Pre-Morbid Risk Factors for Amyotrophic Lateral Sclerosis: Prospective Cohort Study

G David Batty 1
Catharine R Gale 2

1Department of Epidemiology and Public Health, University College London, London, UK; 2Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK

Aim: In the absence of effective treatments for amyotrophic lateral sclerosis (ALS), a neurodegenerative disorder with high case fatality, there is a clear need to identify its primary risk factors.

Methods: UK Biobank is a prospective cohort study in which baseline data were captured between 2006 and 2010 in 502,649 participants aged 37 to 73 years. Follow-up for ALS hospitalisations and death was made via national registries.

Results: Eleven years of event surveillance gave rise to 301 hospitalisations and 279 deaths due to ALS. After adjustment for selected confounding factors, being older (hazard ratio per 10 year increase; 95% confidence interval: 1.92; 1.58, 2.33) and male (1.37; 1.00, 1.87) were associated with elevated rates of hospitalisation for ALS. Similar effects were apparent when death ascribed to the disorder was the outcome of interest. Of the remaining 23 social, biological, and behavioural risk indices, however, there was only a suggestion that taller people experienced an increased risk of hospitalisation (per SD increase: 1.31; 1.09, 1.59).

Conclusion: In the present, large-scale study, other than well known associations, we did not find convincing evidence of links with ALS for other risk indices.

Keywords: risk factors, amyotrophic lateral sclerosis, cohort study, UK Biobank

Introduction

Amyotrophic lateral sclerosis (ALS), also known as motor neuron disease, involves the unabated degeneration of nerve cells responsible for voluntary muscle movement. With there being no effective treatment, in most cases, death, typically from respiratory failure, commonly occurs within 3 years of symptom emergence. 1,2 This brings into sharp focus the need for primary prevention research.

While marked geographical differentials in case occurrence 3 implicates both genetic and non-genetic causes of ALS, a secular increase in rates over three decades 4 – seemingly too rapid to be explained by changes in the gene pool – provides some circumstantial support for a role of environmental determinants. There is strong evidence to suggest being both male and older are associated with an elevation in risk, 5 however, studies examining the role of modifiable environmental characteristics report inconsistent and often counter-intuitive results. Thus, although higher pre-diagnostic body weight appears to confer protection against ALS, 6,7 cigarette smoking has revealed discordant associations, 8,9 as have physical activity, 10 alcohol intake, 11,12 socioeconomic status, 13,14 and physical comorbidities such as diabetes, 15,16 amongst many other factors. 17
The heterogeneity of these findings may, at least in part, be explained by different methodological shortcomings across studies. Given the rarity of ALS, the most time- and cost-efficient method for identifying risk factors is the case-control study, interpretation of findings from which is complicated by recall bias and, potentially, reverse causality, such that biomedical indices, for instance, are captured in the presence of the disorder. Accordingly, we used a well-characterised prospective cohort study of around half a million people with ongoing surveillance for mortality and morbidity from ALS in order to better understand the role, if any, of a wide range of modifiable biological, psychosocial, and behavioural characteristics. In particular, our purpose was to further examine some extant associations and also to relate previously untested characteristics to the occurrence of the disorder. There are, for example, reasons to anticipate associations between ALS and earlier measurements of cognitive function, psychological distress, and lung function, not least because these characteristics may be pre-clinical features of the disorder. Although these indices have shown prognostic value in studies of patient groups – less favourable levels are associated with adverse outcomes including disability severity and reduced life expectancy – they are untested as risk factors for disease occurrence.

Methods

UK Biobank, a UK-wide, closed, prospective cohort study, has been described in detail. In brief, all people aged 40–69 years who were registered with the UK National Health Service and living a maximum of 25 miles from one of the 22 study assessment centres were invited to participate. Of the 9.2 million people asked, 503,325 study members (5.5% response) completed a questionnaire, underwent an interview, and took part in various physical assessments. Ethical approval was obtained from the National Health Service National Research Ethics Service with all participants providing written consent. Using fully anonymised data, the present analyses did not require additional permissions. We followed STROBE guidelines for the presentation of original epidemiological research.

Assessment of Baseline Characteristics

Ethnicity was self-reported and categorised as White, Asian, Black, Chinese, Mixed, or other ethnic group. A social isolation scale was derived from enquiries concerning number of people in the household, visits with friends/family, and social activities. For educational qualifications, we used a two category variable (college/undergraduate degree versus lower). Total annual household income before tax was self-reported. The Townsend deprivation index was our indicator of neighbourhood socio-economic circumstances; a continuously scored variable, higher values denote greater deprivation. Based on national census data, scores comprise information on employment, overcrowding, and car and home ownership. Each participant was assigned a Townsend value corresponding to the postcode of their home address.

