

1 **Vitamin D and COVID-19 Susceptibility and Severity in the COVID-19 Host**
2 **Genetics Initiative: a Mendelian Randomization Study**

3
4 **Authors:**

5 *Guillaume Butler-Laporte^{1,2}, *Tomoko Nakanishi^{1,3,4,5}, Vincent Mooser^{3,6}, David R.
6 Morrison¹, Tala Abdullah¹, Olumide Adeleye¹, Noor Mamlouk¹, Nofar Kimchi^{1,7}, Zaman
7 Afrasiabi¹, Nardin Rezk¹, Annarita Giliberti⁸, Alessandra Renieri^{8,9}, Yiheng Chen¹, Sirui
8 Zhou^{1,2}, Vincenzo Forgetta¹, J Brent Richards^{1,2,3,10}

9
10 *These authors made equal contributions to this manuscript.

11
12 **Affiliations:**

- 13 1) Lady Davis Institute, Jewish General Hospital, McGill University, Montréal,
14 Québec, Canada
15 2) Department of Epidemiology, Biostatistics and Occupational Health, McGill
16 University, Montréal, Québec, Canada
17 3) Department of Human Genetics, McGill University Montréal, Québec, Canada
18 4) Kyoto-McGill International Collaborative School in Genomic Medicine, Graduate
19 School of Medicine, Kyoto University, Kyoto, Japan
20 5) Japan Society for the Promotion of Science
21 6) Canada Excellence Research Chair in Genomic Medicine, McGill University,
22 Montréal, Québec, Canada
23 7) Ruth and Bruce Rappaport Faculty of Medicine, Technion Israel Institute of
24 Technology, Israel
25 8) Medical Genetics, University of Siena, Siena, Italy
26 9) Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy
27 10) Department of Twin Research, King's College London, London, United Kingdom
28

29 **Corresponding author:**

30 Brent Richards, Professor of Medicine
31 McGill University
32 Senior Lecturer, King's College London (Honorary)

33 Contact:

34 Pavilion H-413, Jewish General Hospital
35 3755 Côte-Ste-Catherine
36 Montréal, Québec, Canada, H3T 1E2
37 T: +1 514 340 8222 x24362 F: +1 514 340 7529

38 E: brent.richards@mcgill.ca

39 www.mcgill.ca/genepi

40
41 **Funding:**

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

52 Welcome Trust, Medical Research Council, European Union, the National Institute for
53 Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical
54 Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership
55 with King's College London. VM is supported by the Canada Excellence Research Chair
56 Program. The funders had no role in study design, data collection and analysis, decision
57 to publish, or preparation of the manuscript.

58

59 **Competing interests:**

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

62

63 **Article type:** Research

64 **Word count:** 3480 (Introduction, Methods, Results, and Discussion)

65 **Key words:** Vitamin D, COVID-19, SARS-CoV-2, Mendelian Randomization, Outcome

66

67

68 **Abstract**

69 *Background*

70 Increased vitamin D levels, as reflected by 25OHD measurements, have been proposed
71 to protect against COVID-19 disease based on *in-vitro*, observational, and ecological
72 studies. However, vitamin D levels are associated with many confounding variables and
73 thus associations described to date may not be causal. Vitamin D Mendelian
74 randomization (MR) studies have provided results that are concordant with large-scale
75 vitamin D randomized trials. Here, we used two-sample MR to assess evidence
76 supporting a causal effect of circulating 25OHD levels on COVID-19 susceptibility and
77 severity.

78

79 *Methods and findings*

80 Genetic variants strongly associated with 25OHD levels in a genome-wide association
81 study (GWAS) of 443,734 participants of European ancestry (including 401,460 from the
82 UK Biobank) were used as instrumental variables. GWASs of COVID-19 susceptibility,
83 hospitalization, and severe disease from the COVID-19 Host Genetics Initiative were
84 used as outcome GWASs. These included up to 14,134 individuals with COVID-19, and
85 1,284,876 without COVID-19, from 11 countries. SARS-CoV-2 positivity was determined
86 by laboratory testing or medical chart review. Population controls without COVID-19
87 were also included in the control groups for all outcomes, including hospitalization and
88 severe disease. Analyses were restricted to individuals of European descent when
89 possible. Using inverse-weighted MR, genetically increased 25OHD levels by one
90 standard deviation on the logarithmic scale had no clear association with COVID-19
91 susceptibility (OR = 0.97; 95% CI: 0.95, 1.10; P=0.61), hospitalization (OR = 1.11; 95%
92 CI: 0.91, 1.35; P=0.30), and severe disease (OR = 0.93; 95% CI: 0.73, 1.17; P=0.53).
93 We used an additional 6 meta-analytic methods, as well as sensitivity analyses after
94 removal of variants at risk of horizontal pleiotropy and obtained similar results. These
95 results may be limited by weak instrument bias in some analyses. Further, our results do
96 not apply to individuals with vitamin D deficiency.

97

98 *Conclusion*

99 In this two-sample MR study, we did not observe evidence to support an association
100 between 25OHD levels and COVID-19 susceptibility, severity, or hospitalization. Hence,
101 vitamin D supplementation as a mean of protecting against worsened COVID-19
102 outcomes is not supported by genetic evidence. Other therapeutic or preventative
103 avenues should be given higher priority for COVID-19 randomized controlled trials.

104

105

106 **Author Summary**

107

108 • *Why was this study done?*

109 - Vitamin D levels have been associated with COVID-19 outcomes in
110 multiple observational studies, though confounders are likely to bias these
111 associations.

112 - By using genetic instruments which limit such confounding, Mendelian
113 randomization studies have consistently obtained results concordant with
114 vitamin D supplementation randomized trials. This provides rationale to
115 undertake vitamin D Mendelian randomization studies for COVID-19
116 outcomes.

117 • *What did the researchers do and find?*

118 - We used the genetic variants obtained from the largest consortium of
119 COVID-19 cases and controls, and the largest study on genetic
120 determinants of vitamin D levels.

121 - We used Mendelian randomization to estimate the effect of increased
122 vitamin D on COVID-19 outcomes, while limiting confounding.

123 - In multiple analyses, our results consistently showed no evidence for an
124 association between genetically predicted vitamin D levels and COVID-19
125 susceptibility, hospitalization, or severe disease.

126 • *What do these findings mean?*

127 - Vitamin D is a highly confounded variable, and traditional observational
128 studies are at high risk of biased estimates.

