

ASSOCIATION STUDIES ARTICLE

Causal effects of blood lipids on amyotrophic lateral sclerosis: a Mendelian randomization study

Ping Zeng^{1,*} and Xiang Zhou^{2,*}¹Department of Epidemiology and Biostatistics, Xuzhou Medical University, Xuzhou, Jiangsu 221004, China²Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA

* To whom correspondence should be addressed at: Ping Zeng, Department of Epidemiology and Biostatistics, Xuzhou Medical University, Xuzhou, Jiangsu 221004, China. Tel: +86 51683262659; Fax: +86 51683262658; Email: zpstat@xzhmu.edu.cn and Xiang Zhou, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA. Tel: +1 7347645722; Fax: +1 7346157068; Email: xzhousph@umich.edu

Abstract

Amyotrophic lateral sclerosis (ALS) is a late-onset fatal neurodegenerative disorder that is predicted to increase across the globe by ~70% in the following decades. Understanding the disease causal mechanism underlying ALS and identifying modifiable risks factors for ALS hold the key for the development of effective preventative and treatment strategies. Here, we investigate the causal effects of four blood lipid traits that include high-density lipoprotein, low-density lipoprotein (LDL), total cholesterol and triglycerides on the risk of ALS. By leveraging instrument variables from multiple large-scale genome-wide association studies in both European and East Asian populations, we carry out one of the largest and most comprehensive Mendelian randomization analyses performed to date on the causal relationship between lipids and ALS. Among the four lipids, we found that only LDL is causally associated with ALS and that higher LDL level increases the risk of ALS in both the European and East Asian populations. Specifically, the odds ratio of ALS per 1 standard deviation (i.e. 39.0 mg/dL) increase of LDL is estimated to be 1.14 [95% confidence interval (CI), 1.05–1.24; $P = 1.38E-3$] in the European population and 1.06 (95% CI, 1.00–1.12; $P = 0.044$) in the East Asian population. The identified causal relationship between LDL and ALS is robust with respect to the choice of statistical methods and is validated through extensive sensitivity analyses that guard against various model assumption violations. Our study provides important evidence supporting the causal role of higher LDL on increasing the risk of ALS, paving ways for the development of preventative strategies for reducing the disease burden of ALS across multiple nations.

Introduction

Amyotrophic lateral sclerosis (ALS) is a late-onset fatal neurodegenerative disorder that is characterized by progressive degeneration of upper and lower motor neurons in adults (1). Patients with ALS typically experience signs and symptoms of progressive muscle atrophy and weakness, dysphagia, dysarthria and, eventually, respiratory failure. About 10% of ALS cases are familial with a Mendelian pattern of inheritance, and the remaining 90% of the ALS cases are sporadic (2). The median survival time

of ALS patients is only 3–5 years from the onset of symptoms with peak age of onset being 58–63 years for sporadic ALS and 47–52 years for familial ALS (1,3,4). The number of ALS cases across the globe is expected to increase by ~70% in the next two and a half decades, creating enormous potential social and health burden in the coming years (5). Both genetic and environmental factors contribute to the pathogenesis of ALS. Several disease genes have been identified to be associated with familial ALS and sporadic ALS. The familial ALS-associated genes include *SOD1*, *TARDBP*, *FUS*, *ANG* and *OPTN* (1), each individually

Received: July 24, 2018. Revised: September 7, 2018. Accepted: November 3, 2018

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

accounts for 20%, 5–10%, 5% and 1% of familial ALS cases, respectively. The sporadic ALS-associated genes include *ABCG1*, *C9orf72*, *CPNE4*, *SARM1*, *MOBP*, *ITPR2*, *DPP6* and *UNC13A*, with all genetic variants explaining a total of 8.5% of ALS phenotypic variation (6,7). Besides genetic factors, metabolic traits or environmental factors, such as smoking, exposure to pesticides, lead, organic toxins or electromagnetic radiation, and socioeconomic status are also associated with ALS (8–15).

Among the various genetic and environmental factors that have been identified to be associated with ALS, the association between blood lipid metabolites and ALS has attracted considerable recent attention. The associations between lipids and ALS are strong and are comparable in strength to many previously identified key ALS risk factors (Supplementary Material, Table S1). Lipids are the major source of energy for muscles, and muscle energy alteration is frequently observed in ALS patients (8,16–19). Specifically, ALS patients suffer from high prevalence of malnutrition, increased resting energy expenditure and weight loss, all of which affect the quality of life for ALS patients. Previous observational studies have shown that ALS patients frequently have dyslipidemia (8). Dyslipidemia is characterized by abnormal levels of high-density lipoprotein (HDL), low-density lipoprotein (LDL), total cholesterol (TC) and triglycerides (TG). The positive association between dyslipidemia and ALS suggests that high non-HDL lipid levels may play a protective role in the progression of ALS. Consistent with the human observational studies, researches with ALS mice models have also shown that the overall survival of ALS mice is reduced under caloric restriction (20). However, the relationship between dyslipidemia and ALS has become controversial in recent years. Specifically, conflicting results have been reported for basal serum lipid levels, the cause of dyslipidemia, and the relationship between serum lipid levels and ALS disease progression (21). For example, many follow-up observational studies of ALS did not observe any association between dyslipidemia and ALS (22–24). In addition, some studies have shown that the patients with ALS often have hypolipidemia—which is mainly characterized by low LDL levels—in both male and female ALS patients (25,26). The association between hypolipidemia and ALS is further confirmed in a mice ALS model (21). The conflicting results on the relationship between lipid levels and ALS may be partly due to the relatively small sample sizes employed in previous studies and partly due to uncontrolled confounding factors that are inevitable in observational studies.

Understanding the causal impact of lipids on ALS is important from both medical and public health perspectives. Because lipids are modifiable factors, a better characterization of the causal effects of lipids on ALS can facilitate the development of effective strategies for reducing ALS occurrence and the associated society burden, for treating ALS disease progression and for improving the quality of life for ALS patients. However, determining the causal impact of lipids on ALS through traditional randomized controlled trial studies is challenging, as such studies necessarily require long-term follow-ups, are expensive and often times are unethical to carry out (27). Therefore, it is desirable to determine the causal relationship between lipids and ALS through observational studies. A powerful statistical tool to examine causal relationship and estimate causal effects in observational studies is Mendelian randomization. It adapts the commonly used instrumental variable analysis approach developed in the field of causal inference to settings where genetic variants [i.e. single nucleotide polymorphisms (SNPs)] are served as instruments (28,29). In particular, Mendelian randomization employs genetic variants as proxy indicators

(i.e. instrumental variables) for the modifiable exposure (i.e. lipids) and uses these genetic variants to assess the causal effect of the exposure on the outcome variable of interest (i.e. ALS; Supplementary Material, Fig. S1) (30).

Compared with alternative approaches for causal inference, Mendelian randomization has the following desirable advantages: (i) because genetic variants are measured with high accuracy and can potentially capture long-term exposure effects, the results of Mendelian randomization analysis are often not susceptible to estimation bias due to measurement errors that are commonly encountered in randomized clinical studies (31); and (ii) because the two alleles of an SNP are randomly segregated during gamete formation and conception under Mendel's law and because such segregation is independent of many known or unknown confounders, the results from Mendelian randomization are also less susceptible to reverse causation and confounding factors compared with other study designs (32). As a result, Mendelian randomization has become a popular and cost-effective analytic tool for causal inference in observational studies, alleviating the need to record and control for all possible confounding factors present in these studies (33).

