in the USA take vitamin D supplements[9]. Further, 25OHD supplementation is inexpensive and reasonably safe—thus providing a potential avenue to lessen the burden of the SARS-CoV-2 pandemic.

However, observational studies on 25OHD are prone to confounding and reverse causation bias. Confounding happens when the relationship between exposure (25OHD) and the outcome (Covid-19) is influenced by an unobserved, or improperly controlled common causes. Reverse causation happens when the outcome itself is a cause of the exposure. Likewise, conclusions drawn from *in-vitro* may not be applicable *in-vivo*. Accordingly, randomized controlled trials (RCTs) on 25OHD supplementation have been undertaken to test their effect on disease outcomes where observational studies have supported a role for 25OHD level. However, across endocrinology, respirology, cardiology, and other specialties, these trials have most often have failed to show clinical benefits[10–12]. Some RCTs have even shown detriment to 25OHD supplementation[13]. In the field of infectious diseases, a trial of 25OHD supplementation showed increased risk of upper respiratory infection in the intervention arm[14], and a recent trial on sepsis obtained a numerically higher mortality rate in patients who received 25OHD supplementation[15]. However, while RCTs can control for confounding and provide unbiased estimates of the effect of 25OHD supplementation in Covid-19, they require considerable time, as well as financial and human resources.

Presently, there are more than 15 registered on-going RCTs testing the effect of vitamin D on Covid-19 outcomes on clinicaltrials.gov. Given difficulties recruiting Covid-19 patients into trials and the probability that epidemiological estimates of vitamin D effects are likely to be confounded, data are urgently needed to assess the causal role of vitamin D on Covid-19 outcomes.

Mendelian randomization (MR) is a genetic epidemiology method that uses genetic variants as instrumental variables to infer the causal effect of an exposure (in this case 25OHD level) on an outcome (in this case, Covid-19 susceptibility and severity)[16]. MR overcomes confounding bias since genetic alleles are randomized to the individual at conception, thereby breaking associations with most confounders. Similarly, since genetic alleles are always assigned prior to disease onset, they are not influenced by reverse causation. In the case of vitamin D, MR has been able to provide causal effect estimates consistently in line with those obtained from RCTs[10,17–21], and would therefore support investments in 25OHD supplementation trials in Covid-19, if a benefit was

shown. Further, since MR results can be generated rapidly, such evidence may provide interim findings while awaiting RCT results.

However, MR relies on several core assumptions[22]. First, genetic variants must be associated with the exposure of interest. Second, they should not affect the outcome except through effects on the exposure (also known as lack of horizontal pleiotropy). Third, they should not associate with the confounders of the exposure-outcome relationship. Of these, the most problematic is the second assumption. Yet, in the case of 25OHD, many of its genetic determinants reside at loci that harbour genes whose roles in 25OHD production, metabolism and transport are well known[23]. Leveraging this known physiology can help to prevent the incorporation of genetic variants that could lead to horizontal pleiotropy.

Here, we used genetic determinants of serum 25OHD from a recent genome-wide association study (GWAS) and meta-analysis of more than 443,734 participants of European ancestry[24] in an MR study to test the effect of increased 25OHD level on Covid-19 susceptibility and severity.

Methods

We used a two-sample MR approach to estimate the effect of 25OHD levels on Covid-19 susceptibility and severity. In two-sample MR[25], the effect of genetic variants on 25OHD and on Covid-19 outcomes are estimated in separate GWASs from different populations. This allows for increased statistical power by increasing the sample size in both the exposure and outcome cohorts.

Choice of 25OHD genetic instruments

To find genetic variants explaining 25OHD levels[24], we used a GWAS from our group, which is the largest published GWAS of 25OHD levels. We selected SNPs whose effect on 25OHD level was genome-wide significant ($P<5x10^{-8}$), minor allele frequency was more than 1%, and which were identified using conditional and joint (COJO) analysis[26]. COJO identifies conditionally independent SNPs and allows for the addition of more genetic instruments without generating falsely narrow confidence intervals due to genetic linkage disequilibrium. For SNPs that were not available in the outcome GWAS or with palindromic alleles of intermediate frequency (between 42% and 58%), we used the LDlink tool[27] to find genetic proxies in the European 1000 Genomes dataset using linkage disequilibrium threshold (r^2) of 90%.

Covid-19 outcome definitions and GWASs

We used the Covid-19 Host Genetics Initiative (Covid-19 HGI) phenotype definitions and GWAS summary statistics for Covid-19 susceptibility, hospitalization, and severe disease outcomes[28]. We meta-analyzed summary statistics from cohorts of European ancestry to obtain the effects of the 25OHD-associated genetics variants on Covid-19 outcomes. Fixed effect meta-analysis models were used to combine summary statistics, with the METAL package[29].

