



Exploring Diseases/Traits and Blood Proteins Causally Related to Expression of ACE2, the Putative Receptor of SARS-CoV-2: A Mendelian Randomization Analysis Highlights Tentative Relevance of Diabetes-Related Traits

Shitao Rao,¹ Alexandria Lau,¹ and Hon-Cheong So^{1,2,3,4,5,6}

Diabetes Care 2020;43:1416–1426 | <https://doi.org/10.2337/dc20-0643>

OBJECTIVE

COVID-19 has become a major public health problem. There is good evidence that ACE2 is a receptor for SARS-CoV-2, and high expression of ACE2 may increase susceptibility to infection. We aimed to explore risk factors affecting susceptibility to infection and prioritize drug repositioning candidates, based on Mendelian randomization (MR) studies on ACE2 lung expression.

RESEARCH DESIGN AND METHODS

We conducted a phenome-wide MR study to prioritize diseases/traits and blood proteins causally linked to ACE2 lung expression in GTEx. We also explored drug candidates whose targets overlapped with the top-ranked proteins in MR, as these drugs may alter ACE2 expression and may be clinically relevant.

RESULTS

The most consistent finding was tentative evidence of an association between diabetes-related traits and increased ACE2 expression. Based on one of the largest genome-wide association studies on type 2 diabetes mellitus (T2DM) to date ($N = 898,130$), T2DM was causally linked to raised ACE2 expression ($P = 2.91E-03$; MR-IVW). Significant associations (at nominal level; $P < 0.05$) with ACE2 expression were observed across multiple diabetes data sets and analytic methods for T1DM, T2DM, and related traits including early start of insulin. Other diseases/traits having nominal significant associations with increased expression included inflammatory bowel disease, (estrogen receptor-positive) breast cancer, lung cancer, asthma, smoking, and elevated alanine aminotransferase. We also identified drugs that may target the top-ranked proteins in MR, such as fostamatinib and zinc.

CONCLUSIONS

Our analysis suggested that diabetes and related traits may increase ACE2 expression, which may influence susceptibility to infection (or more severe infection). However, none of these findings withstood rigorous multiple testing corrections (at false discovery rate < 0.05). Proteome-wide MR analyses might help uncover mechanisms underlying ACE2 expression and guide drug repositioning. Further studies are required to verify our findings.

¹School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong

²Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

³KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research of Common Diseases, Kunming Institute of Zoology and The Chinese University of Hong Kong, Shatin, Hong Kong

⁴Department of Psychiatry, The Chinese University of Hong Kong, Shatin, Hong Kong

⁵Margaret K.L. Cheung Research Centre for Management of Parkinsonism, The Chinese University of Hong Kong, Shatin, Hong Kong

⁶Brain and Mind Institute, The Chinese University of Hong Kong, Shatin, Hong Kong

Corresponding author: Hon-Cheong So, hcsso@cuhk.edu.hk

Received 26 March 2020 and accepted 10 May 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.12279131>.

This article is part of a special article collection available at <https://care.diabetesjournals.org/collection/diabetes-and-COVID19>.

S.R. and A.L. contributed equally to this work.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has resulted in a pandemic affecting more than 100 countries worldwide (1–3). More than 2 million confirmed cases have been reported worldwide as of 22 April 2020 (4), while many mild or asymptomatic cases may remain undetected. Considering the severity of the outbreak, it is urgent to seek solutions to control the spread of the disease to susceptible groups and to identify effective treatments. A better understanding of its pathophysiology is also urgently needed.

Notably, recent studies showed that more than one-quarter of confirmed cases had a history of comorbid conditions, such as hypertension, diabetes, cardiovascular disease, and respiratory diseases (2,3,5) (Supplementary Table 1). In addition, the severity of disease is likely to be higher in patients with chronic conditions (2). However, it is unclear whether such comorbidities are *causally* related to increased susceptibility and, if so, what the underlying mechanisms may be. Confounding bias (e.g., by age, sex, comorbidities, medications received, smoking/drinking history, etc.) may lead to spurious associations that preclude conclusions about causality. Establishing causality is important, as this is closely related to the effectiveness of interventions. If a risk factor is causally related to an outcome, then interventions on the risk factor will lead to reduced risks of the outcome, which may not be true for associations per se.

Based on analysis of potential receptor usage and the released sequences of SARS-CoV-2, Wan et al. (6) proposed that the host receptor of SARS-CoV-2 is ACE2. Virus infectivity studies on HeLa cell lines further confirmed that ACE2 is a cellular entry receptor for SARS-CoV-2 (7). Another line of evidence came from structural study of SARS-CoV-2. Wrapp et al. (8) observed that the ACE2 protein could bind to the SARS-CoV-2 spike ectodomain with high affinity. Importantly, ACE2 has previously been established as a receptor for severe acute respiratory syndrome coronavirus (SARS-CoV) (9,10). Taken together, there is strong evidence that ACE2 is a key receptor of the novel coronavirus.

A number of studies have looked into the relationship between ACE2 expression level and coronavirus infection. For example, it was found that overexpression of ACE2 protein leads to more efficient SARS-

CoV replication, which was blocked by anti-ACE2 antibodies in a dose-dependent manner (9). Two further studies also showed that susceptibility to SARS-CoV infection was correlated with ACE2 expression in cell lines (11,12). It is therefore reasonable to hypothesize that ACE expression also affects susceptibility to SARS-CoV-2 infection. Revealing diseases/traits causally associated with altered ACE2 expression may shed light on why certain individuals are more susceptible to SARS-CoV-2 infection (or more severe infections) and the underlying mechanisms (whether the increased susceptibility is mediated via ACE2).

In this study, we wish to answer the following question: what conditions or traits may lead to increased ACE2 expression, which may in turn result in higher susceptibility to SARS-CoV-2 infection? Here, we conducted a phenome-wide Mendelian randomization (MR) study to explore diseases/traits that may be *causally* linked to increased ACE2 lung expression. Our study is different from most existing MR studies: instead of considering a disease as outcome, the outcome measure is ACE2 expression, interpreted as a surrogate for susceptibility to infection, and the exposures tested are diseases/traits. While a number of tissues may also be affected by SARS-CoV-2 (13), pneumonia is a common and major complication of the disease (3); hence, we focused on lung expression in this study. Regarding our study approach, phenome-wide MR is a data-driven approach that has been used in other contexts as a powerful way to uncover unknown causal risk factors for diseases (14–16). This approach allows multiple risk factors or outcomes to be studied simultaneously. MR makes use of genetic variants as “instruments” to represent the exposure of interest and infers causal relationship between the exposure and outcome (17). In general, MR is not affected by reverse causality (18), as genetic variants are fixed at conception (which precedes the outcome). MR is also less susceptible to confounding bias compared with conventional case-control/cohort studies, as genetic instruments are usually less strongly associated with environmental exposures than ordinary risk factors (19) (please also refer to Supplementary Text for more detailed descriptions).

