Received: 5 May 2022 | Accepted: 26 August 2022

DOI: 10.1002/jmv.28104

RESEARCH ARTICLE

JOURNAL OF MEDICAL VIROLOGY WILEY

Genetic predisposition to blood cell indices in relation to severe COVID-19

Yao Yao[1] | Hongfei Song[2] | Fanshuang Zhang[3] | Jibin Liu[2] | Dong Wang[2] | Quansheng Feng[2] | Shuquan Rao[1] | Cen Jiang[2]

[1]State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China

[2]Traditional Chinese Medicine and Inflammation Regulation Research Group, School of Basic Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu, P. R. China

[3]Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China

Correspondence

Cen Jiang and Shuquan Rao
Email: jiangcen517@163.com and raoshuquan@ihcams.ac.cn

Funding information

CAMS Innovation Fund for Medical Sciences, Grant/Award Number: 2021-I2M-1-041; National medical and health science technology strategic platform and system construction project, Grant/Award Number: No.2022-I2M-2-003; the Fundamental Research Funds for the Central Universities, Grant/Award Number: 2021YFA1102300; Chengdu University of Traditional Chinese Medicine, Grant/Award Number: 2021-I2M-1-041; Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences, Grant/Award Number: 2021-RC310-015

Abstract

Despite considerable variation in disease manifestations observed among coronavirus disease 2019 (COVID-19) patients infected with severe acute respiratory syndrome coronavirus 2, the risk factors predicting disease severity remain elusive. Recent studies suggest that peripheral blood cells play a pivotal role in COVID-19 pathogenesis. Here, we applied two-sample Mendelian randomization (MR) analyses to evaluate the potential causal contributions of blood cell indices variation to COVID-19 severity, using single-nucleotide polymorphisms (SNPs) as instrumental variables for 17 indices from the UK Biobank and INTERVAL genome-wide association studies (N = 173 480). Data on the associations between the SNPs and very severe respiratory confirmed COVID-19 were obtained from the COVID-19 host genetics initiative (N = 8779/1 001 875). We observed significant negative association between hematocrit (HCT; odds ratio, OR = 0.775, 95% confidence interval, CI = 0.635–0.915, p = 3.48E−04) or red blood cell count (OR = 0.830, 95% CI = 0.728–0.932, p = 2.19E−03) and very severe respiratory confirmed COVID-19, as well as nominal negative association of hemoglobin concentration (OR = 0.808, 95% CI = 0.673–0.943, p = 3.95E−03) with very severe respiratory confirmed COVID-19 (no effect survived multiple correction). In conclusion, the MR study supports a protective effect of high HCT and red blood cell count from very severe respiratory confirmed COVID-19, suggesting potential strategies to ameliorate/treat clinical conditions in very severe respiratory confirmed COVID-19.

KEYWORDS

blood cell indices, COVID-19, GWAS summary data, MR analyses, severe patients

1 | INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a major global health threat. As of July 27, 2022, there were more than 570 million confirmed cases of COVID-19, including more than 6 million deaths worldwide, according to the reports from World Health Organization.[1] Similar to other human-tropic viral infections, COVID-19 has varied manifestations, with the large majority of infected persons having only mild symptoms or even no symptoms.[2]

Yao Yao, Hongfei Song, and Fanshuang Zhang contributed equally to this study.

Mortality is predominantly caused by severe respiratory failure related to interstitial pneumonia in both lungs and acute respiratory distress syndrome (ARDS).[3] While it is reported the severity of the symptoms is strongly associated with the patient's age in some studies,[4] little is known about the causal influences of severe illness and even mortality in people infected with SARS-CoV-2.

Variation in blood cell indices has been linked to diseases with high population burdens, such as autoimmune disease and susceptibility to viral infection. Multiple studies have reported reduced numbers of natural killer cells in the peripheral blood of COVID-19 patients, which is associated with the severity of the disease.[5] Similarly, analyses of peripheral blood mononuclear cells from symptomatic COVID-19 patients have shown a significant influx of activated $CD4^+$ T cells and inflammatory monocytes.[6] Moreover, a higher neutrophil-to-lymphocyte ratio may predict increased mortality in COVID-19 patients.[7] Nevertheless, it remains largely unclear whether variations in blood cell indices reflect etiological roles of hematological pathways or a consequence of COVID-19.