The susceptibility phenotype compared confirmed Covid-19 cases, defined as laboratory confirmed of SARS-CoV-2 infection based on RNA or serology tests, or by electronic health records (using International Classification of Diseases or physician notes), with controls defined as laboratory tested negative for SARS-CoV-2 infection (for all tests if multiple were performed) or self-reported test negative. The hospitalized phenotype compared cases defined as hospitalized patients with Covid-19, and controls as Covid-

19 positive non-hospitalized individuals. The severe disease phenotype cases were defined as hospitalized individuals with Covid-19 and requiring respiratory support. Respiratory support was defined as intubation, CPAP, BiPAP, continuous external negative pressure, or high flow nasal cannula. Controls were also non-hospitalized Covid-19 infected individuals. Details of the three phenotypes are found in **Table 1** and **Supplement 1**. Details of the UKB GWAS are found in **Supplement 2**.

Primary MR analysis

The effect of 25OHD level on Covid-19 outcomes was obtained for each SNP by using the Wald ratio method. The effect of each SNP was given in standardized log-transformed 25OHD level. Each estimate was first meta-analyzed using the IVW method, and then using MR Egger to provide an estimate of directional pleiotropic effects. Allele harmonization and computations were performed using the TwoSampleMR package[30].

Horizontal pleiotropy sensitivity analysis

We undertook extensive analysis to assess the risk of horizontal pleiotropy (in violation of the second MR assumption). First, we used MR Egger methods. MR Egger allows for an additional intercept (alpha) term which provides an estimate of directional horizontal pleiotropy.

Second, we restricted our choices of SNPs to those whose closest gene is directly involved in the Vitamin D pathway. These genes have an established role in vitamin D regulation through its synthesis (*DHCR7*/*NADSYN1* and *CYP2R1*), transportation (*GC*), and degradation (*CYP24A1*) (**Figure 1**). This decreases the risk of selecting a genetic variant that effects Covid-19 outcomes independent of their effect on 25OHD levels.

Third, we used the Phenoscanner tool[31,32] on the remaining SNPs to check for variants associated (at a genome-wide significant threshold of $p=5x10^{-8}$) with phenotypes at risk of affecting Covid-19 outcomes independent of 25OHD, making them at higher risk of horizontal or vertical pleiotropy. Note that vertical pleiotropy, which happens when the Covid-19 outcome is influenced by a phenotype directly in the causal pathway between 25OHD level and Covid-19 outcome, does not violate MR assumptions.

Extended Phenotypes

Finally, we also used an extended susceptibility, an extended hospitalized, and an extended severe disease phenotype whereas controls were defined as all non-cases in the included cohorts. MR analyses were performed as described above. Details of these phenotypes and the MR results are found in **Supplement 2**.

Research Ethics

Each cohort included in this study received their respective institutional research ethics board approval to enroll patients. All information used for this study are publicly available as deidentified GWAS summary statistics.

Results

Choice of 25OHD genetic instruments

We obtained our 25OHD genetic instruments from our previously published GWAS[24] on circulating 25OHD levels in 401,460 white British participants in the UK Biobank (UKB)[33], which was meta-analyzed with a GWAS on 25OHD levels of 42,274 participants of European ancestry[34]. From this, we used single nucleotide variants (SNPs) from the

conditional and joint (COJO) analysis[26]. Of the 138 reported conditionally independent SNPs (explaining 4.9% of the 25OHD variance), 100 had a minor allele frequency of more than 1%. These explained 3.9% of the variance in 25OHD serum levels.

Covid-19 outcome definitions and GWASs

Using the Covid-19 Host Genetics Initiative (Covid-19 HGI) results restricted to cohorts of European ancestry, we used a total of 3,432 cases and 41,285 controls to define Covid-19 susceptibility, 867 cases and 2,229 controls to define Covid-19 hospitalization, and 293 cases and 607 controls to define Covid-19 severe disease. **Table 1** summarizes the definition and sample size of both the exposure and outcome GWASs. Since the UKB was used in the two phases of the MR study, some overlap between the exposure and the outcome GWASs was unavoidable (**Supplement 1**).

For the susceptibility and severe disease phenotypes, 98 of the 138 conditionally independent SNPs influencing 25OHD level were available in the outcome GWAS. For the severe disease outcome, 97 SNPs were available in the outcome GWAS. There was a total of 4 palindromic alleles with intermediate frequencies. Genetic proxies (linkage disequilibrium r^2 > 90%) found in the European sub-cohort of the 1000 Genomes dataset were also unavailable in the outcome GWAS, and no proxies were used for the analyses below. **Table 2** shows the total number of SNP instruments used for each analysis.