129 - We did not find evidence that vitamin D supplementation would improve
130 COVID-19 outcomes.

131 - Given Mendelian randomization's past track-record of anticipating the
132 results of vitamin D randomized controlled trials, other therapeutic and
133 preventative avenues should be prioritized for COVID-19 trials.

134 **Introduction**

135 SARS-CoV-2 infection has killed millions of individuals and has led to the largest
136 economic contraction since the Great Depression[1]. Therefore, therapies are required
137 to treat severe COVID-19 disease and to prevent its complications. Therapeutic
138 development, in turn, requires well-validated drug targets to lessen COVID-19 severity.

139
140 Recently, vitamin D status, as reflected by 25-hydroxy-vitamin D (25OHD) level has
141 been identified as a potentially actionable drug target in the prevention and treatment of
142 COVID-19[2]. As the pre-hormone to the biologically active calcitriol, 25OHD has been
143 epidemiologically linked to many health outcomes[3,4]. Given calcitriol's recognized *in-*
144 *vitro* immunomodulatory role[5], as well as observational and ecological studies
145 associating measured 25OHD blood levels with COVID-19[6,7], the vitamin D pathway
146 might be a biologically plausible target in COVID-19. This could be of public health
147 importance, given that the prevalence of vitamin D insufficiency is high in most countries,
148 and that more than 37% of elderly adults in the USA take vitamin D supplements[8].
149 Further, 25OHD supplementation is inexpensive and reasonably safe—thus providing a
150 potential avenue to lessen the burden of the SARS-CoV-2 pandemic.

151
152 However, observational studies on 25OHD are prone to confounding and reverse
153 causation bias. Confounding happens when the relationship between exposure (25OHD)
154 and the outcome (COVID-19) is influenced by unobserved, or improperly controlled
155 common causes. Reverse causation happens when the outcome itself is a cause of the
156 exposure. Likewise, conclusions drawn from *in-vitro* may not be applicable *in-vivo*.
157 Accordingly, randomized controlled trials (RCTs) on 25OHD supplementation have been
158 undertaken to test their effect on disease outcomes where observational studies have
159 supported a role for 25OHD level. However, across endocrinology, respiratory,
160 cardiology, and other specialties, these trials have most often not demonstrated
161 statistically significant benefits[9–11]. Some RCTs have even shown detriment to
162 25OHD supplementation[12]. In the field of infectious diseases, an individual patient data
163 meta-analysis of randomized controlled trial of 25OHD supplementation[13] showed
164 some benefit to prevent respiratory tract infections (OR 0.80, 95% CI: 0.69 to 0.93).
165 However, this effect was driven by generally benign upper respiratory tract infections,
166 was not observed in lower respiratory tract disease (OR: 0.96, 95% CI: 0.83 to 1.10) and
167 even showed numerically worse all-cause mortality (OR: 1.39, 95% CI: 0.85 to 2.27).
168 Likewise, a recent trial on sepsis obtained a numerically higher mortality rate in patients
169 who received 25OHD supplementation[14]. At present, we are aware of two RCTs
170 testing the role of vitamin D supplementation on COVID-19 outcomes, both using high-
171 dose vitamin D given at time of hospital admission for COVID-19. The first [15] was a
172 small trial (n=75) showing less intensive care unit admissions in the vitamin D treated
173 arm. However, the follow-up time for mortality varied, and the open-label design put it at
174 high risk of bias. The second[16] was a larger study (n=240) using a double-blind
175 design, and showed no effect on mortality, risk of mechanical ventilation, and length of
176 stay. Nevertheless, questions remain on the use of pre-illness vitamin D
177 supplementation and its effect on disease susceptibility. While RCTs can control for
178 confounding and provide unbiased estimates of the effect of 25OHD supplementation in
179 COVID-19, large well-designed RCTs require considerable resources and time.

180
181 Mendelian randomization (MR) is a genetic epidemiology method that uses genetic
182 variants as instrumental variables to infer the causal effect of an exposure (in this case
183 25OHD level) on an outcome (in this case, COVID-19 susceptibility and severity)[17].
184 MR overcomes confounding bias since genetic alleles are randomized to the individual

185 at conception, thereby breaking associations with most confounders. Similarly, since
186 genetic alleles are always assigned prior to disease onset, they are not influenced by
187 reverse causation. MR has been used in conjunction with proteomics and metabolomics
188 to [15]prioritize drug development and repurposing, and support investment in RCTs
189 which have a higher probability of success[18,19]. In the case of vitamin D, MR has
190 been able to provide causal effect estimates consistently in line with those obtained from
191 RCTs[9,20–24], or support the use of vitamin D supplementation in preventing diseases
192 in at risk individuals (most notably multiple sclerosis [25]). Hence, MR may support
193 investments in 25OHD supplementation trials in COVID-19, if a benefit was shown.
194 Further, since MR results can be generated rapidly, such evidence may provide interim
195 findings while awaiting RCT results.

196
197 However, MR relies on several core assumptions[26]. First, genetic variants must be
198 associated with the exposure of interest. Second, they should not affect the outcome
199 except through effects on the exposure (also known as lack of horizontal pleiotropy).
200 Specifically, MR also assumes that the relationship between the exposure and the
201 outcome is linear. However, this assumption still provides a valid test of the null
202 hypothesis when studying population-level effects[27], as MR then measures the
203 population-averaged effect on the outcome of a shift in the distribution of the exposure.
204 Third, genetic variants should not associate with the confounders of the exposure-
205 outcome relationship. Of these, the most problematic is the second assumption. Yet, in
206 the case of 25OHD, many of its genetic determinants reside at loci that harbour genes
207 whose roles in 25OHD production, metabolism and transport are well known[25].
208 Leveraging this known physiology can help to prevent the incorporation of genetic
209 variants that could lead to horizontal pleiotropy.

210
211 Here, we used genetic determinants of serum 25OHD from a recent genome-wide
212 association study (GWAS) and meta-analysis of more than 443,734 participants of
213 European ancestry[28] in an MR study to test the effect of increased 25OHD level on
214 COVID-19 susceptibility and severity.

