



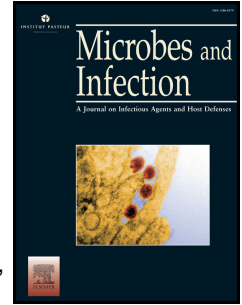
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# Journal Pre-proof

A systematical association analysis of 25 common virus infection and genetic susceptibility of COVID-19 infection

Na Zhang, Yujing Chen, Chun'e Li, Xiaoyue Qin, Dan He, Wenming Wei, Yijing Zhao, Qingqing Cai, Sirong Shi, Xiaoge Chu, Yan Wen, Yumeng Jia, Feng Zhang



PII: S1286-4579(23)00073-4

DOI: <https://doi.org/10.1016/j.micinf.2023.105170>

Reference: MICINF 105170

To appear in: *Microbes and Infection*

Received Date: 26 January 2023

Revised Date: 2 June 2023

Accepted Date: 7 June 2023

Please cite this article as: N. Zhang, Y. Chen, C.'e Li, X. Qin, D. He, W. Wei, Y. Zhao, Q. Cai, S. Shi, X. Chu, Y. Wen, Y. Jia, F. Zhang, A systematical association analysis of 25 common virus infection and genetic susceptibility of COVID-19 infection, *Microbes and Infection*, <https://doi.org/10.1016/j.micinf.2023.105170>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

1 **A systematical association analysis of 25 common virus infection and genetic susceptibility**  
2 **of COVID-19 infection**

3 Na Zhang<sup>1#</sup>, Yujing Chen<sup>1#</sup>, Chun'e Li<sup>1</sup>, Xiaoyue Qin<sup>1</sup>, Dan He<sup>1</sup>, Wenming Wei<sup>1</sup>, Yijing Zhao<sup>1</sup>,  
4 Qingqing Cai<sup>1</sup>, Sirong Shi<sup>1</sup>, Xiaoge Chu<sup>1</sup>, Yan Wen<sup>1</sup>, Yumeng Jia<sup>1</sup>, Feng Zhang<sup>1,2\*</sup>

5

6 1. Key Laboratory of Trace Elements and Endemic Diseases of National Health and Family  
7 Planning Commission, Key Laboratory of Environment and Genes Related to Diseases of  
8 Ministry of Education of China, Key Laboratory for Disease Prevention and Control and Health  
9 Promotion of Shaanxi Province, School of Public Health, Health Science Center, Xi'an Jiaotong  
10 University, Xi'an, China.

11 2. Department of Psychiatry, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an,  
12 China.

13

14 #The two authors contributed equally to this work

15

16 \*Corresponding author:

17 Feng Zhang

18 Key Laboratory of Trace Elements and Endemic Diseases, National Health Commission of the  
19 People's Republic of China. School of Public Health, Health Science Center, Xi'an Jiaotong  
20 University

21 Email: fzhxjtu@mail.xjtu.edu.cn

22 Xi'an, P. R. China 710061

## 23 Abstract

24 **Objectives:** Previous studies identified a number of diseases were associated with 2019  
25 coronavirus disease (COVID-19). However, the associations between these diseases related  
26 viral infections and COVID-19 remains unknown now.

27 **Methods:** In this study, we utilized single nucleotide polymorphisms (SNPs) related to COVID-  
28 19 from genome-wide association study (GWAS) and individual-level genotype data from the  
29 UK biobank to calculate polygenic risk scores (PRS) of 487,409 subjects for eight COVID-19  
30 clinical phenotypes. Then, multiple logistic regression models were established to assess the  
31 correlation between serological measurements (positive/negative) of 25 viruses and the PRS of  
32 eight COVID-19 clinical phenotypes. And we performed stratified analyses by age and gender.

33 **Results:** In whole population, we identified 12 viruses associated with the PRS of COVID-19  
34 clinical phenotypes, such as *VZV seropositivity for Varicella Zoster Virus*  
35 (Unscreened/Exposed\_Negative:  $\beta = 0.1361$ ,  $P = 0.0142$ ; Hospitalized/Unscreened:  $\beta = 0.1167$ ,  
36  $P = 0.0385$ ) and *MCV seropositivity for Merkel Cell Polyomavirus*  
37 (Unscreened/Exposed\_Negative:  $\beta = -0.0614$ ,  $P = 0.0478$ ). After age stratification, we  
38 identified seven viruses associated with the PRS of eight COVID-19 clinical phenotypes. After  
39 gender stratification, we identified five viruses associated with the PRS of eight COVID-19  
40 clinical phenotypes in the women group.

41 **Conclusion:** Our study findings suggest that the genetic susceptibility to different COVID-19  
42 clinical phenotypes is associated with the infection status of various common viruses.

43 **Keywords:** common viral infections, coronavirus disease 2019, positive serological  
44 measurements, genetic susceptibility, genome-wide association study

## 45 **1. Introduction**

46 The coronavirus disease 2019 (COVID-19) pandemic has resulted in over 603 million  
47 infections and more than 6.4 million deaths worldwide as of July 31, 2022, causing a significant  
48 disease burden for countries around the world[1]. COVID-19 is caused by the highly contagious  
49 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which presents with a wide  
50 range of clinical symptoms, from asymptomatic infection or mild illness to serious illness  
51 requiring hospitalization and mechanical ventilation[2]. Previous study has found that clinical  
52 variation in COVID-19 severity and symptom presentation may be due to differences in host  
53 genetic factors associated with immune response[3]. Therefore, changes in immune responses  
54 may be linked to different clinical manifestations of COVID-19. As research on COVID-19  
55 infection has progressed, several single nucleotide polymorphisms (SNPs) and genes associated  
56 with different aspects of susceptibility to infection or disease severity have been identified[4],  
57 suggesting that genetic factors in the host could influence its susceptibility and severity to the  
58 virus[4, 5].