Primary MR analysis

We first used inverse-weighted variance (IVW) meta-analysis to combine effect estimates from each genetic instrument. For a standard deviation increase in log-transformed 25OHD level, we observed no clear effect upon odds of susceptibility (OR = 0.88; 95% CI: 0.69, 1.11; P = 0.26). However, we observed an increased risk of hospitalization per standard deviation increase in log-transformed 25OHD (OR = 2.34; 95% CI: 1.33, 4.11; P = 0.003) and trend towards increased risk of severe disease (OR = 2.21; 95% CI: 0.88, 5.55; P = 0.09) (**Table 2**). **Figure 2** shows the result of these analyses.

Horizontal pleiotropy sensitivity analysis

While MR Egger has less statistical power to detect an effect than IVW meta-analysis, we still found similar effect estimates for increased 25OHD levels on susceptibility (OR = 0.61; 95% CI: 0.37, 1.01; P=0.7), hospitalization (OR = 1.97; 95% CI: 0.66, 5.84; P=0.24) and severe disease (OR = 7.68; 95% CI: 1.35, 43.5; P=0.03). MR Egger did not show evidence of directional pleiotropic effects for these analyses.

Then, we restricted SNPs which reside close to the four genes directly involved in 25OHD metabolism. This left 24 SNPs, explaining 2.7% of 25OHD variation. Using these SNPs, each standard deviation increase in log-transformed 25OHD was not associated with a clear effect on Covid-19 susceptibility (OR = 0.90; 95% CI: 0.64, 1.27; P = 0.55). However, we again observed a trend towards increased odds of hospitalization for each standard deviation increase in 25OHD levels on the logarithmic scale (OR = 2.52 [95% CI: 1.28, 4.95]; P = 0.007) and severe disease (OR = 3.04; 95% CI: 1.06, 8.76; P = 0.04). For the three phenotypes, the intercept term confidence interval crossed the null, again suggesting a lack of directional horizontal pleiotropy.

Lastly, we used the Phenoscanner[31,32] tool to check if the SNPs used in the MR study were associated with other phenotypes. Using Phenoscanner, rs11723621 was associated with white blood cell level, and rs6127099 was associated with glomerular

filtration rate. In both cases, the effect on each phenotype was mild compared to their effect on 25OHD level, as rs11723621 explained less than 0.03% of the variance in white blood cell counts, and rs6127099 explained less than 0.001% of the glomerular filtration rate variance[35,36]. Removing these SNPs from the 24 SNPs above further decreased the proportion of 25OHD variance explained to 1.6%. While confidence intervals widened, effect estimates when restricting our analysis to these 22 SNPs remained similar for susceptibility (0.77; 95% CI: 0.48, 1.23; P=0.27), hospitalization (2.89; 95% CI: 1.18, 7.06; P=0.02), and severe disease (2.52; 95% CI: 0.63, 10.0; P=0.19).

Extended Phenotypes

Next, we used an extended definition for all three outcomes, where controls were defined as all non-cases in each cohort, regardless of whether the Covid-19 status was known or not. For these we obtained 6,182 cases and 960,186 controls for the extended susceptibility phenotype, 2,710 cases and 813,234 controls for the extended hospitalization phenotype, and 540 cases and 366,840 controls for the extended severe disease phenotype. MR analysis using the extended phenotypes showed similar results for Covid-19 susceptibility (**Supplement 2**), but their confidence intervals and effect estimates were closer to the null in the extended hospitalization phenotype. For example, for the analysis restricted to genes directly involved in vitamin D metabolism, we obtained a much smaller OR (1.08; 95% CI: 0.69, 1.69; P = 0.74). However, the three extended severe disease analyses showed worse outcome with increasing 25OHD levels.

Discussion

In this large-scale MR study, we found that genetically increased 25OHD levels did not protect against Covid-19 susceptibility, or severity, and in some analyses was associated with worsened outcomes. These findings therefore do not support vitamin D supplementation to prevent Covid-19 outcomes. These results provide guidance on the use of vitamin D supplementation, given that a large proportion of the elderly population currently takes vitamin D supplements. Further, these results should inform the design of vitamin D supplementation trials, which should include attempt to identify signals of harm in interim analyses.

These results are not consistent with those reported in observational studies on vitamin D and Covid-19 outcomes. The discrepancy between our results and traditional epidemiological studies on Covid-19 and 25OHD can likely be explained by the fact that 25OHD is a highly confounded variable. It is consistently associated with an older, more vulnerable, and more socially isolated population, all strong determinants of Covid-19 disease susceptibility and severity. Therefore, even when using advanced statistical adjustments, observational epidemiological studies will likely obtain biased estimates of the effect of vitamin D on Covid-19 in favor of increased 25OHD level. These same limitations may explain the discrepancy between 25OHD observational epidemiology outcomes and RCTs of vitamin D supplementation for common diseases, such as coronary heart disease, cancer, fracture and type 2 diabetes. Further, while vitamin D deficiency has been associated with enhanced inflammation and cytokine release[37], the concept of “cytokine storm” in sepsis is poorly defined, and is currently being reconsidered, including in the case of Covid-19[38].