One efficient way to study the causality of an association is Mendelian randomization (MR), in which genetic variants associated with a modifiable exposure are used as instrumental variables to estimate the causal effect of the exposure on an outcome.[8] MR analysis uses the random allocation of alleles at conception to obtain an unconfounded estimate of the association between a risk factor and an outcome, thereby avoiding the potential residual confounding and reverse causation in observational association studies. Here, we applied the MR design to evaluate the potential causal contributions of 17 blood cell indices levels to COVID-19 severity. As pharmacological modulation of blood cell indices advances, we believe identifying shared causal pathways between these indices and COVID-19, especially severe illness of COVID-19, could provide new therapeutic opportunities.

2 | MATERIALS AND METHODS

2.1 | Data source

Summary-level data (i.e., beta coefficients and standard errors) of the associations between single-nucleotide polymorphisms (SNPs) associated with blood cell indices (exposures) and very severe respiratory confirmed COVID-19 (outcome) were extracted from large-scale genome-wide association studies for these phenotypes (Figure 1).

Recently, a genome-wide association analysis was performed, testing 29.5 million genetic variants for association with 36 red cell, white cell, and platelet properties on 173 480 European-ancestry participants.[9] All these individuals are from three large-scale UK studies: INTERVAL, approved by Cambridge (East) Research Ethics Committee, UK Biobank, and UK BiLEVE (a selected subset of the UK Biobank cohort), both approved by the North West Multi-Centre Research Ethics Committee.[9] This genome-wide association study (GWAS) identified a total of 6736 independent trait-variant paired genetic significant associations for blood cell indices ($p < 8.31E\text{-}09$).[9] Due to the high degree of genetic correlation between the blood cell indices, in particular, due to the presence of calculated and compound indices, we selected 17 blood cell indices (Supporting Information: Table S1) to

