Polygenic associations and causal inferences between serum immunoglobulins and amyotrophic lateral sclerosis

Xu Chen[1,2]*, Xiaojun Shen[2], Xuzhuo Zhang[1], Yiqiang Zhan[3,4] and Fang Fang[4]

[1]Department of Central Laboratory, Shenzhen Baoan Women's and Children's Hospital, Jinan University, Shenzhen, China.
[2]Institute of Maternal-Fetal Medicine, Shenzhen Baoan Women's and Children's Hospital, Jinan University, Shenzhen, China.
[3]German Center for Neurodegenerative Diseases (DZNE), Ulm, Germany.
[4]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

Correspondence:
Dr. Xu Chen. Department of Central Laboratory, Shenzhen Baoan Women's and Children's Hospital, Jinan University, 56 YuLv Road, Shenzhen 518100, China. Phone: +86-15989385199; Email: dr.xuchen@outlook.com

Abstract

Chronic inflammation might contribute to the development of amyotrophic lateral sclerosis (ALS), the relationship between serum immunoglobulins and risk of ALS remains however unclear. In order to overcome limitations like reverse causation and residual confounding commonly seen in the observational studies, we applied molecular epidemiological analyses to examine the polygenic and causal associations between serum immunoglobulins and ALS. Summary statistics from the large-scale genome-wide association studies (GWAS) among European ancestry populations (~15000 individuals for serum immunoglobulins, and more than 36000 individuals for ALS) were accessed from different consortia. The relationships between three types of serum immunoglobulins (IgA, IgM, and IgG) and ALS were investigated in a discovery phase and then in a replication phase. Polygenic risk score (PRS) analysis was performed with PRSice package to test the polygenic association, and Mendelian randomization (MR) analysis was performed with TwoSampleMR package to infer the causality. 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 GWAS of ALS, and no causal association was observed for either IgA-ALS or IgM-ALS. A positive polygenic 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 association between IgG and ALS was also suggested in both the discovery (OR=1.06, 95%CI: 1.02-1.10, P=0.009) and replication (OR=1.07, 95%CI: 0.90-1.24, P=0.420) analyses, although the latter was not statistically significant. This study suggests a shared polygenic risk between serum IgG (as a biomarker for chronic inflammation) and ALS.

Keywords

Immunoglobulins, Amyotrophic lateral sclerosis, Polygenic risk score, Causal inference, Mendelian randomization

Introduction

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

By using the summary statistics from large-scale genome-wide association studies (GWAS), polygenic risk score (PRS) and Mendelian randomization (MR) analyses have added novel evidence to support polygenic and causal relationships between several traits and ALS, corroborating importantly findings from previous observational studies [4-6]. In this study, we used GWAS summary statistics to assess the associations of genetically predicted levels of immunoglobulins with the genetically predicted risk of ALS, and to test whether such associations are causal. Since IgG is a biomarker for chronic inflammation, we hypothesized that there is a positive and causal relationship between IgG and ALS. We additionally studied another two serum immunoglobulins (IgA and IgM, two biomarkers that have rarely been studied in ALS) to assess the specificity of the IgG result.

Methods

Summary statistics

We used publicly available summary statistics of several GWAS within the European ancestry population in this study. GWAS of serum immunoglobulins (IgA, IgM, and IgG) were conducted among ~15000 individuals (from Iceland and Sweden), from which single nucleotide

polymorphisms (SNPs) with the association P-values $<1.0\times10^{-6}$ were available [7]. GWAS summary statistics of ALS were obtained from two studies (available in the IEU GWAS database; https://gwas.mrcieu.ac.uk, ID for ALS_2016: ieu-a-1085; ID for ALS_2018: ebi-a-GCST005647). In the discovery analysis, summary statistics of ALS (ALS_2016) were obtained from a GWAS performed among 36052 individuals (including 12577 patients with ALS and 23475 controls) by the Project MinE group [8]. In the replication analysis, summary statistics of ALS (ALS_2018) were obtained from a meta-analysis incorporating the results from ALS_2016 and another GWAS including 40598 individuals (8229 patients with ALS and 32369 controls) [9]. Because we only used GWAS summary statistics rather than individual-level data in this study, participant inform consent and ethical review permit were waived according to the ethical review board of Shenzhen Baoan Women's and Children's Hospital.