Here, we perform a two-sample summary data Mendelian randomization study to comprehensively investigate the causal relationship between lipid levels and sporadic ALS. The primary goal of our study is to investigate whether HDL, LDL, TC or TG has any causal effect on ALS and whether such causal relationship holds in two different human populations that include both the European and East Asian populations. ALS incidence rates vary greatly across different populations and can be two to several times lower in the East Asian population than in the European population (4,7,34–36). Therefore, examining the causal relationship between lipids and ALS in both European and East Asian populations can shed light on the similarity or difference in the genetic underpinning of ALS in different populations, facilitating coordination of ALS presentation strategies across nations. To investigate the causal effects of lipids on ALS, we carry out one of the largest and most comprehensive Mendelian randomization analyses to date by using summary statistics obtained from large-scale genome-wide association studies (GWASs) with ~190 000 individuals for lipids and ~42 000 individuals for ALS in the European population as well as ~130 000 individuals for lipids and ~4100 individuals for ALS in the East Asian population.

Results

Causal effects of HDL, LDL, TC and TG on ALS in the European population

We first selected SNPs that can serve as valid instrumental variables for each of the four lipid traits (HDL, LDL, TC and TG) in the European population based on association summary statistics from the Global Lipids Genetics Consortium (GLGC) study (37). The GLGC study is the largest GWAS meta-analysis performed thus far on four lipid traits using 188 577 individuals of European ancestry. From the GLGC study, we obtained a total of 88, 78, 86 and 54 index SNPs to serve as instrumental variables for HDL, LDL, TC and TG (Supplementary Material, Tables S2–S5), respectively. These SNPs were obtained through the clumping procedure of PLINK (38) by following (39), are independent of each other and are all associated with the corresponding lipid trait at the genome-wide significance level of $P = 5.00E-8$. The selected SNPs in total explained 6.15, 8.75, 8.05 or 4.92% of phenotypic variance [proportion of variance explained (PVE)] for HDL, LDL, TC or TG, respectively. Importantly, the F statistics for all these SNPs are >10 (Supplementary Material, Tables S2–S5),

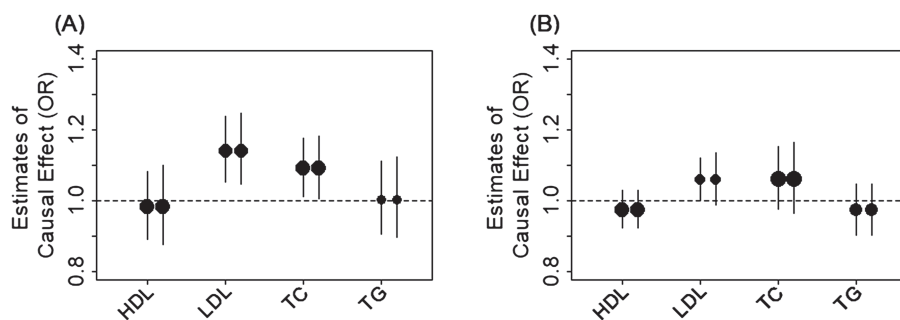


Figure 1. Causal effect estimates and 95% CIs for four lipids on ALS. Estimates are shown in both the European population (A) and the East Asian population (B). In both plots, the center of the dot represents estimated causal effects and the magnitude of the dot is in proportion to the number of the instrumental variables used in the analysis. For each lipid, the left dot shows results using the fixed-effects IVW method and the right dot shows causal results using the random-effects IVW method.

suggesting that the selected SNPs have sufficiently strong effect sizes to serve as valid instruments and that weak instrument bias is unlikely to occur (40,41).

With the identified instruments, we performed Mendelian randomization analyses to investigate the causal effects of each of the four lipid traits on ALS. To do so, we obtained association summary statistics from an ALS GWAS in the European population for the selected instrumental variables of the lipid traits. The ALS GWAS consists association results based on 41 689 individuals of European ancestry (7). Among the SNPs that were selected as instrumental variables, we removed SNPs that are associated with ALS with a marginal P -value below the 0.05 level after Bonferroni correction (i.e. $P < 0.05/k$, where k is the number of the instrumental variables available for the given lipid trait). With the remaining instruments, for each lipid in turn, we applied both the fixed-effects and the random-effects versions of the inverse-variance weighted (IVW) methods to perform Mendelian randomization analysis. These IVW methods pool results from individual SNPs to produce an estimate of the overall causal effect of each lipid on ALS. Because there is no evidence of effect size heterogeneity across SNP instruments for any of the four lipid traits (the P -values for these Q statistics are 0.017, 0.174, 0.161 and 0.123 for HDL, LDL, TC and TG, respectively; and I^2 statistics are 25.0, 11.9, 12.1 and 17.1% for HDL, LDL, TC and TG, respectively) and because the random-effects IVW yielded almost identical results to the fixed-effects IVW method (Fig. 1A), we report the results mainly from the fixed-effects IVW method in the following paragraph.

We found that, among the four lipid traits (Fig. 1A), only LDL is causally associated with ALS at a Bonferroni-corrected P -value threshold ($0.05/4 = 1.25E-2$). Specifically, the odds ratio (OR) of ALS per 1 standard deviation (SD) increase of LDL is estimated to be 1.14 [95% confidence interval (CI), 1.05–1.24; $P = 1.38E-3$], suggesting that high LDL level is causally associated with an increased risk of ALS. Among the other lipid traits, we found that TC is causally associated with ALS at a nominal P -value threshold of 0.05; the OR of ALS per 1 SD increase of TC is estimated to be 1.09 (95% CI, 1.01–1.18; $P = 2.53E-2$). In contrast, the OR of ALS per 1 SD increase of HDL is 0.98 (95% CI, 0.89–1.08; $P = 0.716$) and the OR of ALS per 1 SD increase of TG is 1.00 (95% CI, 0.91–1.11; $P = 0.948$); both are not statistically significant. The causal effect estimates above are obtained based on the unit of milligrams per deciliter; the ORs of ALS for HDL, LDL, TC and TG are estimated to be 0.51 (95% CI, 0.01–18.63; $P = 0.716$), 135.71 (95% CI, 6.70–2747.14; $P = 1.38E-3$), 24.67 (95% CI, 1.49–408.87; $P = 2.53E-2$) and 1.13 (95% CI, 0.03–47.98; $P = 0.948$), respectively, if the unit of millimoles per liter were employed. Note that the P -values are unchanged and the difference between the two types

of estimates are due to the scale difference of the two types of units (i.e. milligrams per deciliter versus millimoles per liter).

Finally, to examine result robustness, instead of using the SNP instruments identified by the clumping procedure, we performed an additional Mendelian randomization analysis using independent association SNPs reported in the original study (37) as instruments (71 instruments for HDL, 58 for LDL, 74 for TC and 40 for TG). With the new set of instruments, we obtained almost identical estimation results. For example, with fixed-effects IVW, the ORs of ALS per 1 SD for HDL, LDL, TC and TG are estimated to be 0.97 (95% CI, 0.87–1.07; $P = 0.525$), 1.11 (95% CI, 1.01–1.24; $P = 4.51E-2$), 1.09 (95% CI, 0.99–1.19; $P = 7.76E-2$) and 1.00 (95% CI, 0.90–1.11; $P = 0.990$), respectively.