215 216 **Methods**

217 We used a two-sample MR approach to estimate the effect of 25OHD levels on COVID-
218 19 susceptibility and severity. In two-sample MR[29], the effect of genetic variants on
219 25OHD and on COVID-19 outcomes are estimated in separate GWASs from different
220 populations. This allows for increased statistical power by increasing the sample size in
221 both the exposure and outcome cohorts. This study is reported as per the Strengthening
222 the Reporting of Observational Studies in Epidemiology (STROBE) guideline[30]
223 (**Supplement 1**).

224
225 Our study did not employ a prospective protocol. Analyses were first planned and
226 performed in July 2020 and updated following peer-review in December 2020. Three
227 major changes were made during the update. First, we used the most up to date COVID-
228 19 Host Genetics Initiative (COVID-19 HGI) GWAS summary statistics. These were
229 made available during the peer-review process. Second, to alleviate potential selection
230 and collider bias, we modified the outcome phenotypes to include population controls.
231 We also performed additional MR sensitivity analyses to check for the robustness of our
232 results. The latter two modifications were made at the request of peer-reviewers. Finally,
233 minor changes to the results' interpretations were made following further peer-review in
234 February 2021.

235

236 *Choice of 25OHD genetic instruments*

237 To find genetic variants explaining 25OHD levels[28], we used a GWAS from our group,
238 which is the largest published GWAS of 25OHD levels, to the best of our knowledge.
239 Importantly, this meta-analysis controlled for season of vitamin D measurement to obtain
240 genetic variants significantly associated with 25OHD levels. From the list of conditionally
241 independent variants provided, we further selected SNPs whose effect on 25OHD level
242 was genome-wide significant ($P < 5 \times 10^{-8}$), minor allele frequency was more than 1%, and
243 with linkage disequilibrium coefficients (r^2) of less than 5%. For SNPs that were not
244 available in the outcome GWAS or with palindromic alleles of intermediate frequency
245 (between 42% and 58%), we used the LDlink tool[31] to find genetic proxies in the
246 European 1000 Genomes dataset (excluding Finnish populations) using linkage
247 disequilibrium r^2 of 90% or more.

248

249 *COVID-19 outcome definitions and GWASs*

250 We used the COVID-19 HGI outcome definitions and GWAS summary statistics for
251 COVID-19 susceptibility, hospitalization, and severe disease outcomes[32]. For all
252 outcomes, a COVID-19 infection was defined as a positive SARS-CoV-2 laboratory test
253 (e.g. RNA RT-PCR or serology tests) or electronic health record evidence of SARS-CoV-
254 2 infection (using International Classification of Diseases or physician notes). The
255 susceptibility phenotype compared COVID-19 cases, with controls which were defined
256 as any individuals without a history of COVID-19. The hospitalized outcome compared
257 cases defined as hospitalized patients with COVID-19, and controls as any individuals
258 not experiencing a hospitalization for COVID-19, which includes those without COVID-
259 19. The severe disease outcome cases were defined as hospitalized individuals with
260 COVID-19 who required respiratory support. Respiratory support was defined as
261 intubation, CPAP, BiPAP, continuous external negative pressure, or high flow nasal
262 cannula. Controls for the severe COVID-19 outcome were defined as individuals without
263 severe COVID-19 (including those without COVID-19). The inclusion of COVID-19
264 negative participants as controls in each outcome decreases the possibility of collider
265 bias[33] and allows for better population level comparisons. These three outcome
266 phenotypes are referred to as C2, B2, and A2, respectively, in the COVID-19 HGI
267 documentation.

268

269 For our study, we used the October 20th 2020 (v4) COVID-19 HGI fixed effect meta-
270 analysis of GWAS from up to 22 cohorts, performed in up to 11 countries. Every
271 participating cohort was asked to provide summary statistics from a GWAS on the above
272 three outcomes, and including the following non-genetic covariates: age, sex, age*age,
273 age*sex, 20 genetic principal components, as well as any locally relevant covariates at
274 the discretion of participating studies (e.g. hospital, genotype panel, etc.). Cohorts were
275 asked to follow common sample and variant quality control, and only performed analysis
276 if they enrolled 100 cases or more. Analyses were done separately for each major
277 ancestry group to further control for population stratification. For the purposes of our
278 study, we used the meta-analysis results from European ancestry cohorts, except for the
279 severe COVID-19 outcome, for which this meta-analysis was not available. Further
280 details on the three phenotypes and participating cohorts are found in **Table 1** and
281 **Supplement 2**.

282

283 *Primary MR analysis*

284 The effect of 25OHD level on COVID-19 outcomes was obtained for each SNP by using
285 the Wald ratio method. The effect of each SNP was given in standardized log-
286 transformed 25OHD level. Each estimate was meta-analyzed using the inverse-variance

287 weighted (IVW) method, and we performed variant heterogeneity tests to check
288 robustness of IVW results. Allele harmonization and computations were performed using
289 the TwoSampleMR package[34].

290

291 *Horizontal pleiotropy sensitivity analysis*

292 We undertook extensive analysis to assess the risk of horizontal pleiotropy (a violation of
293 the second MR assumption). First, we used MR Egger method, which allows for an
294 additional intercept (alpha) term which also provides an estimate of directional horizontal
295 pleiotropy. This method relies upon the assumption that the size of the direct effects of
296 the genetic variants on the outcome that do not operate through the exposure are
297 independent of the variant's effect on the exposure. Given possible instability in MR
298 Egger estimates[35], we also used the bootstrap MR Egger method to meta-analyze the
299 causal effect estimates from each SNP instrument. Further, we used four additional
300 meta-analysis methods known to be more robust to presence of horizontal pleiotropy (at
301 the expense of statistical power): penalised weighted median, simple mode, weighted
302 median, and weighted mode[36].

303

304 Second, we restricted our choices of SNPs to those whose closest gene is directly
305 involved in the vitamin D pathway. These genes have an established role in vitamin D
306 regulation through its synthesis (*DHCR7/NADSYN1* and *CYP2R1*), transportation (*GC*),
307 and degradation (*CYP24A1*) (**Supplement 3**). This decreases the risk of selecting a
308 genetic variant that affects COVID-19 outcomes independent of its effect on 25OHD
309 levels.

310

311 Third, we used the Phenoscanner tool[37,38] on the remaining SNPs to check for
312 variants associated (at a genome-wide significant threshold of $p=5 \times 10^{-8}$) with
313 phenotypes at risk of affecting COVID-19 outcomes independent of 25OHD, making
314 them at higher risk of horizontal or vertical pleiotropy. Note that vertical pleiotropy, which
315 happens when the COVID-19 outcome is influenced by a phenotype directly in the
316 causal pathway between 25OHD level and COVID-19 outcome, does not violate MR
317 assumptions.