Most importantly, 25OHD MR studies have given results consistent with RCTs in the colorectal cancer[10,12,19], diabetes[11,18], bone density and fractures[20,39,40], coronary artery

disease[10,21], and many other health outcomes. Our study would also be consistent with a 2015 RCT on 25OHD supplementation in the prevention of respiratory tract infection[7], which showed more upper respiratory tract infections and longer duration of symptoms in those randomized to 25OHD supplements. While this RCT was not powered for more severe outcomes and may not apply to the case of the Covid-19 pandemic, this further highlights the potential role of MR studies to guide future 25OHD supplementation trials.

To date, there are more than 15 trials registered on clinicaltrials.gov assessing the role of 25OHD supplementation in Covid-19, with a wide variety of study designs and primary outcomes. The two largest RCTs are both placebo-controlled trial but have not yet started recruitment. The first will be trial of a 2,414 participants, for which supplementation will be given to healthy Canadian healthcare workers[41], while the second is planned for 1,265 participants in Argentina admitted for Covid-19, but who have not yet developed severe disease[42]. The two trials' primary outcome therefore corresponds to our susceptibility and severe disease phenotypes, and our results should be able to better inform their design and ensure their participants' safety.

Our findings also have important implication for large proportion of the elderly population who are already taking 25OHD supplements[9]. As the evidence for 25OHD supplementation is growing weaker for many other diseases, and with potential adverse events during the Covid-19 pandemic, official public health recommendations should be adjusted to discourage liberal use of 25OHD supplements to protect against Covid-19 outcomes.

Our study has limitations. First, we obtained discrepant results in the hospitalization and extended hospitalization phenotypes. That is, the extended hospitalization MRs did not show any increased risk from a genetically higher 25OHD level. However, hospitalization is a "softer", more subjective outcome than severe disease, and we expect large variations in hospitalization criteria between cohorts, given differences in admission criteria by country. Further, by extending the hospitalization control definition to all individuals without a record of Covid-19 (including those who did not have a test or whose result is unavailable), we expect that the cohorts included in the extended hospitalization phenotype were even more heterogeneous. Hence, the extended hospitalization phenotype results were likely biased towards the null. Nevertheless, our results did not show any benefit from increasing 25OHD levels, and our concerns about vitamin D supplementation remain valid for severe disease outcomes. Second, the UKB was the largest case contributor in the hospitalization phenotype, and the only contributor of the severe disease phenotype. Therefore, for these phenotypes, our MR analyses were closer to one-sample MRs. This distinction is important in analyses suffering from weaker genetic instruments, as the estimate will be biased. In two-sample MR, this bias is towards the null hypothesis. However, in one-sample MR, the bias is in the direction of the epidemiologic association[25]. Yet, univariate epidemiological association studies between 25OHD and Covid-19 severe disease have demonstrated a slight benefit to higher 25OHD levels in the UKB cohort[43,44]. Therefore, if the one-sample MR design introduced any bias, it would be to pull our results closer towards the beneficial effects. Nevertheless, in the extended severe disease phenotype, which incorporated data from an additional cohort with 327 cases, and the results were similar to the severe disease phenotype from UKB with a severe disease odds ratio of 2.39 (95% CI: 1.15, 4.96; P = 0.02) for every standard deviation increase in log transformed 25OHD (for the analysis restricted to genes directly involved in the 25OHD metabolism).

Third, we were only able to include individuals of European ancestry, and it remains possible that 25OHD levels might have different effects on Covid-19 outcomes in other populations. However, previous RCTs on vitamin D supplementation have given similar results in populations of various ancestries[39,40]. Our study did not consider frank vitamin D deficiency, and it remains possible that vitamin D supplementation may remain beneficial in this population. Nevertheless, most proposed RCTs of vitamin D supplementation for Covid-19 (including the largest two) do not target deficient populations[41,42]. Lastly, while we performed multiple sensitivity analyses to attempt to detect horizontal pleiotropy, this source of potential bias cannot be entirely excluded.

In conclusion, using a method that has consistently replicated RCT results for 25OHD outcomes in large sample sizes, we find no evidence to support a protective role for vitamin D supplementation in Covid-19 outcomes. Our findings also suggest that all on-going RCTs studying vitamin D supplementation should have mechanisms in place to detect early signs of worsened outcomes, and that public health recommendations should advise more caution about taking 25OHD supplements.

Acknowledgement

We thank the patients and investigators who contributed to the Covid-19 HGI (**Supplement 4**) and the Vitamin D GWAS consortium. Members of the GEN-COVID study are acknowledged in **Supplement 5**.