59 As COVID-19 prevention measures become normalized, identifying co-infection with one  
60 or more respiratory viruses can help us understand the infection status of SARS-CoV-2. During  
61 the COVID-19 pandemic, there have been increasing reports of co-infection with other  
62 pathogens in COVID-19 patients[6, 7]. A retrospective study found that 242 (94.2%) patients  
63 had co-infected with one or more pathogens, the most frequent of which were the Epstein virus  
64 Barr virus(EBV), the rhinovirus, and the adenovirus[8]. Herpes simplex virus 1 (HSV-1) and  
65 varicella zoster virus (VZV) are DNA viruses of the neurotropic alpha human herpesvirus  
66 subfamily (HHV)[9]. The virus remains dormant in the body after recovery from the initial

67 infection and reactivates when the immune system is compromised, causing significant damage  
68 to the host organism[10]. Although few case reports of VZV and HSV in patients with COVID-  
69 19 have been published, studies have suggested that VZV may be an indicator of potential  
70 COVID-19 infection[11]. Previous studies have identified associations between host  
71 polymorphisms in genes related to cell entry, cytokine production, and immune response with  
72 multiple viruses, and antibody responses have been found to be highly heritable (32%–48%)  
73 [12]. The HLA-B\*46:01 allele in East Asian patients was associated with infection severity  
74 during the 2003 severe acute respiratory syndrome (SARS) outbreak caused by the SARS-CoV-  
75 2-related  $\beta$  coronavirus[13].

76 Researchers have discovered an association between the prevalence of COVID-19 and  
77 other viral infections[14]. In addition to age, obesity, hypertension and other common risk  
78 factors associated with increased COVID-19 severity[15], a study found that cytomegalovirus  
79 (CMV) seropositivity was associated with more than twice the risk of hospitalization due to  
80 SARS-CoV-2 infection[14]. Potential CMV infection influences future infection with other  
81 viruses and shapes the distribution of adaptive immune cell populations[16, 17]. In SARS-CoV-  
82 2 infection, CMV seropositivity results in a severe immunological signature, which is the  
83 activation of T<sub>EMRA</sub> cells[18].

84 The aim of our study was to investigate which viral infections are associated with the  
85 genetic susceptibility to COVID-19 by using polygenic risk scores (PRS) for COVID-19 eight  
86 clinical phenotypes. PRS is a score that aggregates genetic variants to predict disease risk. In  
87 this work, we used data from the UK Biobank cohort study and established multiple logistic  
88 regression models to assess the correlation between serological results of 25 viral infections

89 and the PRS of COVID-19 eight clinical phenotypes.

## 90 **2. Material and methods**

### 91 **2.1 UK Biobank Cohort**

92 The UK Biobank cohort is a large prospective cohort study that recruited approximately  
93 500,000 participants aged 40–69 between 2006 and 2010 ([https://www.ukbiobank.ac.uk/learn-](https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank)  
94 [more-about-uk-biobank](https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank)). Participant characteristics and other health-related indicators were  
95 collected through touchscreen questionnaires, short interviews, and a series of body  
96 measurements at 22 assessment centers in the United Kingdom[19]. The UK Biobank study  
97 was approved by the National Health Service National Research Ethics Service(11/NW/0382),  
98 and all participants provided written informed consent to participate[20]. This study was  
99 conducted with the permission from the UK Biobank (application number: 46,478).

100 Multi-batch genotyping was performed using two slightly different arrays, the Applied  
101 Biosystems UK BiLEVE Axiom Array from Affymetrix and the Applied Biosystems UK  
102 Biobank Axiom Array. For quality control, sex mismatches, departures from Hardy–Weinberg  
103 equilibrium, missing genotype rate  $>0.05$  or imputation accuracy score  $<0.3$  were excluded.  
104 Samples identified as outliers for heterozygosity and missing rates were removed. Detailed  
105 array design, genotyping and quality control procedures can be found in previous studies[19].

### 106 **2.2 Common Virus Serological Measurements**

107 In this study, we selected serological measurements of 25 common viruses from the UK  
108 Biobank (UK Biobank data fields: 23,050–23,071, 23,073–23,075). Detailed information was  
109 shown in the Supplemental tables (Table S1). The serological measurements were defined as  
110 positive or negative for each virus.

### 111 **2.3 Genome-wide Association Study (GWAS) Data of COVID-19 Clinical Phenotypes**

112 The GWAS data of COVID-19 were derived from a genetic association study of the expanded  
 113 phenotypic definition of COVID-19, including eight clinical phenotypes associated with  
 114 COVID-19 outcomes[5]. Detailed information of eight clinical phenotypes was presented in  
 115 Supplementary tables (Table S2). Power analysis of case-control discrete traits was performed  
 116 using the Purcell power calculator. Array-based genotyping and SNP calling were performed  
 117 by Illumina with the GenotypeStudio platform or Quest/Athena Diagnostics. Genetic principal  
 118 components were calculated to include in the association studies to control residual population  
 119 structure and were computed using FlashPCA 2.0.2. Inbred-related participants were removed  
 120 by using AncestryDNA. Details of the array design, genotyping, and quality control procedures  
 121 have been described previously[5].