Cigarette smoking, physical activity, and alcohol consumption were all self-reported. Smoking status was categorised into never, former, and current; alcohol intake was grouped into daily/almost daily, three to four times a week, once or twice a week, one to three times a month, special occasions only and never; and physical activity was captured using the International Physical Activity Questionnaire short form. Height and weight were recorded directly and body mass index was calculated using the usual formula (weight, kg/height, m²); overweight and obesity were denoted by values ≥25kg/m². Self-reported physician diagnosis was collected for ALS, vascular or heart problems, diabetes, and cancer. We defined hypertension according to existing guidelines: measured systolic/diastolic blood pressure ≥140/90 mmHg (two assessments) and/or self-reported use of anti-hypertensive medication.

Forced expiratory volume in one second, a measure of pulmonary function, was quantified using spirometry with the best of three technically satisfactory exhalations used in our analyses. Handgrip strength was measured using a hydraulic hand dynamometer with the participant maximally squeezing the handle of the dynamometer while seated for 3 seconds; an average of readings from the right and left hand was used. Non-fasting venous blood, available in a sub-sample, was drawn with assaying for high-density lipoprotein cholesterol and glycated haemoglobin A1c conducted at a single laboratory.

Study members were asked if they had ever been under the care of a psychiatrist for any mental health problem. Symptoms of psychological distress – anxiety, worrying, anhedonia, and depression – were measured using the four item version of the Patient Health Questionnaire (PHQ-4) in which individual items are rated on a Likert scale from 0 (“not at all”) to 3 (“nearly every day”) such that total scores range from 0 to 12 (higher scores denote greater distress). A score of 3 or above was used to indicate high distress in....
the present analyses.\textsuperscript{31} Neuroticism was measured with the 12-item Eysenck Personality Questionnaire-Revised Short Form;\textsuperscript{34,35} again, higher scores denote higher levels. Verbal and numerical reasoning was captured using a computerized 13-item multiple-choice test with a two-minute time limit; the score was the number of correct answers.\textsuperscript{36} This test was introduced after the beginning of the baseline assessment period so data are only available for a subset of study members (N=180,914).\textsuperscript{24,37} Reaction time was measured using a computerized Go/No-Go “Snap” game.\textsuperscript{37} Participants were presented with electronic images of two cards. If symbols on the cards were identical, participants were instructed to immediately push the button-box using their dominant hand. The first five pairs were used as a practice with the remaining seven pairs, containing four identical cards, forming the assessment. Reaction time score was the mean time to press the button when each of these four pairs was presented. Choice reaction time correlates strongly with single mental tests that involve complex reasoning and knowledge.\textsuperscript{38}

**Ascertainment of Amyotrophic Lateral Sclerosis Cases**

Study participants were linked to the National Health Service’s Central Registry which provided vital status data on study members and, where applicable, cause of death. Linkage was also made to hospital in-patients records via the Hospital Episode Statistics, a registry of all hospitalisations in the UK.\textsuperscript{39} Using both databases, ALS was denoted by International Classification of Disease code G122.

**Derivation of Analytical Sample and Statistical Analyses**

To addresses concerns regarding reverse causality – the notion that ALS might influence the exposure of interest (eg, physical activity, lung function) rather than the opposite – we excluded 56 people who self-reported ALS at baseline medical examination. Additionally, to capture study members with potentially subclinical (undiagnosed) ALS, we left-censored study members such that those who were hospitalised for or died from the condition within the first 3 years of baseline were also excluded (N=72) – as described, death commonly occurs within this period after symptom emergence.\textsuperscript{1,2} This resulted in an analytical sample of 502,599 people (273,454 women). To summarise risk factor–ALS associations, we computed hazard ratios with accompanying 95% confidence intervals using Cox proportional hazards regression.\textsuperscript{40} In these analyses, we used calendar period as the time scale and study members were censored at date of hospitalisation or death from ALS, or end of follow-up (4th October 2020 for mortality, 24th November 2020 for hospitalisation) – whichever came first. We used Stata version 15 for all analyses.