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 predicted level for each type of immunoglobulin in the present study.

PRS analyses

Polygenic associations were investigated by PRS analyses in Linux with PRSice package (version 1.25) [10]. For the summary statistics of each exposure, correlated SNPs were pruned by linkage 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 p1=p2=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_T, ranging from 0 to 1.0×10^{-6}, in steps of 1.0×10^{-11} per quantile). The P_T of the quantile that explained the largest variance of the target was defined as the best-fitted P_T.

MR analyses

Utilizing the instrumental variable (IV) methods (Fig. 1), MR analyses were performed in R (version 3.6.1) with TwoSampleMR package (version 0.5.2) [11]. Causal inference by MR analysis requires three basic assumptions: 1) relevance, i.e., IVs have causal effects on the exposure (e.g., IgG); 2) exclusion restriction, i.e., IVs affect the outcome (e.g., ALS), only through the exposure; and 3) exchangeability or independence, i.e., no common causes exist between IVs and the outcome [12]. Lead SNPs from the genetic loci reported to be associated with each exposure were extracted from the GWAS of immunoglobulins [7]. The associations of the SNPs with other traits (associations with a P-value $<5\times10^{-8}$) were checked in GWAS Catalog (https://www.ebi.ac.uk/gwas) and excluded, leaving independent lead SNPs after LD clumping as the IVs (Additional file: Table S1). The causal relationship of each exposure-outcome pair was primarily examined by the inverse variance weighted (IVW) method and complemented with another four methods (MR Egger regression, weighted median, simple and weighted mode) as sensitivity analyses.

Results and discussion

Inverse polygenic associations were discovered for IgA-ALS and IgM-ALS, neither of the associations was however replicated (Table 1, Additional file: Figs. 1-2). A positive polygenic association between IgG and ALS was found both in the discovery and replication analyses (Table 1, Additional file: Fig. 3). The OR was 1.18 (95% CI: 1.12-1.25, $P=5.9\times10^{-7}$) in the discovery analysis and 1.13 (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 (Additional file: Figs. 4-5 and Table S2). For IgG-ALS (Table 2, Additional file: Fig. 6), 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).

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 involved in the different phases of ALS [13]. Physiological barriers (e.g., blood-brain barrier) and 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 [14]. B cells play important roles in adaptive immunity, through producing immunoglobulins and presenting antigens to T cells [15]. IgA constitutes 10-20% of serum immunoglobulins and is mainly involved in 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 immunoglobulins and can be transmitted through placental and blood-brain barriers [16].

Several studies have indicated a link between IgG and ALS. Deposit of IgG has been found in the spinal cord and motor cortex of some ALS patients [17]. Prevalence of serum monoclonal immunoglobulins (60%, including 44% IgG and 16% IgM) was found to be higher among ALS patients than that of the control group (13%) [18]. High levels of 20 IgG antibodies were also found to be help diagnose ALS cases [19]. In our 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 [2]. 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). However, because genotypes are determined at birth and unchangeable during the life course, our results about a potentially causal relationship between genetically predicted IgG level and genetically predicted ALS risk are unlikely to be dependent on early or late stage of ALS.

By using summary statistics from several large-scale GWAS, in the present study, we 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 analysis. Although statistically significant result was only obtained in the discovery analysis, similar point

estimate was noted in the replication analysis. The differences between the discovery (ALS_2016) and replication (ALS_2018) samples might have contributed to the lack of statistical significance in the replication analysis. For instance, in the discovery GWAS (ALS_2016), patients with ALS and controls were matched for age, sex, and geographic regions, whereas more than 50% of participants in the replication GWAS (ALS_2018) were only matched for race/ethnicity.

In contrast to IgG, no clear polygenic association nor causal relationship was noted for IgA-ALS and IgM-ALS. A recent observational study, including 489 patients with ALS and 17475 neurologically healthy controls, also failed to detect an association between total serum IgA and ALS [20]. The different result pattern between IgA, IgM and IgG might be partly explained by the fact that IgG is a biomarker for chronic inflammation and bypasses the blood-brain-barrier, which might be more relevant for neurodegenerative diseases including ALS, compared to the other immunoglobulins [21].

