1	Polygenic Associations and Causal Inferences between Serum
2	Immunoglobulins and Amyotrophic Lateral Sclerosis
3	Xu Chen, PhD <sup>1,2</sup> ; Xiaojun Shen, BSc <sup>2</sup> ; Yiqiang Zhan, PhD <sup>3,4</sup> ; Fang Fang, MD, PhD <sup>3</sup>
4	From the <sup>1</sup> Department of Central Laboratory and <sup>2</sup> Institute for Maternal-Fetal Medicine, Baoan
5	Women's and Children's Hospital, Jinan University, Shenzhen, China; <sup>3</sup> Institute of Environmental
6	Medicine, Karolinska Institutet, Stockholm, Sweden; <sup>4</sup> German Center for Neurodegenerative
7	Diseases (DZNE), Ulm, Germany.
8	
9	Corresponding author: Dr. Xu Chen. Department of Central Laboratory, Baoan Women's and
10	Children's Hospital, No.56 YuLv Road, 518100 Shenzhen, China. Phone: +86-15989385199. Email:
11	dr.xuchen@outlook.com.
12	
13	Running head: Immunoglobulins and ALS
14	Number of words in the abstract: 212
15	Number of words in the main text: 1505
16	Number of figures: 0
17	Number of tables: 2
18	

#### PAGE 2

# 19 Abstract

Chronic inflammation might contribute to the development of amyotrophic lateral sclerosis (ALS). 20 The relationship between serum immunoglobulins and risk of ALS remains however greatly 21 unknown. In order to overcome limitations like reverse causation and residual confounding 22 commonly seen in observational studies, we applied polygenic risk score (PRS) and Mendelian 23 randomization (MR) analyses on summary statistics from the large-scale genome-wide association 24 studies (GWAS), to examine the polygenic and causal associations between three serum 25 immunoglobulins (IgA, IgM, and IgG) and risk of ALS (first in a discovery phase and then in a 26 27 replication phase). An inverse polygenic association was discovered between IgA and ALS as well as between IgM and ALS. Such associations were however not replicated using a larger ALS GWAS 28 and no causal association was observed for either IgA-ALS or IgM-ALS. A positive polygenic 29 30 association was both discovered [odds ratio (OR) = 1.18; 95% confidence interval (CI): 1.12-1.25,  $P=5.9\times10^{-7}$ ] and replicated (OR=1.13, 95% CI: 1.06-1.20, P=0.001) between IgG and ALS. A causal 31 association between IgG and ALS was also suggested in both the discovery (OR=1.06, 95%CI: 1.02-32 1.10, P=0.009) and replication (OR=1.07, 95%CI: 0.90-1.24, P=0.420) analyses, although the latter 33 was not statistically significant. This study suggests a shared polygenic risk between serum IgG (as a 34 biomarker for chronic inflammation) and ALS. 35

#### PAGE 3

# 36 Introduction

Amyotrophic lateral sclerosis (ALS) is a relatively rare but fatal neurodegenerative disease. Chronic 37 inflammation and altered immune responses have been suggested to contribute to the pathogenesis 38 of ALS.<sup>1, 2</sup> The relationship between serum immunoglobulins, as known biomarkers for 39 inflammation and immune responses, and risk of ALS remains however greatly unknown. In a 40 previous study, we did not find a statistically significant positive association between 41 immunoglobulin G (IgG) and risk of ALS [hazard ratio=1.04; 95% confidence interval (CI): 0.89-42 1.22; per 2.99 g/L increase of IgG].<sup>3</sup> In the same study, we found also that during the 20 years before 43 diagnosis, patients with ALS had more rapidly declining levels of IgG compared to others.<sup>3</sup> We 44 speculate that the noted lack of association between IgG and ALS might be attributable to both a real 45 positive association between IgG and ALS (i.e., higher levels of IgG lead to higher risk of ALS) and 46 reverse causation (i.e., ALS results in declined levels of IgG). In addition to reverse causation, 47 observational studies are also prone to other systemic errors, which collectively make it often 48 difficult to infer causality for associations noted in observational studies.<sup>4</sup> 49

By using the summary statistics from large-scale genome-wide association studies (GWAS), 50 polygenic risk score (PRS) and Mendelian randomization (MR) analyses have added novel evidence 51 to support polygenic and causal relationships between several traits and ALS, corroborating 52 importantly findings from previous observational studies.<sup>5-7</sup> In this study, we aimed to examine the 53 polygenic association and infer the causality between IgG and risk of ALS, using summary statistics 54 55 from GWAS of serum immunoglobulins and ALS. Because IgG is a biomarker for chronic inflammation, we hypothesized that there is a positive and causal relationship between IgG and ALS. 56 57 We additionally studied another two serum immunoglobulins (IgA and IgM) to assess the specificity 58 of the IgG result. The roles of these two biomarkers have rarely been studied in ALS.

