Assessing the Causal Role of Selenium in Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study

Di He (dh528@medschl.cam.ac.uk)  
Peking Union Medical College Hospital  
https://orcid.org/0000-0002-0755-2348

Liying Cui  
Peking Union Medical College Hospital

Research Article

Keywords: amyotrophic lateral sclerosis, two-sample Mendelian randomization, selenium, GWAS

Posted Date: March 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-358559/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. 
Read Full License
Abstract

Background: The relation between selenium overexposure and increased risk of amyotrophic lateral sclerosis (ALS) has been subject to considerable interest. Epidemiologic studies have reported suggestive associations between selenium and ALS, although the causal inference between selenium and ALS remains to be established. Here we conducted a two-sample Mendelian randomization (MR) analysis to analyze the causal role of selenium on ALS risk.

Methods: Variants associated with selenium levels were obtained from the GWAS meta-analysis of circulating selenium levels (n = 5,477) and toenail selenium levels (n = 4,162). Outcome data were from the largest ALS GWAS dataset with 20,806 ALS cases and 59,804 controls. Inverse variance weighted (IVW) method was used as the main analysis, with an array of sensitivity analyses performed to detect potential violations of MR assumptions.

Results: Inverse variance weighted (IVW) analysis indicated no evidence of a causal role for selenium levels in ALS development (odds ratio [OR] = 1.02, 95% confidence interval [CI] = 0.96–1.08). Similar results were observed for sensitivity analyses (OR = 1.00, 95% CI 0.95–1.07 for weighted median; OR = 1.07, 95% CI = 0.87–1.32 for MR-Egger), with no pleiotropy detected.

Conclusion: Although selenium was found associated with ALS according to earlier epidemiologic studies, current evidence does not support the causal effect of selenium on ALS risk. Correcting overall selenium levels in general population will unlikely reduce ALS incidence.

Introduction

Amyotrophic lateral sclerosis (ALS) is a paralytic disorder progressively affecting both upper and lower motor neurons [1, 2]. It is considered a complex genetic disease with Mendelian inheritance pattern observed in some familial cases, and the cause remains unclear in most sporadic patients [3, 4]. Although multiple ALS risk variants have been identified during the past two decades, likely due to incomplete penetrance, these implicated genotypes do not necessarily lead to disease phenotypes [5, 6]. Alternatively, it has been suggested that the manifestation of ALS is a stepwise process, in which predisposing variants carried by individuals interact with multiple environmental triggers [7, 8]. This multistep model empathizes the relevance of studying both genetic and environmental risk factors in ALS [9].

Among these environmental factors, studies in the past decades have highlighted the potential role of ionic homeostasis in the etiopathogenesis of ALS [10, 11]. In particular, suggestive epidemiologic evidence seem to support an association between increased ASL incidence and selenium exposure [12–14]. Such relation is further supported by evidence from biological research that certain selenium species may be detrimental to neurons [15, 16], which is the pathological feature of ALS. However, although the etiological role of environmental factors has been frequently investigated, to what extent the pathogenesis of ALS can be ascribed to these environmental risk exposures remains inconclusive [17]. For example, in retrospective studies, the concentrations of the suspected risk factors were usually
measured after disease onset, whereas the exposure might have taken place years before the onset. Such study design was thus limited by its inability to rule out reverse causality, in which the observed differences might be the consequence of disease progression. In addition, the questionnaire-based observational studies relying on self-reported information for the assessment of exposures are also subject to recall and selection biases [14]. The prospective case-control study design, on the other hand, is usually restricted by the modest number of cases enrolled, partly due to the low prevalence of ALS in general population [18]. Therefore, given the rarity of the disease and ethical issues, it is difficult to conduct unbiased environmental studies of ALS.

Two-sample Mendelian randomization (MR) analysis offers us the unique opportunity to probe the question of causality via exploiting the massive wealth of the ever-growing number of Genome-Wide Association Studies (GWAS). Analogous to the randomized controlled trial (RCT), two-sample MR uses genetic variants as unbiased proxies for random assignment, thereby enabling us to estimate the causal effect of exposures on the outcome of interest [19]. Two-sample MR is based on the natural genetic variation effect sizes on the exposure cohort and the outcome cohort, the statistics of which can be derived from their respective summarized GWAS dataset. If the exposure influences the outcome, then the influence of these valid genetic proxies on the outcome is proportional to their effect on the exposure. Since genetic variants are fixed at conception and temporally precede the outcome, MR is less likely biased by reverse causation and confounding [20]. In the present study, we evaluated the causal effects of selenium exposure on ALS risk by conducting a two-sample MR analysis with publicly available GWAS summary statistics.