Contributions

Conception and design: GBL, TN, JBR. Data acquisition and standardization: AR, AG, DRM, TA, OA, NM, NK, ZA. Data analyses: GBL and TN. Interpretation: GBL, TN, VM, DRM, TA, OA, NM, NK, ZA, AR, AG, SZ, YC, VF, JBR. Computational resources and support: VF, JBR. Writing original draft: GBL, TN, JBR. All authors were involved in reviewing the manuscript and critically reviewed its content. All authors gave final approval of the version to be published. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data availability

Covid-19 outcome GWAS summary statistics are freely available for download through the Covid-19 HGI website (https://www.covid19hg.org/).

Funding source: The Richards research group is supported by the Canadian Institutes of Health Research (CIHR), the Lady Davis Institute of the Jewish General Hospital, the Canadian Foundation for Innovation, the NIH Foundation, Cancer Research UK and the Fonds de Recherche Québec Santé (FRQS). TN is supported by Research Fellowships of Japan Society for the Promotion of Science (JSPS) for Young Scientists and JSPS Overseas Challenge Program for Young Researchers. JBR is supported by a FRQS Clinical Research Scholarship. TwinsUK is funded by the Wellcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. These funding agencies had no role in the design, implementation or interpretation of this study. This research has been conducted using the UK Biobank Resource (project number: 27449).

Competing interests: JBR has served as an advisor to GlaxoSmithKline and Deerfield Capital. These companies had no role in the design, implementation or interpretation of this study.

Transparency statement: The manuscript's guarantors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Supplementary files captions

Supplement 1: Cohorts used for each outcome phenotype for the Covid-19 Host Genetics Initiative.
Supplement 2: Genome-wide association studies for the susceptibility and the severity of SARS-CoV-2 infections in UK Biobank.
Supplement 3: Extended phenotypes definitions and MR results.
Supplement 4: Acknowledgement to data contributors and the Covid-19 Host Genetics Initiative.
Supplement 5: GEN-COVID Multicenter Study

References

1. McKee, M. & Stuckler, D. If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nat. Med.* **26**, 640–642 (2020).
2. Koirala, A., Joo, Y. J., Khatami, A., Chiu, C. & Britton, P. N. Vaccines for COVID-19: The current state of play. *Paediatr. Respir. Rev.* (2020) doi:10.1016/j.prrv.2020.06.010.
3. Mansur, J. L. *et al.* Vitamin D high doses supplementation could represent a promising alternative to prevent or treat COVID-19 infection. *Clin. e Investig. en Arterioscler. Publ. Of. la Soc. Esp. Arterioscler.* (2020) doi:10.1016/j.arteri.2020.05.003.
4. Charoenngam, N. & Holick, M. F. Immunologic Effects of Vitamin D on Human Health and Disease. *Nutrients* **12**, (2020).
5. Marcos-Pérez, D. *et al.* Low Vitamin D Levels and Frailty Status in Older Adults: A Systematic Review and Meta-Analysis. *Nutrients* **12**, (2020).
6. Skrobot, A., Demkow, U. & Wachowska, M. Immunomodulatory Role of Vitamin D: A Review. *Adv. Exp. Med. Biol.* **1108**, 13–23 (2018).
7. Martineau, A. R. & Forouhi, N. G. Vitamin D for COVID-19: a case to answer? *Lancet Diabetes Endocrinol.* (2020) doi:10.1016/S2213-8587(20)30268-0.
8. Munshi, R. *et al.* Vitamin D insufficiency as a potential culprit in critical COVID-19 patients. *J. Med. Virol.* (2020) doi:10.1002/jmv.26360.
9. Rooney, M. R. *et al.* Trends in Use of High-Dose Vitamin D Supplements Exceeding 1000 or 4000 International Units Daily, 1999-2014. *JAMA* **317**, 2448–2450 (2017).
10. Manson, J. E. *et al.* Vitamin D Supplements and Prevention of Cancer and Cardiovascular Disease. *N. Engl. J. Med.* **380**, 33–44 (2019).
11. Pittas, A. G. *et al.* Vitamin D Supplementation and Prevention of Type 2 Diabetes. *N. Engl. J. Med.* **381**, 520–530 (2019).
12. Wactawski-Wende, J. *et al.* Calcium plus vitamin D supplementation and the risk of colorectal cancer. *N. Engl. J. Med.* **354**, 684–696 (2006).
13. Sanders, K. M. *et al.* Annual High-Dose Oral Vitamin D and Falls and Fractures in Older Women: A Randomized Controlled Trial. *JAMA* **303**, 1815–1822 (2010).
14. Martineau, A. R. *et al.* Double-blind randomised controlled trial of vitamin D3 supplementation for the prevention of acute respiratory infection in older adults and their carers (ViDiFlu). *Thorax* **70**, 953 LP – 960 (2015).
15. Ginde, A. A. *et al.* Early High-Dose Vitamin D(3) for Critically Ill, Vitamin D-Deficient Patients. *N. Engl. J. Med.* **381**, 2529–2540 (2019).
16. Davies, N. M., Holmes, M. V & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
17. Hysinger, E. B. *et al.* Mendelian randomization analysis demonstrates that low vitamin D is unlikely causative for pediatric asthma. *J. Allergy Clin. Immunol.* **138**, 1747-1749.e4 (2016).
18. Ye, Z. *et al.* Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol.* **3**, 35–42 (2015).
19. He, Y. *et al.* Exploring causality in the association between circulating 25-hydroxyvitamin D and colorectal cancer risk: a large Mendelian randomisation study. *BMC Med.* **16**, 142 (2018).
20. Trajanoska, K. *et al.* Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *BMJ* **362**, k3225 (2018).