Because we only detected a non-zero causal effect of LDL on ALS, we examined whether the lack of detectable non-zero causal effects of the other lipid traits are due to a lack of statistical power for these traits in the Mendelian randomization analysis. To do so, for each lipid in turn, we performed simulations and power analyses using 41 689 individuals. We estimated the statistical power to detect an OR of 1.10 (or 0.90) in the risk of ALS per unit change of the lipid following the method in (42). The statistical power based on a Bonferroni-adjusted threshold to detect such amount of causal effects from HDL, LDL, TC and TG on ALS are estimated to be 44, 62, 57 and 34%, respectively. The power results suggest that we would have comparable and reasonable power across the four lipids.

Causal effects of HDL, LDL, TC and TG on ALS in the East Asian population

We next estimated the causal effects of HDL, LDL, TC and TG on ALS in the East Asian population. To do so, we obtained association summary statistics from an East Asian lipid GWAS with 162 255 Japanese individuals (43) and corresponding association summary statistics from another East Asian ALS GWAS on 4084 Chinese individuals (4). We first obtained a set of 31, 22, 32 and 26 lipid-associated SNPs to serve as instrumental variables for HDL, LDL, TC and TG, respectively (Supplementary Material, Tables S6–S9). The estimated PVE by the selected SNPs are 7.75, 3.89, 2.67 and 5.02% for HDL, LDL, TC and TG, respectively. All selected SNPs have an F statistic >10 and are thus considered as strong instruments (Supplementary Material, Tables S6–S9) (40,41).

With the identified instruments, we performed Mendelian randomization analyses using both the fixed-effects and random-effects IVW methods to investigate the causal effects of each of the four lipid traits on ALS. Like what we observed in the European data, we did not detect any SNP effect size

heterogeneity on any of the four lipids in the East Asian population; the *P*-values for the *Q* statistics are 0.494, 0.066, 0.161 and 0.769 for HDL, LDL, TC and TG, respectively; the *I*² statistics are 0, 27.3, 14.9 and 0% for HDL, LDL, TC and TG, respectively. In addition, the results from both the fixed-effects and random-effects IVW analysis are largely consistent with each other (Fig. 1B). Specifically, in the fixed-effects IVW analyses, we found that high LDL is causally associated with an increased risk of ALS at the significance level of 0.05 (Fig. 1B). The OR of ALS per 1 SD increase of LDL is estimated to be 1.06 (95% CI, 1.00–1.12; *P* = 0.044). In contrast, the ORs of ALS per 1 SD increase of HDL, TC and TG are 0.97 (95% CI, 0.92–1.03; *P* = 0.361), 1.06 (95% CI, 0.98–1.15; *P* = 0.165) and 0.97 (95% CI, 0.90–1.05; *P* = 0.470), respectively.

Next, we examined whether the lack of detectable non-zero causal effects of the other non-LDL lipid traits in the East Asian population were due to a lack of statistical power for these traits. To do so, for each lipid in turn, we performed simulations and power analyses using 4084 individuals. We estimated the statistical power to detect an OR of 1.10 (or 0.90) in the risk of ALS per unit change of the lipid. The statistical powers based on a Bonferroni-adjusted threshold to detect such amount of causal effects from HDL, LDL, TC and TG on ALS were estimated to be 4, 3, 2 and 3%, respectively (or 13, 8, 8 and 10% based on a nominal *P*-value threshold of 0.05). Simulation results suggest that we have a small amount of power to detect non-zero lipid effects due to the small sample size in the East Asian ALS GWAS. Therefore, larger sample size is likely needed to validate our causal effect estimates in the East Asian population.

Finally, we note that we did not observe statistically significant causal effect size heterogeneity between the two populations for any of the lipid on ALS. Specifically, the causal effect estimates of each lipid on ALS from the two populations share the same directionality and their 95% CIs are overlapped (*P* = 0.115 with a *z*-test). Therefore, we used the IVW meta-analysis method to combine the causal effect estimates from both European and East Asian populations to obtain a joint causal effect estimate. By assuming identical causal effects of lipids on ALS in both populations, we estimated the ORs of ALS per 1 SD increase of HDL, LDL, TC and TG to be 0.98 (95% CI, 0.93–1.02; *P* = 0.165), 1.08 (95% CI, 0.1.03–1.14; *P* = 2.54E-4), 1.08 (95% CI, 1.02–1.14; *P* = 4.81E-3) and 0.98 (95% CI, 0.93–1.04; *P* = 0.292), respectively. The combined results confirm our main conclusion that LDL is positively causally associated with ALS.

Sensitivity analyses confirm the causal effects of LDL on ALS

We now validate the causal relationship between LDL and ALS inferred in both European and East Asian populations through various sensitivity analyses. First, to guard against the possibility that the associated instruments for one lipid may exhibit their causal effects on ALS through the other lipids, we performed Mendelian randomization for each lipid using lipid-specific instruments (Supplementary Material, Fig. S2). These lipid-specific instruments were obtained by removing those instruments that were also selected as instruments for the other three lipids (i.e. based on the same clumping strategy with a significance level of *P* = 5.00E-8). The Mendelian

randomization results using these lipid-specific instruments are consistent with the main results in both European and East Asian populations (Supplementary Material, Table S10).

Next, for LDL, we also excluded potential pleiotropic instrumental effects by performing a residual-based (or joint) multivariable Mendelian randomization analysis with HDL and TG treated as covariates (44,45). Note that we cannot include TC in addition to HDL and TG as covariate in the model because LDL level can be completely determined by the other three lipids (46) and is highly correlated with TC (correlation coefficients *r* = 0.992 and 0.787 in the European and East Asian populations, respectively). The multivariate Mendelian randomization analysis tests for the causal effects of LDL on ALS while controlling for the other lipid effects, thus the bias of the causal effects of LDL due to the instrumental pleiotropy of other lipids is minimized. The estimation results from the residual-based (or joint) multivariable Mendelian randomization analysis are again consistent with those in the main text (Supplementary Material, Table S10), supporting the observation that LDL has a positive causal effect on ALS. Specifically, after removing the effects of HDL and TG and applying the residual-based multivariable regression, the OR of ALS per 1 SD increase of LDL is estimated to be 1.13 (95% CI, 1.03–1.23; *P* = 8.84E-3) in the European population, albeit less significant in the East Asian population presumably due to the smaller number of instruments used there (Supplementary Material, Table S10). The results above imply that the causal effect of LDL on ALS is unlikely acted through the other lipids.

Next, we examined whether there exist potential SNP instrument outliers and whether these outliers influence Mendelian randomization analysis results. To do so, we visualized all SNP instruments by plotting the SNP effect sizes on LDL versus their effect sizes on ALS in Figure 2. In the European population, three SNPs (rs11591147, rs7254892 and rs12721109) have effect sizes on LDL larger than 0.40 and these SNPs appear to be potential outliers (Fig. 2A). However, these potential outliers do not influence the causal effect estimation in the Mendelian randomization analysis. Indeed, after removing these three SNPs, the OR of ALS per 1 SD increase of LDL is estimated to be 1.15 (95% CI, 1.05–1.25; *P* = 1.04E-3; Fig. 2B), almost identical to that obtained using all instruments (i.e. OR = 1.14; 95% CI, 1.05–1.24; *P* = 1.38E-3). In the East Asian population, one SNP, rs769446, has a large effect size of -0.51 (95% CI, -0.54 to -0.49; *P* = 2.98E-322) on LDL and appears to be an outlier on the effect size scatter plot (Fig. 2C). After removing this SNP, the OR of ALS per 1 SD increase of LDL is estimated to be 1.24 (95% CI, 1.11–1.39; *P* = 1.54E-4; Fig. 2D), which is larger and more significantly different from zero than the estimate obtained using all instruments (i.e. OR = 1.06; 95% CI, 1.00–1.12; *P* = 0.044). Therefore, even though rs769446 may be a potential outlier, it only supports our main conclusion that LDL has a statistically significantly positive causal effect on ALS in the East Asian population.