**Results**

Event surveillance gave rise to 301 hospitalisations (mean 11.8 years of follow-up) and 279 deaths (11.4 years) due to ALS. In Tables 1 (hospitalisations) and 2 (deaths) we show the relationships between twenty-five baseline characteristics and the ALS outcomes. After adjustment for selected confounding factors, being older (hazard ratio per 10 year increase; 95% confidence interval: 1.92; 1.58, 2.33) and male (1.37; 1.00, 1.87) were associated with elevated rates of hospitalisation for ALS. Similar effects were apparent when death ascribed to the disorder was the outcome of interest. Of the remaining 23 potential risk indices, however, there was only a suggestion that taller people experienced an increased risk of hospitalisation (per SD increase: 1.31; 1.09, 1.59); weaker effects were apparent in the mortality analyses (1.21; 0.98, 1.48). While there was some indication of an elevated rate of hospitalisations in people who reported cigarette smoking, confidence intervals included unity and there was marked attenuation after adjustment for covariates which included socioeconomic status and health behaviours (1.10; 0.73, 1.66).

**Discussion**

The main finding of the present study was that there was no clear evidence of an association between ALS and the twenty-five risk indicators. The exceptions were the two most established risk factors – age and sex; that these associations were recapitulated here gives us some confidence in the generally negative results for the remainder of the baseline characteristics. These null findings are in keeping with observations from systematic reviews of risk factors for ALS\textsuperscript{17,20,41,42} and those for other neurodegenerative disorders such as dementia.\textsuperscript{43–45}

The positive association between height and ALS – also apparent for cancer\textsuperscript{46,47} – potentially implicates early life exposures in the occurrence of this condition. The same has been speculated for dementia,\textsuperscript{48} although a reverse gradient to that seen herein has been reported.\textsuperscript{49} While under a degree of genetic control, relative to their shorter counterparts, taller people have, on average, been exposed to a more favourable pre-adult environment, which includes an optimal diet, fewer bouts of illness,
more affluent social circumstances, and a lower stress household,\textsuperscript{50} though we are unable to identify which, if any, of these correlates may be exerting an influence on ALS risk.

**Study Strengths and Weaknesses**

The present study has its strengths, including its size which facilitates the accumulation of a sufficiently high number of cases for analyses of a rare health outcome alongside left-censoring to take into account reverse causality, and the well characterised nature of the study participants. Further, in the aetiological literature, investigators often focus on the predictive value of a single characteristic rather than take a comprehensive approach, as we have, so facilitating cross-comparison. Inevitably, however, our work has several weaknesses. First, the present study sample comprises only the 5.5% of the target population who agreed to participate.\textsuperscript{51} As has been demonstrated,\textsuperscript{52,53} the data material is therefore inappropriate for estimation of risk factor or disease prevalence