21. Despoina, M., E., M. L., Stephanie, R., David, G. & Brent, R. J. Mendelian Randomization Studies Do Not Support a Role for Vitamin D in Coronary Artery Disease. *Circ. Cardiovasc. Genet.* **9**, 349–356 (2016).
22. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int. J. Epidemiol.* **32**, 1–22 (2003).
23. Mokry, L. E. *et al.* Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLOS Med.* **12**, e1001866 (2015).
24. Manousaki, D. *et al.* Genome-wide Association Study for Vitamin D Levels Reveals 69 Independent Loci. *Am. J. Hum. Genet.* **106**, 327–337 (2020).
25. Lawlor, D. A. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int. J. Epidemiol.* **45**, 908–915 (2016).
26. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369-S3 (2012).
27. Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555–3557 (2015).
28. Initiative, T. C.-19 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 (2020).
29. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
30. Walker, V. M. *et al.* Using the MR-Base platform to investigate risk factors and drug targets for thousands of phenotypes. *Wellcome open Res.* **4**, 113 (2019).
31. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).
32. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
33. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
34. Manousaki, D. *et al.* Low-Frequency Synonymous Coding Variation in CYP2R1 Has Large Effects on Vitamin D Levels and Risk of Multiple Sclerosis. *Am. J. Hum. Genet.* **101**, 227–238 (2017).
35. Astle, W. J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **167**, 1415-1429.e19 (2016).
36. Wuttke, M. *et al.* A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).
37. Daneshkhah, A. *et al.* The Possible Role of Vitamin D in Suppressing Cytokine Storm and Associated Mortality in COVID-19 Patients. *medRxiv* 2020.04.08.20058578 (2020) doi:10.1101/2020.04.08.20058578.
38. Sinha, P., Matthay, M. A. & Calfee, C. S. Is a "Cytokine Storm" Relevant to COVID-19? *JAMA Intern. Med.* (2020) doi:10.1001/jamainternmed.2020.3313.
39. Aloia, J. F., Talwar, S. A., Pollack, S. & Yeh, J. A randomized controlled trial of vitamin D3 supplementation in African American women. *Arch. Intern. Med.* **165**, 1618–1623 (2005).
40. Aspray, T. J. *et al.* Randomized controlled trial of vitamin D supplementation in older people to optimize bone health. *Am. J. Clin. Nutr.* **109**, 207–217 (2019).
41. Preventing COVID-19 With High-dose Vitamin D Supplements. https://clinicaltrials.gov/show/NCT04483635.
42. Cholecalciferol to Improve the Outcomes of COVID-19 Patients.

https://clinicaltrials.gov/show/NCT04411446.

43. Hastie, C. E. *et al.* Vitamin D concentrations and COVID-19 infection in UK Biobank. *Diabetes Metab. Syndr.* **14**, 561–565 (2020).

44. Hastie, C. E., Pell, J. P. & Sattar, N. Short Communication: Vitamin D and COVID-19 infection and mortality in UK Biobank. *medRxiv* 2020.06.26.20140921 (2020) doi:10.1101/2020.06.26.20140921.