To formally rule out the possibility that a single instrument may strongly influence the point estimates and the statistical test results, we performed a leave-one-out (LOO) Mendelian randomization analysis. In the European population, the LOO analysis results show that no single instrument acts as an outlier and influences the causal effect estimates (Supplementary Material, Fig. S3A). In the East Asian population, the only detected outlier is rs769446 (Supplementary Material, Fig. S3B). However, as described in the previous paragraph, removing rs769446 does not influence the conclusion that LDL has a positive causal effect on ALS.

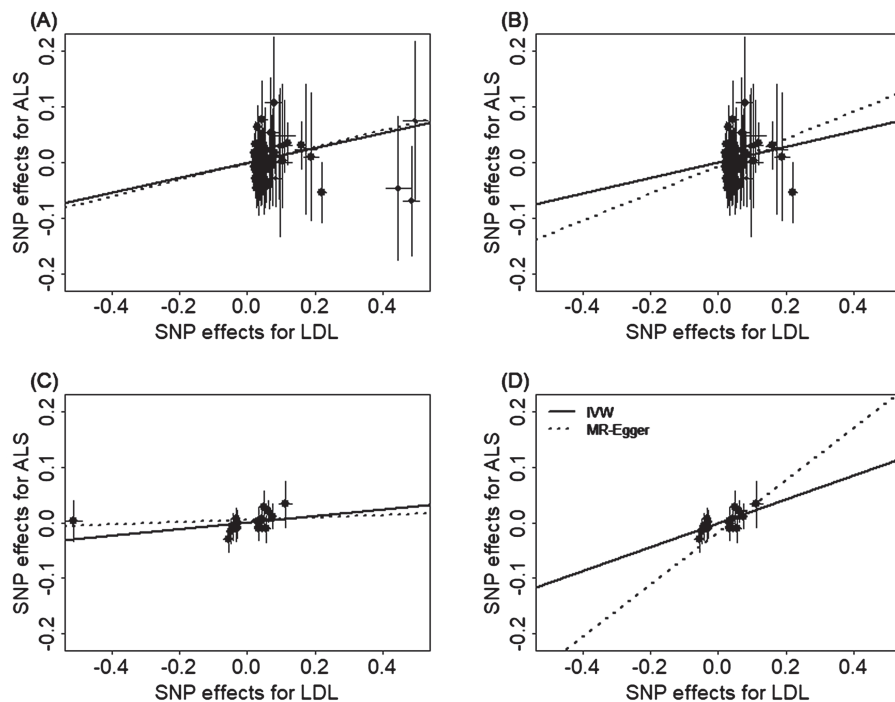


Figure 2. Relationship between the SNP effect size estimates on LDL (x-axis), and the corresponding effect size estimates on ALS (y-axis) for SNPs that serve as instrumental variables. Results are shown in the European population for all SNP instruments (A) and for the remaining instrumental SNPs after removing three potential outliers (B). Results are also shown in the East Asian population for all SNP instruments (C) and for the remaining instrumental SNPs after removing one potential outlier (D). On each plot, 95% CIs for the estimated SNP effect sizes on ALS are shown as vertical black lines, while the 95% CIs for the estimated SNP effect sizes on LDL are shown as horizontal black lines. The slope of fitted lines represents the estimated causal effects of LDL on ALS obtained using either the IVW methods (solid lines) or the MR-Egger regression (dotted lines). Because of the inclusion of an intercept in the MR-Egger regression, the fitted lines by MR-Egger regression (dotted lines) do not necessarily pass the origin.

To guard against the possibility that some instruments are invalid, we conducted a Mendelian randomization analysis using the weighted median method (47). The weighted median method yields qualitatively similar results as our main analysis. In particular, the ORs of ALS per 1 SD of LDL are estimated to be 1.17 (95% CI, 1.03–1.33; $P = 1.60E-2$) in the European population and 1.29 (95% CI, 1.11–1.49; $P = 6.29E-4$) in the East Asian population (after removing the outlier rs769446). The weighted median analysis results suggest that invalid instruments unlikely bias our main results.

To guard against the possibility of horizontal pleiotropy of these instruments, we performed Mendelian randomization using the MR-Egger regression (48,49). The results from the MR-Egger regression analysis are again largely consistent with our main results. For example, in the European population, MR-Egger regression estimates the OR of ALS per 1 SD increase of LDL to be 1.16 (95% CI, 1.02–1.32; $P = 0.026$) when all instruments are used. When the three SNP outliers (rs11591147, rs7254892 and rs12721109) are removed, MR-Egger regression estimates the OR of ALS per 1 increase SD of LDL to be 1.23 (95% CI, 1.05–1.45; $P = 9.19E-3$). In the East Asian population, MR-Egger regression estimates the OR of ALS per 1 SD increase of LDL to be 1.02 (95% CI, 0.94–1.11; $P = 0.645$) when all instruments are used. When the SNP outlier (rs12927205) is removed, MR-Egger regression estimates the OR of ALS per 1 SD increase of LDL to be 1.60 (95% CI, 1.25–2.06; $P = 2.41E-4$). Examining the intercept estimates from MR-Egger regression also leads to the same conclusion that horizontal pleiotropy unlikely influence results in either population. For example, in the European population, the MR-Egger regression intercept was estimated to be -0.001 (95% CI,

-0.009 to 0.007 ; $P = 0.767$) when all instruments were used and was estimated to be -0.005 (95% CI, -0.015 to 0.004 ; $P = 0.289$) when three instrument outliers were removed. In the East Asian population, the MR-Egger regression intercept was estimated to be 0.006 (95% CI, -0.003 to 0.014 ; $P = 0.180$) when all instruments were used and was estimated to be -0.016 (95% CI, -0.029 to -0.003 ; $P = 0.020$) when one instrument outlier is removed. Furthermore, funnel plots also display symmetric pattern of effect size variation around the causal effect point estimate (Fig. 3). Together, funnel plots and MR-Egger regression analyses suggest that horizontal pleiotropy unlikely bias our results.

To examine the possibility that our results are driven by reverse causality, we performed IVW analysis in the reverse direction to examine the causal effects of ALS on LDL. In total, we identified five and one SNPs to serve as instruments for ALS in the European and East Asian populations, respectively. Results show that there is no detectable causal effect of ALS on LDL; the causal effect of ALS on LDL is estimated to be 0.003 (95% CI, -0.022 to 0.027 ; $P = 0.834$) in the European population and 0.094 (95% CI, -0.145 to 0.334 ; $P = 0.439$) in the East Asian population, respectively.