### 122 **2.4 Polygenic Risk Scores (PRS) Analysis**

123 In this study, we calculated the PRS of eight COVID-19 clinical phenotypes for each subject.  
 124 SNPs of  $P < 1.00 \times 10^{-5}$  were selected. The PRS of eight COVID-19 clinical phenotype was  
 125 calculated by PLINK2.0[21], according to the formula[22]:

$$126 \quad PRS_m = \sum_{i=1}^n \beta_i SNP_{im}$$

127  $PRS_m$  represents the PRS value of COVID-19 clinical phenotypes of the  $m$ th subject;  $n$   
 128 denotes the total number of sample size;  $\beta_i$  is the effect parameter of risk allele of the  $i$ th  
 129 significant SNP associated with COVID-19 clinical phenotypes, which was obtained from the  
 130 GWAS of COVID-19 clinical phenotypes; and  $SNP_{im}$  denotes the dosage (0, 1, 2) of the risk  
 131 allele of the  $i$ th SNP for the  $m$ th individual[22, 23]. The PRS of eight COVID-19 clinical  
 132 phenotypes were used as instrumental variables to participate in the subsequent statistical



133 analysis.

## 134 **2.5 Statistical Analysis**

135 We used logistic regression models to evaluate the correlation between common viral infections  
136 and genetic predisposition to COVID-19. Serological measurements of common viruses were  
137 used as outcome variables, while the calculated PRS of eight COVID-19 clinical phenotypes  
138 was used as instrumental variable. Age, sex, Townsend deprivation index(TDI), frequency of  
139 alcohol drinking per week, frequency of smoking per day, body mass index (BMI) and 10  
140 principal components of population structure were used as covariates. We also conducted a  
141 stratified analysis of age and gender. We set a threshold of  $P < 0.05$  for suggestive significance.  
142 All statistical analyses were performed using R software.

## 143 **3. Results**

### 144 **3.1 Descriptive characteristics of study participants**

145 For *C. trachomatis* Definition II seropositivity for *Chlamydia trachomatis*, 3887 subjects were  
146 selected, 50.14% of them were woman, mean age was 57.30 years, standard deviation(SD) was  
147 7.91. For *H. pylori* Definition I seropositivity for *Helicobacter pylori*, 2394 subjects were  
148 selected, 53.88% of them were woman, mean age was 56.81 (SD:7.88) years. For the other  
149 viruses, 4800 subjects were selected, 52.92% of them were women, mean age was 56.95 (SD:  
150 7.94) years.

### 151 **3.2 Viruses Associated with COVID-19 Phenotypes in the Whole Population**

152 We first performed the logistic analysis in the whole population and identified 12 viruses  
153 associated with the PRS of COVID-19 clinical phenotypes. For example, *VZV seropositivity for*  
154 *Varicella Zoster Virus* was associated with *Unscreened/Exposed\_Negative* ( $\beta = 0.1361$ ,  $P =$

155 0.0142), and *C. trachomatis* Definition II seropositivity for *Chlamydia trachomatis* was  
156 associated with *Positive/Unscreened* ( $\beta = 0.2174$ ,  $P = 0.0185$ ). The detailed results were shown  
157 in Table 1.

### 158 3.3 Viruses Associated with COVID-19 Phenotype After Age Stratification

159 In the age  $\leq 65$  years group, we identified seven viruses associated with PRS of the COVID-19  
160 clinical phenotypes. For example, the association between *HSV-2 seropositivity for Herpes*  
161 *Simplex virus-2* and *Exposed\_Positive/Exposed\_Negative* ( $\beta = 0.1017$ ,  $P = 0.0177$ ) was  
162 significant. Besides, *HSV-2 seropositivity for Herpes Simplex virus-2* infection was also  
163 associated with *Positive/Negative* ( $\beta = -0.0987$ ,  $P = 0.0242$ ). In the age  $> 65$  years group, we  
164 found 10 viruses associated with PRS of the COVID-19 clinical phenotypes. For example, *HSV-*  
165 *2 seropositivity for Herpes Simplex virus-2* was associated with *Continuous\_Severity\_Score* ( $\beta$   
166  $= 0.3132$ ,  $P = 0.0084$ ). In addition, *HSV-2 seropositivity for Herpes Simplex virus-2* was also  
167 associated with *Symptomatic/Paucisymptomatic* ( $\beta = 0.2436$ ,  $P = 0.0426$ ). The detailed results  
168 were shown in Table 2.

### 169 3.4 Viruses Significantly Associated with COVID-19 Phenotype After Gender 170 Stratification

171 In the women group, we found five viruses associated with the PRS of COVID-19 clinical  
172 phenotypes. For example, *HPV 16 Definition II seropositivity for Human Papillomavirus type-*  
173 *16* was associated with *Exposed\_Positive /Exposed\_Negative* ( $\beta = -0.2766$ ,  $P = 0.0026$ ), and  
174 *HPV 18 seropositivity for Human Papillomavirus type-18* was associated with  
175 *Continuous\_Severity\_Score* ( $\beta = -0.3710$ ,  $P = 0.0054$ ). In the men group, we found 14 viruses  
176 associated with the PRS of COVID-19 clinical phenotypes, such as the association between

177 *HSV-2 seropositivity for Herpes Simplex virus-2 and Positive/Negative* ( $\beta = -0.1881$ ,  $P =$   
178  $0.0017$ ), and the association between *JCV seropositivity for Human Polyomavirus JCV and*  
179 *Exposed\_Positive/Exposed\_Negative* ( $\beta = 0.1237$ ,  $P = 0.0023$ ). The detailed results were  
180 shown in Table 3.