In addition to diseases, as a secondary analysis we also studied serum/plasma

proteins as exposure, as they may point to potential molecular mechanisms underlying ACE2 expression and may serve as potential predictive or prognostic biomarkers. Such proteome-wide studies may help to reveal drug repositioning candidates (20) through the search for drugs that target the top-ranked proteins. For example, if a protein causally increases the risk of a disease, then by the definition of causality, blocking the protein will lead to reduced disease risks. By finding plasma/serum proteins *causally* linked to ACE2 expression, one may find drugs altering ACE2 expression, which in turn may be useful for treatment.

RESEARCH DESIGN AND METHODS

Genome-Wide Association Study Data

All genome-wide association study (GWAS) data are extracted from publicly available databases, detailed below.

Exposure Data

Most GWAS data used were based on predominantly European samples, and proper correction for population stratification has been performed. Please also refer to Supplementary Tables 2A and 2B for details on the ethnic composition and methods to account for population stratification for GWAS included in this work.

To perform the phenome-wide study, here we made use of the latest MRC Integrative Epidemiology Unit (IEU) (University of Bristol) GWAS database (<https://gwas.mrcieu.ac.uk/>), which contains up to 111,908,636,549 genetic associations from 31,773 GWAS summary data sets (as at 26 February 2020). Details of each GWAS study may be retrieved from <https://gwas.mrcieu.ac.uk/datasets/>. The database was retrieved via the R package TwoSampleMR (version 0.5.1). MR analysis was conducted with the same package. Due to the extremely huge number of traits in the database, we performed some preselection to the list of traits/diseases before full analysis.

Briefly, we selected the following categories of traits: 1) traits listed as priority 1 (high priority) and labeled as “disease” or “risk factor” (81 and 71 items, respectively), 2) traits labeled as “protein” (3,371 items originally studied in 21,22); and 3) (selected) traits from the UK Biobank (UKBB), as it is one of the largest sources of GWAS data worldwide ($N = \sim 500,000$). We considered that a proportion of traits have presumably low prior probability of association with respiratory infections, and

others are less directly clinically relevant. For reduction of computational burden and for ease of interpretation, a proportion of UKBB traits were filtered. More specifically, we excluded GWAS data of diseases or traits related to the following: eye or hearing problems, orthopedic and trauma-related conditions (except autoimmune diseases), skin problems (except systemic or autoimmune diseases), perinatal and obstetric problems, operation history, medication history (as confounding by indication is common and may affect the validity of results [23]), diet/exercise habits (as accuracy of information cannot be fully guaranteed and recall bias may be present), and socioeconomic features (such as type of jobs). A total of 425 UKBB traits were retained for final analysis under the third category. GWAS of blood proteins and UKBB traits were restricted to European samples.

GWAS of UKBB were based on analysis results from the Neale laboratory (<https://www.nealelab.is/uk-biobank>) and from MRC IEU. GWAS analysis was performed using linear models with adjustment for population stratification (details of the analytic approach: references 24–26). For binary outcomes, we converted the regression coefficients obtained from the linear model to those under a logistic model, based on methodology previously presented (27). The SE under a logistic model was derived by the delta method (equation 37 in reference 27).

Outcome Data

The outcome was pulmonary expression of ACE2. While ideally one should study the protein expression in the lung, such data are scarce and corresponding genotype data (required for MR) are not available. Here we focus on the gene expression of ACE2 in the lung ($N = 515$). We retrieved GWAS summary data from the Genotype-Tissue Expression (GTEx) database (with API); it is one of the largest databases to date with both genotype and expression data for a large variety of tissues. The majority of the GTEx samples are European in ancestry (~85%); other ancestries included African Americans, Asians, and American Indians (Supplementary Table 2A). Population stratification was controlled by inclusion of principal components in genetic association analysis. For further details of GTEx please refer to reference 28; the expression quantitative trait loci (eQTL) analysis procedure is described in reference 29.

MR Analysis

Here we performed two-sample MR, in which the instrument-exposure and instrument-outcome associations were estimated in different samples.

Instrument Single Nucleotide Polymorphism Selection

MR was performed on (approximately) independent single nucleotide polymorphisms (SNPs) with r^2 threshold of 0.001, following default settings in the R package TwoSampleMR. SNPs passing genome-wide significance ($P < 5e-8$) were included as instruments. Clinical traits or blood proteins were treated as exposures, and we used the “extract_instruments” function in TwoSampleMR to retrieve SNPs for each trait from corresponding GWAS. The source GWAS for each exposure are listed in Supplementary Table 2. Only SNPs with available SNP-exposure and SNP-outcome association data were retained.

MR Methods

We conducted MR primarily with the inverse variance-weighted (MR-IVW) (30) and Egger regression (MR-Egger) (31) approaches, which are among the most widely used MR methods. For exposure with only one instrument, the Wald ratio method was used. For analysis with fewer than three genetic instruments, we used MR-IVW only since MR-Egger cannot reliably be performed. The intercept from MR-Egger was used to evaluate presence of significant directional (imbalanced) horizontal pleiotropy.

For selected traits with at least nominally significant associations by MR-IVW or MR-Egger ($P < 0.05$), we also performed further analysis by GSMR (generalized summary data-based MR), weighted median (an “implicit” outlier-removal method [32]), and MR robust adjusted profile score (MR-RAPS). GSMR also accounts for correlated SNPs and removes likely pleiotropic outliers (33).

We tried several r^2 thresholds (0.001, 0.05, 0.1, 0.15, and 0.2) for GSMR analysis on diabetes based on the work of Mahajan et al. (34) (see RESULTS and Table 2). SNP correlations were derived from 1000 Genomes European samples. MR-RAPS (35) is another methodology that takes into account multiple weak instruments by a robust procedure; we used a more relaxed P value threshold for SNP selection (0.01) for this method. One of the major concerns of MR is horizontal pleiotropy, in which the

genetic instruments have effects on the outcome other than through effects on the exposure. MR-Egger, GSMR, weighted median, and MR-RAPS are able to provide valid MR estimates under pleiotropy subject to certain assumptions (see Hemani et al. [32] and Supplementary Text).

Heterogeneity among the MR estimates across individual SNPs may indicate problems related to violation of instrumental variable assumptions. One of the most notable problems is that one or more SNPs may be showing horizontal pleiotropy (32,36). The Cochran Q statistic and the MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) global test (37) were used to test for heterogeneity for nominally significant MR findings.

Interpretation of Effect Sizes From MR

Regarding the effect sizes of causal associations, if the exposures were binary, the regression coefficients (β) from MR may be roughly interpreted as average change in the outcome (per SD increase in normalized ACE2 expression levels) per 2.72-fold increase in the prevalence of the exposure (38). For continuous exposures, the MR estimates are average changes in outcome per unit increase of exposure (see Supplementary Table 2A for the units).