Finally, we note that several previous studies have shown that the ratio of LDL/HDL, instead of LDL alone, is also correlated with ALS (8,9). In our results, we have shown that HDL is negatively associated with ALS (though not statistically significant) while LDL is positively associated with ALS, suggesting that the ratio of LDL/HDL is likely positively associated with ALS. While it is tempting to directly examine the causal effect of the ratio of LDL/HDL on ALS, we acknowledge that it would be challenging to do so because there is currently no GWAS

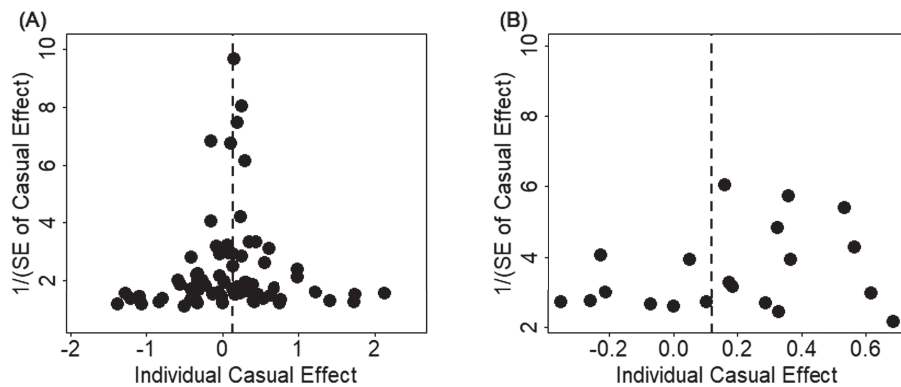


Figure 3. Funnel plot displays causal effect estimates of LDL on ALS obtained using individual SNP instruments in the European population (A) and the East Asian population (B). Center of the dots represents the estimated causal effect (log OR) for each instrumental variable. The vertical dotted lines represent the estimated causal effect obtained using all instrumental variables.

summary statistics available for the ratio of LDL/HDL. Instead of directly investigating the causal effects of LDL/HDL on ALS, we indirectly examined this relationship by estimating the causal effect of LDL on ALS while controlling for HDL. Intuitively, if LDL/HDL is high for an individual, then the residual value of LDL after controlling for HDL would also be high for the individual. Therefore, if LDL/HDL is associated with ALS status, then LDL after controlling for HDL will also likely be associated with ALS. With multivariable Mendelian randomization analysis (50), we found that, after removing the effect of LDL/HDL, the OR of ALS per 1 SD increase of LDL is nearly unchanged and is estimated to be 1.13 (95% CI, 1.04–1.23; $P = 7.58E-3$) in the European population and increases to 1.26 (95% CI, 1.13–1.41; $P = 5.65E-4$) in the East Asian population.

Discussion

We have presented a comprehensive Mendelian randomization analysis to investigate the causal relationship between lipids and ALS using summary statistics from large GWASs. The large sample size used in our study allows us to fully establish a positive causal effect of the modifiable factor LDL on ALS in both European and East Asian populations. The inferred causal relationship between LDL and ALS is robust with respect to the choice of statistical methods and is carefully validated through various sensitivity analyses. Overall, our study provides important evidence supporting the causal role of higher LDL on increasing the risk of ALS.

The positive causal effect of LDL on ALS suggests that future development of strategies to reduce LDL levels would likely reduce disease burden of ALS (51). LDL is a modifiable risk factor whose levels can be reduced through various intervention strategies. For example, dietary changes such as increased fiber consumption, increased phytosterol consumption and increased nut consumption can all lead to a reduction in LDL levels (52). Restrictions in dietary cholesterol intake, restrictions in high-carbohydrate diets and restrictions in trans-fatty acid consumption can also reduce LDL levels (53). Beside lifestyle and dietary changes, lowering LDL can be achieved by drug-based therapy. For example, various drugs such as cholestyramine, niacin, ezetimibe and statins can each significantly reduce LDL levels (54). The future development of LDL reduction strategies and the making of public policies to promote such strategies would likely lead to a reduction of ALS disease burden in the society.

Our results dovetail a recently published Mendelian randomization study which relies on instrumental variables associated

with both LDL and TC to investigate their joint causal effect on ALS (55). The estimated causal effect of LDL on ALS in our study is consistent with the estimated causal effect of LDL and TC on ALS in Chen *et al.* (55) both in terms of magnitude and directionality (OR = 1.14 for LDL on ALS in our study versus OR = 1.16 for LDL in Chen *et al.*; OR = 1.09 for TC on ALS in our study versus OR = 1.15 for TC on ALS in Chen *et al.* for the European population). However, compared with (55), our study has several unique advantages. First, the samples used in our study to select the instruments for lipids are proximately two times larger than that in Chen *et al.* (55) (i.e. ~190 000 versus ~100 000). Therefore, we are able to identify more and stronger instruments than that in Chen *et al.*, leading to a potential power gain of Mendelian randomization analysis (30,41,56,57). Second, unlike Chen *et al.*, we have performed extensive sensitivity analyses (e.g. LOO, median-based and multivariable analyses, Egger regression and guarding against horizontal pleiotropic effects etc.) (58) to ensure the robustness of our results and to guard against any possible model assumption violation in the Mendelian randomization analysis. Third, by examining both European and East Asian populations, we are able to establish a causal effect of higher LDL on increasing risk of ALS in both populations. The inferred causal relationship between LDL and ALS in multiple populations suggests that strategies to reduce LDL levels have the potential to lower ALS in multiple populations and ethnic groups.

The causal relationship between lipids and ALS identified in the European and East Asian populations are inferred using a different set of instruments in the two populations. Indeed, the overlap between the index SNPs used in the European population and the index SNPs used in the East Asian population is small; only 2, 1, 2 and 1 index SNPs are overlapped between the 2 populations for HDL, LDL, TC and TG, respectively. However, most index SNPs from one population are in linkage disequilibrium with index SNPs from another population, suggesting similar genetic architectures underlying lipid traits in different populations. In addition, we found that the causal effects of LDL and TC on ALS estimated in the European population are comparable with those estimated in the East Asian population (e.g. 1.14 versus 1.06 for LDL and 1.09 versus 1.06 for TC), allowing us to combine causal effect estimates from the two populations via meta-analysis. The similarity in the causal effects of LDL on ALS in the two populations suggests potential commonality of ALS etiology in multiple populations. However, we also recognize that there remains substantial unexplained diversity underlying ALS etiology in the two populations in terms of, for example, the

different prevalence rates of ALS. Future studies are needed to understand the genetic and environmental basis underlying ALS prevalence difference across populations.

Our results on the causal effects of lipids on ALS in the European population are obtained using a European lipid GWAS and a European ALS GWAS. However, our results in the East Asian population are obtained using a Japanese lipid GWAS and a Chinese ALS GWAS. Because our East Asian results are based on using two different ethnic groups, one might concern that the genetic difference between the two ethnic groups may bias our causal effect estimates. However, it is well known that the major East Asian ethnic groups (e.g. Japanese, Chinese and Korean) are genetically closely related to each other with similar genetic architectures (59,60). Indeed, recent studies have shown that the genetic difference among multiple East Asian ethnic groups is comparable to the genetic difference among multiple European ethnic groups, both in terms of time to the most recent common ancestor or divergence time (61). Consequently, GWAS meta-analyses in East Asian are commonly performed by combining association results from multiple ethnic groups (62–66), as we also do here. Therefore, our causal effect estimation results in the East Asian population are unlikely biased by the genetic difference between the two ethnic groups. However, to further alleviate concerns, we have conducted additional Mendelian randomization analyses to estimate the causal effects of lipids on ALS using a Chinese lipid GWAS (Supplementary Material, Tables S11 and S12) (67), which has a smaller sample size than the Japanese lipid GWAS (23 083 versus 94 356). The results are largely consistent with the results we have presented in the main text (Supplementary Material, Fig. S4). For example, the OR of ALS per 1 SD increase of LDL is estimated to be 1.18 (95% CI, 1.04–1.35; $P = 9.44E-3$) using the Chinese lipid GWAS, close to what we have obtained using the Japanese lipid GWAS. The results based on the Chinese lipid GWAS further confirm a positive causal effect of LDL on the risk of ALS in the East Asian population.