#### 181 **4. Discussion**

182 A previous study found that the severity of COVID-19 infection is influenced by host genetic  
183 factors[24]. We are curious if there is a potential correlation between COVID-19-related genetic  
184 information and the infection status of other viruses. In this work, we used PRS to represent an  
185 individual's genetic susceptibility to COVID-19. Logistic regression models were used to  
186 assess the association between multiple common viral infections and the PRS of COVID-19  
187 clinical phenotypes. Our study aimed to detect which viral infections were associated with  
188 genetic susceptibility to COVID-19.

189 Our study builds upon previous research by adding covariates to logistic regression models  
190 for correction, in order to explore whether genetic susceptibility to COVID-19 influences the  
191 risk of infection by other common viruses. We found that VZV had a significant association  
192 with *Unscreened/Exposed\_Negative* in the whole population. Herpes zoster is a viral skin  
193 disease in which herpes zoster remains dormant in the dorsal root ganglion of the cutaneous  
194 nerve after chickenpox infection[25]. Among reported COVID-19 cases, infected patients  
195 exhibited diverse skin manifestations, with varicella-like lesions being one of the major skin  
196 manifestations during the COVID-19 outbreak[26]. Cases of herpes zoster infection have been  
197 identified in recent symptomatic COVID-19 infections[27]. It is possible that this is SARS-  
198 CoV-2 could directly infect lymphocytes and promote apoptosis of lymphocytes, leading to

199 lymphopenia and impaired antiviral response, which may further favor herpes virus  
200 recurrence[28].

201 It has been found that some viruses exhibit a change in incidence with age. For example,  
202 the zoster virus demonstrates a steady increase in incidence starting at age 50 years, with the  
203 higher incidence in people over 65 years[29]. Therefore, in our study we performed a stratified  
204 analysis by controlling for age. After controlling for age variables, we found that the significant  
205 associations were not the same between the age  $\leq 65$  years group and the age  $> 65$  years group.  
206 In the age  $\leq 65$  years group, the most significant association was found between *HSV-2*  
207 *seropositivity for Herpes Simplex virus-2* and *Exposed\_Positive/Exposed\_Negative*. However,  
208 in the age  $> 65$  years group, the most significant association was found between *HSV-2*  
209 *seropositivity for Herpes Simplex virus-2* and *Continuous\_Severity\_Score*. HSV-2 causes  
210 ulcerative lesions in adults and primarily affects the genital region through sexual  
211 transmission[30]. Following primary infection, Herpes simplex virus enters the latent state in  
212 the ganglion and may emerge later, leading to recurrent active infection[31]. Recent studies  
213 been found that individuals infected with HSV could affect SARS-CoV-2 IgM/IgG serologic  
214 results due to direct binding of IgM antibodies to otherwise detected surface-modified  
215 polystyrene particles[32].

216 The prevalence of common viruses is also gender-dependent. For example, the  
217 epidemiology of HSV-2 differs between women and men, with a greater probability of  
218 transmission from male-to-female than female-to-male[33]. Therefore, we conducted a gender  
219 stratified analysis. After stratifying, we found more virus in the men group may be affected by  
220 the genetic susceptibility of COVID-19. For example, 14 significant associations were found

221 in the men group, such as HSV-2 and *Positive/Negative*, which also consist with previous  
222 studies[32]. In the women group, human papillomavirus (HPV)-associated virus infection is  
223 associated with genetic susceptibility to COVID-19. In addition, we found human herpesvirus  
224 (HHV) is also associated with genetic susceptibility to COVID-19. HHV reactivation was  
225 considered a positive polymerase chain reaction result taken at the time of COVID-19  
226 infection[34]. The reactivation of HSV is associated with an increased risk of hospital-acquired  
227 pneumonia/ventilator-associated pneumonia (HAP/VAP)[35]. Overall, co-infection with herpes  
228 viruses leads to poor clinical outcomes, particularly in critically ill COVID-19 patients[11, 36].

229 This is a new study that explores which viral infections are associated with genetic  
230 susceptibility to COVID-19. Our study finally identified several viral infections that are  
231 associated with genetic susceptibility to COVID-19. However, there are some limitations to  
232 consider when interpreting these findings. First, our data comes from a UK biobank, which only  
233 includes information from people of European descent. Therefore, our conclusions are limited  
234 in their applicability to other racial and ethnic populations. Second, our work is only exploratory,  
235 and the results can only demonstrate correlation rather than causation. Third, although we  
236 controlled for confounding factors, there may still be potential confounding factors that we did  
237 not account for. Therefore, the association between viral infections and genetic susceptibility  
238 should be interpreted with caution. Finally, more large-scale prospective and biological studies  
239 are needed to confirm our results and elucidate the specific mechanisms involved.

240 In summary, our work identified viral infections that are associated with genetic  
241 susceptibility to COVID-19. These findings may help clinicians to prevent and detect the

242 recurrence of other viruses closely related to COVID-19 in a timely manner. Moreover, this  
243 association might assist clinicians in identifying patients with a poorer prognosis.