Plasma/Serum Proteins as Exposure and Further Analysis

In addition to MR analysis on individual plasma/serum proteins, we also performed pathway analysis by ClueGO (39). Hypergeometric tests were conducted on the top-ranked proteins (with $P < 0.05$). As an exploratory analysis, we also searched for drugs with targets overlapping with the top-ranked proteins. Drug targets were defined based on the DrugBank database. Our aim is to uncover drug candidates leading to alteration of ACE2 expression, which may be therapeutically relevant.

Multiple Testing Correction

We employed a false discovery rate (FDR) approach to multiple testing correction. It controls the expected proportion of false positives among the hypotheses declared significant. FDR is also valid under positive dependency of tests (40).

The FDR in fact depends on the overall fraction of truly null hypotheses, or π_0 . It can also be considered as the prior probability that a null hypothesis is true. In reality, π_0 may vary for different subgroups of hypotheses. For instance, in our analyses, one may expect different

π_0 for diseases/exposures of different kinds. Previous studies (see Supplementary Table 1) suggested that some chronic disease patients are more likely affected by the infection. To address the above problem, we adopted an FDR control procedure that accounts for varying prior probabilities of association (i.e., different π_0) among different types of hypotheses. The procedure is “objective” in the sense that it estimates π_0 based on the data automatically, without the need to specify π_0 by the researcher. We used the methodology “FDR regression” proposed in 41 and the R program by the author (FDRreg, version 0.2). In brief, we divided our hypothesis based on the type of exposure/disease (e.g., respiratory, cardiovascular diseases, etc.). These categories served as predictors or covariates, which can be used as input by FDRreg in a regression to estimate the π_0 of each hypothesis test. We also computed the significance of each predictor; it indicates which categories predicted nonnull associations better than chance. For input into FDRreg, we took the results from MR-IVW unless the Egger intercept had $P < 0.05$.

RESULTS

MR Analysis for Diseases and Clinically Relevant Traits

MR results are presented in Tables 1 and 2 (full results shown in Supplementary Tables 3 and 4). Traits were shown in main tables if MR-IVW or MR-Egger showed nominally significant ($P < 0.05$) results and three or more instrument SNPs are available (such that pleiotropy can be assessed and results are more informative).

Overall, 25 traits showed associations with *ACE2* expression at FDR < 0.2 and 10 had FDR < 0.1 (Supplementary Table 4). No MR results showed FDR < 0.05 . There were 68 nominally significant ($P < 0.05$) associations based on MR-IVW and 9 based on MR-Egger. Many significant findings were concentrated on traits related to diabetes.

Diabetes-Related Traits

Remarkably, a number of top-ranked results were related to diabetes. We observed five diabetes-related traits that showed nominally significant MR results with FDR < 0.1 ; they were all positively associated with *ACE2* expression. Three are related to diagnosis of diabetes (including both type 1 and 2) in the UKBB. Both doctor-diagnosed diabetes and self-

reported cases of diabetes in the UKBB, which were presumably comprised of mainly type 2 diabetes mellitus (T2DM), were significantly associated with higher *ACE2* expression (MR-IVW $P = 0.0152$ and 0.0343 ; FDR = 0.0547 and 0.0667 respectively). Another finding (identifier: ieu-a-23) was based on a transethnic meta-analysis on T2DM in 2014 (42) (MR-IVW $P = 0.0421$; FDR = 0.0748), which had no overlap with the UKBB sample. The finding of a (nominally) significant result in this data set can therefore be considered as an independent replication of the UKBB result.

We also observed that starting insulin within 1 year of diagnosis, which was only assessed among patients with diabetes, was causally associated with increased *ACE2* expression (MR-IVW $P = 0.031$; FDR = 0.061). Early use of insulin may indicate type 1 diabetes mellitus (T1DM) as the underlying diagnosis or more severe/late-stage disease for T2DM patients (43). We also observed that as a whole, diabetes-related traits were significantly associated with higher probability of having nonnull associations with *ACE2* expression ($P = 0.026$) (Supplementary Table 7), based on FDRreg. No evidence of significant directional pleiotropy was observed in the above results (Egger intercept $P > 0.05$). We therefore primarily reported the results from MR-IVW, as generally the SE of causal estimates is larger with MR-Egger (44) (resulting in weaker power).

In view of the consistent causal associations with diabetes or related traits, we further searched for GWAS summary statistics that have not been included in the IEU GWAS database. We found another publicly available data set from the DIABetes Genetics Replication And Meta-analysis (DIAGRAM) consortium, based on a recent meta-analysis of T2DM by Mahajan et al. (34) based on European samples ($N = 898,130$). For a more in-depth analysis, we also used GSMR at various r^2 thresholds and MR-RAPS in addition to IVW and MR-Egger. The full results are presented in Table 2 (also see Supplementary Figures). Reassuringly, with the exception of MR-Egger (which is less powerful [44]), all other methods showed (at least nominally) significant results. GSMR, which accounts for correlated SNPs, showed significant results consistently across different r^2 thresholds (lowest $P = 9.74E-18$; r^2 threshold = 0.2). While this study (34) has

partial overlap with the transethnic analysis in 2014 (42), the consistent associations provide further support to a causal link between diabetes and expression of *ACE2*.

We note that the Egger intercept P value was borderline ($P = 0.0545$), which may raise some concern for pleiotropy. However, we have conducted multiple tests for directional pleiotropy, so false positive findings are possible. The corresponding FDR was 0.999 for this test if multiple testing was taken into account (573 items).

We did not find any evidence of heterogeneity based on Cochran Q (heterogeneity $P_{IVW} = 0.431/P_{Egger} = 0.486$) or MR-PRESSO global test ($P = 0.418$). To further compare MR-IVW and MR-Egger models, we followed the “Rucker framework” proposed in (32,45) and computed the improvement in model heterogeneity by using MR-Egger. The difference was small and nonsignificant ($Q_{IVW} = 197.77$; $Q_{Egger} = 196.06$; difference = 1.71 ; $P = 0.191$), indicating MR-IVW is a reasonably good fit for the data.

For T2DM or self-reported cases of diabetes from UKBB (which presumably comprised mainly T2DM), the causal estimates ranged from ~ 0.162 to 0.210 . The causal estimate from T1DM was slightly lower and estimated to be ~ 0.1006 .