Like many other Mendelian randomization analyses, our results are not without limitations. First, we emphasize again that the validity of Mendelian randomization analyses relies on three key modeling assumptions listed in Supplementary Material Fig. S1 (30,33,57). While we have been extremely careful in selecting instrumental variables to satisfy these assumptions and have performed various sensitivity analyses to guard against various model assumption violations, we still would like to follow the standard practice of Mendelian randomization analysis to caution against the overinterpretation of the causality results obtained in the present study (29,41,56,68,69). Second, like other Mendelian randomization applications, we have assumed a linear relationship between lipids and ALS in the Mendelian randomization model. While a linear relationship can be viewed as a first-order approximation to any non-linear relationship, we acknowledge that modeling a linear relationship can be suboptimal in terms of power if the true relationship is non-linear. Third, because of the use of GWAS summary statistics, we cannot carry out stratified analyses by gender to estimate the causal effects of lipids on ALS in males and females separately. Therefore, we are unable to estimate or validate previously estimated gender-specific effects of lipids on ALS in observational studies (25,26).

Materials and Methods

We provide an overview of the materials and methods used in the present study here, with details available in Supplementary Material.

GWAS data sources

In the European population, we obtained association summary statistics for SNPs in terms of their marginal effect size estimates and standard errors for each of the four lipid traits (HDL, LDL, TC and TG) from the GLGC study (Supplementary Material, Table S13) (37). The GLGC study is the largest GWAS performed to date on lipid traits and is a meta-analysis that analyzed ~2.5 million genotyped and imputed SNPs on 188 577 individuals. To examine the causal effect of lipids on ALS in Europe, we also obtained association summary statistics from an European ALS GWAS that analyzed ~8.7 million genotyped and imputed SNPs on 41 689 individuals (14 791 cases and 26 898 controls; Supplementary Material, Table S14) (7). In the East Asian population, we obtained association summary statistics for lipid traits from a Japanese lipid GWAS. The Japanese lipid GWAS analyzed a total of ~6.0 million genotyped and imputed SNPs on 70 657 individuals for HDL, 72 866 individuals for LDL, 128 305 individuals for TC and 105 597 individuals for TG (Supplementary Material, Table S15) (43). To examine the causal effect of lipids on ALS in East Asia, we also obtained association summary statistics from a Chinese ALS GWAS that analyzed ~6.6 million genotyped and imputed SNPs on up to 4084 individuals (1234 cases and 2850 controls) (4).

Instrumental variable selection

In the European population, from the GLGC study (37), we selected 88, 78, 86 and 54 independent index association SNPs to serve as instrumental variables for HDL, LDL, TC and TG (Supplementary Material, Tables S2–S5), respectively. These index SNPs were obtained using the clumping procedure of the PLINK software (38) following previous work (39). To estimate the causal effect of each of the four lipids on ALS, we also extracted association summary statistics with ALS for these index SNPs from the European ALS GWAS (7). In the East Asian population, from the Japanese lipid GWAS (43), we selected 30, 21, 31 and 26 index association SNPs to serve as instrumental variables for HDL, LDL, TC and TG (Supplementary Material, Tables S6–S9), respectively. To estimate the causal effect of each of the four lipids on ALS, we also extracted association summary statistics of these index SNPs for ALS from the Chinese ALS GWAS (4).

Estimation of causal effect with Mendelian randomization

In either population, we excluded index SNPs that show pleiotropic associations for ALS to ensure the validity of Mendelian randomization analysis (30,56). However, none of the index SNPs show pleiotropic associations with ALS in both populations in our main text analyses, thus we kept all of them in the analysis. For each index SNP in turn, we estimated the proportion of PVE for the corresponding lipid trait explained by the index SNP using summary statistics following the method presented in (70). We further computed F statistic to ensure strong instruments (40,41). We then performed Mendelian randomization analysis for one lipid at a time to estimate its causal effect on ALS in terms of the OR for one-unit change of lipid level, where the unit is defined as SD of the trait, estimated to be 15.1 mg/dL for HDL, 39.0 mg/dL for LDL, 40.6 mg/dL for TC and 79.3 mg/dL for TG. Our two-sample Mendelian randomization analyses (48,71) were performed using either the fixed-effects version and the random-effects version of

the IVW methods (41,72). Compared with the fixed-effects version, the random-effects version of IVW accounts for causal effect size heterogeneity across instruments and often yields more conservative results. Besides applying the random-effects version of IVW, we also computed both Q and I^2 statistics to directly measure effect size heterogeneity in the data (71,73).

Sensitivity and other complementary analyses

In both populations, to ensure results robustness and guard against various modeling misspecifications, we performed extensive sensitivity analyses. These analyses include (i) Mendelian randomization analysis using lipid-specific instruments to guard against the possibility that some associated instruments for one lipid may exhibit their causal effects on ALS through the other lipids; (ii) LOO cross-validation analysis to examine whether there are instrumental outliers and whether these outliers affect results (39); (iii) median-based method to directly exclude outlier influence (47); (iv) MR-Egger regression to examine the assumption of directional pleiotropic effects (48,49); (v) reverse causation analysis to exclude the possibility that ALS affects lipids; and (vi) in addition to the sensitivity analyses, we also performed a complementary Mendelian randomization analysis to indirectly test the causal effect of the ratio of LDL/HDL on ALS (8,9) by examining the causal effect of LDL on ALS while controlling for LDL/HDL (44,50). We also considered removing the potential mediation effect of HDL and/or TG by performing residual-based multivariable Mendelian randomization analysis as done in (44,45).

Supplementary Material

Supplementary Material is available at HMG online.

Acknowledgements

We thank all the GWAS consortium studies for making the summary data publicly available and are grateful of all the investigators and participants contributed to those studies.

Conflict of Interest statement. None declared.

Funding

National Institutes of Health (R01HG009124 to X.Z.); National Science Foundation (DMS1712933 to X.Z.); Natural Science Foundation of Jiangsu (BK20181472 to P.Z.); Youth Foundation of Humanity and Social Science funded by Ministry of Education of China (18YJC910002 to P.Z.); China Postdoctoral Science Foundation (2018M630607 to P.Z.); Jiangsu QingLan Research Project for Outstanding Young Teachers (to P.Z.); National Natural Science Foundation of China (81402765 to P.Z.); Statistical Science Research Project from National Bureau of Statistics of China (2014LY112 to P.Z.); Priority Academic Program Development of Jiangsu Higher Education Institutions for Xuzhou Medical University (to P.Z.).