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
 & \textbf{Age- and Sex-Adjusted} & & \textbf{Multiply-Adjusted} & \\
 & \textbf{N Events/N Risk} & \textbf{Hazard Ratio (95% Confidence Interval)} & \textbf{N Events/N Risk} & \textbf{Hazard Ratio (95% Confidence Interval)} \\
\hline
\textbf{Demographic factors} & & & & \\
Age (yrs), per 10 year increase & 301/502,524 & 1.96 (1.67, 2.31) & 258/446,321 & 1.92 (1.58, 2.33) \\
Male & 301/502,524 & 1.23 (0.98, 1.54) & 258/446,321 & 1.37 (1.00, 1.87) \\
Non-white ethnicity & 300/499,634 & 0.65 (0.32, 1.31) & 258/446,321 & 0.42 (0.17, 1.04) \\
Socially isolated & 301/502,524 & 0.91 (0.60, 1.36) & 258/446,321 & 0.82 (0.52, 1.31) \\
\hline
\textbf{Lifestyle factors} & & & & \\
Current smoker & 300/499,571 & 1.71 (0.81, 1.68) & 258/446,321 & 1.10 (0.73, 1.66) \\
No physical activity & 298/495,378 & 0.89 (0.55, 1.45) & 258/446,321 & 0.88 (0.51, 1.53) \\
Drinks alcohol daily/almost daily & 300/501,021 & 0.87 (0.65, 1.15) & 257/446,013 & 0.89 (0.66, 1.22) \\
Obese/overweight & 296/499,469 & 0.80 (0.62, 1.03) & 257/446,013 & 0.86 (0.65, 1.14) \\
\hline
\textbf{Comorbidities} & & & & \\
Vascular or heart disease & 300/500,300 & 1.13 (0.89, 1.44) & 257/445,599 & 1.08 (0.83, 1.42) \\
Hypertension & 296/493,912 & 0.99 (0.77, 1.27) & 256/443,514 & 1.03 (0.78, 1.35) \\
Diabetes & 299/499,907 & 1.10 (0.70, 1.72) & 258/446,321 & 1.10 (0.67, 1.82) \\
Cancer & 200/499,749 & 1.17 (0.81, 1.68) & 257/445,179 & 1.27 (0.85, 1.88) \\
\hline
\textbf{Biomarkers} & & & & \\
Lung function, per SD (0.8 L) increase & 298/453,787 & 0.91 (0.74, 1.11) & 258/446,321 & 0.92 (0.77, 1.09) \\
Hand grip strength, per SD (11.3 kg) increase & 299/500,164 & 0.97 (0.82, 1.16) & 258/445,857 & 1.03 (0.84, 1.25) \\
High-density lipoprotein, 25th vs 75th centile & 279/466,404 & 0.92 (0.65, 1.30) & 243/418,663 & 1.10 (0.73, 1.64) \\
HbA1C, 25th vs 75th centile & 262/429,780 & 1.16 (0.81, 1.68) & 224/385,554 & 1.18 (0.78, 1.77) \\
Height, per SD (9.2 cm) increase & 296/499,990 & 1.17 (0.99, 1.38) & 257/445,898 & 1.31 (1.09, 1.59) \\
\hline
\textbf{Psychological factors} & & & & \\
Psychiatric consultation & 300/498,839 & 1.12 (0.79, 1.58) & 257/444,419 & 1.03 (0.70, 1.52) \\
High psychological distress & 267/448,959 & 1.25 (0.94, 1.65) & 232/401,583 & 1.27 (0.93, 1.72) \\
Neuroticism, per SD (3.28 points) increase & 154/425,652 & 1.03 (0.88, 1.21) & 132/381,324 & 1.03 (0.86, 1.23) \\
Reasoning, per SD (2.16 points) increase & 120/180,869 & 0.78 (0.64, 0.93) & 103/162,497 & 0.85 (0.69, 1.05) \\
Reaction time, per SD (118.2 ms) increase & 299/496,751 & 1.09 (0.98, 1.21) & 256/442,583 & 1.06 (0.94, 1.19) \\
\hline
\textbf{Socioeconomic factors} & & & & \\
No college/university degree & 293/492,384 & 1.11 (0.86, 1.44) & 254/442,098 & 1.15 (0.87, 1.54) \\
Low annual household income & 250/425,341 & 0.99 (0.74, 1.31) & 218/383,913 & 0.88 (0.64, 1.23) \\
Deprivation score, per point disadvantage & 299/501,897 & 1.01 (0.97, 1.05) & 258/446,321 & 1.01 (0.97, 1.06) \\
\hline
\multicolumn{4}{l}{Notes: Multiple adjustment is adjustment for age, sex, ethnicity, smoking status, physical activity, diabetes, lung function, and deprivation score.} \\
\end{tabular}
\end{table}
for event incidence, including presumably ALS. These observations do not, however, seem to influence reproducibility of the association of established risk factors for non-communicable disease such as vascular disease and selected cancers, and other health endpoints such as suicide and selected cancers.

Second, while we set out to further test the relation of selected risk factors (e.g., cigarette smoking, diabetes), and potentially identify novel ones (e.g., cognitive function, psychological distress), several other factors (e.g., alcohol intake, extant cancer, blood lipids) were not hypothesis-driven. Third, baseline data are certainly time-varying in the period between study induction and end of follow-up. This is a perennial issue in cohort studies and one we were able to investigate using data from a resurvey that took place around 8 years after baseline examination in a representative sub-sample of participants.

### Table 2 Hazard Ratios for the Association Between Baseline Characteristics and Death from Amyotrophic Lateral Sclerosis, UK Biobank 2006 to 2020