Journal Pre-proof

244 **Funding Information**

245 This work was supported by the National Natural Scientific Foundation of China [grant  
246 numbers:81922059] and the Natural Science Basic Research Program in Shaanxi Province of  
247 China:[grant numbers:2021JCW-08].

248 **Institutional Review Board Statement**

249 This study has been approved by UK Biobank (Application number: 46478) and obtained  
250 health-related records of participants.

251 **Data Availability Statement**

252 The data that support the findings of this study are available from the corresponding author  
253 upon reasonable request.

254 **Disclosure**

255 All authors report no biomedical financial interests or potential conflicts of interest.

256 **Author contributions**

257 **Na Zhang:** Writing – original draft, Conceptualization, Formal analysis. **Yujing Chen:** Writing  
258 – review & editing, Formal analysis. **Chun'e Li:** Formal analysis. **Xiaoyue Qin:** Writing –  
259 review & editing. **Dan He:** Methodology. **Wenming Wei:** Validation. **Yijing Zhao:** Software.  
260 **Qingqing Cai:** Software. **Sirong Shi:** Visualization. **Xiaoge Chu:** Visualization. **Yan Wen:**  
261 Investigation. **Yumeng Jia:** Investigation. **Feng Zhang:** Supervision, Project administration.

262

263 **References**

- 264 [1] W.H. Organization, Coronavirus disease (COVID-19) Weekly Epidemiological Update and  
265 Weekly Operational Update, (2022) 1-25.
- 266 [2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng,  
267 T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G.  
268 Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019  
269 novel coronavirus in Wuhan, China, *Lancet* 395(10223) (2020) 497-506.
- 270 [3] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, W. Tan, Detection of SARS-CoV-2 in  
271 Different Types of Clinical Specimens, *Jama* 323(18) (2020) 1843-1844.
- 272 [4] T.P. Velavan, S.R. Pallerla, J. Rüter, Y. Augustin, P.G. Kremsner, S. Krishna, C.G. Meyer,  
273 Host genetic factors determining COVID-19 susceptibility and severity, *EBioMedicine* 72  
274 (2021) 103629.
- 275 [5] G.H.L. Roberts, R. Partha, B. Rhead, S.C. Knight, D.S. Park, M.V. Coignet, M. Zhang, N.  
276 Berkowitz, D.A. Turrisini, M. Gaddis, S.R. McCurdy, M. Pavlovic, L. Ruiz, C. Sass, A.K. Haug  
277 Baltzell, H. Guturu, A.R. Girshick, C.A. Ball, E.L. Hong, K.A. Rand, Expanded COVID-19  
278 phenotype definitions reveal distinct patterns of genetic association and protective effects, *Nat*  
279 *Genet* 54(4) (2022) 374-381.
- 280 [6] Q. Wu, Y. Xing, L. Shi, W. Li, Y. Gao, S. Pan, Y. Wang, W. Wang, Q. Xing, Coinfection and  
281 Other Clinical Characteristics of COVID-19 in Children, *Pediatrics* 146(1) (2020).
- 282 [7] D. Lin, L. Liu, M. Zhang, Y. Hu, Q. Yang, J. Guo, Y. Guo, Y. Dai, Y. Xu, Y. Cai, X. Chen,  
283 Z. Zhang, K. Huang, Co-infections of SARS-CoV-2 with multiple common respiratory  
284 pathogens in infected patients, *Sci China Life Sci* 63(4) (2020) 606-609.
- 285 [8] X. Zhu, Y. Ge, T. Wu, K. Zhao, Y. Chen, B. Wu, F. Zhu, B. Zhu, L. Cui, Co-infection with  
286 respiratory pathogens among COVID-2019 cases, *Virus Res* 285 (2020) 198005.
- 287 [9] W.J. Ouwendijk, K.J. Laing, G.M. Verjans, D.M. Koelle, T-cell immunity to human  
288 alphaherpesviruses, *Curr Opin Virol* 3(4) (2013) 452-60.
- 289 [10] B.M. Mitchell, D.C. Bloom, R.J. Cohrs, D.H. Gilden, P.G. Kennedy, Herpes simplex virus-  
290 1 and varicella-zoster virus latency in ganglia, *J Neurovirol* 9(2) (2003) 194-204.
- 291 [11] P. Le Balc'h, K. Pinceaux, C. Pronier, P. Seguin, J.M. Tadié, F. Reizine, Herpes simplex