PAGE 4

### 59 Materials and Methods

### 60 Summary statistics

We used publicly available summary statistics of several GWAS within the European ancestry 61 population in this study. GWAS of serum immunoglobulins (IgA, IgM, and IgG) were conducted 62 among ~15000 individuals, from which single nucleotide polymorphisms (SNPs) with the 63 association P-values <1.0×10<sup>-6</sup> were available.<sup>8</sup> GWAS summary statistics of ALS were obtained 64 from two studies (available in the IEU GWAS database; https://gwas.mrcieu.ac.uk, ID for 65 ALS 2016: ieu-a-1085; ID for ALS 2018: ebi-a-GCST005647). In the discovery analysis, summary 66 statistics of ALS (ALS 2016) were obtained from a GWAS performed among 36052 individuals 67 (including 12577 patients with ALS and 23475 controls) by the Project MinE group.<sup>9</sup> In the 68 replication analysis, summary statistics of ALS (ALS 2018) were obtained from a meta-analysis 69 incorporating the results from ALS 2016 and another GWAS including 44558 individuals (8229 70 patients with ALS and 32369 controls).<sup>10</sup> 71

Using IgA, IgM, and IgG as the exposures (base traits) and ALS as the outcome (target trait), the polygenic and causal associations of six exposure-outcome pairs (three in the discovery analysis and three in the replication analysis) were investigated by PRS and MR analyses, respectively. Because the values of immunoglobulins were standardized in the GWAS, we estimated accordingly odds ratio (OR) per standard deviation of genetically determined level for each type of immunoglobulin in the present study.

78

79 *PRS analyses* 

Polygenic associations were investigated by PRS analyses in Linux with PRSice package (version
1.25).<sup>11</sup> For the summary statistics of each exposure, correlated SNPs were pruned by linkage

#### PAGE 5

disequilibrium (LD) clumping with default parameter settings. Genotypes of the HapMap\_ceu\_all (release 22, 60 individuals, 3.96 million SNPs) were used as the reference panel, clumping threshold  $p_1=p_2=0.5$ , LD threshold  $r^2=0.05$ , and the distance threshold of 300Kb. Independent SNPs were then extracted and included in 100000 quantiles with gradually increasing P-value thresholds (P<sub>T</sub>, ranging from 0 to  $1.0 \times 10^{-6}$ , in steps of  $1.0 \times 10^{-11}$  per quantile). The P<sub>T</sub> of the quantile that explained the largest variance of the target was defined as the best-fitted P<sub>T</sub>.

88

## 89 MR analyses

Causal inferences were performed by MR analyses in R (version 3.6.1) with TwoSampleMR 90 package (version 0.5.2).<sup>12</sup> Lead SNPs from the genetic loci reported to be associated with each 91 exposure were extracted from the GWAS of immunoglobulins.<sup>8</sup> The associations of the SNPs with 92 other traits (associations with a P-value  $<5\times10^{-8}$ ) were checked in GWAS Catalog 93 (https://www.ebi.ac.uk/gwas) and excluded, leaving independent lead SNPs after LD clumping as 94 the instrumental variables (IVs, Supplementary Table 1). The causal relationship of each exposure-95 outcome pair was primarily examined by the inverse variance weighted (IVW) method and 96 complemented with another four methods (MR Egger regression, weighted median, simple and 97 weighted mode) as sensitivity analyses. 98

99

100

### 101 **Results**

Inverse polygenic associations were discovered for IgA-ALS and IgM-ALS (Table 1). Neither of the associations was however replicated (Supplementary Figures 1-2). A positive polygenic association between IgG and ALS was found both in the discovery and replication analyses (Supplementary

#### PAGE 6

105	Figure 3). The OR was 1.18 (95% CI: 1.12-1.25, $P=5.9\times10^{-7}$ ) in the discovery analysis and 1.12
106	(95%CI: 1.06-1.20, P=0.001) in the replication analysis.