References

- Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., Burrell, J.R. and Zoing, M.C. (2011) Amyotrophic lateral sclerosis. *Lancet*, **377**, 942–955.
- Riva, N., Agosta, F., Lunetta, C., Filippi, M. and Quattrini, A. (2016) Recent advances in amyotrophic lateral sclerosis. *J. Neurol.*, **263**, 1241–1254.
- Chiò, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., Traynor, B.G. and Eurals Consortium (2009) Prognostic factors in ALS: a critical review. *Amyotroph. Lateral Scler.*, **10**, 310–323.
- Benyamin, B., He, J., Zhao, Q., Gratten, J., Garton, F., Leo, P.J., Liu, Z., Mangelsdorf, M., Al-Chalabi, A. and Anderson, L. (2017) Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis. *Nat. Commun.*, **8**, 611.
- Arthur, K.C., Calvo, A., Price, T.R., Geiger, J.T., Chiò, A. and Traynor, B.J. (2016) Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat. Commun.*, **7**, 12408.
- MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J. et al. (2017) The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.*, **45**, D896–D901.
- van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A., Vösa, U., de Jong, S. Robinson, M.R. et al. (2016) Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.*, **48**, 1043–1048.
- Dupuis, L., Corcia, P., Fergani, A., Gonzalez De Aguilar, J.L., Bonnefont-Rousselot, D., Bittar, R., Seilhean, D., Hauw, J.J., Lacomblez, L., Loeffler, J.P. et al. (2008) Dyslipidemia is a protective factor in amyotrophic lateral sclerosis. *Neurology*, **70**, 1004–1009.
- Chiò, A., Calvo, A., Ilardi, A., Cavallo, E., Moglia, C., Mutani, R., Palmo, A., Galletti, R., Marinou, K., Papetti, L. et al. (2009) Lower serum lipid levels are related to respiratory impairment in patients with ALS. *Neurology*, **73**, 1681–1685.
- Mitchell, J.D. and Borasio, G.D. (2007) Amyotrophic lateral sclerosis. *Lancet*, **369**, 2031–2041.
- Corcia, P. and Meininger, V. (2008) Management of amyotrophic lateral sclerosis. *Drugs*, **68**, 1037–1048.
- Malek, A.M., Barchowsky, A., Bowser, R., Heiman-Patterson, T., Lacomis, D., Rana, S., Youk, A., Stickler, D., Lackland, D.T. and Talbott, E.O. (2014) Environmental and occupational risk factors for amyotrophic lateral sclerosis: a case-control study. *Neurodegener. Dis.*, **14**, 31–38.
- Andrew, A.S., Caller, T.A., Tandan, R., Duell, E.J., Henegan, P.L., Field, N., Bradley, W.G. and Stommel, E.W. (2017) Environmental and occupational exposures and amyotrophic lateral sclerosis (ALS) in New England. *Neurodegener. Dis.*, **17**, 110–116.
- Alonso, A., Logroscino, G. and Hernán, M.A. (2010) Smoking and the risk of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *J. Neurol. Neurosurg. Psychiatry*, **81**, 1249–1252.
- Calvo, A., Canosa, A., Bertuzzo, D., Cugnasco, P., Solero, L., Clerico, M., De Mercanti, S., Bersano, E., Cammarosano, S., Ilardi, A. et al. (2016) Influence of cigarette smoking on ALS outcome: a population-based study. *J. Neurol. Neurosurg. Psychiatry*, **87**, 1229–1233.
- Dupuis, L., Pradat, P.F., Ludolph, A.C. and Loeffler, J.P. (2011) Energy metabolism in amyotrophic lateral sclerosis. *Lancet Neurol.*, **10**, 75–82.
- Funalot, B., Desport, J.C., Sturtz, F., Camu, W. and Couratier, P. (2009) High metabolic level in patients with

- familial amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.*, **10**, 113–117.
18. Desport, J.C., Preux, P.M., Bouteloup-Demange, C., Clavelou, P., Beaufrère, B., Bonnet, C. and Couratier, P.P. (2003) Validation of bioelectrical impedance analysis in patients with amyotrophic lateral sclerosis. *Am. J. Clin. Nutr.*, **77**, 1179–1185.
 19. Schmitt, F., Hussain, G., Dupuis, L., Loeffler, J.P. and Henriques, A. (2014) A plural role for lipids in motor neuron diseases: energy, signaling and structure. *Front. Cell. Neurosci.*, **8**, 25.
 20. Hamadeh, M.J., Rodríguez, M.C., Kaczor, J.J. and Tarnopolsky, M.A. (2005) Caloric restriction transiently improves motor performance but hastens clinical onset of disease in the Cu/Zn-superoxide dismutase mutant G93A mouse. *Muscle Nerve*, **31**, 214–220.
 21. Kim, S.M., Kim, H., Kim, J.E., Park, K.S., Sung, J.J., Kim, S.H. and Lee, K.W. (2011) Amyotrophic lateral sclerosis is associated with hypolipidemia at the presymptomatic stage in mice. *PLoS One*, **6**, e17985.
 22. Paganoni, S., Deng, J., Jaffa, M., Cudkowicz, M.E. and Wills, A.M. (2011) Body mass index, not dyslipidemia, is an independent predictor of survival in amyotrophic lateral sclerosis. *Muscle Nerve*, **44**, 20–24.
 23. Huang, R., Guo, X., Chen, X., Zheng, Z., Wei, Q., Cao, B., Zeng, Y. and Shang, H. (2015) The serum lipid profiles of amyotrophic lateral sclerosis patients: a study from south-west China and a meta-analysis. *Amyotroph. Lateral Scler. Frontotemporal Degener.*, **16**, 359–365.
 24. Rafiq, M.K., Lee, E., Bradburn, M., McDermott, C.J. and Shaw, P.J. (2015) Effect of lipid profile on prognosis in the patients with amyotrophic lateral sclerosis: insights from the olesoxime clinical trial. *Amyotroph. Lateral Scler. Frontotemporal Degener.*, **16**, 478–484.
 25. Sutedja, N.A., van der Schouw, Y.T., Fischer, K., Sizoo, E.M., Huisman, M.H., Veldink, J.H. and Van den Berg, L.H. (2011) Beneficial vascular risk profile is associated with amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry*, **82**, 638–642.
 26. Yang, J.W., Kim, S.M., Kim, H.J., Kim, J.E., Park, K.S., Kim, S.H., Lee, K.W. and Sung, J.J. (2013) Hypolipidemia in patients with amyotrophic lateral sclerosis: a possible gender difference? *J. Clin. Neurol.*, **9**, 125–129.
 27. Armon, C. (2003) An evidence-based medicine approach to the evaluation of the role of exogenous risk factors in sporadic amyotrophic lateral sclerosis. *Neuroepidemiology*, **22**, 217–228.
 28. Angrist, J.D., Imbens, G.W. and Rubin, D.B. (1996) Identification of causal effects using instrumental variables. *J. Am. Stat. Assoc.*, **91**, 444–455.
 29. Greenland, S. (2000) An introduction to instrumental variables for epidemiologists. *Int. J. Epidemiol.*, **29**, 722–729.
 30. Sheehan, N.A., Didelez, V., Burton, P.R. and Tobin, M.D. (2008) Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med.*, **5**, e177.
 31. Haycock, P.C., Burgess, S., Wade, K.H., Bowden, J., Relton, C. and Smith, G.D. (2016) Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.*, **103**, 965–978.
 32. Smith, G.D. and Ebrahim, S. (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.*, **32**, 1–22.
 33. Sleiman, P.M. and Grant, S.F. (2010) Mendelian randomization in the era of genomewide association studies. *Clin. Chem.*, **56**, 723–728.
 34. Deng, M., Wei, L., Zuo, X., Tian, Y., Xie, F., Hu, P., Zhu, C., Yu, F., Meng, Y., Wang, H. et al. (2013) Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis. *Nat. Genet.*, **45**, 697–700.
 35. Chiò, A., Logroscino, G., Traynor, B.J., Collins, J., Simeone, J.C., Goldstein, L.A. and White, L.A. (2013) Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology*, **41**, 118–130.
 36. Doi, Y., Atsuta, N., Sobue, G., Morita, M. and Nakano, I. (2014) Prevalence and incidence of amyotrophic lateral sclerosis in Japan. *J. Epidemiol.*, **24**, 494–499.
 37. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L., Mora, S. et al. (2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.*, **45**, 1274–1283.
 38. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
 39. Noyce, A.J., Kia, D.A., Hemani, G., Nicolas, A., Price, T.R., De Pablo-Fernandez, E., Haycock, P.C., Lewis, P.A., Foltynie, T., Davey Smith, G. et al. (2017) Estimating the causal influence of body mass index on risk of Parkinson disease: a Mendelian randomisation study. *PLoS Med.*, **14**, e1002314.
 40. Cragg, J.G. and Donald, S.G. (1993) Testing identifiability and specification in instrumental variable models. *Econ. Theory*, **9**, 222–240.
 41. Burgess, S., Small, D.S. and Thompson, S.G. (2017) A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.*, **26**, 2333–2355.
 42. Brion, M.J., Shakhbazov, K. and Visscher, P.M. (2013) Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.*, **42**, 1497–1501.
 43. Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K. et al. (2018) Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat. Genet.*, **50**, 390–400.
 44. Burgess, S., Thompson, D.J., Rees, J.M.B., Day, F.R., Perry, J.R. and Ong, K.K. (2017) Dissecting causal pathways using Mendelian randomization with summarized genetic data: application to age at menarche and risk of breast cancer. *Genetics*, **207**, 481–487.
 45. Do, R., Willer, C.J., Schmidt, E.M., Sengupta, S., Gao, C., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J. et al. (2013) Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat. Genet.*, **45**, 1345–1352.
 46. Martin, S.S., Blaha, M.J., Elshazly, M.B., Toth, P.P., Kwiterovich, P.O., Blumenthal, R.S. and Jones, S.R. (2013) Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile. *JAMA*, **310**, 2061–2068.
 47. Bowden, J., Davey Smith, G., Haycock, P.C. and Burgess, S. (2016) Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.*, **40**, 304–314.
 48. Bowden, J., Del Greco, M.F., Minelli, C., Davey Smith, G., Sheehan, N.A. and Thompson, J.R. (2016) Assessing the suitability of summary data for two-sample Mendelian random-