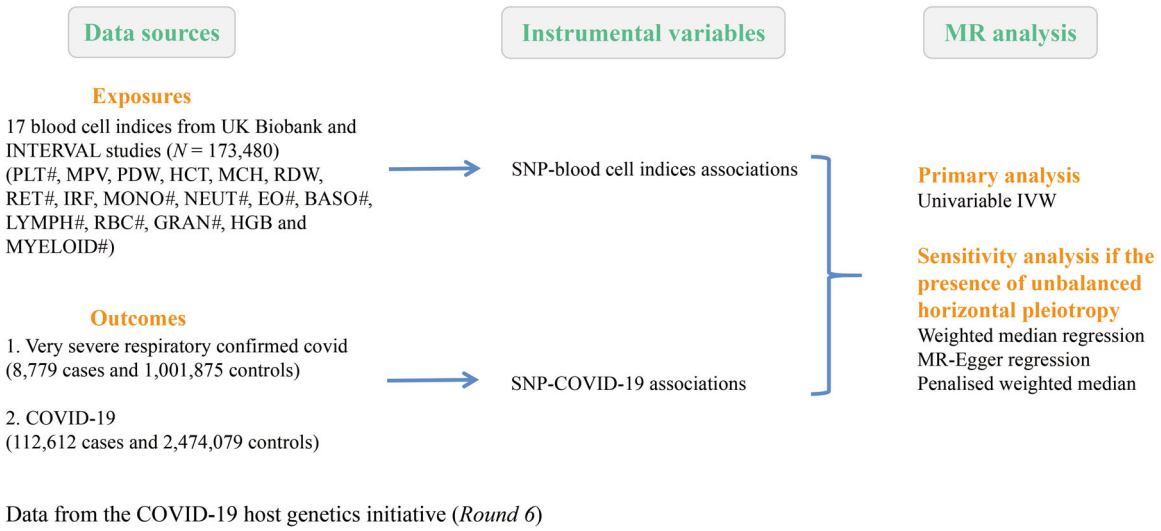

FIGURE 1 Summary of data sources and flowchart of study design. SNP-blood cell indices associations at genome-wide significance ($p < 8.31E{-}09$) were used as instrumental variables. SNP-blood cell indices associations at genome-wide significance ($p < 8.31E{-}09$) were used as instrumental variables. BASO#, basophil count; COVID-19, coronavirus disease 2019; EO#, eosinophil count; GRAN#, granulocyte count; HCT, hematocrit; HGB, hemoglobin concentration; IRF, immature fraction of reticulocytes; IVW, inverse-variance weighted; LYMPH#, lymphocyte count; MCH, mean corpuscular hemoglobin; MONO#, monocyte count; MPV, mean platelet volume; MR, Mendelian randomization; MYELOID#, myeloid white cell count; NEUT#, neutrophil count; PDW, platelet distribution width; PLT#, platelet count; RBC#, red blood cell count; RDW, red cell distribution width; RET#, reticulocyte count; SNP, single-nucleotide polymorphism.

represent all 36 indices. These 17 blood cell indices are platelet count (PLT#), mean platelet volume (MPV), platelet distribution width (PDW), hematocrit (HCT), mean corpuscular hemoglobin (MCH), red cell distribution width (RDW), reticulocyte count (RET#), an immature fraction of reticulocytes (IRF), monocyte count (MONO#), neutrophil count (NEUT#), eosinophil count (EO#), basophil count (BASO#), lymphocyte count (LYMPH#), red blood cell count (RBC#), granulocyte count (GRAN#), hemoglobin concentration (HGB), and myeloid white cell count (MYELOID#).

The COVID-19 host genetics initiative (COVID-19 HGI) is a global initiative, trying to elucidate the role of host genetic factors in the susceptibility and severity of the SARS-CoV-2 virus pandemic.[10] Genetic variants associated with COVID-19 susceptibility, severity, and outcomes were identified by GWAS. The round 6 very severe respiratory confirmed COVID-19 (total cases 8,779) versus population (total controls 1 001 875) GWAS summary-level data released by COVID-19 HGI (COVID19-hg GWAS meta-analyses round 6 (covid19hg.org)) was used in the current study. As a control, GWAS statistics from COVID-19 (i.e., both hospitalized and nonhospitalized) versus negative control population (112 612/2 474 079) was chosen. Detailed information such as phenotype definition, diagnosis criteria, sample size, and ancestry can be found in Supporting Information: Table S1.

2.2 | Two-sample MR analysis

In this design, the exposure phenotype of interest was 17 blood cell indices described above, and the severity of COVID-19 was defined as the outcome. A total of 3289 genetic significant SNPs ($p < 8.31E-09$) of 17 blood traits were acquired from the UK study as instrumental variables, and 2462 were kept after harmonizing (Supporting Information: Table S2). To further investigate the causal relationships between the exposure and outcome and also test the validity of the genetic score as an instrument, two-sample MR approaches were used to detect and accommodate violations of the MR assumptions, specifically horizontal pleiotropy. Briefly, for each SNP instrument indexing the exposure, the ratio estimate is calculated as the beta coefficient for the SNP-exposure association divided by the beta coefficient for the SNP-outcome association. These estimates are then combined across SNPs using the inverse-variance-weighted (IVW) method under a multiplicative random-effects model,[11] alongside other methods to overcome the violations of specific instrumental variable assumptions, as no single method controls for all statistical properties that may impact MR estimates, including MR-Egger (Egger),[12] weighted median,[13] and penalized weighted median[13] approaches.

To further determine whether pleiotropy is unbalanced, we tested the significance of the MR-Egger intercept. A significant test indicates the presence of unbalanced pleiotropy, and sensitivity analyses are thus required to validate the results. To account for 17 tests, we applied a Bonferroni correction and considered associations with $p < 2.94E-03$ (i.e., 0.05/17) to be significant, and nominal association with a p-value between 0.05 and 2.94E−03. MR-Egger regression was used to assess horizontal pleiotropy.