Other Diseases/Traits

As shown in Table 1, a number of other diseases/traits also showed (nominally) significant results. Several neoplasms, such as breast and lung cancer, may be associated with increased *ACE2* expression. We also observed that several autoimmune disorders, especially inflammatory bowel diseases, may be causally associated with *ACE2* expression. Interestingly, asthma and tobacco use also showed nominal significant associations with higher *ACE2* expression. As for other traits, high alanine aminotransferase (ALT), commonly associated with liver diseases, may be related to elevated *ACE2* expression. Other commonly measured blood measures that may lead to altered *ACE2* expression included red cell distribution width (often associated with iron deficiency, folate, or vitamin B₁₂ deficiency anemia), basophil percentage (inverse relationship), calcium level, urate level, and HDL and LDL cholesterol (inverse relationship). Note that the FDR is dependent on the category to which a trait belongs; for example, diabetes-related

Table 1—Overall MR analysis results achieving nominal significance ($P < 0.05$), with diseases/traits as exposure and ACE2 lung expression as outcome

Identifier	Trait	nsnps	b _{IVW}	P _{IVW}	b _{Egger}	P _{Egger}	Egger intercept	P _{intercept}	b _{median}	P _{median}	b _{GSMR}	P _{GSMR}	FDR	P _{het} IVW	P _{het} Egger	P _{global} PRESSO
Diseases as exposure																
Diabetes related																
ukb-b-10753	Diabetes diagnosed by doctor	58	0.175	0.015	0.002	0.988	0.020	0.225	0.146	0.183	0.171	0.017	0.055	0.350	0.379	0.304
ukb-b-12948	Noncancer illness code, self-reported: diabetes	49	0.162	0.034	0.203	0.255	−0.005	0.797	0.126	0.280	0.158	0.034	0.067	0.322	0.288	0.306
ukb-b-10694	Diagnosis, secondary, ICD-10 E10.9, T1DM without complications	3	0.101	0.042	0.184	0.428	−0.052	0.654	0.110	0.033	—	—	0.071	0.851	0.659	—
ieu-a-23	T2DM	25	0.210	0.042	0.160	0.696	0.005	0.899	0.261	0.063	0.187	0.076	0.075	0.849	0.811	0.859
ukb-b-8388	Started insulin within 1 year diagnosis of diabetes	7	0.076	0.031	0.077	0.352	−0.001	0.991	0.084	0.041	—	—	0.061	0.988	0.967	0.990
Neoplasms																
ukb-d-D12	Diagnosis, main, ICD-10 D12, benign neoplasm of colon, rectum, anus, and anal canal	12	−0.286	0.015	−0.058	0.904	−0.031	0.621	−0.314	0.047	−0.331	0.007	0.930	0.865	0.815	0.873
ukb-d-C3	Malignant neoplasm of respiratory system and intrathoracic organs	3	0.514	0.009	0.711	0.511	−0.056	0.821	0.416	0.047	—	—	0.509	0.066	0.022	—
ieu-a-1134	ER ⁺ breast cancer	7	0.176	0.020	0.539	0.128	−0.094	0.259	0.122	0.224	—	—	0.272	0.547	0.645	0.497
ieu-a-1013	Glioma	3	0.200	0.037	0.438	0.659	−0.071	0.800	0.205	0.081	—	—	0.502	0.942	0.910	—
Autoimmune disorders																
ukb-b-18194	Noncancer illness code, self-reported: ankylosing spondylitis	3	0.087	0.016	0.017	0.991	0.059	0.964	0.095	0.023	—	—	0.112	0.944	0.735	—
ieu-a-32	Ulcerative colitis	29	0.026	0.667	0.586	0.003	−0.105	0.003	0.099	0.240	0.026	0.626	0.054	0.171	0.629	0.165
ieu-a-292	Inflammatory bowel disease	107	0.003	0.938	0.246	0.019	−0.031	0.011	0.023	0.728	0.010	0.827	0.126	0.724	0.846	0.700
ieu-a-30	Crohn disease	41	0.011	0.786	0.196	0.035	−0.044	0.027	0.055	0.352	0.011	0.781	0.159	0.651	0.826	0.623
ieu-a-31	Inflammatory bowel disease	49	−0.014	0.783	0.287	0.036	−0.051	0.019	0.024	0.742	−0.011	0.814	0.170	0.724	0.846	0.700
Other diseases																
ukb-b-17219	Diagnosis, secondary, ICD-10 J45.9, asthma, unspecified	19	0.264	0.035	0.877	0.056	−0.064	0.152	0.295	0.092	0.269	0.035	0.500	0.769	0.826	0.770
ukb-b-5115	Diagnosis, secondary, ICD-10 Z72.0, tobacco use	3	0.918	0.016	0.624	0.856	0.024	0.930	0.915	0.059	—	—	0.510	0.415	0.187	—
ukb-d-19	Diseases of veins, lymphatic vessels, and lymph nodes, not elsewhere classified	14	−0.269	0.024	−0.138	0.656	−0.017	0.645	−0.174	0.265	−0.264	0.028	0.934	0.886	0.858	0.900
Other risk factors or clinically relevant traits as exposure																
ukb-d-30620_raw	ALT (units/L)	91	0.047	0.007	0.081	0.041	−0.012	0.341	0.069	0.015	0.051	0.004	0.223	0.717	0.716	0.722
ukb-d-30070_irmt	Red blood cell (erythrocyte) distribution width (SD)	225	0.243	0.020	0.349	0.066	−0.004	0.500	0.208	0.191	0.239	0.023	0.555	0.524	0.514	0.537
ukb-d-30220_irmt	Basophil percentage (SD)	77	−0.504	0.027	−0.584	0.169	0.003	0.821	−0.473	0.179	−0.468	0.045	0.935	0.535	0.504	0.558
ukb-d-30780_raw	LDL direct (mmol/L)	126	−0.272	0.105	−0.758	0.005	0.020	0.021	−0.469	0.076	−0.179	0.299	0.934	0.520	0.633	0.541
ukb-d-30680_raw	Calcium (mmol/L)	152	0.202	0.915	−10.380	0.011	0.031	0.003	−1.675	0.545	0.626	0.728	0.935	0.135	0.259	0.136

Continued on p. 1421

Table 1—Continued

Identifier	Trait	nsnps	b_{IVW}	P_{IVW}	b_{Egger}	P_{Egger}	Egger intercept	$P_{intercept}$	b_{median}	P_{median}	b_{GSMR}	P_{GSMR}	FDR	$P_{het_{IVW}}$	$P_{het_{Egger}}$	P_{global_PRESSO}
ukb-a-30830_raw	SHBG (mmol/L)	163	0.004	0.443	0.022	0.042	−0.016	0.055	0.012	0.166	0.004	0.441	0.934	0.219	0.267	0.209
ieu-a-793	Urate (mg/dL)	4	0.026	0.027	0.026	0.167	−0.008	0.884	0.028	0.037	—	—	0.608	0.792	0.603	0.382
ieu-a-1034	Height (SD)	4	0.636	0.047	2.060	0.806	−0.118	0.865	0.599	0.124	—	—	0.553	0.647	0.445	0.643
ieu-a-299	HDL cholesterol (SD)	84	0.084	0.515	0.563	0.022	−0.026	0.020	0.065	0.758	0.086	0.508	0.581	0.575	0.714	0.592

Some items are missing, as the number of SNPs is insufficient. FDR refers to the P value from MR-IVW (if Egger intercept $P > 0.05$) or MR-Egger. Values of FDR < 0.1 are in boldface type. b , β (causal estimate); ER, estrogen receptor; FDR, derived from FDR regression; median, weighted median approach; nsnps, number of SNPs; P_{global_PRESSO} , P value from the global test of MR-PRESSO (used to assess heterogeneity of MR estimates); P_{het} , heterogeneity P value; SHBG, sex hormone-binding globulin.