In the MR analyses, causal association was not found for either IgA-ALS or IgM-ALS (Table 2), regardless of the use of MR methods (Supplementary Table 2 and Supplementary Figures 4-5). For IgG-ALS, a statistically significant causal association was found in the discovery analysis (OR=1.06, 95% CI: 1.02-1.10, P=0.009). A causal positive association was also suggested in the replication analysis, although not statistically significant (OR=1.07, 95%CI: 0.90-1.24, P=0.420).

113

## 114 **Discussion**

Utilizing PRS and MR analyses of large-scale GWAS of immunoglobulins and ALS, we provided new evidence supporting a potentially causal positive association between IgG, but not IgA or IgM, and risk of ALS.

Accumulating evidence supports that inflammation and altered immune responses are 118 involved in the different phases of ALS.<sup>13</sup> Physiological barriers (e.g., blood-brain barrier) and 119 120 innate immunity cells (e.g., macrophages) are fundamental in the maintenance and modulation of the central nervous system, whereas chronic inflammation might lead to dysfunction of the neurons.<sup>14</sup> B 121 cells play important roles in adaptive immunity, through producing immunoglobulins and presenting 122 antigens to T cells.<sup>15</sup> IgA constitutes 10-20% of serum immunoglobulins and is mainly involved in 123 124 the mucosal immune response. IgM constitutes 5-10% of serum immunoglobulins and is the earliest released antibody during adaptive immune response. IgG constitutes ~75% of serum 125 immunoglobulins and can be transmitted through placental and blood-brain barriers.<sup>16</sup> Few studies 126 127 have however examined the roles of these immunoglobulins in ALS to our knowledge. In our

#### PAGE 7

previous cohort study including more than half a million participants with >20 years of follow-up, we found that per 2.99 g/L increase of IgG levels in blood, there was a 4% increased risk of ALS, although the association was not statistically significant.<sup>3</sup> The lack of statistical significance might be multifactorial, including first a potentially real positive association between IgG and ALS risk, a lack of statistical power due to the relatively small number of ALS cases identified (N=152), and also reverse causation (i.e., patients with ALS had more rapidly declining IgG levels, compared to individuals not developing ALS, during the 20 years before ALS diagnosis).<sup>3</sup>

By using summary statistics from several large-scale GWAS, in the present study, we 135 136 identified a statistically significant positive polygenic association between IgG and ALS in both the discovery and replication analyses. The association is further suggested to be causal in the MR 137 analysis. Although statistically significant result was only obtained in the discovery analysis, similar 138 point estimate was noted in the replication analysis. The differences between the discovery 139 (ALS 2016) and replication (ALS 2018) samples might have contributed to the lack of statistical 140 significance in the replication analysis. For instance, in the discovery GWAS (ALS 2016), patients 141 with ALS and controls were matched for age, sex, and geographic regions, whereas in the replication 142 GWAS (ALS 2018) >50% of participants were only matched for race/ethnicity. In contrast to IgG, 143 no clear polygenic association nor causal relationship was noted for IgA-ALS and IgM-ALS. This 144 different result pattern might be partly explained by the fact that IgG is a biomarker for chronic 145 inflammation and bypasses the blood-brain-barrier, which might be more relevant for 146 neurodegenerative diseases including ALS, compared to the other immunoglobulins.<sup>17</sup> 147

Causal inference by MR analysis requires three basic assumptions: 1) relevance, i.e., IVs have causal effects on the exposure such as IgG; 2) exclusion restriction, i.e., IVs affect the outcome, such as ALS, only through the exposure; and 3) exchangeability or independence, i.e., no common