3 | RESULTS

3.1 | Identification of genetic instruments

For each of the 17 blood cell indices (PLT#, MPV, PDW, HCT, MCH, RDW, RET#, IRF, MONO#, NEUT#, EO#, BASO#, LYMPH#, RBC#, GRAN#, HGB, and MYELOID#), SNPs surpassing genome-wide significant p-value ($p < 8.31E-09$) were chosen as genetic instruments.[9] Final analyses included uncorrelated ($r^2 < 0.001$; validated in European British panel using LDlink) SNPs available in GWAS of very severe respiratory confirmed COVID-19. The number of genetic instruments ranged from 64 for BASO# to 235 for MPV (Supporting Information: Table S2).

3.2 | Associations of blood cell indices with severe COVID-19

Using IVW method, under a multiplicative random-effects model, we detected significant negative association between HCT (OR= 0.775, 95% CI = 0.635–0.915, $p = 3.48E-04$) or RBC# (OR = 0.830, 95% CI = 0.728–0.932, $p = 2.19E-03$) and very severe respiratory confirmed COVID-19 (Figure 2 and Supporting Information: Table S3).

In addition, nominal negative associations of very severe respiratory confirmed COVID-19 were observed for HGB (OR = 0.808, 95% CI = 0.673–0.943, $p = 3.95E-03$), which did not survive multiple Bonferroni correction ($p < 2.94E-03$, i.e., 0.05/17) (Figure 2 and Supporting Information: S1, Supporting Information: Table S3). Weak pleiotropy was detected for EO# ($p = 0.029$), MYELOID# ($p = 0.021$), and IRF($p = 0.012$), which may bias the IVW MR estimate. There is no evidence of heterogeneity and unbalanced horizontal pleiotropy (all $p > 0.05$) was observed for the left 14 traits, as indicated by the MR-Egger intercept test (Table S4).

As a negative control, we further performed similar MR analyses on COVID-19 (i.e., both hospitalized and nonhospitalized). Surprisingly, no evidence of significant association was observed between COVID-19 and blood cell indices as mentioned above (all $p > 2.94E-03$; Supporting Information: Table S5), suggesting that these associations are specific to severe COVID-19.

4 | DISCUSSION

Everyone is at risk of getting COVID-19 if they are exposed to the virus. Some people are more likely than others to become severely ill, which means that they may require hospitalization, intensive care, or a ventilator to help them breathe, or they may even die. Understanding the risk factors conferring risk for severe COVID-19 illness will provide