318

319 *Research Ethics*

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

323

324 **Results**

325 *Choice of 25OHD genetic instruments*

326 We obtained our 25OHD genetic instruments from our previously published GWAS[28]
327 on circulating 25OHD levels in 401,460 white British participants in the UK Biobank
328 (UKB)[39], which was meta-analyzed with a GWAS on 25OHD levels of 42,274
329 participants of European ancestry[40]. Of the 138 reported conditionally independent
330 SNPs (explaining 4.9% of the 25OHD variance), 100 had a minor allele frequency of
331 more than 1%, of which 78 were directly available in the COVID-19 HGI GWAS
332 summary statistic and had linkage disequilibrium coefficient of less than 5%.
333 Additionally, 3 more variants had good genetic proxies ($r^2 > 90\%$) and were therefore
334 added to our instrument lists, for a total of 81 variants. These explained 4.3% of the
335 variance in 25OHD serum levels. The full list of SNPs used can be found in **Supplement**

336 **4.**

337

338 *COVID-19 outcome definitions and GWASs*

339 Using the COVID-19 HGI results restricted to cohorts of European ancestry, we used
340 total of 14,134 cases and 1,284,876 controls to define COVID-19 susceptibility, 6,406
341 cases and 902,088 controls to define COVID-19 hospitalization, and 4,336 cases and
342 623,902 controls to define COVID-19 severe disease. **Table 1** summarizes the definition
343 and sample size of both the exposure and outcome GWASs. Since the UKB was used in
344 the two phases of the MR study, some overlap between the exposure and the outcome
345 GWASs was unavoidable (**Supplement 2**).

346

347 *Primary MR analysis*

348 We first used IVW meta-analysis to combine effect estimates from each genetic
349 instrument. For a standard deviation increase in log-transformed 25OHD level, we
350 observed no statistically significant effect upon odds of susceptibility (OR = 0.97; 95%
351 CI: 0.85, 1.10; P = 0.61). Of note, in the UKB, the distribution of 25OHD levels has a
352 mean of 48.6 nmol/L and a standard deviation of 21.1 nmol/L. This standard deviation is
353 comparable to what can be achieved with vitamin D supplementation, especially over
354 short therapeutic courses[41]. Similarly, we observed no significant difference in risk of
355 hospitalization (OR = 1.11; 95% CI: 0.91, 1.35; P = 0.30) or risk of severe disease (OR =
356 0.93; 95% CI: 0.73, 1.17; P = 0.53) associated with a standard deviation increase in log-
357 transformed 25OHD level (**Table 2** and **Figure 1**).

358

359 *Horizontal pleiotropy assessment and sensitivity analysis*

360 Using the MR Egger intercept terms, we did not observe evidence of horizontal
361 pleiotropy. While they have less statistical power than IVW meta-analysis, the 6
362 sensitivity meta-analyses we used also showed no evidence of an association between
363 25OHD levels and COVID-19 susceptibility, hospitalization, and severe disease, with
364 each confidence interval crossing the null in the primary analysis using all SNPs (**Figure**
365 **1** and **Supplement 5**). Our results are therefore unlikely to be strongly biased by
366 horizontal pleiotropy.

367

368 Second, we restricted our analysis to SNPs which reside close to the four genes directly
369 involved in 25OHD metabolism. This left 12 SNPs, explaining 3.2% of 25OHD variation.
370 Using IVW, each standard deviation increase in log-transformed 25OHD was again not
371 associated with COVID-19 susceptibility (OR = 0.96; 95% CI: 0.83, 1.11; P = 0.59),
372 hospitalization (OR = 1.07 [95% CI: 0.78, 1.47]; P = 0.67) and severe disease (OR =
373 0.87; 95% CI: 0.63, 1.19; P = 0.38). For the three phenotypes, the MR Egger intercept
374 term did not support bias from directional horizontal pleiotropy.

375

376 Lastly, we used the Phenoscanner[37,38] tool to check if the SNPs used in the MR study
377 were associated with other phenotypes. Using Phenoscanner, rs11723621 was
378 associated with white blood cell level, and rs6127099 was associated with glomerular
379 filtration rate. In both cases, the association with each phenotype was mild compared to
380 their effect on 25OHD level, as rs11723621 explained less than 0.03% of the variance in
381 white blood cell counts, and rs6127099 explained less than 0.001% of the glomerular
382 filtration rate variance[42,43]. Removing these SNPs from the 12 SNPs above further
383 decreased the proportion of 25OHD variance explained to 1.7%. While confidence
384 intervals widened, effect estimates when restricting our analysis to these SNPs
385 remained null for susceptibility (0.80; 95% CI: 0.77, 1.23; P=0.80), hospitalization (1.09;
386 95% CI: 0.68, 1.75; P=0.71), and severe disease (0.91; 95% CI: 0.54, 1.55; P=0.73).

387

388 *Genetic instruments heterogeneity*

389 Overall, our results showed little evidence of heterogeneity of effect between our genetic
390 instruments (**Table 2**). We nonetheless observed that for at least one of the three
391 analyses, we would have rejected the null hypothesis of homogeneous genetic effects in
392 the COVID-19 hospitalization phenotype. However, given the large number of
393 hypotheses tested, this may be due to chance.

394

395 Discussion

396 In this large-scale MR study, we did not find evidence to support that genetically
397 increased 25OHD levels are protective against COVID-19 susceptibility, hospitalization,
398 or severity. This lack of evidence was consistent across phenotypes, sensitivity
399 analyses, and choice of genetic instruments. These findings highlight the confounded
400 association between vitamin D and COVID-19 due to factors such as older age,
401 institutionalization, or medical comorbidities, that are all linked to lower vitamin D levels
402 and cannot be controlled for even when using advanced statistical adjustments. While
403 our study assessed the association between genetically determined levels of 25OHD
404 and COVID-19, this can still inform us on the role of vitamin D supplementation.
405 Specifically, in contrast to observational studies, our findings do not support an
406 association between higher 25OHD level and better COVID-19 outcome, and therefore
407 do not support the use of vitamin D supplementation to prevent COVID-19 outcomes.
408 Further, while a randomized trial[15] showed benefit of vitamin D supplementation using
409 an endpoint at risk of bias due to the unblinded intervention (admission to the critical
410 care unit) and a small sample size (n=75), a larger randomized trial[16] of 240 patients
411 showed no effect of a single high dose of vitamin D3 on mortality, length of stay, or risk
412 of mechanical ventilation. These findings from the largest randomized trial are thus
413 concordant with our MR results.

