**Long-term traffic-related air pollutant exposure and amyotrophic lateral sclerosis diagnosis in Denmark using Bayesian Hierarchical Regression Modelling**

*Robbie M Parks, PhD*

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA

The Earth Institute, Columbia University, New York, New York, USA

*Arin Balilian, MD, MPH*

Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA

*Yanelli Nunez, PhD*

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA

*Johnni Hansen, PhD*

Danish Cancer Society Research Center, Copenhagen, Denmark

*Ole Raaschou-Nielsen, PhD*

Danish Cancer Society Research Center, Copenhagen, Denmark

*Matthias Ketzel, PhD*

Department of Environmental Science, Aarhus University, Roskilde, Denmark

*Jibran Khan, PhD*

Department of Environmental Science, Aarhus University, Roskilde, Denmark

*Marc G. Weisskopf, ScB, PhD, ScD*

Department of Environmental Health, T. H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA

*Roel Vermeulen, PhD*

Institute for Risk Assessment Sciences, Universiteit Utrecht, Utrecht, the Netherlands

*Susan Peters, PhD*

Institute for Risk Assessment Sciences, Universiteit Utrecht, Utrecht, the Netherlands

*Diane B. Re, PhD*

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA

*Marianthi-Anna Kioumourtzoglou, ScD*

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA

**Corresponding Author:**

Robbie M Parks

Department of Environmental Health Sciences

Columbia University Mailman School of Public Health

722 West 168th Street, #1104

New York, New York, 10032

Email: [robbie.parks@columbia.edu](mailto:robbie.parks@columbia.edu)

**Word Count:**

Abstract: XX words

Main Text: XX words

**Key Points**

**Question:** How are traffic-related pollutant concentrations associated with diagnosis of amyotrophic lateral sclerosis (ALS)?

**Findings:** In this largest study case-control study of ALS diagnosis to date that included 4,011 diagnoses of ALS in Denmark, we observed that a standard deviation increase of 5-year concentration of traffic-related pollutants was associated with a non-significant increased risk of ALS diagnosis, with an 86.5% posterior probability of a positive association.

**Meaning:** Our results indicate a potential positive association between ALS diagnosis and traffic-related pollution. Further work is needed to understand the role of air pollution on ALS pathogenesis and timing of onset.

**Abstract**

**Importance:** Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disease. There is some limited evidence to suggest ALS onset is associated with exposure to air pollution, and specifically to traffic-related pollution.

**Objective:** To determine whether exposure to components of traffic-related pollution is associated with ALS diagnosis.

**Design:** In this case-control study, we used Bayesian Hierarchical Regression Modelling (BHRM) in a conditional logistic model. We used prospectively collected data from the Danish National Registers system from 4,011 ALS cases diagnosed between 1989 – 2013 and matched on age, sex, and vital status to 20,055 controls. We used predictions from a validated spatio-temporal model to assign 5-year average exposures prior to diagnosis to combined traffic-related pollutants, as well as nitrogen oxides (NOx), carbon monoxide (CO), elemental carbon (EC), fine particles (PM2*.*5), and ozone (O3) at residential addresses of study participants.

**Setting:** We used prospectively collected ALS diagnosis case data from the Danish National Patient Register and control data from the Danish Civil Registration System.

**Participants:** All adults in Denmark between 1989 – 2013.

**Main Outcome Measure:** ALS diagnosis in Denmark during 1989 – 2013.

**Results:** We found that for a standard deviation (SD) increase in 5-year average concentrations, the joint effect of included traffic-related pollutants (NOx, CO, EC) was associated with an increase in odds of ALS diagnosis (XX%; 95% credible interval [CrI]:XX, XX%), with an increase in odds for elemental carbon (SD=XX µg/m3) (XX%; 95%CrI: -XX%, XX%). Overall, there was an 86.5% posterior probability of a positive association between the joint effect of included traffic-related pollutants and ALS diagnosis.

**Conclusion:** Our results indicate a potential positive association between ALS diagnosis and traffic-related pollution. Further work is needed to understand the role of air pollution on ALS pathogenesis and timing of onset.