**Material And Methods**

**Exposure dataset and genetic instruments**

Summary statistics for the genetic variants showing genome-wide significant association ($p < 5 \times 10^{-8}$) with selenium levels were obtained from the GWAS meta-analysis of circulating selenium levels ($n = 5,477$) and toenail selenium levels ($n = 4,162$) in European-ancestry individuals [21, 22]. Of note, since the units of toenail and blood selenium level were not comparable, the Z score were translated from $\beta$ (SE) for the analysis. The variants were clumped based on 1000 Genomes Project linkage disequilibrium (LD) structure ($R^2 < 0.3$ with any other associated SNP within 10 Mb) to ensure that the selected instrumental variables (IVs) were independently predicting the exposure. The proportion of phenotypic variance explained (PVE) by IVs as well as the F statistics were calculated to test the strength of the instruments.

**Outcome dataset**

The largest publicly available GWAS summary statistics for ALS involving 20,806 ALS cases and 59,804 controls of European ancestry was used as outcome data [23], which was comparable to the exposure dataset given the composition of population ethnicity. The analyses were restricted to ethnically homogeneous group to avoid population stratification [24]. Harmonization step was undertaken to rule
out strand mismatches [25]. Since only summarized statistics from publicly available GWAS was used, and no individual-level data was involved, ethical approval was not sought for the present study.

**Statistical analysis**

To estimate the causal effect of selenium exposure on ALS, individual Wald-type ratios for each of the IVs were meta-analyzed using the inverse–variance weighted (IVW) approach, with Cochran's Q statistic calculated for heterogeneity. Additionally, extensive sensitivity tests were performed to guard against potential violation of the model assumptions in MR analysis. Specifically, because the IVW estimate is not guarded against any SNPs violating the IV assumptions, weighted median method, which only requires the majority of variants being valid instruments, was included as complementary test [26, 27], whereas MR-Egger regression was performed to account for the bias caused by directional horizontal pleiotropy [28]. Outliers substantially influence causal effect were checked by leave-one-out (LOO) analysis and MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) [24]. Notably, because the summary statistics for selenium variants were expressed in Z-score units per allele [22], which were converted to beta and standard error values for the purpose of MR analysis, neither the effect sizes from MR analysis nor the beta values for associations of SNPs with selenium levels have interpretable units. All statistical analyses were conducted using R package TwoSampleMR (version 0.4.26)

**Results**

In total, 12 independent SNPs were selected as IVs (Table 1). No heterogeneity of effects was detected using Cochran's Q test (P = 0.08). The genetic instruments explained 0.32% – 1.76% of the variation in circulating and toenail selenium levels, and the F statistics were larger than 10 for all included IVs, which indicate that the instruments used in MR analysis were unlikely to suffer from weak instrument bias. The MR analysis did not support an association between selenium levels and ALS risk using the IVW method (OR = 1.02, 95% CI = 0.96–1.08) (Fig. 1A). Association estimates from sensitivity analyses such as weighted median and MR-Egger methods were consistent with that reported by IVW analysis, as summarized in Table 2.
Table 1
Characteristics of the instrumental variables used in two-sample MR analysis.