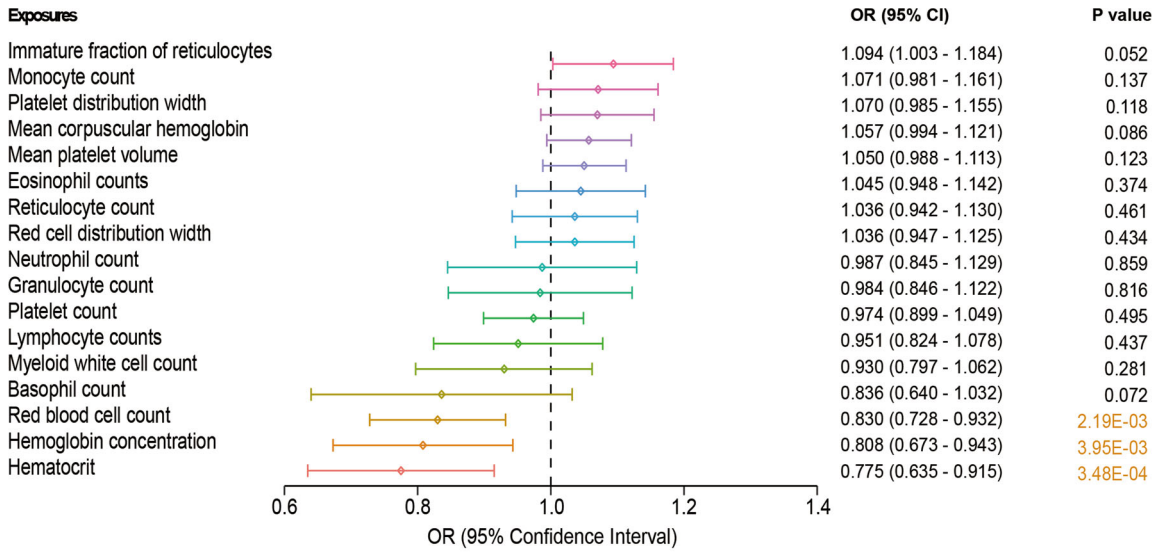

FIGURE 2 Associations of genetic predisposition to blood cell indices with severe COVID-19. The OR corresponds to the increase of one standard deviation in blood cell indicator level. Estimates are from the multiplicative random-effects inverse variance-weighted method. Bonferroni corrected associations with $p < 2.94E{-}03$ (i.e., 0.05/17) were considered significant. CI, confidence interval; COVID-19, coronavirus disease 2019; OR, odds ratio.

vital information to help public health officials make decisions to protect those vulnerable populations. Previous studies have revealed causal roles of blood cell indices in diseases with high population burdens, such as autoimmune disease and respiratory and cardiovascular illnesses, which have been implicated in COVID-19 pathogenesis.[9,14] Here, we used MR inference to unravel causal mechanisms underlying blood cell indices-COVID-19 correlations. We observed significant negative association between HCT or RBC# and severe COVID-19, as well as nominal negative association with HGB (OR = 0.808, 95% CI = 0.673–0.943, $p = 3.95E{-}03$). Our genetic evidence may suggest protective roles of HCT or RBC# from severe COVID-19.

Although most of the clinical manifestations of COVID-19 are thought to be caused by cytokine storm,[15] leading to an exaggerated inflammatory response in the body, SARS-CoV-2 is also known to attack one of the four sites on hemoglobin that binds to oxygen, thus decreasing the oxygen-carrying capacity and finally leading to ARDS. Supporting it, significantly lower hemoglobin levels were observed in COVID-19 patients, especially in severely ill patients.[16] We also observed a trend of a negative association between hemoglobin concentration and severe COVID-19 (OR = 0.808 and $p = 3.95E{-}03$), according to our MR analyses.

Most of the critically ill patients, who presented with hypoxia due to ARDS, have a poor prognosis, especially in people with underlying co-morbidities.[17] Ejigu et al.[18] reported a COVID-19 case, who was intubated and mechanically ventilated for acute respiratory failure. Despite the patient being presented with multiple comorbidities and a cardiac arrest, he improved drastically after being given five units of packed RBC# within 2 days, in comparison with the median duration of mechanical ventilation (around 10 days).[19] These lines of evidence, combined with our results, strongly suggested a protective role of RBC# from severe COVID-19, especially those with ARDS. Since respiratory and ventilator support, the current standard therapeutic treatment of COVID-19, is associated with high mortality rates, strategies to increase HCT and/or RBC levels can be considered to treat/ameliorate respiratory conditions associated with COVID-19. However, more work is needed to ascertain the mechanistic role played by red blood cells in SARS-CoV-2 control and pathogenesis.

This study uses a robust quasi-experimental approach, based on high-quality GWAS data from international consortia using large samples. The genetic instrument for blood cell indices comprised multiple SNPs robustly associated with each blood cell indicator, thereby providing a strong genetic instrument. Second, the majority of samples are of European descent in UK Biobank, thus minimizing population stratification bias. However, the following limitations must be acknowledged. First, the possibility that the blood cell indices-related SNPs affect very severe respiratory confirmed COVID-19 outcomes through other causal pathways than through blood cell indices levels can't be entirely ruled out. Second, genetic variants associated with very severe respiratory confirmed COVID-19 are derived from a comparison between very severe respiratory confirmed COVID-19 and general populations. Limited information is available about the SARS-CoV-2 infection status in the control participants and the presence of infected persons in the control group would result in a false negative association with very severe respiratory confirmed COVID-19, which might further affect the MR analyses in the current study. Third, participants of the UK Biobank were included in both the exposure and outcome datasets, which may have introduced some bias in the causal estimates in the direction of the observational association between blood cell indices and COVID-19. However, the genetic variants are reasonably strongly associated with the exposure, meaning that bias due to sample overlap is reasonably small.