#### PAGE 8

151	causes between IVs and the outcome. <sup>18</sup> Even though we used the latest and largest GWAS of
152	immunoglobulins, only two independent SNPs could be selected as IVs for IgG. As a result, we
153	could only use IVW methods in the MR analysis. According to the current GWAS Catalog, no
154	genome-wide significant association was identified between the two IVs of IgG and other non-ALS
155	traits. Interestingly, the two genes near these two IVs (FCGR2B near rs7554873 and HLA-DQB1
156	near rs2133037) are both biologically related to immune responses, neuroinflammation, and
157	neurodegeneration. <sup>19, 20</sup>
158	In conclusion, our study suggests a shared polygenic risk between serum IgG and ALS. The
159	causality between IgG and ALS needs to be further validated when more instrumental variables are
160	available and when different methods of causal inference are possible to use.
161	
162	
163	Acknowledgment
164	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the
164 165	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research
164 165 166	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic
164 165 166 167	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).
164 165 166 167 168	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).
164 165 166 167 168 169	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).
164 165 166 167 168 169 170	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).  Author Contributions X.C. and F.F. contributed to the conception and design of the study; X.C., X.S., and Y.Z. contributed
164 165 166 167 168 169 170 171	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).  Author Contributions X.C. and F.F. contributed to the conception and design of the study; X.C., X.S., and Y.Z. contributed to acquisition and analyses of data; X.C., X.S., Y.Z., and F.F. contributed to drafting and revising the
164 165 166 167 168 169 170 171 172	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology). Author Contributions X.C. and F.F. contributed to the conception and design of the study; X.C., X.S., and Y.Z. contributed to acquisition and analyses of data; X.C., X.S., Y.Z., and F.F. contributed to drafting and revising the manuscript.
<ol> <li>164</li> <li>165</li> <li>166</li> <li>167</li> <li>168</li> <li>169</li> <li>170</li> <li>171</li> <li>172</li> <li>173</li> </ol>	The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS. This project is partly supported by the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).  Author Contributions X.C. and F.F. contributed to the conception and design of the study; X.C., X.S., and Y.Z. contributed to acquisition and analyses of data; X.C., X.S., Y.Z., and F.F. contributed to drafting and revising the manuscript.

PAGE 9

# 174 **Potential Conflicts of Interest**

175 Noting to report.

176

177 **References** 

178 1. Beers DR, Appel SH. Immune dysregulation in amyotrophic lateral sclerosis: mechanisms 179 and emerging therapies. Lancet Neurol. 2019 Feb;18(2):211-20.

Hammond TR, Marsh SE, Stevens B. Immune Signaling in Neurodegeneration. Immunity.
 2019 Apr 16;50(4):955-74.

Yazdani S, Mariosa D, Hammar N, et al. Peripheral immune biomarkers and
 neurodegenerative diseases: A prospective cohort study with 20 years of follow-up. Ann Neurol.
 2019 Dec;86(6):913-26.

Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology:
 the need for a pluralistic approach. Int J Epidemiol. 2016 Dec 1;45(6):1776-86.

187 5. Chen X, Yazdani S, Piehl F, Magnusson PKE, Fang F. Polygenic link between blood lipids and
 188 amyotrophic lateral sclerosis. Neurobiol Aging. 2018 Jul;67:202 e1- e6.

189 6. Bandres-Ciga S, Noyce AJ, Hemani G, et al. Shared polygenic risk and causal inferences in 190 amyotrophic lateral sclerosis. Ann Neurol. 2019 Apr;85(4):470-81.

1917.Zhan Y, Fang F. Smoking and amyotrophic lateral sclerosis: A mendelian randomization192study. Ann Neurol. 2019 Apr;85(4):482-4.

193 8. Jonsson S, Sveinbjornsson G, de Lapuente Portilla AL, et al. Identification of sequence
 194 variants influencing immunoglobulin levels. Nat Genet. 2017 Aug;49(8):1182-91.

195 9. van Rheenen W, Shatunov A, Dekker AM, et al. Genome-wide association analyses identify
196 new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet. 2016
197 Sep;48(9):1043-8.

10. Nicolas A, Kenna KP, Renton AE, et al. Genome-wide Analyses Identify KIF5A as a Novel ALS
Gene. Neuron. 2018 Mar 21;97(6):1268-83 e6.

200 11. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics.
2015 May 1;31(9):1466-8.

Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal
 inference across the human phenome. Elife. 2018 May 30;7.

### PAGE 10

13. Molteni M, Rossetti C. Neurodegenerative diseases: The immunological perspective. J
 Neuroimmunol. 2017 Dec 15;313:109-15.