<table>
<thead>
<tr>
<th>Instrumental variables</th>
<th>Position (GRCh38.p13)</th>
<th>EA</th>
<th>EAF</th>
<th>Association with the exposure</th>
<th>Association with the outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>z-score</td>
<td>β (S.E.)</td>
</tr>
<tr>
<td>rs672413</td>
<td>5:78982406</td>
<td>A</td>
<td>0.32</td>
<td>7.53</td>
<td>0.16 (0.02)</td>
</tr>
<tr>
<td>rs705415</td>
<td>5:78996137</td>
<td>T</td>
<td>0.14</td>
<td>-6.23</td>
<td>-0.20 (0.03)</td>
</tr>
<tr>
<td>rs3797535</td>
<td>5:79004574</td>
<td>T</td>
<td>0.08</td>
<td>7.94</td>
<td>0.30 (0.04)</td>
</tr>
<tr>
<td>rs11951068</td>
<td>5:79008491</td>
<td>A</td>
<td>0.07</td>
<td>6.72</td>
<td>0.27 (0.04)</td>
</tr>
<tr>
<td>rs921943</td>
<td>5:79020653</td>
<td>T</td>
<td>0.29</td>
<td>13.14</td>
<td>0.29 (0.02)</td>
</tr>
<tr>
<td>rs10944</td>
<td>5:79090022</td>
<td>T</td>
<td>0.49</td>
<td>12.65</td>
<td>0.26 (0.02)</td>
</tr>
<tr>
<td>rs567754</td>
<td>5:79120593</td>
<td>T</td>
<td>0.34</td>
<td>-9.11</td>
<td>-0.20 (0.02)</td>
</tr>
<tr>
<td>rs558133</td>
<td>5:79129365</td>
<td>A</td>
<td>0.69</td>
<td>-6.55</td>
<td>-0.14 (0.02)</td>
</tr>
<tr>
<td>rs6859667</td>
<td>5:79449219</td>
<td>T</td>
<td>0.96</td>
<td>-6.92</td>
<td>-0.36 (0.05)</td>
</tr>
<tr>
<td>rs6586282</td>
<td>21:43058387</td>
<td>T</td>
<td>0.17</td>
<td>-5.89</td>
<td>-0.16 (0.03)</td>
</tr>
<tr>
<td>rs1789953</td>
<td>21:43062826</td>
<td>T</td>
<td>0.14</td>
<td>5.52</td>
<td>0.16 (0.03)</td>
</tr>
<tr>
<td>rs234709</td>
<td>21:43066854</td>
<td>T</td>
<td>0.45</td>
<td>-5.84</td>
<td>-0.12 (0.02)</td>
</tr>
</tbody>
</table>

EA, effect allele; EAF, effect allele frequency; β, per allele effect on the exposure; SE, standard error; P value, P value for the genetic association.
Table 2

Mendelian randomization results for the selenium levels and the risk in developing ALS.

<table>
<thead>
<tr>
<th>Method</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVW</td>
<td>1.02 (0.96–1.08)</td>
<td>0.61</td>
</tr>
<tr>
<td>Weighted median</td>
<td>1.00 (0.95–1.07)</td>
<td>0.88</td>
</tr>
<tr>
<td>MR Egger</td>
<td>1.07 (0.87–1.32)</td>
<td>0.52</td>
</tr>
</tbody>
</table>

The robustness of the results was confirmed by various sensitivity tests. The test for directional pleiotropy by MR-Egger did not give evidence for pleiotropy in the causality investigated as the intercept did not differ from zero (P = 0.60). This was supported by the funnel plot, which displayed symmetric pattern of effect size variation around the point estimate (Fig. 1B). The MR-PRESSO analysis detected no potential instrumental outlier (P = 0.12), and the LOO analysis also suggested that no single instrumental variable could disproportionally influence the estimated causal effect (Fig. 1C).

**Discussion**

According to previous epidemiology studies, the neurotoxic effects of excess selenium exposure may contribute to ALS etiology [29, 30]. However, observational studies are prone to reverse causation and various confounders, in which case incorrect causal inference might be made even with careful study design and statistical adjustment [31–33]. Here we leveraged the summary statistics from recent large-scale GWAS datasets to probe the association between selenium exposure and the risk for ALS. The current evidence did not support any causal relationship between the two, which is in accordance with the null association found between ALS and erythrocyte-bound selenium level in a recent prospective case-control study [18]. However, given the modest number of valid IVs available for this analysis and the relatively low percentage of variance in selenium level explained by these IVs, the statistical power to detect any postulated causal association might be limited. Therefore, until more genome-wide significant selenium variants are identified from future large scale GWAS studies, we cannot completely rule out the possibility that selenium exposure may influence the risk for ALS.