- ization analyses using MR-Egger regression: the role of the I2 statistic. *Int. J. Epidemiol.*, **45**, 1961–1974.
49. Burgess, S. and Thompson, S.G. (2017) Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.*, **32**, 377–389.
 50. Burgess, S. and Thompson, S.G. (2015) Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am. J. Epidemiol.*, **181**, 251–260.
 51. Wadhera, R.K., Steen, D.L., Khan, I., Giugliano, R.P. and Foody, J.M. (2016) A review of low-density lipoprotein cholesterol, treatment strategies, and its impact on cardiovascular disease morbidity and mortality. *J. Clin. Lipidol.*, **10**, 472–489.
 52. Brown, L., Rosner, B., Willett, W.W. and Sacks, F.M. (1999) Cholesterol-lowering effects of dietary fiber: a meta-analysis. *Am. J. Clin. Nutr.*, **69**, 30–42.
 53. Williams, P.T., Krauss, R.M., Kindel-Joyce, S., Dreon, D.M., Vranizan, K.M. and Wood, P.D. (1986) Relationship of dietary fat, protein, cholesterol, and fiber intake to atherogenic lipoproteins in men. *Am. J. Clin. Nutr.*, **44**, 788–797.
 54. Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova, B., Emberson, J., Blackwell, L., Keech, A., Simes, J., Barnes, E.H., Voysey, M., Gray, A., Collins, R. and Baigent, C. (2012) The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet*, **380**, 581–590.
 55. Chen, X., Yazdani, S., Piehl, F., Magnusson, P.K.E. and Fang, F. (2018) Polygenic link between blood lipids and amyotrophic lateral sclerosis. *Neurobiol. Aging*, **67**, 202.e1–202.e6
 56. Lawlor, D.A., Harbord, R.M., Sterne, J.A., Timpson, N. and Davey Smith, G. (2008) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.*, **27**, 1133–1163.
 57. Paternoster, L., Tilling, K. and Davey Smith, G. (2017) Genetic epidemiology and Mendelian randomization for informing disease therapeutics: conceptual and methodological challenges. *PLoS Genet.*, **13**, e1006944.
 58. Burgess, S., Bowden, J., Fall, T., Ingelsson, E. and Thompson, S.G. (2017) Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology*, **28**, 30–42.
 59. The HUGO Pan-Asian SNP Consortium, Abdulla, M.A., Ahmed, I., Assawamakin, A., Bhak, J., Brahmachari, S.K., Calacal, G.C., Chaurasia, A., Chen, C.H., Chen, J. et al. (2009) Mapping human genetic diversity in Asia. *Science*, **326**, 1541–1545.
 60. Wang, Y., Lu, D., Chung, Y.J. and Xu, S. (2018) Genetic structure, divergence and admixture of Han Chinese, Japanese and Korean populations. *Hereditas*, **155**, 19.
 61. Ingman, M., Kaessmann, H., Pääbo, S. and Gyllensten, U. (2000) Mitochondrial genome variation and the origin of modern humans. *Nature*, **408**, 708–713.
 62. Wen, W., Zheng, W., Okada, Y., Takeuchi, F., Tabara, Y., Hwang, J.Y., Dorajoo, R., Li, H., Tsai, F.J., Yang, X. et al. (2014) Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum. Mol. Genet.*, **23**, 5492–5504.
 63. Foo, J.N., Tan, L.C., Irwan, I.D., Au, W.L., Low, H.Q., Prakash, K.M., Ahmad-Annuar, A., Bei, J., Chan, A.Y. and Chen, C.M. (2016) Genome-wide association study of Parkinson's disease in East Asians. *Hum. Mol. Genet.*, **26**, 226–232.
 64. He, M., Xu, M., Zhang, B., Liang, J., Chen, P., Lee, J.Y., Johnson, T.A., Li, H., Yang, X., Dai, J. et al. (2015) Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. *Hum. Mol. Genet.*, **24**, 1791–1800.
 65. Okada, Y., Sim, X., Go, M.J., Wu, J.Y., Gu, D., Takeuchi, F., Takahashi, A., Maeda, S., Tsunoda, T., Chen, P. et al. (2012) Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. *Nat. Genet.*, **44**, 904–909.
 66. Wen, W., Cho, Y.S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., Chen, C.H., Delahanty, R.J., Okada, Y., Tabara, Y. et al. (2012) Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.*, **44**, 307–311.
 67. Lu, X., Huang, J., Mo, Z., He, J., Wang, L., Yang, X., Tan, A., Chen, S., Chen, J., Gu, C.C. et al. (2016) Genetic susceptibility to lipid levels and lipid change over time and risk of incident hyperlipidemia in Chinese populations. *Circ. Cardiovasc. Genet.*, **9**, 37–44.
 68. VanderWeele, T.J., Tchetgen Tchetgen, E.J., Cornelis, M. and Kraft, P. (2014) Methodological challenges in Mendelian randomization. *Epidemiology*, **25**, 427–435.
 69. Nitsch, D., Molokhia, M., Smeeth, L., DeStavola, B.L., Whitaker, J.C. and Leon, D.A. (2006) Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. *Am. J. Epidemiol.*, **163**, 397–403.
 70. Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M. and Stephens, M. (2015) A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS One*, **10**, e0120758.
 71. Yavorska, O.O. and Burgess, S. (2017) MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.*, **46**, 1734–1739.
 72. Hartwig, F.P., Davey Smith, G. and Bowden, J. (2017) Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.*, **46**, 1985–1998.
 73. Thompson, S.G. and Sharp, S.J. (1999) Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat. Med.*, **18**, 2693–2708.