**Introduction**

XX

**Methods**

*Study population and Outcome Assessment*

We used data from the Danish National Registers system during 1989 – 2013, through which details on demographic characteristics and certain health outcomes of all Danish residents can be linked based on a 10-digit unique personal identifier.1 The Danish National Registers system was established in 1977 and is a comprehensive patient register, including nationwide clinical and administrative records for all somatic inpatient data. Outpatient data have also been included in the Danish National Registers system since 1995. In a previous validation study, we found that Danish National Registers system data for ALS ascertainment are highly reliable, with a positive predictive value of 0.93 (95% CI, 0.88-0.96).2

We identified ALS cases based on their International Classification of Diseases (ICD) discharge diagnoses, i.e., ICD-8 code 348.0 (ALS) until 1993 and ICD-10 code G12.2 (motor neuron disease) thereafter. For the diagnosis date, we used the date of the first relevant code. We only included patients who were at least 20 years old when diagnosed. We found controls through the Danish Civil Registration System, which was established in 1968 and includes administrative records (e.g., date and place of birth, vital status, and history of civil status and addresses) on all persons living in Denmark; records are kept even when a person dies or emigrates.3 We identified controls as any person with no mention – up to the diagnosis date for the matched case – of ICD-8 code 348.0 or ICD-10 G12.2 in the Danish National Registers system. We randomly matched five controls per case by age, sex and date of birth. Controls also had to be alive at the time of first mention of ALS in the Danish National Registers system of the matched case.

We obtained all addresses of cases and controls from January 1st 1979 onwards from the Danish Civil Registration System, including the dates of moving to and leaving from each address, prior to the case diagnosis date. We then obtained the geographical co-ordinates at the door of the house of each house of the residential history of the participants. A previous study has demonstrated the high accuracy of the geocoding of addresses in Denmark.4

*Exposure data*

We obtained predictions on monthly concentrations of nitrogen oxides (NOx), carbon monoxide (CO), elemental carbon (EC), fine particles (PM2*.*5), and ozone (O3) at residential addresses of study participants from a validated spatio-temporal model with full space and time coverage over our study period, described in detail elsewhere.5,6 The predictions in pollutant concentrations have been extensively used in previous air pollution epidemiologic studies in Denmark.4,7–9 Based on the residential history of each case or control, we then calculated 1-, 5-, and 10-year average exposure to each pollutant before the date of the associated case’s ALS diagnosis. A small number of the Danish Civil Registration System lack a complete address history, indicated when the address history for a subject is incomplete (typically lack of house number: 1.7% of addresses). We therefore set the following criteria for including cases and controls across the length of exposure averages: (i) 1-year averages: 9 out of 12 months with complete exposure records, and at least one measurement in each season; (ii) 5-year averages: at least 30 out of 60 months with complete exposure records; and (iii) 10-year averages: at least 60 out of 120 months with complete exposure records.

*Covariate data*

We included a set of covariates to account for potential variation between the matched cases and controls. We used the five-category socioeconomic status (SES) definitions developed by the Danish Institute of Social Sciences, which are based on job titles and income tax forms. The group with the highest status (group 1) includes corporate managers and academics; group 2 includes proprietors, managers of small businesses and teachers; group 3 includes technicians and nurses; group 4 includes skilled workers; and group 5 includes unskilled workers. We also included an additional group for unemployed participants (group 9). If a participant were married and information was available, we used the higher of the couple’s individual SES ranks. We also used information on civil status (never married, married, divorced and widowed), place of residence (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland) and place of birth (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland, foreign, unknown) to adjust for other family-specific, location-specific and early-life potential confounders.

*Statistical analysis*

We analyzed the association between ALS diagnosis (binary outcome; 1 for diagnosed cases and 0 for matched controls) and exposure to traffic-related pollutants by applying a Bayesian formulation of the conditional logistic model, with Bayesian Hierarchical Regression Modelling (BHRM) on the traffic-related pollutants. The logistic regression model is appropriate for binary outcomes regressed against continuous or discrete variables. The conditional approach automatically accounts for matching factors (age, sex and date of birth) between cases and controls within each strata, which here are the groupings of case and matched controls. Bayesian inference allows for full distributional estimation of the parameters of interest. The BHRM formulation on the traffic-related pollutants is a mixture method which allows a group effect, as well as the individual component effects, while accounting for the associated variance-covariance structure between the highly-correlated exposures. We included a linear term for each included pollutant. We also adjusted by SES, civil status, place of residence and place of birth.

Specifically, via a logit function, we modelled ALS diagnosis, as follows:

where denotes whether there was an ALS diagnosis for subject in group ; the common intercept; the group-specific intercepts for each subject (not estimated in the conditional logistic model); ,,the pollutant-specific coefficients (log-odds) per standard deviation increase in concentration of , , respectively; and the rest as subject-specific covariates.