Phenotype	Source of genetic variants	
	Cohort	**Participants**
25OHD circulating levels	Manousaki *et al*[24]	Meta-analysis of two 25OHD GWAS: - 401,460 adult white British participants form the UKB[33] - 42,274 from an international consortium of adult individuals of European ancestry[34]
Covid-19 susceptibility	Susceptibility	Meta-analysis of 6 GWAS performed in individuals of European ancestry from 4 countries: - **Cases:** 3,432 individuals with Covid-19 by laboratory confirmation, chart review, or self-report - **Controls:** 41,285 individuals without Covid-19 by laboratory confirmation or self-report
Covid-19 severity	Hospitalized	Meta-analysis of 3 GWAS performed in individuals of European ancestry from 3 countries: **Cases:** 867 hospitalized individuals with Covid-19 **Controls:** 2,229 non-hospitalized individuals with Covid-19
	Severe Disease	GWAS from white British UKB participants: - **Cases:** 293 Covid-19 infected hospitalized individuals who died or required respiratory support (intubation, CPAP, BiPAP, continuous external negative pressure, high flow nasal cannula). - **Controls:** 607 non-hospitalized individuals with Covid-19

Table 1: Sources of data for the analysis. Covid-19 susceptibility and severity outcomes are taken from the Covid-19 HGI[28]. See **Supplement 1** for details on cohorts of Covid-19 susceptibility and severity phenotypes.

Outcome	nSNPs	IVW OR (95% CI)	IVW p-value	Egger alpha	Alpha p-value	Egger OR (95% CI)	Egger p-value
25OHD primary analysis							
Susceptibility	96	0.88 (0.69, 1.11)	0.26	0.005 (-0.006, 0.02)	0.39	0.79 (0.57, 1.10)	0.16
Hospitalization	95	2.34 (1.33, 4.11)	0.003	-0.005 (-0.03, 0.02)	0.70	2.61 (1.18, 5.78)	0.02
Severe disease	96	2.21 (0.88, 5.55)	0.09	-0.03 (-0.07, 0.02)	0.27	3.78 (1.01, 14.1)	0.05
25OHD sensitivity analysis restricted to genes in the vitamin D pathway							
Susceptibility	21	0.90 (0.64, 1.27)	0.55	0.04 (-0.0001, 0.08)	0.07	0.61 (0.37, 1.01)	0.07
Hospitalization	21	2.52 (1.28, 4.95)	0.007	0.02 (-0.06, 0.11)	0.58	1.97 (0.66, 5.84)	0.24
Severe disease	21	3.04 (1.06, 8.76)	0.04	-0.09 (-0.21, 0.04)	0.20	7.68 (1.35, 43.5)	0.03
25OHD sensitivity analysis after removal of SNPs identified by Phenoscanner							
Susceptibility	19	0.77 (0.48, 1.23)	0.27	0.06 (0.01, 0.10)	0.02	0.38 (0.19, 0.75)	0.01
Hospitalization	19	2.89 (1.18, 7.06)	0.02	0.0005 (-0.10, 0.10)	0.99	2.87 (0.64, 12.9)	0.19
Severe disease	19	2.52 (0.63, 10.0)	0.19	-0.14 (-0.29, 0.006)	0.08	17.1 (1.50, 195)	0.04

Table 2: MR results. nSNPs: number of SNPs retained for this analysis. CI: confidence interval.