Table 2—Further MR analysis results for T2DM based on work by Mahajan et al. (34)

Method	β	SE	P	Egger intercept	Intercept P	n_pleio	nsnps
MR-IVW	0.177	0.060	2.91E-03	—	—	—	196
MR-Egger	-0.039	0.126	0.758	0.0159	0.0545	—	196
GSMR†							
$r^2 = 0.001$	0.170	0.060	4.46E-03	—	—	0	194
$r^2 = 0.05$	0.140	0.035	7.12E-05	—	—	0	289
$r^2 = 0.1$	0.177	0.054	9.62E-04	—	—	0	332
$r^2 = 0.15$	0.146	0.027	5.93E-08	—	—	0	392
$r^2 = 0.2$	0.197	0.023	9.74E-18	—	—	0	448
MR-RAPS§	0.064	0.030	3.43E-02	—	—	—	3,737

We did not find any evidence of heterogeneity based on Cochran Q (heterogeneity $P_{IVW} = 0.431$ / $P_{Egger} = 0.486$) or MR-PRESSO global test ($P = 0.418$). We also computed the improvement in model heterogeneity by using MR-Egger over IVW following the Rucker framework. The difference was small and nonsignificant ($Q_{IVW} = 197.77$; $Q_{Egger} = 196.06$; difference = 1.71; $P = 0.191$). The exposure GWAS data set on T2DM was based on work of Mahajan et al. (34). Instrument SNPs were only selected if they passed genome-wide significance ($P < 5e-8$) (except for MR-RAPS). If not otherwise specified, SNPs were clumped at $r^2 = 0.001$. n_pleio, number of pleiotropic SNPs identified by GSMR; nsnps, number of SNPs. †GSMR can account for correlation among SNPs. We performed GSMR based on SNPs clumped at different r^2 clumping thresholds. We consider the association to be more robust if significant results are observed across multiple r^2 thresholds. §MR-RAPS is an MR methodology designed for the inclusion of multiple weak instruments. A more relaxed P value threshold (0.01) was used for SNP instrument selection.

and autoimmune diseases showed lower FDR, likely because these types of diseases had more significant associations in general. As a tradeoff, other traits/diseases, although having nominally significant results, may have higher FDR. FDR provides an additional reference to guide prioritization of the findings; however, FDR estimation is subject to variability and should not be considered as an absolute guide. Other traits with at least nominal significance may still be worthy of further studies, especially with support by clinical observation or other evidence.

For traits showing nominally significant findings (Table 1), we have performed other additional analyses. We do not observe significant heterogeneity in MR estimates across SNPs (by IVW/Egger) for most traits, except one related to lung cancer (ukb-d-C3). The MR-PRESSO global test was also nonsignificant for all traits, supporting a lack of heterogeneity. This lack of heterogeneity suggests that substantial horizontal pleiotropy is not very likely. The weighted median estimator supports associations for a subset of traits, including three diabetes-related traits (ukb-b-10694, ieu-a-23, and ukb-b-8388). The GSMR method, which removes pleiotropic outliers, is generally consistent with IVW findings (SNPs clumped at $r^2 = 0.001$ for both GSMR/IVW).

MR Results With Plasma/Serum Proteins as Exposure

Full results are shown in Supplementary Tables 3 and 4, and the enriched pathways

are shown in Table 3 and Supplementary Table 5. Since a large number of proteins are involved, we only highlight a few top pathways here. Some of the top pathways include cytokine-and-cytokine receptor interaction, VEGFA-VEGF2 signaling pathway, JAS-STAT signaling pathway, etc. Table 4 and Supplementary Table 6 show the list of drugs with targets that overlap with the top-ranked proteins. Note that the tables do not explicitly discern the direction of effects of the drugs. A few drugs target more than one protein. If they are ranked by the number of proteins targeted, the top drugs are fostamatinib, copper, zinc, and zonisamide, which target three or more proteins.

CONCLUSIONS

In this study, we have used MR to uncover diseases/traits that may be causally linked to ACE2 expression in the lung, which in turn may influence susceptibility to the infection. MR is a relatively well-established technique in evaluating causal relationships, and the wide availability of GWAS data enables many different exposures to be studied at the same time.

Diseases/Traits Causally Linked to ACE2 Expression

From our analysis, the most consistent finding was the tentative causal link between diabetes (and related traits) with ACE2 expression, which was supported by multiple data sets and different analytic approaches. Other results were more tentative but may be worthy of further

studies. For example, several neoplasms (e.g., breast and lung cancers) and autoimmune diseases, elevated ALT, asthma, and smoking all showed nominally significant and positive associations with ACE2 expression.

Some of these findings were supported by previous studies. A number of COVID-19 cases (~5.4% from Supplementary Table 1) were comorbid with diabetes. This proportion is only a rough estimate, since mild or asymptomatic cases may remain undetected. Notably, diabetes has been reported to be associated with poorer outcomes among infected patients (46). Similarly, diabetes was also common in patients infected with MERS-CoV (47,48). Kulcsar et al. (49) built a mouse model susceptible to MERS-CoV infection and induced T2DM using a high-fat diet. They found that, if affected by the virus, these diabetic mice suffered from a prolonged phase of disease and delayed recovery, possibly due to a dysregulated immune response. Regarding comorbidity with cancers, Liang et al. (50) recently carried out a nationwide analysis of 1,590 patients with confirmed COVID-19 and suggested that patients with cancer have higher infection and complication risks than those without cancer.

We highlight a few research directions of interest if our findings are confirmed in future studies. For example, as far as treatment is concerned, if certain conditions (e.g., diabetes) increase susceptibility to infection or severe infections via ACE2, drugs targeting this gene/protein may be particularly useful for this patient subgroup. For example, human recombinant ACE2 has been proposed as a treatment and is under clinical trial (51,52). It will be interesting to see if the drug may be more beneficial for patients with diabetes. More generally speaking, if diabetes is causally linked to elevated ACE2 and potentially increased susceptibility to infection, then antidiabetes drugs or improved glycemic control may ameliorate the process. Interestingly, a recent work highlighted metformin as one of the top repositioning candidates for COVID-19, based on a different mechanism as an MRC1 inhibitor (53). From a public health perspective, identification of at-risk populations may guide prevention strategies, e.g., prioritization of groups to receive vaccination. Nevertheless, all the above require substantial additional research before clinical applications.