5 | CONCLUSION

In conclusion, the MR study supports a protective effect of high HCT and/or RBC# from very severe respiratory confirmed COVID-19, suggesting potential strategies to ameliorate/treat clinical conditions in very severe respiratory confirmed COVID-19.

AUTHOR CONTRIBUTIONS

Shuquan Rao, Cen Jiang, and Yao Yao contributed to the conception and study design. Yao Yao and Hongfei Song contributed to data collection and analysis. Fanshuang Zhang, Jibin Liu, and Dong Wang contributed to drafting the text and preparing figures. Quansheng Feng, Cen Jiang, and Shuquan Rao contributed to data interpretation and wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank Prof. Andre Franke (member of the Covid-19 Host Genetics Initiative, University Hospital of Schleswig-Holstein) for valuable discussion and suggestions. We would also like to thank the COVID-19 host genetics initiative for generously sharing the genome-wide association study summary statistics. This work was supported by CAMS Innovation Fund for Medical Sciences (CIFMS, No. 2021-I2M-1-041), National medical and health science technology strategic platform and system construction project (No. 2022-I2M-2-003) , the Fundamental Research Funds for the Central Universities (No. 3332021031) and Chengdu University of Traditional Chinese Medicine (XGZX2006, XSGG2019004).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Genome-wide association study summary statistics data of 17 blood cell indices are available from http://www.bloodcellgenetics.org. Data on the associations between the SNPs and COVID-19 were obtained from the COVID-19 host genetics initiative (https://www.covid19hg.org/results/r6/). The results of our study are provided in the Supporting Information Tables.

ORCID

Quansheng Feng http://orcid.org/0000-0003-3437-8987

REFERENCES

1. WHO. *WHO coronavirus (COVID-19) dashboard*. Accessed August 3, 2022. https://covid19.who.int/,WHOCoronavirus
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. 2020;323(13):1239-1242.
3. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York city area. *JAMA*. 2020;323(20):2052-2059.
4. Chen Y, Klein SL, Garibaldi BT, et al. Aging in COVID-19: vulnerability, immunity and intervention. *Ageing Res Rev*. 2021;65:101205.
5. Zheng M, Gao Y, Wang G, et al. Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cell Mol Immunol*. 2020;17(5):533-535.
6. Zhang D, Guo R, Lei L, et al. Frontline science: COVID-19 infection induces readily detectable morphologic and inflammation-related phenotypic changes in peripheral blood monocytes. *J Leukoc Biol*. 2021;109(1):13-22.
7. Liu Y, Du X, Chen J, et al. Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. *J Infect*. 2020;81(1):e6-e12.
8. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Smith GD. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133-1163.
9. Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell*. 2016;167(5):1415-1429.
10. The COVID-19 Host Genetics Initiative. The COVID-19 host genetics initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Human Genet*. 2020;28(6):715-718.
11. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658-665.
12. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32(5):377-389.
13. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304-314.
14. Siedlinski M, Jozefczuk E, Xu X, et al. White blood cells and blood pressure: a Mendelian randomization study. *Circulation*. 2020;141(16):1307-1317.
15. Mehta P, McAuley DF, Brown M, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet*. 2020;395(10229):1033-1034.
16. Lippi G, Mattiuzzi C. Hemoglobin value may be decreased in patients with severe coronavirus disease 2019. *Hematol Transfus Cell Ther*. 2020;42(2):116-117.
17. Chen R, Liang W, Jiang M, et al. Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. *Chest*. 2020;158(1):97-105.
18. Ejigu T, Patel N, Sharma A, Vanjarapu JMR, Nookala V. Packed red blood cell transfusion as a potential treatment option in COVID-19 patients with hypoxemic respiratory failure: a case report. *Cureus*. 2020;12(6):e8398.
19. Bhatraju PK, Ghassemieh BJ, Nichols M, et al. Covid-19 in critically ill patients in the Seattle region—case series. *N Engl J Med*. 2020;382(21):2012-2022.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Yao Y, Song H, Zhang F, et al. Genetic predisposition to blood cell indices in relation to severe COVID-19. *J Med Virol* 2022;1-5. doi:10.1002/jmv.28104