14. Novellino F, Sacca V, Donato A, et al. Innate Immunity: A Common Denominator between
 Neurodegenerative and Neuropsychiatric Diseases. Int J Mol Sci. 2020 Feb 7;21(3).

- Sabatino JJ, Jr., Probstel AK, Zamvil SS. B cells in autoimmune and neurodegenerative
   central nervous system diseases. Nat Rev Neurosci. 2019 Dec;20(12):728-45.
- 210 16. Chen K, Magri G, Grasset EK, Cerutti A. Rethinking mucosal antibody responses: IgM, IgG
   211 and IgD join IgA. Nat Rev Immunol. 2020 Feb 3.
- Watson N, Ding B, Zhu X, Frisina RD. Chronic inflammation inflammaging in the ageing
   cochlea: A novel target for future presbycusis therapy. Ageing Res Rev. 2017 Nov;40:142-8.
- Armon C. Smoking is a cause of amyotrophic lateral sclerosis. High low-density lipoprotein
   cholesterol levels? Unsure. Ann Neurol. 2019 Apr;85(4):465-9.
- 216 19. Choi YR, Kang SJ, Kim JM, et al. FcgammaRIIB mediates the inhibitory effect of aggregated
   217 alpha-synuclein on microglial phagocytosis. Neurobiol Dis. 2015 Nov;83:90-9.
- 218 20. Hallberg P, Smedje H, Eriksson N, et al. Pandemrix-induced narcolepsy is associated with 219 genes related to immunity and neuronal survival. EBioMedicine. 2019 Feb;40:595-604.
- 220

PAGE 11

# 221 TABLES

Base	Target	<b>Best-fitted P</b> <sub>T</sub>	nsnp	OR	SE	$r^2(\%)$	P-value	-
IgA	ALS_2016	8.6000×10 <sup>-10</sup>	54	0.861	0.032	0.059	4.2×10 <sup>-6</sup>	-
	ALS_2018	7.0001×10 <sup>-7</sup>	43	0.948	0.031	0.004	0.085	
IgM	ALS_2016	4.8000×10 <sup>-10</sup>	47	0.905	0.032	0.028	0.002	
	ALS_2018	3.5010×10 <sup>-8</sup>	42	1.051	0.028	0.004	0.074	
IgG	ALS_2016	5.1001×10 <sup>-7</sup>	49	1.181	0.033	0.069	5.9×10 <sup>-7</sup>	
	ALS_2018	5.7001×10 <sup>-7</sup>	22	1.130	0.036	0.014	0.001	

### 222 TABLE 1. Polygenic Associations between Serum Immunoglobulins and ALS in PRS Analyses

PRS=polygenic risk score; ALS=amyotrophic lateral sclerosis; Best-fitted  $P_T$ =association P-value threshold of the quantile that explained the largest variance of the target;  $n_{SNP}$ =numbers of independent single nucleotide polymorphisms (SNPs) included in the best-fitted quantile; OR: odds ratio per standard deviation of genetically determined level for each type of immunoglobulin; SE=standard error; r<sup>2</sup>= the proportion of the target variation explained by SNPs in the best-fitted quantile.

229

### PAGE 12

230

Base	Target	n <sub>SNP</sub>	OR <sub>IVW</sub>	SEIVW	P <sub>IVW</sub>	Het_P <sub>IVW</sub>	
IgA	ALS_2016	6	0.968	0.022	0.148	0.206	
	ALS_2018	6	0.880	0.100	0.201	0.125	
IgM	ALS_2016	5	0.996	0.015	0.787	0.574	
	ALS_2018	5	1.021	0.054	0.706	0.874	
IgG	ALS_2016	2	1.056	0.021	0.009	0.617	
	ALS_2018	2	1.073	0.087	0.420	0.407	

# 231 TABLE 2. Causal Inferences between Serum Immunoglobulins and ALS in the MR Analyses

232 MR=mendelian randomization; ALS=amyotrophic lateral sclerosis; n<sub>SNP</sub>= numbers of independent

233 single nucleotide polymorphisms (SNPs) uses as instrumental variables; OR: odds ratio per standard

234 deviation of genetically determined level for each type of immunoglobulin; SE=standard error;

235 IVW= inverse variance weighted method; Het\_P=P-value of heterogeneity.