414

415 Our study's main strength is MR's track record in the study of vitamin D, with multiple
416 studies like ours having provided results consistent with RCTs in multiple medical
417 conditions [9–11,21–24,44,45]. Our study also leverages the largest cohort of COVID-19
418 cases and controls currently available (even outside of genetic studies) and the largest
419 study on genetic determinants of 25OHD levels to date. Using these data sources, we
420 were able to obtain results robust to multiple sensitivity analysis and MR methodology.

421

422 Our study still has limitations. First, our results do not take true vitamin D deficiency into
423 account, and it remains possible that truly deficient patients may benefit from
424 supplementation for COVID-19 related purposes. However, when comparing previous
425 results from MR studies and RCTs, we do not expect a large impact even in this
426 population. Second, our study may suffer from weak instrument bias, especially the
427 sensitivity analyses restricted to smaller sets of genetic instruments. In two-sample MR,
428 this bias would tend to make estimates closer to the null. Nonetheless, similar studies
429 have been able to use MR to establish an association between 25OHD levels and other
430 diseases (most notably multiple sclerosis[25]), suggesting that these instruments are
431 strong enough to find associations. Further, given the large percentage of shared
432 individuals from the UKB between the vitamin D exposure GWAS[28] and the severe
433 COVID-19 phenotype, this analysis is close to a one-sample MR, which would show bias
434 towards the observational study association. Given that this analysis also shows largely
435 null effects, we do not suspect that weak instruments bias is a significant issue in our
436 results. Third, given that vitamin D levels are affected by season (with higher levels after
437 sunlight exposure), even if our SNP-instruments were obtained from a GWAS that
438 controlled for season of blood draw, there is still a chance for effect attenuation by
439 averaging the effect of 25OHD levels on COVID-19 over all seasons. Nevertheless, a

440 recent study in a Finnish cohort (where sun exposure greatly varies with seasons)
441 showed that genetic determinants of 25OHD level were able to discriminate between
442 individuals with predisposition to varying levels of 25OHD, regardless of the season[46].
443 Therefore, while the cyclical nature of 25OHD level is not completely modelled by MR,
444 the size of this bias is likely small. Fourth, our MR analyses assume a linear exposure-
445 outcome relationship. While this may slightly bias our results, simulation studies have
446 previously shown that this assumption provides adequate results when looking at a
447 population effect[27]. Therefore, for the purpose of vitamin D supplementation in the
448 general population, our conclusions should still be valid. Lastly, as we only studied the
449 effect of 25OHD and COVID-19 in individuals of European ancestry, it remains possible
450 that 25OHD levels might have different effects on COVID-19 outcomes in other
451 populations. However, previous RCTs on vitamin D supplementation have given similar
452 results in populations of various ancestries[44,45].

453

454 In conclusion, using a method that has consistently replicated RCT results from vitamin
455 D supplementation studies in large sample sizes, we find no evidence to support a
456 protective role for higher 25OHD on COVID-19 outcomes. Specifically, vitamin D
457 supplementation as a public health measure to improve COVID-19 outcomes is not
458 supported by this MR study. Most importantly, our results suggest that investment in
459 other therapeutic or preventative avenues should be prioritized for COVID-19 RCTs.

460

461 **Acknowledgements**

462 We thank the patients and investigators who contributed to the COVID-19 HGI
463 (**Supplement 6**) and the Vitamin D GWAS consortium. Members of the GEN-COVID
464 study are acknowledged in **Supplement 7**. This research has been conducted using the
465 UK Biobank Resource (project number: 27449).

466

467 **Contributions**

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

476

477 **Supplementary files captions**

478 **Supplement 1:** STROBE case-control study checklist

479 **Supplement 2:** Cohorts used for each outcome phenotype for the COVID-19 Host
480 Genetics Initiative.

481 **Supplement 3:** Vitamin D metabolism pathway and genes involved.

482 **Supplement 4:** Genetic instruments summary statistics.

483 **Supplement 5:** results from Mendelian randomization sensitivity analyses.

484 **Supplement 6:** Acknowledgement to data contributors and the COVID-19 Host Genetics
485 Initiative.

486 **Supplement 7:** GEN-COVID Multicenter Study.