Since two-sample MR assumes that the SNPs influence the outcome because the hypothesized exposure does (vertical pleiotropy), three assumptions need to be satisfied for valid MR analysis: the genetic variants used as IVs are associated with the exposure (the relevance assumption); the genetic variants were not associated with any confounders (the independence assumption); and the genetic variants influence the risk of ALS only through the pathway of the exposure (the exclusion assumption) [20, 34]. Thus, to validate the IV assumptions, two alternative mechanisms need to be ruled out: IVs also being in LD with a causal variant for the outcome; IVs influencing the outcome through a pathway other than the exposure (horizontal pleiotropy) [35]. After LD-based clumping and pruning, multiple independent genetic variants reaching the conventional genome-wide significance level (thereby validating the relevance assumption) were meta-analyzed via IVW for an overall estimate of their effect on the outcome in our
study. However, although using multiple genetic variants can enhance the statistical power of MR analysis, the causal estimate would be liable to bias with inflated type I error rates if invalid IVs are included [24]. Thus, no variant having potential pleiotropic associations with ALS (defined by an ALS association p value below the genome-wide suggestive significance level of $10^{-5}$) was included as IV in the current MR analysis. Since the second and third IV assumptions are not fully testable in practice, we compared the estimates from a range of sensitivity analyses, which were in accordance with the IVW result.

Nonetheless, since metal homeostasis is critical for normal brain function, an excess of metal levels has been postulated as potential risk factors for a variety of neurodegenerative disorders [36]. Accordantly, the concentration of trace metals in Alzheimer's disease patients' hair and nails were found related to the clinical course of the disease [37]. It has been found that the concentration of selenium in urine and scalp hair was elevated in men, which is consistent with the epidemiologic findings that ALS is more common in men than in women [6, 38, 39]. The neurotoxic effects of selenium might be mediated by inducing oxidation of thiol-containing protein and promoting translocation of copper/zinc superoxide dismutase (SOD1) into mitochondria [40]. However, the biomarkers currently used to assess selenium exposure have various inherent limitations, and the reliability of these assessment methods in reflecting the long-term cumulative exposure of selenium has been debated and challenged [41]. In addition, peripheral indicators of selenium exposure may not necessarily correspond to its CNS content, given the independence of selenium level in paired serum and cerebrospinal fluid (CSF) samples [42]. Thus, despite the negative results from the MR analysis, further functional studies investigating the association between selenium and ALS are still warranted.

The study is subject to a number of limitations. First, although there is evidence supporting the existence of variation in the concentration of metals/metalloids by age and gender [38], we cannot decide whether there is any age- or gender-specific effect of selenium exposure on ALS, as individual-level GWAS datasets are not accessible. Second, to avoid population stratification, we focused on subjects of European ethnicity. Whether the findings may be extended to other populations remains unclear. Finally, although the MR-Egger regression results did not support horizontal pleiotropy, it is difficult to completely rule out pleiotropy or alternative causal pathways in MR analyses. In addition, MR analysis assumed linearity and homogeneity between the exposure, the genetic variants, and the risk for ALS, which may not represent the true associations in nature. This could potentially limit us from identifying putative thresholds of exposure above or below which the exposure can induce specific effects.

In conclusion, using summary statistics from GWAS, we did not find strong evidence for the causal inference of selenium on the risk of ALS in the present study. Such findings might be informative for epidemiologic studies of ALS in the future.

Declarations

Acknowledgement
We thank the research groups of the cited GWAS studies for making their data available and study participants.

**Funding statement**

This work was supported by Chinese Academy of Medical Science Innovation Fund for Medical Sciences (Grant number: 2016-I2M-1-004); the National Key Research and Development Program of China (Grant number: 2016YFC0905100 and 2016YFC0905103).

**Author contributions**

D.H. and L.C. contributed to the conception and design of the study; D.H., contributed to the analysis of data and drafting the manuscript. Both authors read and approved the manuscript.

**Data availability statement**

The present analysis was conducted using publicly available GWAS datasets provided by the original authors of the respective studies.

**Compliance with ethical standards**

**Potential conflicts of interest**

No conflict of interest to be disclosed.

**Consent to participate**

Owing to the use of publicly available deidentified GWAS data, this study did not require institutional review board approval. Ethical approval had been obtained in the original studies cited.

**Consent for Publication**

Not applicable.

**References**


Figures

A. 

B. 

C. 

Figure 1
Association estimates from sensitivity analyses such as weighted median and MR-Egger methods were consistent with that reported by IVW analysis, as summarized in Table 2. The robustness of the results was confirmed by various sensitivity tests. The test for directional pleiotropy by MR-Egger did not give evidence for pleiotropy in the causality investigated as the intercept did not differ from zero (P = 0.60). This was supported by the funnel plot, which displayed symmetric pattern of effect size variation around the point estimate (Figure 1B). The MR-PRESSO analysis detected no potential instrumental outlier (P = 0.12), and the LOO analysis also suggested that no single instrumental variable could disproportionally influence the estimated causal effect (Figure 1C).