Table 3—Top 10 enriched pathways for (nominally) significant proteins in MR analysis

GO ID	GO term	Ontology source	Term <i>P</i>	Term <i>P</i> (Bonf)	Associated Genes
KEGG:04060	Cytokine-cytokine receptor interaction	KEGG_27.02.2019	1.82E−06	7.29E−05	<i>CCL25, CTF1, CX3CL1, CXCL12, IL15RA, IL22, IL34, IL37, LTA, LTBR, OSM, TNFSF4, TNFSF8</i>
WP:3888	VEGFA-VEGFR2 signaling pathway	WikiPathways_27.02.2019	7.79E−06	3.11E−04	<i>ACACB, BCL2L1, CFL1, EEA1, F3, IGFBP7, JAG1, KDR, PIK3CA, PTPN1, TXNIP</i>
WP:254	Apoptosis	WikiPathways_27.02.2019	1.07E−04	4.29E−03	<i>BCL2L1, BIRC5, DIABLO, IGF1, LTA, MCL1</i>
WP:3614	Photodynamic therapy-induced HIF-1 survival signaling	WikiPathways_27.02.2019	2.99E−04	1.19E−02	<i>BCL2L1, BIRC5, HK1, MCL1</i>
R-HSA:399954	Sema3A PAK dependent axon repulsion	REACTOME_Pathways_27.02.2019	3.40E−04	1.36E−02	<i>CFL1, PAK3, PLXNA1</i>
R-HSA:2173782	Binding and uptake of ligands by scavenger receptors	REACTOME_Pathways_27.02.2019	4.89E−04	1.96E−02	<i>FTH1, HP, STAB1, STAB2</i>
KEGG:04630	JAK-STAT signaling pathway	KEGG_27.02.2019	5.61E−04	2.24E−02	<i>BCL2L1, CTF1, IL15RA, IL22, MCL1, OSM, PIK3CA</i>
KEGG:04672	Intestinal immune network for IgA production	KEGG_27.02.2019	8.84E−04	3.53E−02	<i>CCL25, CXCL12, IL15RA, LTBR</i>
WP:3657	Hematopoietic stem cell gene regulation by GABP alpha/beta complex	WikiPathways_27.02.2019	9.00E−04	3.60E−02	<i>BCL2L1, FLT3, MCL1</i>
WP:3872	Regulation of apoptosis by parathyroid hormone-related protein	WikiPathways_27.02.2019	0.00090	0.03602	<i>BCL2L1, MCL1, PIK3CG</i>

Bonf, Bonferroni correction; GO, Gene Ontology; ID, identifier; Term *P*, *P* value for the GO term.

On ACE2 Expression and Pulmonary Complications

As discussed above, increased expression of *ACE2* appears to correlate with susceptibility to SARS-CoV and SARS-CoV-2 infection. Nevertheless, the consequences of altered *ACE2* expression on pulmonary complications may be rather

complex. Kuba et al. (10) reported that SARS-CoV downmodulated *ACE2* expression, which may lead to heightened risks of acute lung injury (ALI). Another study (54) suggested that *ACE2* may protect against ALI by blocking the renin-angiotensin pathway. However, whether the same applies to SARS-CoV-2 is unknown. If this is the

case, one may hypothesize that for unaffected individuals or those without (or with minimal) lung involvement yet, lower *ACE2* pulmonary expression may be beneficial in reducing susceptibility to more sustained infection by reducing viral entry. However, for patients with severe lung involvement or at risk for ALI, higher *ACE2* expression may prevent acute respiratory failure. Therefore, it may be clinically relevant to identify both risk factors and drugs leading to increased and decreased *ACE2* expression. Further studies are warranted to clarify the role of *ACE2* in COVID-19 and related complications.

Another related controversy concerns the use of ACE inhibitors (ACEI) and angiotensin II receptor blockers (ARB) (55,56), although the current study does not directly address this issue. There is some evidence that ACEI/ARB may upregulate *ACE2* expression in the heart (57), kidney (58), and aorta (59) in animal models; however, how these drugs affect pulmonary *ACE2* levels in humans is still unclear (60). In addition, it is possible that patients' other underlying conditions may affect *ACE2* expression. It is worthy to further investigate how ACEI/ARB together with other chronic conditions affects the risks and severity of infection.

Table 4—Drugs with targets overlapping with (nominally) significant proteins from MR analysis

Drug	No. of proteins targeted	Targets (that overlap with proteins with at least significance in MR analysis)						
Fostamatinib	7	ZAP70	FLT3	HIPK3	KDR	MST1R	PAK3	PIK3CG
Copper	6	CFL1	S100A2	PARK7	AHSG	APOD	CBX5	
Zinc	4	S100A2	AHSG	C8A	APLP2			
Zonisamide	3	CA4	CA9	CA10				
Benzthiazide	2	CA4	CA9					
Hyaluronic acid	2	LAYN	STAB2					
Hydroflumethiazide	2	CA4	CA9					
Isosorbide	2	BCL2L1	MCL1					
Midostaurin	2	KDR	FLT3					
Nintedanib	2	KDR	FLT3					
Ponatinib	2	FLT3	KDR					
Sodium carbonate	2	CA4	CA9					
Sorafenib	2	KDR	FLT3					
Sunitinib	2	KDR	FLT3					

Direction and magnitude of the drugs' effects on *ACE2* expression cannot be determined from our analysis alone and hence are not indicated here.

Highlight of Tentative Repositioning Candidates Based on Blood Proteins Potentially Linked to ACE2 Expression

The drugs we highlighted in this study may help researchers to prioritize repositioning candidates for further studies, given the huge cost and long time frame in developing a brand-new drug. Nevertheless, the overall direction and magnitude of effect of each drug could not be determined from our analysis alone; hence, further studies are required. Here we briefly highlight a few top candidates. Fostamatinib targets the largest number (seven) of proteins potentially linked to ACE2 expression. According to DrugBank, it serves as an inhibitor for all these proteins, and all were linked to elevated ACE2 expression except one. Interestingly, a recent computational repositioning study (61,62) identified baricitinib, a JAK1/2 and AAK1 inhibitor approved for rheumatoid arthritis as a top candidate. Fostamatinib is a spleen tyrosine kinase inhibitor but also inhibits JAK1/2 and AAK1 (from DrugBank) (63) and can be used to treat rheumatoid arthritis (64). JAK-STAT signaling was also among the top 10 pathways enriched for top proteins affecting ACE2 expression. Interestingly, fostamatinib was reported to be effective for T1DM (65). Another candidate, highlighted in 61, sunitinib, was also top listed by our analysis. Zinc is also a top-listed candidate and was previously reported to reduce risks of lower respiratory tract infections (66), but the evidence is not firm. Interestingly, a study in rat tissues showed reduction of ACE2 activity by zinc (67). Zinc and zinc-ionophores may inhibit SARS-CoV as shown in experimental studies (68). Zinc was recently suggested for clinical trials for COVID-19, although there is no clinical evidence yet (ClinicalTrials.gov, NCT04342728, NCT04326725, and NCT04351490 [69]). As for the enriched pathways for top-ranked proteins affecting ACE2 expression, they are discussed in Supplementary Text.