487 **References**

- 488 1. McKee M, Stuckler D. If the world fails to protect the economy, COVID-19 will
489 damage health not just now but also in the future. *Nat Med.* 2020;26: 640–642.
490 doi:10.1038/s41591-020-0863-y
- 491 2. Mansur JL, Tajer C, Mariani J, Insera F, Ferder L, Manucha W. Vitamin D high
492 doses supplementation could represent a promising alternative to prevent or treat
493 COVID-19 infection. *Clin e Investig en Arterioscler Publ Of la Soc Esp*
494 *Arterioscler.* 2020. doi:10.1016/j.arteri.2020.05.003
- 495 3. Charoengam N, Holick MF. Immunologic Effects of Vitamin D on Human Health
496 and Disease. *Nutrients.* 2020;12. doi:10.3390/nu12072097
- 497 4. Marcos-Pérez D, Sánchez-Flores M, Proietti S, Bonassi S, Costa S, Teixeira JP,
498 et al. Low Vitamin D Levels and Frailty Status in Older Adults: A Systematic
499 Review and Meta-Analysis. *Nutrients.* 2020;12. doi:10.3390/nu12082286
- 500 5. Skrobot A, Demkow U, Wachowska M. Immunomodulatory Role of Vitamin D: A
501 Review. *Adv Exp Med Biol.* 2018;1108: 13–23. doi:10.1007/5584_2018_246
- 502 6. Martineau AR, Forouhi NG. Vitamin D for COVID-19: a case to answer? *Lancet*
503 *Diabetes Endocrinol.* 2020. doi:10.1016/S2213-8587(20)30268-0
- 504 7. Munshi R, Hussein MH, Toraih EA, Elshazli RM, Jardak C, Sultana N, et al.
505 Vitamin D insufficiency as a potential culprit in critical COVID-19 patients. *J Med*
506 *Virol.* 2020. doi:10.1002/jmv.26360
- 507 8. Rooney MR, Harnack L, Michos ED, Ogilvie RP, Sempos CT, Lutsey PL. Trends
508 in Use of High-Dose Vitamin D Supplements Exceeding 1000 or 4000
509 International Units Daily, 1999–2014. *JAMA.* 2017;317: 2448–2450.
510 doi:10.1001/jama.2017.4392
- 511 9. Manson JE, Cook NR, Lee I-M, Christen W, Bassuk SS, Mora S, et al. Vitamin D
512 Supplements and Prevention of Cancer and Cardiovascular Disease. *N Engl J*
513 *Med.* 2019;380: 33–44. doi:10.1056/NEJMoa1809944
- 514 10. Pittas AG, Dawson-Hughes B, Sheehan P, Ware JH, Knowler WC, Aroda VR, et
515 al. Vitamin D Supplementation and Prevention of Type 2 Diabetes. *N Engl J Med.*
516 2019;381: 520–530. doi:10.1056/NEJMoa1900906
- 517 11. Wactawski-Wende J, Kotchen JM, Anderson GL, Assaf AR, Brunner RL,
518 O’Sullivan MJ, et al. Calcium plus vitamin D supplementation and the risk of
519 colorectal cancer. *N Engl J Med.* 2006;354: 684–696.
520 doi:10.1056/NEJMoa055222
- 521 12. Sanders KM, Stuart AL, Williamson EJ, Simpson JA, Kotowicz MA, Young D, et
522 al. Annual High-Dose Oral Vitamin D and Falls and Fractures in Older Women: A
523 Randomized Controlled Trial. *JAMA.* 2010;303: 1815–1822.
524 doi:10.1001/jama.2010.594
- 525 13. Martineau AR, Jolliffe DA, Hooper RL, Greenberg L, Aloia JF, Bergman P, et al.
526 Vitamin D supplementation to prevent acute respiratory tract infections:
527 systematic review and meta-analysis of individual participant data. *BMJ.*
528 2017;356: i6583. doi:10.1136/bmj.i6583
- 529 14. Ginde AA, Brower RG, Caterino JM, Finck L, Banner-Goodspeed VM, Grissom
530 CK, et al. Early High-Dose Vitamin D(3) for Critically Ill, Vitamin D-Deficient
531 Patients. *N Engl J Med.* 2019;381: 2529–2540. doi:10.1056/NEJMoa1911124
- 532 15. Entrenas Castillo M, Entrenas Costa LM, Vaquero Barrios JM, Alcalá Díaz JF,
533 López Miranda J, Bouillon R, et al. “Effect of calcifediol treatment and best
534 available therapy versus best available therapy on intensive care unit admission
535 and mortality among patients hospitalized for COVID-19: A pilot randomized
536 clinical study.” *J Steroid Biochem Mol Biol.* 2020/08/29. 2020;203: 105751.
537 doi:10.1016/j.jsbmb.2020.105751

- 538 16. Murai IH, Fernandes AL, Sales LP, Pinto AJ, Goessler KF, Duran CSC, et al.
539 Effect of Vitamin D Supplementation vs Placebo on Hospital Length of Stay in
540 Patients with Severe COVID-19: A Multicenter, Double-blind, Randomized
541 Controlled Trial. medRxiv. 2020; 2020.11.16.20232397.
542 doi:10.1101/2020.11.16.20232397
- 543 17. Davies NM, Holmes M V, Davey Smith G. Reading Mendelian randomisation
544 studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362: k601.
545 doi:10.1136/bmj.k601
- 546 18. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al.
547 Phenome-wide Mendelian randomization mapping the influence of the plasma
548 proteome on complex diseases. *Nat Genet*. 2020. doi:10.1038/s41588-020-0682-
549 6
- 550 19. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al.
551 Connecting genetic risk to disease end points through the human blood plasma
552 proteome. *Nat Commun*. 2017;8: 14357. doi:10.1038/ncomms14357
- 553 20. Hysinger EB, Roizen JD, Mentch FD, Vazquez L, Connolly JJ, Bradfield JP, et al.
554 Mendelian randomization analysis demonstrates that low vitamin D is unlikely
555 causative for pediatric asthma. *J Allergy Clin Immunol*. 2016/08/20. 2016;138:
556 1747-1749.e4. doi:10.1016/j.jaci.2016.06.056
- 557 21. Ye Z, Sharp SJ, Burgess S, Scott RA, Imamura F, Langenberg C, et al.
558 Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes:
559 a mendelian randomisation study. *Lancet Diabetes Endocrinol*. 2015;3: 35–42.
560 doi:10.1016/S2213-8587(14)70184-6
- 561 22. He Y, Timofeeva M, Farrington SM, Vaughan-Shaw P, Svinti V, Walker M, et al.
562 Exploring causality in the association between circulating 25-hydroxyvitamin D
563 and colorectal cancer risk: a large Mendelian randomisation study. *BMC Med*.
564 2018;16: 142. doi:10.1186/s12916-018-1119-2
- 565 23. Trajanoska K, Morris JA, Oei L, Zheng H-F, Evans DM, Kiel DP, et al.
566 Assessment of the genetic and clinical determinants of fracture risk: genome wide
567 association and mendelian randomisation study. *BMJ*. 2018;362: k3225.
568 doi:10.1136/bmj.k3225
- 569 24. Despoina M, E. ML, Stephanie R, David G, Brent RJ. Mendelian Randomization
570 Studies Do Not Support a Role for Vitamin D in Coronary Artery Disease. *Circ*
571 *Cardiovasc Genet*. 2016;9: 349–356. doi:10.1161/CIRCGENETICS.116.001396
- 572 25. Mokry LE, Ross S, Ahmad OS, Forgetta V, Smith GD, Leong A, et al. Vitamin D
573 and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLOS Med*.
574 2015;12: e1001866. Available: <https://doi.org/10.1371/journal.pmed.1001866>
- 575 26. Davey Smith G, Ebrahim S. ‘Mendelian randomization’: can genetic epidemiology
576 contribute to understanding environmental determinants of disease?*. *Int J*
577 *Epidemiol*. 2003;32: 1–22. doi:10.1093/ije/dyg070
- 578 27. Burgess S, Davies NM, Thompson SG. Instrumental variable analysis with a
579 nonlinear exposure-outcome relationship. *Epidemiology*. 2014;25: 877–885.
580 doi:10.1097/EDE.000000000000161
- 581 28. Manousaki D, Mitchell R, Dudding T, Haworth S, Harroud A, Forgetta V, et al.
582 Genome-wide Association Study for Vitamin D Levels Reveals 69 Independent
583 Loci. *Am J Hum Genet*. 2020;106: 327–337.
584 doi:<https://doi.org/10.1016/j.ajhg.2020.01.017>
- 585 29. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities
586 and challenges. *Int J Epidemiol*. 2016/07/17. 2016;45: 908–915.
587 doi:10.1093/ije/dyw127
- 588 30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP.