- 292 virus and cytomegalovirus reactivations among severe COVID-19 patients, *Crit Care* 24(1)  
293 (2020) 530.
- 294 [12] A.D. Kenney, J.A. Dowdle, L. Bozzacco, T.M. McMichael, C. St Gelais, A.R. Panfil, Y.  
295 Sun, L.S. Schlesinger, M.Z. Anderson, P.L. Green, C.B. López, B.R. Rosenberg, L. Wu, J.S.  
296 Yount, Human Genetic Determinants of Viral Diseases, *Annu Rev Genet* 51 (2017) 241-263.
- 297 [13] M. Lin, H.K. Tseng, J.A. Trejaut, H.L. Lee, J.H. Loo, C.C. Chu, P.J. Chen, Y.W. Su, K.H.  
298 Lim, Z.U. Tsai, R.Y. Lin, R.S. Lin, C.H. Huang, Association of HLA class I with severe acute  
299 respiratory syndrome coronavirus infection, *BMC Med Genet* 4 (2003) 9.
- 300 [14] C. Alanio, A. Verma, D. Mathew, S. Gouma, G. Liang, T. Dunn, D.A. Oldridge, J. Weaver,  
301 L. Kuri-Cervantes, M.B. Pampena, M.R. Betts, R.G. Collman, F.D. Bushman, N.J. Meyer, S.E.  
302 Hensley, D. Rader, E.J. Wherry, Cytomegalovirus Latent Infection is Associated with an  
303 Increased Risk of COVID-19-Related Hospitalization, *J Infect Dis* 226(3) (2022) 463-473.
- 304 [15] J.T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P.M. de Salazar, B.J. Cowling,  
305 M. Lipsitch, G.M. Leung, Estimating clinical severity of COVID-19 from the transmission  
306 dynamics in Wuhan, China, *Nat Med* 26(4) (2020) 506-510.
- 307 [16] M.J. Cannon, D.S. Schmid, T.B. Hyde, Review of cytomegalovirus seroprevalence and  
308 demographic characteristics associated with infection, *Rev Med Virol* 20(4) (2010) 202-13.
- 309 [17] D. Furman, V. Jovic, S. Sharma, S.S. Shen-Orr, C.J. Angel, S. Onengut-Gumuscu, B.A.  
310 Kidd, H.T. Maecker, P. Concannon, C.L. Dekker, P.G. Thomas, M.M. Davis, Cytomegalovirus  
311 infection enhances the immune response to influenza, *Sci Transl Med* 7(281) (2015) 281ra43.
- 312 [18] E. Patin, M. Hasan, J. Bergstedt, V. Rouilly, V. Libri, A. Urrutia, C. Alanio, P. Scepovic,  
313 C. Hammer, F. Jönsson, B. Beitz, H. Quach, Y.W. Lim, J. Hunkapiller, M. Zepeda, C. Green, B.  
314 Piasecka, C. Leloup, L. Rogge, F. Huetz, I. Peguillet, O. Lantz, M. Fontes, J.P. Di Santo, S.  
315 Thomas, J. Fellay, D. Duffy, L. Quintana-Murci, M.L. Albert, Natural variation in the  
316 parameters of innate immune cells is preferentially driven by genetic factors, *Nat Immunol* 19(3)  
317 (2018) 302-314.
- 318 [19] C. Bycroft, C. Freeman, D. Petkova, G. Band, L.T. Elliott, K. Sharp, A. Motyer, D.  
319 Vukcevic, O. Delaneau, J. O'Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. McVean,  
320 S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK Biobank resource with deep phenotyping  
321 and genomic data, *Nature* 562(7726) (2018) 203-209.

- 322 [20] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J.  
323 Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T.  
324 Peakman, R. Collins, UK biobank: an open access resource for identifying the causes of a wide  
325 range of complex diseases of middle and old age, *PLoS Med* 12(3) (2015) e1001779.
- 326 [21] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P.  
327 Sklar, P.I. de Bakker, M.J. Daly, P.C. Sham, PLINK: a tool set for whole-genome association  
328 and population-based linkage analyses, *Am J Hum Genet* 81(3) (2007) 559-75.
- 329 [22] S. Cheng, X. Qi, M. Ma, L. Zhang, B. Cheng, C. Liang, L. Liu, P. Li, O.P. Kafle, Y. Wen,  
330 F. Zhang, Assessing the Relationship Between Gut Microbiota and Bone Mineral Density, *Front*  
331 *Genet* 11 (2020) 6.
- 332 [23] F. Dudbridge, Polygenic Epidemiology, *Genet Epidemiol* 40(4) (2016) 268-72.
- 333 [24] D. Zhu, R. Zhao, H. Yuan, Y. Xie, Y. Jiang, K. Xu, T. Zhang, X. Chen, C. Suo, Host Genetic  
334 Factors, Comorbidities and the Risk of Severe COVID-19, *J Epidemiol Glob Health* (2023)  
335 1-13.
- 336 [25] R.R. Dayan, R. Peleg, Herpes zoster - typical and atypical presentations, *Postgrad Med*  
337 129(6) (2017) 567-571.
- 338 [26] U. Wollina, A.S. Karadağ, C. Rowland-Payne, A. Chiriac, T. Lotti, Cutaneous signs in  
339 COVID-19 patients: A review, *Dermatol Ther* 33(5) (2020) e13549.
- 340 [27] A. Ferreira, T.T. Romão, Y.S. Macedo, C. Pupe, O.J.M. Nascimento, COVID-19 and herpes  
341 zoster co-infection presenting with trigeminal neuropathy, *Eur J Neurol* 27(9) (2020) 1748-  
342 1750.
- 343 [28] H.Y. Zheng, M. Zhang, C.X. Yang, N. Zhang, X.C. Wang, X.P. Yang, X.Q. Dong, Y.T.  
344 Zheng, Elevated exhaustion levels and reduced functional diversity of T cells in peripheral  
345 blood may predict severe progression in COVID-19 patients, *Cell Mol Immunol* 17(5) (2020)  
346 541-543.
- 347 [29] K. Kawai, B.G. Gebremeskel, C.J. Acosta, Systematic review of incidence and  
348 complications of herpes zoster: towards a global perspective, *BMJ Open* 4(6) (2014) e004833.
- 349 [30] A.J. Nahmias, B. Roizman, Infection with herpes-simplex viruses 1 and 2. II, *N Engl J*  
350 *Med* 289(14) (1973) 719-25.
- 351 [31] D.T. Fleming, G.M. McQuillan, R.E. Johnson, A.J. Nahmias, S.O. Aral, F.K. Lee, M.E. St