Limitations

We wish to emphasize that we consider this work as largely an exploratory rather than confirmatory study. As such, the findings might not be immediately applicable clinically. Our main purpose is to prioritize diseases, traits, or proteins with potential causal links with ACE2 expression. There are several limitations in our analysis. A major limitation is that the sample size for GTEx is relatively modest ($N = 515$), which limits the power of MR

analysis. However, to our knowledge, GTEx is one of the largest databases with both genotype and expression data. We note that many associations were relatively modest, with no results showing $FDR < 0.05$, although 25 had $FDR < 0.2$. On the other hand, we examined the consistency of the observed associations across different data sets and considered those supported by more than one set of data (e.g., diabetes-related traits) as relatively more robust—similar to the approach in 70. However, our findings will require further support by further studies. Besides, some results could be false negatives owing to limited power. Also, while most GWAS were based on predominantly European samples, subjects of other ethnicities were included in some samples. It is possible for genetic associations to differ across ethnicities, which may affect the causal estimates of MR, e.g., if some SNP-exposure or SNP-outcome associations are stronger in one ethnic group than another. Apart from the above, this study does not address what factors may aggravate or ameliorate coronavirus-induced changes in ACE2 levels. Also, we studied ACE2 mRNA expression as the outcome; associations of the reported traits with protein expression levels remain to be investigated.

Finally, from a methodological point of view, we have used MR in a manner different from that of most other studies. Usually MR is used to identify causal risk factors with a disease as the outcome, for which GWAS data are available. Here, we presented a novel analytic approach: we made use of existing knowledge of a key receptor of an infectious agent to uncover causal risk factors and repositioning candidates. This analytic framework may also be applied to other diseases, especially when a target can be identified but genomic data for the disease is limited or if one is interested in the underlying disease mechanism of the risk factor.

Conclusion

Notwithstanding the limitations, we have identified several diseases and traits that may be causally related to ACE2 expression in the lung, which in turn may mediate susceptibility to SARS-CoV-2 infection. In addition, our proteome-wide MR analysis revealed proteins that may lead to changes in ACE2 expression. Subsequent drug repositioning analysis highlighted several candidates that may

warrant further investigations. We stress that most of the findings require validation in further studies, especially the part on repositioning. Nevertheless, we believe this work is of value in view of the urgency to address the outbreak of COVID-19.

Acknowledgments. The authors thank Prof. Stephen Tsui (School of Biomedical Sciences, Chinese University of Hong Kong) for computing support. The authors also thank Carlos Chau (School of Biomedical Sciences, Chinese University of Hong Kong) for assistance in part of the analysis.

Funding. This study was partially supported by the Lo Kwee Seong Biomedical Research Fund, a National Natural Science Foundation of China grant (81971706), and a Chinese University of Hong Kong Direct Grant.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. H.-C.S. (lead) conceived and designed the study, with input from S.R. H.-C.S. supervised the study. H.-C.S. (lead), S.R., and A.L. contributed to data analysis. H.-C.S., S.R., and A.L. contributed to data interpretation. H.-C.S. drafted the manuscript, with input from A.L. and S.R.

References

- Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020;382:1199–1207
- Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua liu xing bing xue za zhi* 2020;41:145–151 [in Chinese]
- Guan WJ, Ni ZY, Hu Y, et al.; China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708–1720
- World Health Organization. Coronavirus Disease 2019 (COVID-19) Situation Report – 93 [Internet], 2020. Available from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200422-sitrep-93-covid-19.pdf?sfvrsn=35cf80d7_4. Accessed 22 April 2020
- Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506
- Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J Virol* 2020;94:e00127-20
- Zhou P, Yang X-L, Wang X-G, et al. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. 23 January 2020 [preprint]. *bioRxiv*: 2020.2020.01.22.914952
- Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. 15 February 2020 [preprint]. *bioRxiv*: 2020.02.11.944462
- Li W, Moore MJ, Vasilieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 2003;426:450–454
- Kuba K, Imai Y, Rao S, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med* 2005;11:875–879