Even though we used the latest and largest GWAS of immunoglobulins, only two independent SNPs could be selected as IVs for IgG. As a result, we could only use IVW methods in the MR analysis. According to the current GWAS Catalog, no genome-wide significant association was identified between the two IVs of IgG and other non-ALS traits. Interestingly, the two genes near these two IVs are both biologically related to immune responses, neuroinflammation, and neurodegeneration. Genetic variant rs7554873 is close to the *FCGR2B* (Fc fragment of IgG receptor IIb) gene on chromosome 1. Protein FcγRIIB encoded by *FCGR2B* gene is a low-affinity receptor for the Fc region of IgG. FcγRIIB involves in the regulation of antibodies produced by B-cells and in the phagocytosis of immune complexes, it is also reported to mediate the inhibitory effect of aggregated α-synuclein on microglial phagocytosis [22]. Genetic variant rs2133037 is close to the HLA-DQB1 (major histocompatibility complex, class II, DQ beta 1) gene on chromosome 6. Its encoded protein HLA-DQB1 is also reported to be associated with immunity and neuronal survival [23].

In this study, we had only access to GWAS summary statistics data when testing the polygenic and causal relationship between serum immunoglobulins (IgA, IgM, IgG) and ALS. Using a Mendelian randomization analysis, our study is less prone to methodological limitations commonly seen in observational studies such as confounding and reverse causation. Future studies with individual genotype data are warranted in further understanding the roles of other factors (such as age, sex,

ethnicity, etc.) in the link between different immunoglobulins (especially IgG) and ALS. Further, to test whether chronic inflammation is indeed a mechanism linking together IgG and ALS, the role of IgG on other diseases (both diseases with a known link with chronic inflammation and diseases without such a link) needs to be studied. Finally, as high levels of IgG antibodies were suggested to improve the diagnosis of ALS [19], the potential use of IgG measurement in clinical diagnosis of ALS needs to be further examined.

Conclusions

In conclusion, our study suggests a shared polygenic risk between serum IgG and ALS. The causality between IgG and ALS needs to be further validated when more instrumental variables are available and when different methods of causal inference are possible to use.

Supplementary information

Additional file: Table S1. Selection of instrumental variables for serum immunoglobulins in the Mendelian randomization analyses. **Table S2.** Comparison of results from different methods in Mendelian randomization analyses. **Figure S1.** Results from the polygenic risk score analyses between serum IgA and ALS. **Figure S2.** Results from the polygenic risk score analyses between serum IgM and ALS. **Figure S3.** Results from the polygenic risk score analyses between serum IgG and ALS. **Figure S4.** Results from the Mendelian randomization analyses between serum IgA and ALS. **Figure S5.** Results from the Mendelian randomization analyses between serum IgM and ALS. **Figure S6.** Results from the Mendelian randomization analyses between serum IgG and ALS.

List of abbreviations

ALS: Amyotrophic lateral sclerosis; Ig: immunoglobulin; CI: confidence interval; GWAS: genome-wide association study; PRS: polygenic risk score; MR: Mendelian randomization; SNP: single nucleotide polymorphism; OR: odds ratio; LD: linkage disequilibrium; IVW: inverse variance weighted; FCGR2B: Fc fragment of IgG receptor IIb; HLA-DQB1: Major histocompatibility complex, class II, DQ beta 1.

Declarations

Acknowledgements: The authors would like to thank the Project MinE group and the ALS Variant Server for sharing the summary statistics of GWAS on ALS.

Authors' contributions: X.C. and F.F. contributed to the conception and design of the study; X.C., X.S., X.Z., and Y.Z. contributed to acquisition and analyses of data; X.C., X.S., X.Z.,Y.Z., and F.F. contributed to drafting and revising the manuscript.

Funding: This project is supported by the Science, Technology and Innovation Bureau of Baoan District (Grant No.: 2020JD445), the Swedish Research Council (Grant No.: 2019-01088) and Karolinska Institutet (Senior Researcher Award and Strategic Research Area in Epidemiology).

Availability of data and materials: GWAS summary statistics of ALS were accessed from the Project MinE group (https://www.projectmine.com/research/download-data) and the ALS Variant Server (http://als.umassmed.edu).