<table>
<thead>
<tr>
<th>Demographic factors</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>Age (yrs), per 10 year increase</td>
<td>279/502,524</td>
<td>2.42 (2.02, 2.90)</td>
</tr>
<tr>
<td>Male</td>
<td>279/502,524</td>
<td>1.28 (1.01, 1.62)</td>
</tr>
<tr>
<td>Non-white ethnicity</td>
<td>278/499,634</td>
<td>0.46 (0.19, 1.12)</td>
</tr>
<tr>
<td>Socially isolated</td>
<td>279/502,524</td>
<td>0.94 (0.61, 1.43)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lifestyle factors</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>Current smoker</td>
<td>277/499,571</td>
<td>1.16 (0.79, 1.72)</td>
</tr>
<tr>
<td>No physical activity</td>
<td>274/495,378</td>
<td>0.98 (0.59, 1.62)</td>
</tr>
<tr>
<td>Drinks alcohol daily/almost daily</td>
<td>278/501,021</td>
<td>1.10 (0.84, 1.46)</td>
</tr>
<tr>
<td>Obese/overweight</td>
<td>277/499,469</td>
<td>0.90 (0.69, 1.18)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comorbidities</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>Vascular or heart disease</td>
<td>278/500,300</td>
<td>1.09 (0.85, 1.40)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>274/493,912</td>
<td>1.02 (0.78, 1.32)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>278/499,907</td>
<td>0.91 (0.55, 1.50)</td>
</tr>
<tr>
<td>Cancer</td>
<td>277/499,749</td>
<td>1.02 (0.68, 1.54)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>Lung function, per SD (0.8 L) increase</td>
<td>242/453,729</td>
<td>1.03 (0.87, 1.22)</td>
</tr>
<tr>
<td>Hand grip strength, per SD (11.3 kg) increase</td>
<td>278/500,164</td>
<td>1.02 (0.85, 1.22)</td>
</tr>
<tr>
<td>High-density lipoprotein, 25th vs 75th centile</td>
<td>237/429,780</td>
<td>1.06 (0.72, 1.56)</td>
</tr>
<tr>
<td>HbA1C, 25th vs 75th centile</td>
<td>255/466,404</td>
<td>0.90 (0.62, 1.32)</td>
</tr>
<tr>
<td>Height, per SD (9.28 cm) increase</td>
<td>277/499,990</td>
<td>1.17 (0.99, 1.39)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Psychological factors</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>Psychiatric consultation</td>
<td>277/498,839</td>
<td>0.85 (0.57, 1.27)</td>
</tr>
<tr>
<td>High psychological distress</td>
<td>244/448,959</td>
<td>0.99 (0.72, 1.36)</td>
</tr>
<tr>
<td>Neuroticism, per SD (3.28 points) increase</td>
<td>100/425,652</td>
<td>0.93 (0.76, 1.15)</td>
</tr>
<tr>
<td>Reasoning, per SD (2.16 points) increase</td>
<td>88/180,869</td>
<td>0.82 (0.66, 1.02)</td>
</tr>
<tr>
<td>Reaction time, per SD (118.2 ms) increase</td>
<td>275/496,751</td>
<td>0.97 (0.86, 1.10)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Socioeconomic factors</th>
<th>Age- and Sex-Adjusted</th>
<th>Multiply-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Events/N Risk</td>
<td>Hazard Ratio (95% Confidence Interval)</td>
</tr>
<tr>
<td>No college/university degree</td>
<td>272/492,384</td>
<td>1.09 (0.83, 1.42)</td>
</tr>
<tr>
<td>Low annual household income</td>
<td>227/425,341</td>
<td>0.96 (0.71, 1.29)</td>
</tr>
<tr>
<td>Deprivation score, per point disadvantage</td>
<td>278/501,897</td>
<td>0.99 (0.95, 1.03)</td>
</tr>
</tbody>
</table>

**Notes:** Multiple adjustment is adjustment for age, sex, ethnicity, smoking status, physical activity, diabetes, lung function, and deprivation score.

and for event incidence, including presumably ALS. These observations do not, however, seem to influence reproducibility of the association of established risk factors for non-communicable disease such as vascular disease and selected cancers, and other health endpoints such as suicide and selected cancers. We think the same reasoning can be applied to the present disease outcome.
(r=0.90, p < 0.001, N=34,662), whereas the magnitude was somewhat lower for co-morbidities such as diabetes (r=0.63, P < 0.001, N= 31,037) and serious mental illness (r=0.64, p < 0.001, N=47,291), as it was for cigarette smoking (r=0.60, p < 0.001, N=31,037). Lastly, we did not have data on event onset, only in-patient hospitalisation and death.

In conclusion, other than the known associations for age and sex, we did not find convincing evidence of links with ALS for other risk indices in the present analyses.

Access to Data
Data from UK Biobank (http://www.ukbiobank.ac.uk/) are available to bona fide researchers upon application. This research has been conducted under application 10279.

Acknowledgments
We thank the participants in UK Biobank for their forbearance.

Funding
GDB is supported by the Medical Research Council (MR/P023444/1) and the US National Institute on Aging (1R56AG052519-01; 1R01AG052519-01A1).

Disclosure
The authors report no conflicts of interest in this work.

References