- 352 Louis, Herpes simplex virus type 2 in the United States, 1976 to 1994, *N Engl J Med* 337(16)  
353 (1997) 1105-11.
- 354 [32] L. Vandervore, E. Van Mieghem, V. Nowé, S. Schouwers, C. Steger, P. Abrams, J. Van  
355 Schaeren, A. Meskal, T. Vandamme, False positive Herpes Simplex IgM serology in COVID-  
356 19 patients correlates with SARS-CoV-2 IgM/IgG seropositivity, *Diagn Microbiol Infect Dis*  
357 103(1) (2022) 115653.
- 358 [33] A. Wald, A.G. Langenberg, K. Link, A.E. Izu, R. Ashley, T. Warren, S. Tying, J.M.  
359 Douglas, Jr., L. Corey, Effect of condoms on reducing the transmission of herpes simplex virus  
360 type 2 from men to women, *Jama* 285(24) (2001) 3100-6.
- 361 [34] A. Shafiee, M.M. Teymouri Athar, M.J. Amini, H. Hajishah, S. Siahvoshi, M. Jalali, B.  
362 Jahanbakhshi, S.H. Mozhgani, Reactivation of herpesviruses during COVID-19: A systematic  
363 review and meta-analysis, *Rev Med Virol* 33(3) (2023) e2437.
- 364 [35] A. Meyer, N. Buetti, N. Houhou-Fidouh, J. Patrier, M. Abdel-Nabey, P. Jaquet, S. Presente,  
365 T. Girard, F. Sayagh, S. Ruckly, P.H. Wicky, E. de Montmollin, L. Bouadma, R. Sonnevile, D.  
366 Descamps, J.F. Timsit, HSV-1 reactivation is associated with an increased risk of mortality and  
367 pneumonia in critically ill COVID-19 patients, *Crit Care* 25(1) (2021) 417.
- 368 [36] J. Katz, S. Yue, W. Xue, Herpes simplex and herpes zoster viruses in COVID-19 patients,  
369 *Ir J Med Sci* 191(3) (2022) 1093-1097.

370

371 **Legends for Figures and Tables**

372 **Figure 1:** Forest plots of the results of logistic regression analysis in the whole population.

373 (a) *Continuous\_Severity\_Score*;

374 (b) *Exposed\_Positive/Exposed\_Negative*;

375 (c) *Hospitalized/Not\_Hospitalized*;

376 (d) *Hospitalized/Unscreened*

377 Note: Odds ratio (OR) forest plot of the PRS of COVID-19 clinical phenotypes . Use forest plot  
378 to visualize logistic regression analysis. The outcome variables were serological measurements  
379 of 25 viruses. The instrumental variables were the PRS of COVID-19 clinical phenotypes.

380 **Figure 2:** Forest plots of the results of logistic regression analysis in the whole population.

381 (a) *Positive/Negative*;

382 (b) *Positive/Unscreened*;

383 (c) *Symptomatic/Paucisymptomatic*;

384 (d) *Unscreened/Exposed\_Negative*

385 Note: Odds ratio (OR) forest plot of the PRS of COVID-19 clinical phenotypes . Use forest plot  
386 to visualize logistic regression analysis. The outcome variables were serological measurements  
387 of 25 viruses. The instrumental variables were the PRS of COVID-19 clinical phenotypes.

388 **Table 1:** Viral infections associated with COVID-19 clinical phenotypes in the whole  
389 population

390 **Table 2:** Viral infections associated with COVID-19 clinical phenotypes in the age < 65 years  
391 group and the age > 65 years group

392 **Table 3:** Viral infections associated with COVID-19 clinical phenotypes in the women group

393 and the men group

Journal Pre-proof

Table 1: Viral infections associated with COVID-19 clinical phenotypes in whole population

Outcome Variable	Instrumental Variable	$\beta$	$P$
Unscreened/Exposed_Negative	VZV seropositivity for Varicella Zoster Virus	0.1361	0.0142
Positive/Unscreened	C. trachomatis Definition II seropositivity for Chlamydia trachomatis	0.2174	0.0185
Exposed_Positive/Exposed_Negative	HSV-2 seropositivity for Herpes Simplex virus-2	0.0935	0.0192
Positive/Negative	HSV-1 seropositivity for Herpes Simplex virus-1	0.0718	0.0213
Exposed_Positive/Exposed_Negative	HBV seropositivity for Hepatitis B Virus	0.2722	0.0219
Hospitalized/Unscreened	VZV seropositivity for Varicella Zoster Virus	0.1167	0.0385
Positive/Unscreened	HCV seropositivity for Hepatitis C Virus	0.7220	0.0398
Exposed_Positive/Exposed_Negative	KSHV seropositivity for Kaposi's Sarcoma-Associated Herpesvirus	-0.1044	0.0444
Positive/Negative	HSV-2 seropositivity for Herpes Simplex virus-2	-0.0814	0.0446
Unscreened/Exposed_Negative	MCV seropositivity for Merkel Cell Polyomavirus	-0.0614	0.0478
Unscreened/Exposed_Negative	T. gondii seropositivity for Toxoplasma gondii	0.0677	0.0481
Unscreened/Exposed_Negative	HSV-2 seropositivity for Herpes Simplex virus-2	0.0817	0.0493

Note: The threshold of significance is  $P < 0.05$ .