- 589 The Strengthening the Reporting of Observational Studies in Epidemiology
590 (STROBE) statement: guidelines for reporting observational studies. *PLoS Med.*
591 2007;4: e296. doi:10.1371/journal.pmed.0040296
- 592 31. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring
593 population-specific haplotype structure and linking correlated alleles of possible
594 functional variants. *Bioinformatics.* 2015;31: 3555–3557.
595 doi:10.1093/bioinformatics/btv402
- 596 32. Initiative TC-19 HG. The COVID-19 Host Genetics Initiative, a global initiative to
597 elucidate the role of host genetic factors in susceptibility and severity of the
598 SARS-CoV-2 virus pandemic. *Eur J Hum Genet.* 2020;28: 715–718.
599 doi:10.1038/s41431-020-0636-6
- 600 33. Griffith GJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider
601 bias undermines our understanding of COVID-19 disease risk and severity. *Nat*
602 *Commun.* 2020;11: 5749. doi:10.1038/s41467-020-19478-2
- 603 34. Walker VM, Davies NM, Hemani G, Zheng J, Haycock PC, Gaunt TR, et al. Using
604 the MR-Base platform to investigate risk factors and drug targets for thousands of
605 phenotypes. *Wellcome open Res.* 2019;4: 113.
606 doi:10.12688/wellcomeopenres.15334.2
- 607 35. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-
608 sample Mendelian randomization. *Genet Epidemiol.* 2016; 1–12.
609 doi:10.1002/gepi.21998
- 610 36. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods
611 using summary data. *Genet Epidemiol.* 2020;44: 313–329.
612 doi:https://doi.org/10.1002/gepi.22295
- 613 37. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al.
614 PhenoScanner: a database of human genotype-phenotype associations.
615 *Bioinformatics.* 2016;32: 3207–3209. doi:10.1093/bioinformatics/btw373
- 616 38. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al.
617 PhenoScanner V2: an expanded tool for searching human genotype-phenotype
618 associations. *Bioinformatics.* 2019;35: 4851–4853.
619 doi:10.1093/bioinformatics/btz469
- 620 39. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK
621 Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562:
622 203–209. doi:10.1038/s41586-018-0579-z
- 623 40. Manousaki D, Dudding T, Haworth S, Hsu Y-H, Liu C-T, Medina-Gómez C, et al.
624 Low-Frequency Synonymous Coding Variation in CYP2R1 Has Large Effects on
625 Vitamin D Levels and Risk of Multiple Sclerosis. *Am J Hum Genet.* 2017;101:
626 227–238. doi:https://doi.org/10.1016/j.ajhg.2017.06.014
- 627 41. Żebrowska A, Sadowska-Krępa E, Stanula A, Waśkiewicz Z, Łakomy O, Bezuglov
628 E, et al. The effect of vitamin D supplementation on serum total 25(OH) levels and
629 biochemical markers of skeletal muscles in runners. *J Int Soc Sports Nutr.*
630 2020;17: 18. doi:10.1186/s12970-020-00347-8
- 631 42. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic
632 Landscape of Human Blood Cell Trait Variation and Links to Common Complex
633 Disease. *Cell.* 2016;167: 1415-1429.e19. doi:10.1016/j.cell.2016.10.042
- 634 43. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic
635 loci associated with kidney function from analyses of a million individuals. *Nat*
636 *Genet.* 2019;51: 957–972. doi:10.1038/s41588-019-0407-x
- 637 44. Aloia JF, Talwar SA, Pollack S, Yeh J. A randomized controlled trial of vitamin D3
638 supplementation in African American women. *Arch Intern Med.* 2005;165: 1618–
639 1623. doi:10.1001/archinte.165.14.1618

- 640 45. Aspray TJ, Chadwick T, Francis RM, McColl E, Stamp E, Prentice A, et al.
641 Randomized controlled trial of vitamin D supplementation in older people to
642 optimize bone health. *Am J Clin Nutr.* 2019;109: 207–217.
643 doi:10.1093/ajcn/nqy280
644 46. Sallinen RJ, Dethlefsen O, Ruotsalainen S, Mills RD, Miettinen TA, Jääskeläinen
645 TE, et al. Genetic Risk Score for Serum 25-Hydroxyvitamin D Concentration
646 Helps to Guide Personalized Vitamin D Supplementation in Healthy Finnish
647 Adults. *J Nutr.* 2020. doi:10.1093/jn/nxaa391
648

Phenotype	Source of genetic variants	
	Cohort	Participants
25OHD circulating levels	Manousaki <i>et al</i> [28]	Meta-analysis of two 25OHD GWAS: <ul style="list-style-type: none"> - 401,460 adult white British participants from the UKB[39] - 42,274 from an international consortium of adult individuals of European ancestry[40]
COVID-19 susceptibility	Susceptibility	Meta-analysis of 22 GWAS performed in individuals of European ancestry from 11 countries: <ul style="list-style-type: none"> - Cases: 14,134 individuals with COVID-19 by laboratory confirmation, chart review, or self-report - Controls: 1,284,876 individuals without confirmation or history of COVID-19
COVID-19 severity	Hospitalized	Meta-analysis of 13 GWAS performed in individuals of European ancestry from 11 countries: <ul style="list-style-type: none"> - Cases: 6,406 hospitalized individuals with COVID-19 - Controls: 902,088 without hospitalization with COVID-19
	Severe Disease	Meta-analysis of 12 GWAS performed in individuals of European ancestry from 9 countries: <ul style="list-style-type: none"> - Cases: 4,336 SARS-CoV-2 infected hospitalized individuals who died or required respiratory support (intubation, CPAP, BiPAP, continuous external negative pressure, high flow nasal cannula). - Controls: 623,902 without severe COVID-19

Table 1: Sources of data for the analysis. COVID-19 susceptibility and severity outcomes are taken from the COVID-19 HGI[32]. See **Supplement 2** for details on cohorts of COVID-19 susceptibility and severity phenotypes.