11. Hofmann H, Geier M, Marzi A, et al. Susceptibility to SARS coronavirus S protein-driven infection correlates with expression of angiotensin converting enzyme 2 and infection can be blocked by soluble receptor. *Biochem Biophys Res Commun* 2004;319:1216–1221
12. Jia HP, Look DC, Shi L, et al. ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *J Virol* 2005;79:14614–14621
13. Zou X, Chen K, Zou J, Han P, Hao J, Han Z. Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. *Front Med* 2020;14:185–192
14. Langdon RJ, Richmond RC, Hemani G, et al. A phenome-wide Mendelian randomization study of pancreatic cancer using summary genetic data. *Cancer Epidemiol Biomarkers Prev* 2019;28:2070–2078
15. Li X, Meng X, He Y, et al. Genetically determined serum urate levels and cardiovascular and other diseases in UK Biobank cohort: a phenome-wide Mendelian randomization study. *PLoS Med* 2019;16:e1002937
16. Meng X, Li X, Timofeeva MN, et al. Phenome-wide Mendelian-randomization study of genetically determined vitamin D on multiple health outcomes using the UK Biobank study. *Int J Epidemiol* 2019;48:1425–1434
17. Smith GD, Ebrahim S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1–22
18. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;23:R89–R98
19. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;4:e352
20. Schmidt AF, Finan C, Gordillo-Marañón M, et al. Genetic drug target validation using Mendelian randomization. 25 September 2019 [preprint]. *bioRxiv*: 781039
21. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature* 2018;558:73–79
22. Folkersen L, Fauman E, Sabater-Lleal M, et al.; IMPROVE Study Group. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet* 2017;13:e1006706
23. Bellera C, Proust-Lima C, Joseph L, et al. A two-stage model in a Bayesian framework to estimate a survival endpoint in the presence of confounding by indication. *Stat Methods Med Res* 2018;27:1271–1281
24. UK_Biobank_GWAS / imputed-v2-gwas / [Internet]. Available from https://github.com/Nealelab/UK_Biobank_GWAS/tree/master/imputed-v2-gwas. Accessed 3 March 2020
25. Details and considerations of the UK Biobank GWAS [Internet], 2017. Available from <http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas>. Accessed 3 March 2020
26. MRC IEU UK Biobank GWAS pipeline version 2 [Internet], 2019. Available from <https://doi.org/10.5523/bris.pnoat8cxo0u52p6ynfaeig>. Accessed 3 March 2020
27. Lloyd-Jones LR, Robinson MR, Yang J, Visscher PM. Transformation of summary statistics from linear mixed model association on all-or-none traits to odds ratio. *Genetics* 2018;208:1397–1408
28. Gamazon ER, Segre AV, van de Bunt M, et al.; GTEx Consortium. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nat Genet* 2018;50:956–967
29. Analysis methods [Internet], 2019. Available from <https://gtexportal.org/home/documentation> Page#staticTextAnalysisMethods. Accessed 3 March 2020
30. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658–665
31. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512–525
32. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet* 2018;27:R195–R208
33. Zhu Z, Zheng Z, Zhang F, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun* 2018;9:224
34. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;50:1505–1513
35. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. 29 January 2018 [preprint]. *arXiv*:1801.09652
36. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: a review. *Res Synth Methods* 2019;10:486–496
37. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases [published correction appears in *Nat Genet* 2018;50:1196]. *Nat Genet* 2018;50:693–698
38. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol* 2018;33:947–952
39. Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;25:1091–1093
40. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001;29:1165–1188
41. Scott JG, Kelly RC, Smith MA, Zhou P, Kass RE. False discovery rate regression: an application to neural synchrony detection in primary visual cortex. *J Am Stat Assoc* 2015;110:459–471
42. Mahajan A, Go MJ, Zhang W, et al.; Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Mexican American Type 2 Diabetes (MAT2D) Consortium; Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* 2014;46:234–244
43. Thomas NJ, Jones SE, Weedon MN, Shields BM, Oram RA, Hattersley AT. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol* 2018;6:122–129
44. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017;32:377–389
45. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med* 2017;36:1783–1802
46. Guan W-J, Liang WH, Zhao Y, et al.; China Medical Treatment Expert Group for Covid-19. Comorbidity and its impact on 1590 patients with Covid-19 in China: a nationwide analysis. *Eur Respir J*. 26 March 2020 [Epub ahead of print]. DOI: 10.1183/13993003.00547-2020
47. Banik GR, Alqahtani AS, Booy R, Rashid H. Risk factors for severity and mortality in patients with MERS-CoV: analysis of publicly available data from Saudi Arabia. *Virol Sin* 2016;31:81–84
48. Alqahtani FY, Aleanizy FS, Ali El Hadi Mohamed R, et al. Prevalence of comorbidities in cases of Middle East respiratory syndrome coronavirus: a retrospective study. *Epidemiol Infect* 2019;147:e35
49. Kulcsar KA, Coleman CM, Beck SE, Frieman MB. Comorbid diabetes results in immune dysregulation and enhanced disease severity following MERS-CoV infection. *JCI Insight* 2019;4:e131774
50. Liang W, Guan W, Chen R, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol* 2020;21:335–337
51. Monteil V, Kwon H, Prado P, et al. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell*. 17 April 2020 [Epub ahead of print]. DOI: 10.1016/j.cell.2020.04.004
52. Recombinant Human Angiotensin-converting Enzyme 2 (rhACE2) as a Treatment for Patients With COVID-19 [Internet], 2020. Available from <https://ClinicalTrials.gov/show/NCT04335136>. Accessed 9 May 2020
53. Gordon DE, Jang GM, Bouhaddou M, et al. SARS-CoV-2-human protein-protein interaction map reveals drug targets and potential drug repurposing. *bioRxiv* 2020:2020.2003.2022.002386
54. Imai Y, Kuba K, Rao S, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature* 2005;436:112–116
55. Li G, Hu R, Zhang X. Antihypertensive treatment with ACEI/ARB of patients with COVID-19 complicated by hypertension. *Hypertens Res*. 30 March 2020 [Epub ahead of print]. DOI: 10.1038/s41440-020-0433-1
56. Sommerstein R, Kochen MM, Messerli FH, Gräni C. Coronavirus disease 2019 (COVID-19): do angiotensin-converting enzyme inhibitors/angiotensin receptor blockers have a biphasic effect? *J Am Heart Assoc* 2020;9:e016509
57. Ferrario CM, Jessup J, Chappell MC, et al. Effect of angiotensin-converting enzyme inhibition and angiotensin II receptor blockers on cardiac angiotensin-converting enzyme 2. *Circulation* 2005;111:2605–2610

58. Ferrario CM, Varagic J. The ANG-(1-7)/ACE2/mas axis in the regulation of nephron function. *Am J Physiol Renal Physiol* 2010;298:F1297–F1305
59. Igase M, Strawn WB, Gallagher PE, Geary RL, Ferrario CM. Angiotensin II AT1 receptors regulate ACE2 and angiotensin-(1-7) expression in the aorta of spontaneously hypertensive rats. *Am J Physiol Heart Circ Physiol* 2005;289:H1013–H1019
60. Vuille-dit-Bille RN, Camargo SM, Emmenegger L, et al. Human intestine luminal ACE2 and amino acid transporter expression increased by ACE-inhibitors. *Amino Acids* 2015;47:693–705
61. Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet* 2020;395:e30–e31
62. Stebbing J, Phelan A, Griffin I, et al. COVID-19: combining antiviral and anti-inflammatory treatments. *Lancet Infect Dis* 2020;20:400–402
63. Rolf MG, Curwen JO, Veldman-Jones M, et al. In vitro pharmacological profiling of R406 identifies molecular targets underlying the clinical effects of fostamatinib. *Pharmacol Res Perspect* 2015;3:e00175
64. Kang Y, Jiang X, Qin D, et al. Efficacy and safety of multiple dosages of fostamatinib in adult patients with rheumatoid arthritis: a systematic review and meta-analysis. *Front Pharmacol* 2019;10:897
65. Colonna L, Catalano G, Chew C, et al. Therapeutic targeting of Syk in autoimmune diabetes [published correction appears in *J Immunol* 2011; 187:1516]. *J Immunol* 2010;185:1532–1543
66. Roth DE, Richard SA, Black RE. Zinc supplementation for the prevention of acute lower respiratory infection in children in developing countries: meta-analysis and meta-regression of randomized trials. *Int J Epidemiol* 2010;39:795–808
67. Speth R, Carrera E, Jean-Baptiste M, Joachim A, Linares A. Concentration-dependent effects of zinc on angiotensin-converting enzyme-2 activity (1067.4) (Abstract). *FASEB J* 2014;28(Suppl. 1)
68. te Velthuis AJ, van den Worm SH, Sims AC, Baric RS, Snijder EJ, van Hemert MJ. Zn(2+) inhibits coronavirus and arterivirus RNA polymerase activity in vitro and zinc ionophores block the replication of these viruses in cell culture. *PLoS Pathog* 2010;6:e1001176
69. Skalny AV, Rink L, Ajsuvakova OP, et al. Zinc and respiratory tract infections: perspectives for COVID-19 (Review). *Int J Mol Med*. 14 April 2020 [Epub ahead of print]. DOI: 10.3892/ijmm.2020.4575
70. Pendergrass SA, Brown-Gentry K, Dudek SM, et al. The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. *Genet Epidemiol* 2011;35:410–422