Ethics approval and consent to participate: Because GWAS summary statistics rather than individual-level data were used in this study, both informed consent and ethical approval were waived according to the ethical review board in Shenzhen Baoan Women's and Children's Hospital.

Consent for publication: Not applicable.

Competing interests: The authors declare that they have no competing interests.

References

1. Beers DR, Appel SH: Immune dysregulation in amyotrophic lateral sclerosis: mechanisms and emerging therapies. Lancet Neurol 2019, 18:211-220.

2. Yazdani S, Mariosa D, Hammar N, Andersson J, Ingre C, Walldius G, Fang F: Peripheral immune biomarkers and neurodegenerative diseases: A prospective cohort study with 20 years of follow-up. Ann Neurol 2019, 86:913-926.

3. Vandenbroucke JP, Broadbent A, Pearce N: Causality and causal inference in epidemiology: the need for a pluralistic approach. Int J Epidemiol 2016, 45:1776-1786.

4. Chen X, Yazdani S, Piehl F, Magnusson PKE, Fang F: Polygenic link between blood lipids and amyotrophic lateral sclerosis. Neurobiol Aging 2018, 67: e201-e206.

5. Bandres-Ciga S, Noyce AJ, Hemani G, Nicolas A, Calvo A, Mora G, Consortium I, International ALSGC, Tienari PJ, Stone DJ, et al: Shared polygenic risk and causal inferences in amyotrophic lateral sclerosis. Ann Neurol 2019, 85:470-481.

6. Zhan Y, Fang F: Smoking and amyotrophic lateral sclerosis: A mendelian randomization study. Ann Neurol 2019, 85:482-484.

7. Jonsson S, Sveinbjornsson G, de Lapuente Portilla AL, Swaminathan B, Plomp R, Dekkers G, Ajore R, Ali M, Bentlage AEH, Elmer E, et al: Identification of sequence variants influencing immunoglobulin levels. Nat Genet 2017, 49:1182-1191.

8. van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RA, Vosa U, de Jong S, Robinson MR, et al: Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet 2016, 48:1043-1048.

9. Nicolas A, Kenna KP, Renton AE, Ticozzi N, Faghri F, Chia R, Dominov JA, Kenna BJ, Nalls MA, Keagle P, et al: Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. Neuron 2018, 97:1268-1283.

10. Euesden J, Lewis CM, O'Reilly PF: PRSice: Polygenic Risk Score software. Bioinformatics 2015, 31:1466-1468.

11. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, et al: The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018, 7:e34408.

12. Armon C: Smoking is a cause of amyotrophic lateral sclerosis. High low-density lipoprotein cholesterol levels? Unsure. Ann Neurol 2019, 85:465-469.

13. Molteni M, Rossetti C: Neurodegenerative diseases: The immunological perspective. J Neuroimmunol 2017, 313:109-115.

14. Novellino F, Sacca V, Donato A, Zaffino P, Spadea MF, Vismara M, Arcidiacono B, Malara N, Presta I, Donato G: Innate Immunity: A Common Denominator between Neurodegenerative and Neuropsychiatric Diseases. Int J Mol Sci 2020, 21(3):1115.

15. Sabatino JJ, Jr., Probstel AK, Zamvil SS: B cells in autoimmune and neurodegenerative central nervous system diseases. Nat Rev Neurosci 2019, 20:728-745.

16. Chen K, Magri G, Grasset EK, Cerutti A: Rethinking mucosal antibody responses: IgM, IgG and IgD join IgA. Nat Rev Immunol 2020,20(7):427-441.

17. Donnenfeld H, Kascsak RJ, Bartfeld H: Deposits of IgG and C3 in the spinal cord and motor cortex of ALS patients. J Neuroimmunol 1984, 6(1):51-57.

18. Duarte F, Binet S, Lacomblez L, Bouche P, Preud'homme JL, Meininger V: Quantitative analysis of monoclonal immunoglobulins in serum of patients with amyotrophic lateral sclerosis. J Neurol Sci 1991, 104(1):88-91.

19. May C, Nordhoff E, Casjens S, Turewicz M, Eisenacher M, Gold R, Brüning T, Pesch B, Stephan C, Woitalla D, et al: Highly immunoreactive IgG antibodies directed against a set of twenty human proteins in the sera of patients with amyotrophic lateral sclerosis identified by protein array. PLoS One 2014, 9(2):e89596.