Outcome	nSNPs	IVW OR (95% CI)	IVW p-value	IVW SNP Heterogeneity p-value	Egger alpha	Alpha p-value
25OHD primary analysis with all SNPs						
Susceptibility	81	0.97 (0.85, 1.10)	0.61	0.008	0.003 (-0.004, 0.009)	0.43
Hospitalization	81	1.11 (0.91, 1.35)	0.30	0.066	0.0002 (-0.010, 0.011)	0.97
Severe disease	81	0.93 (0.73, 1.17)	0.53	0.126	0.010 (-0.002, 0.022)	0.11
25OHD sensitivity analysis restricted to genes in the vitamin D pathway						
Susceptibility	12	0.96 (0.83, 1.11)	0.59	0.160	0.005 (-0.021, 0.032)	0.70
Hospitalization	12	1.07 (0.78, 1.47)	0.67	0.004	0.031 (-0.027, 0.088)	0.32
Severe disease	12	0.87 (0.64, 1.19)	0.38	0.105	0.056 (0.006, 0.106)	0.05
25OHD sensitivity analysis after removal of SNPs identified by Phenoscanner						
Susceptibility	10	0.97 (0.77, 1.23)	0.80	0.083	0.004 (-0.032, 0.040)	0.83
Hospitalization	10	1.09 (0.68, 1.75)	0.71	0.010	0.014 (-0.059, 0.088)	0.71
Severe disease	10	0.91 (0.54, 1.55)	0.73	0.095	0.072 (-0.003, 0.146)	0.01

Table 2: MR results. SNP: single nucleotide polymorphism. nSNPs: number of SNPs retained for this analysis. IVW: inverse-variance weighted method. CI: confidence interval. Confidence intervals were obtained using Normal approximations, explaining minor discrepancies with p-values close to the alpha=5% statistical significance threshold.

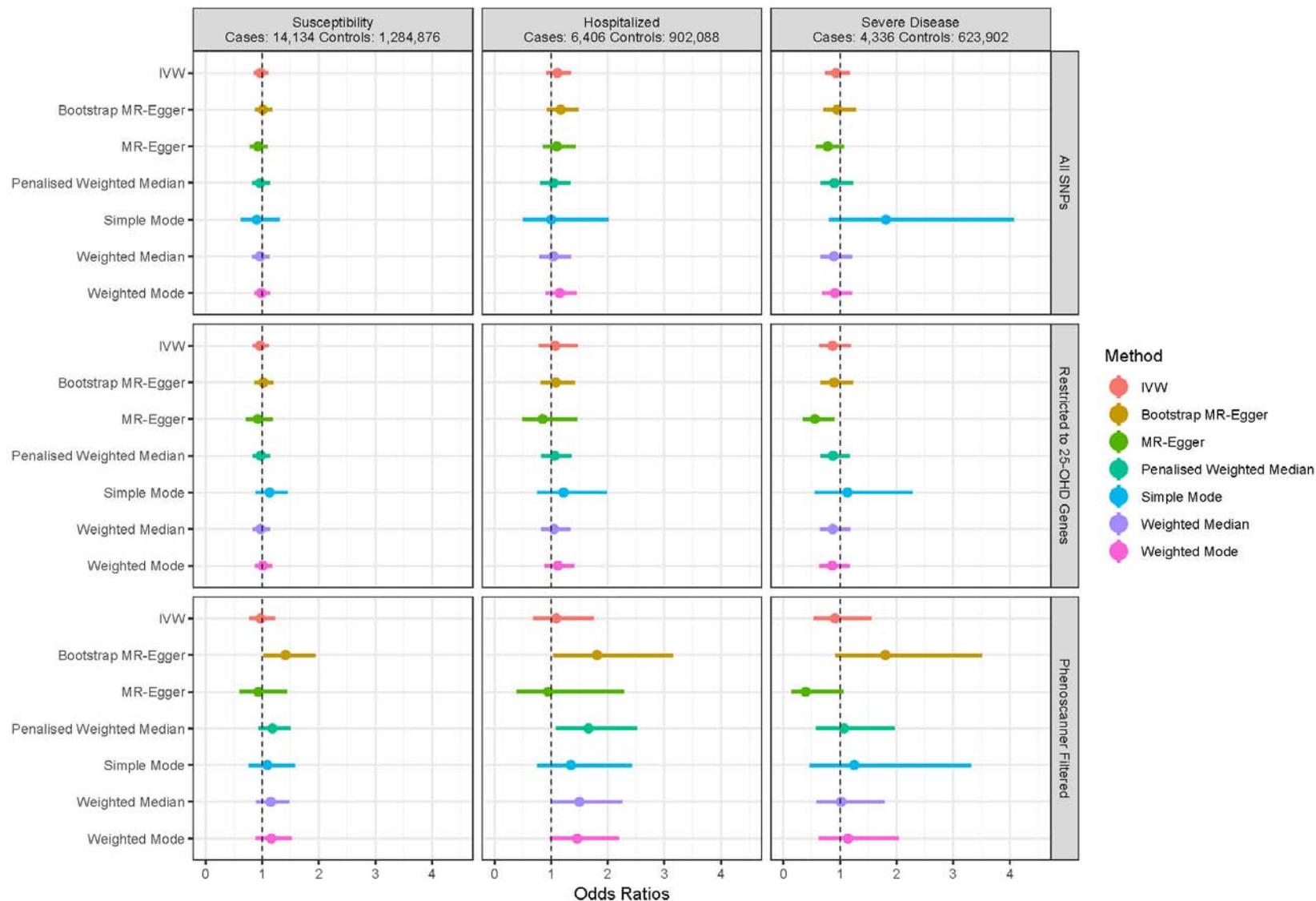


Figure 1: Odds ratio point estimates and 95% confidence intervals for a one standard deviation increase in 25OHD levels (on the log scale) on COVID-19 susceptibility and severity. Restricted to 25-OHD Genes: analysis restricted to SNPs near the 4 genes involved in known vitamin D metabolic pathways. Phenoscaner Filtered: analysis restricted to the 4 genes above, and with removal of SNPs identified to have other associations in Phenoscaner. Full results including odds ratios, confidence intervals, and p-values are available in [Supplement 5](#).