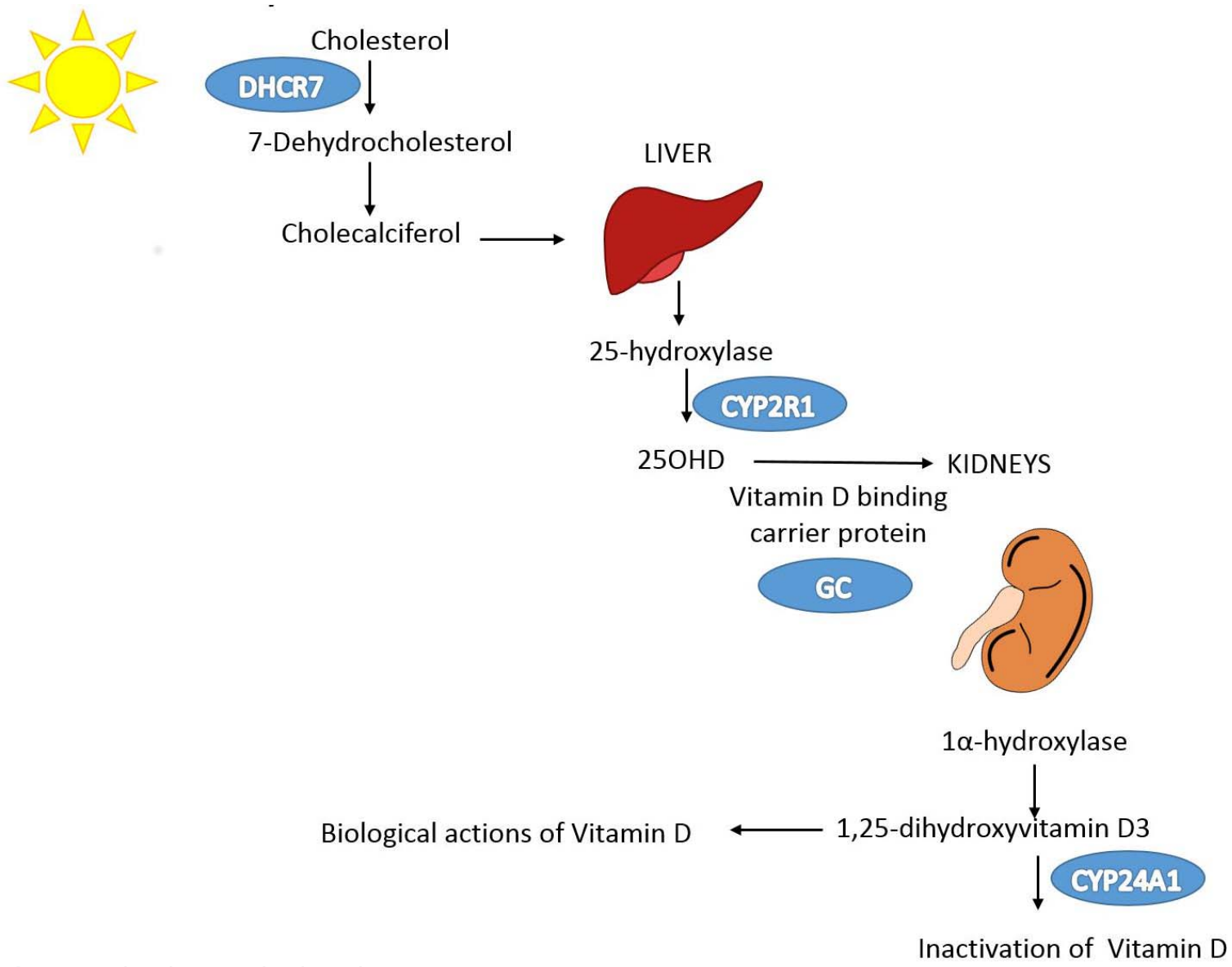

Figure 1: Vitamin D synthesis pathway.

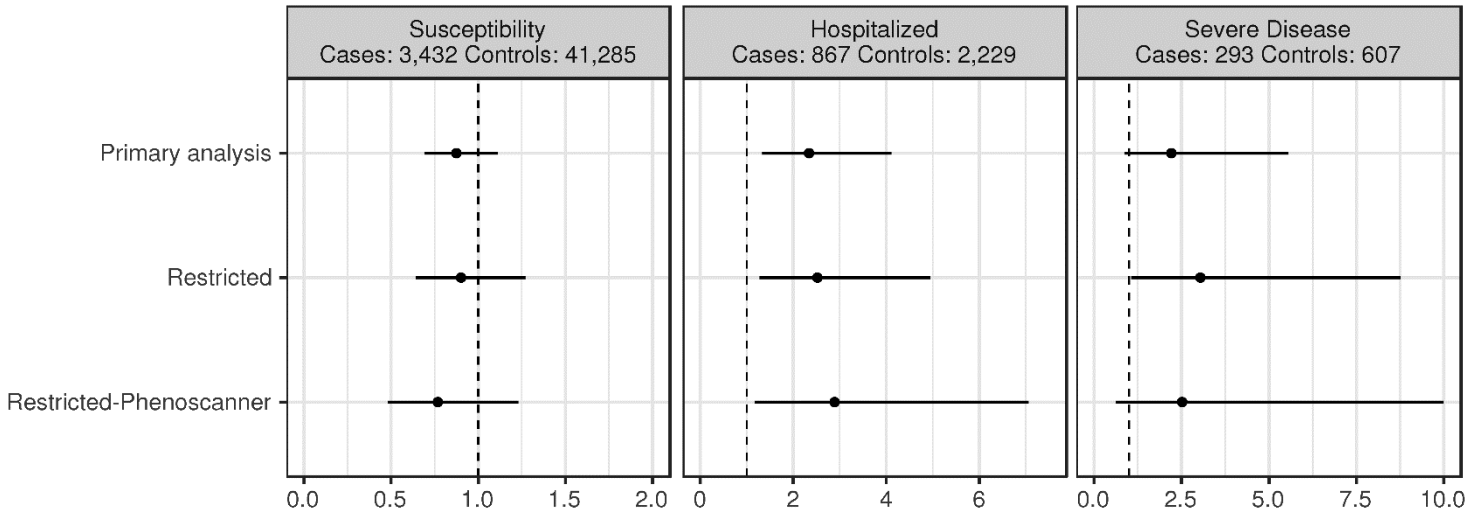

Figure 2: Odds ratio point estimates and 95% confidence intervals for a one standard deviation increase in 25OHD levels on Covid-19 susceptibility and severity. Restricted: analysis restricted to SNPs near the 4 genes involved in known vitamin D metabolic pathways. Restricted-Phenoscanner: restricted analysis, with SNPs identified to have other associations in Phenoscanner removed.