20. Crayle J, Elmallah M, Sleasman J, Bedlack R: Total serum immunoglobulin A in ALS. Amyotroph Lateral Scler Frontotemporal Degener 2020, 1-5.

21. 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, 40:142-148.

22. Choi YR, Kang SJ, Kim JM, Lee SJ, Jou I, Joe EH, Park SM: FcgammaRIIB mediates the inhibitory effect of aggregated alpha-synuclein on microglial phagocytosis. Neurobiol Dis 2015, 83:90-99.

23. Hallberg P, Smedje H, Eriksson N, Kohnke H, Daniilidou M, Ohman I, Yue QY, Cavalli M, Wadelius C, Magnusson PKE, et al: Pandemrix-induced narcolepsy is associated with genes related to immunity and neuronal survival. EBioMedicine 2019, 40:595-604.

Figure 1. Illustration of the Mendelian randomization analysis

Mendelian randomization analysis utilizes the instrumental variable (IV) methods to infer the causality between exposure and outcome. Because alleles are randomly assigned during mitosis, restricted single nucleotide polymorphisms (SNPs) could be used as instrumental variables to test the causal effect of exposure on the outcome. ALS: amyotrophic lateral sclerosis.

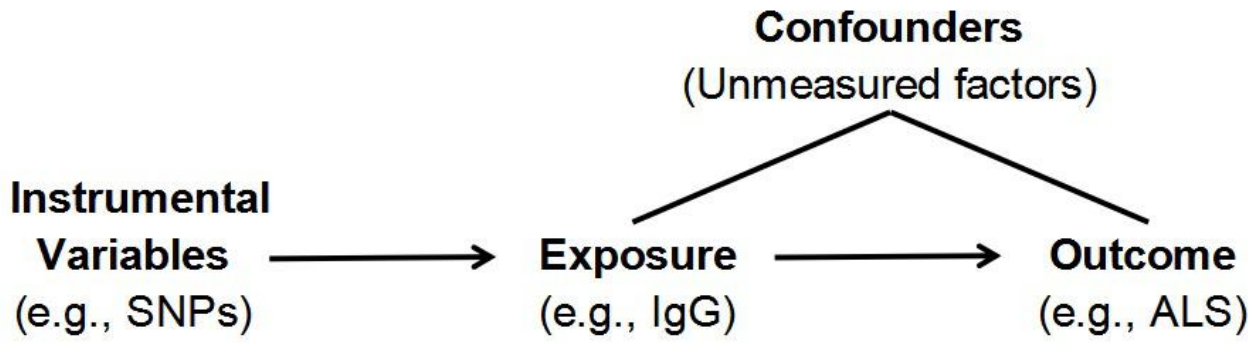

Table 1. Polygenic associations between serum immunoglobulins and ALS in PRS analyses

Base	Target	Best-fitted P_T	n_{SNP}	OR	SE	r^2 (%)	P-value
IgA	ALS_2016	8.6000×10^{-10}	54	0.861	0.032	0.059	4.2×10^{-6}
	ALS_2018	7.0001×10^{-7}	43	0.948	0.031	0.004	0.085
IgM	ALS_2016	4.8000×10^{-10}	47	0.905	0.032	0.028	0.002
	ALS_2018	3.5010×10^{-8}	42	1.051	0.028	0.004	0.074
IgG	ALS_2016	5.1001×10^{-7}	49	1.181	0.033	0.069	5.9×10^{-7}
	ALS_2018	5.7001×10^{-7}	22	1.130	0.036	0.014	0.001

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^2= the proportion of the target variation explained by SNPs in the best-fitted quantile.

Table 2. Causal inferences between serum immunoglobulins and ALS in the MR analyses

Base	Target	n_{SNP}	OR_{IVW}	SE_{IVW}	P_{IVW}	Het_P_{IVW}
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

MR=mendelian randomization; ALS=amyotrophic lateral sclerosis; n_{SNP}= numbers of independent single nucleotide polymorphisms (SNPs) uses as instrumental variables; OR: odds ratio per standard deviation of genetically determined level for each type of immunoglobulin; SE=standard error; IVW= inverse variance weighted method; Het_P=P-value of heterogeneity.