In addition, a hierarchy was placed on the traffic-specific pollutant terms in the model:

XX

We used weakly-informative priors so that parameter estimation was driven by the data. <<<Details of priors here.>>>

We present all results as percentage change in ALS diagnosis per standard deviation increase in pollutant concentration. We conducted statistical analyses using the R Statistical Software, version 4.1.1 (Foundation for Statistical Computing, Vienna, Austria), and R-STAN, version 2.21.2. All code for analysis and visualization presented in this manuscript will be publicly available via GitHub.

<<<Sensitivity analysis. Priors etc.>>>

**Results**

After filtering the original 4,011 cases and 20,055 controls based on completeness of exposure and covariate records, we were left with (i) XX cases and XX controls for 1-year average exposure; (ii) XX cases and XX controls for 5-year average exposure; and (iii) XX cases and XX controls for 5-year average exposure. Descriptive statistics of included cases and controls can be found in Table XX.

* Paragraph about exposure, referring to Table 2 and Figure 1
* Paragraph about correlations, justifying use of BHRM, referring to Figure 2.
* Paragraph about results, referring to Figure 3.

**Discussion**

XX

**Table 1.** Summary of patients characteristics by case/control/not included.

**Table 2.** Summary of pollutant concentrations.

**Figure 1.** Pollutants plot.

**Figure 2**.Correlation plot.

**Figure 3**. Results.

**Acknowledgements**

**Author contributions**: Dr Parks had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Parks, Kioumourtzoglou.

*Acquisition, analysis, or interpretation of the data:* Parks, Kioumourtzoglou, Balilian, Nunez, Hansen, Ketzel, Weisskopf, XX.

*Drafting of the manuscript:* Parks, Kioumourtzoglou.

*Critical revision of the manuscript for important intellectual content:* XX

*Statistical analysis:* Parks, Kioumourtzoglou.

*Obtained funding*: Kioumourtzoglou.

*Administrative, technical, or material support:* XX

*Study Supervision*: Kioumourtzoglou.

**Conflict of interest disclosures:** None reported.

**Funding/Support:** Robbie M Parks was partially supported by the Earth Institute post-doctoral research fellowship at Columbia University. Funding was also provided by the National Institute of Environmental Health Sciences (NIEHS) grants R01 ES030616, R01 ES028805, R01 ES028033, R01 MD012769, R01 AG066793, R01 ES029950, R21 ES028472, P30 ES009089, and P42 ES010349

**References**

1. Frank L. When an entire country is a cohort. *Science*. 2000;287(5462):2398-2399.

2. Kioumourtzoglou M-A, Seals RM, Himmerslev L, Gredal O, Hansen J, Weisskopf MG. Comparison of diagnoses of amyotrophic lateral sclerosis by use of death certificates and hospital discharge data in the Danish population. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2015;16(3-4):224-229.

3. Pedersen CB. The Danish civil registration system. *Scandinavian journal of public health*. 2011;39(7\_suppl):22-25.

4. Ritz B, Lee P-C, Hansen J, et al. Traffic-related air pollution and Parkinson’s disease in Denmark: a case–control study. *Environmental health perspectives*. 2016;124(3):351-356.

5. Ketzel M, Berkowicz R, Hvidberg M, Jensen SS, Raaschou-Nielsen O. Evaluation of AirGIS: a GIS-based air pollution and human exposure modelling system. *International Journal of Environment and Pollution*. 2011;47(1-4):226-238.

6. Ketzel M, Burman M, Nøjgaard JK, Christensen JH, Im U, Brandt J. High resolution modelling of elemental carbon for Denmark. In: *18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, HARMO 2017*. ; 2017.

7. Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, et al. Lung cancer incidence and long-term exposure to air pollution from traffic. *Environmental health perspectives*. 2011;119(6):860-865.

8. Raaschou-Nielsen O, Sørensen M, Ketzel M, et al. Long-term exposure to traffic-related air pollution and diabetes-associated mortality: a cohort study. *Diabetologia*. 2013;56(1):36-46.

9. Sørensen M, Hoffmann B, Hvidberg M, et al. Long-term exposure to traffic-related air pollution associated with blood pressure and self-reported hypertension in a Danish cohort. *Environmental health perspectives*. 2012;120(3):418-424.