medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Polygenic Associations and Causal Inferences between Serum Immunoglobulins and Amyotrophic Lateral Sclerosis

Xu Chen, PhD[1,2]; Xiaojun Shen, BSc[2]; Yiqiang Zhan, PhD[3,4]; Fang Fang, MD, PhD[3]

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

Corresponding author: Dr. Xu Chen. Department of Central Laboratory, Baoan Women's and Children's Hospital, No.56 YuLv Road, 518100 Shenzhen, China. Phone: +86-15989385199. Email: dr.xuchen@outlook.com.

Running head: Immunoglobulins and ALS

Number of words in the abstract: 212

Number of words in the main text: 1505

Number of figures: 0

Number of tables: 2

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Chronic inflammation might contribute to the development of amyotrophic lateral sclerosis (ALS). The relationship between serum immunoglobulins and risk of ALS remains however greatly unknown. In order to overcome limitations like reverse causation and residual confounding commonly seen in observational studies, we applied polygenic risk score (PRS) and Mendelian randomization (MR) analyses on summary statistics from the large-scale genome-wide association studies (GWAS), to examine the polygenic and causal associations between three serum immunoglobulins (IgA, IgM, and IgG) and risk of ALS (first in a discovery phase and then in a 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 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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, 2] The relationship between serum immunoglobulins, as known biomarkers for inflammation and immune responses, and risk of ALS remains however greatly unknown. In a previous study, 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].[3] In the same study, we found also that during the 20 years before diagnosis, patients with ALS had more rapidly declining levels of IgG compared to others.[3] 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.[4]

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.[5-7] In this study, we aimed to examine the polygenic association and infer the causality between IgG and risk of ALS, using summary statistics 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. We additionally studied another two serum immunoglobulins (IgA and IgM) to assess the specificity of the IgG result. The roles of these two biomarkers have rarely been studied in ALS.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Materials and 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 which single nucleotide polymorphisms (SNPs) with the association P-values $<1.0\times10^{-6}$ were available.[8] 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.[9] 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 44558 individuals (8229 patients with ALS and 32369 controls).[10]

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.

PRS analyses

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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

Causal inferences were performed by MR analyses in R (version 3.6.1) with TwoSampleMR package (version 0.5.2).[12] Lead SNPs from the genetic loci reported to be associated with each exposure were extracted from the GWAS of immunoglobulins.[8] 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 instrumental variables (IVs, Supplementary Table 1). 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

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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 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 (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).

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 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] Few studies have however examined the roles of these immunoglobulins in ALS to our knowledge. In our

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.[3] 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).[3]

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 in the replication GWAS (ALS_2018) >50% of participants 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. This different result pattern 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.[17]

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

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

causes between IVs and the outcome.[18] 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 (*FCGR2B* near rs7554873 and *HLA-DQB1* near rs2133037) are both biologically related to immune responses, neuroinflammation, and neurodegeneration.[19, 20]

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.

Acknowledgment

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Potential Conflicts of Interest

Noting to report.

References

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

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

3. 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.

4. 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.

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

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

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

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

9. van Rheenen W, Shatunov A, Dekker AM, et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet. 2016 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.

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

12. 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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).

15. 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.

16. Chen K, Magri G, Grasset EK, Cerutti A. Rethinking mucosal antibody responses: IgM, IgG and IgD join IgA. Nat Rev Immunol. 2020 Feb 3.

17. 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.

18. 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.

19. Choi YR, Kang SJ, Kim JM, et al. FcgammaRIIB mediates the inhibitory effect of aggregated alpha-synuclein on microglial phagocytosis. Neurobiol Dis. 2015 Nov;83:90-9.

20. Hallberg P, Smedje H, Eriksson N, et al. Pandemrix-induced narcolepsy is associated with genes related to immunity and neuronal survival. EBioMedicine. 2019 Feb;40:595-604.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

TABLES

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20057265; this version posted April 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.