Table 2: Viral infections associated with COVID-19 clinical phenotypes in the age  $\leq$  65 years group and the age  $>$  65 years group

Group	Outcome Variable	Instrumental Variable	$\beta$	$P$
Age $\leq$ 65 years	Exposed_Positive/Exposed_Negative	HSV-2 seropositivity for Herpes Simplex virus-2	0.1017	0.0177
	Positive/Negative	HHV-6B seropositivity for Human Herpesvirus-6	-0.0869	0.0226
	Positive/Negative	HSV-2 seropositivity for Herpes Simplex virus-2	-0.0987	0.0242
	Positive/Negative	HSV-1 seropositivity for Herpes Simplex virus-1	0.0748	0.0270
	Exposed_Positive/Exposed_Negative	KSHV seropositivity for Kaposi's Sarcoma-Associated Herpesvirus	-0.1208	0.0353
	Unscreened/Exposed_Negative	VZV seropositivity for Varicella Zoster Virus	0.1206	0.0432
	Hospitalized/Not_Hospitalized	HBV seropositivity for Hepatitis B Virus	-0.2565	0.0464
Age $>$ 65 years	Continuous_Severity_Score	HSV-2 seropositivity for Herpes Simplex virus-2	0.3132	0.0084
	Positive/Negative	HHV-6B seropositivity for Human Herpesvirus-6	0.2150	0.0149
	Hospitalized/Unscreened	KSHV seropositivity for Kaposi's Sarcoma-Associated Herpesvirus	-0.3371	0.0171
	Unscreened/Exposed_Negative	BKV seropositivity for Human Polyomavirus BKV	0.3491	0.0208
	Continuous_Severity_Score	HHV-6 overall seropositivity for Human Herpesvirus-6	-0.2731	0.0323
	Exposed_Positive/Exposed_Negative	HPV 18 seropositivity for Human Papillomavirus type-18	-0.5712	0.0343
	Hospitalized/Not_Hospitalized	C. trachomatis Definition I seropositivity for Chlamydia trachomatis	-0.2309	0.0367
	Symptomatic/Paucisymptomatic	HPV 16 Definition I seropositivity for Human Papillomavirus type-16	-0.5834	0.0368
	Symptomatic/Paucisymptomatic	HSV-2 seropositivity for Herpes Simplex virus-2	0.2436	0.0426
	Positive/Unscreened	HHV-6B seropositivity for Human Herpesvirus-6	0.1801	0.0465

Note: The threshold of significance is  $P < 0.05$ .

Table 3: Viral infections associated with COVID-19 clinical phenotypes in the women group and the men group

Group	Outcome Variable	Instrumental Variable	$\beta$	$P$
Women	Exposed_Positive/Exposed_Negative	HPV 16 Definition II seropositivity for Human Papillomavirus type-16	-0.2766	0.0026
	Continuous_Severity_Score	HPV 18 seropositivity for Human Papillomavirus type-18	-0.3710	0.0054
	Hospitalized/Not_Hospitalized	HHV-7 seropositivity for Human Herpesvirus-7	-0.2380	0.0230
	Symptomatic/Paucisymptomatic	HPV 16 Definition II seropositivity for Human Papillomavirus type-16	-0.2105	0.0252
	Hospitalized/Unscreened	BKV seropositivity for Human Polyomavirus BKV	-0.1802	0.0356
Men	Positive/Negative	HSV-2 seropositivity for Herpes Simplex virus-2	-0.1881	0.0017
	Exposed_Positive/Exposed_Negative	JCV seropositivity for Human Polyomavirus JCV	0.1237	0.0023
	Positive/Negative	HHV-6 overall seropositivity for Human Herpesvirus-6	-0.1986	0.0048
	Positive/Unscreened	C. trachomatis Definition II seropositivity for Chlamydia trachomatis	0.3021	0.0078
	Unscreened/Exposed_Negative	T. gondii seropositivity for Toxoplasma gondii	0.1269	0.0122
	Positive/Negative	KSHV seropositivity for Kaposi's Sarcoma-Associated Herpesvirus	-0.1936	0.0126
	Unscreened/Exposed_Negative	H. pylori Definition I seropositivity for Helicobacter pylori	0.1633	0.0162
	Positive/Unscreened	HPV 16 Definition II seropositivity for Human Papillomavirus type-16	0.2164	0.0256
	Positive/Negative	HSV-1 seropositivity for Herpes Simplex virus-1	0.0989	0.0285
	Positive/Negative	HHV-6A seropositivity for Human Herpesvirus-6	-0.1090	0.0308
	Exposed_Positive/Exposed_Negative	H. pylori Definition I seropositivity for Helicobacter pylori	0.1312	0.0333
	Unscreened/Exposed_Negative	VZV seropositivity for Varicella Zoster Virus	0.1913	0.0364
	Hospitalized/Not_Hospitalized	JCV seropositivity for Human Polyomavirus JCV	-0.0876	0.0417
	Exposed_Positive/Exposed_Negative	HSV-2 seropositivity for Herpes Simplex virus-2	0.1158	0.0445

Note: The threshold of significance is  $P < 0.05$ .



