Epidemiology

Long-term traffic-related air pollutant exposure and amyotrophic lateral sclerosis diagnosis in Denmark: A Bayesian hierarchical analysis

--Manuscript Draft--

Manuscript Number:	EDE21-0753R1
Full Title:	Long-term traffic-related air pollutant exposure and amyotrophic lateral sclerosis diagnosis in Denmark: A Bayesian hierarchical analysis
Article Type:	Original Article
Corresponding Author:	Robbie M Parks, Ph.D. Columbia University NY, New York UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Columbia University
Corresponding Author's Secondary Institution:	
First Author:	Robbie M Parks, Ph.D.
First Author Secondary Information:	
Order of Authors:	Robbie M Parks, Ph.D.
	Yanelli Nunez, PhD
	Arin A Balalian, MD, MPH
	Elizabeth A Gibson, PhD
	Johnni Hansen, PhD
	Ole Raaschou-Nielsen, PhD
	Matthias Ketzel, PhD
	Jibran Khan, PhD
	Jørgen Brandt, PhD
	Roel Vermeulen, PhD
	Susan Peters, PhD
	Jeff Goldsmith, PhD
	Diane B Re, PhD
	Marc G Weisskopf, PhD, ScD
	Marianthi-Anna Kioumourtzoglou, ScD
Order of Authors Secondary Information:	
Abstract:	Background: Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease. Limited evidence suggests ALS diagnosis may be associated with air pollution exposure and specifically traffic-related pollutants. Methods: In this population-based case-control study, we used 3,939 ALS cases from the Danish National Patient Register diagnosed during 1989–2013 and matched on age, sex, year of birth and vital status to 19,298 population-based controls free of ALS at index date. We used validated predictions of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) to assign 1-, 5-, and 10-year average exposures pre-ALS diagnosis at study participants' present and historical residential addresses. We used a Bayesian hierarchical conditional logistic model, adjusting for potential confounders, to estimate overall and joint associations for the traffic-related pollutants (NO_x, CO, EC), and pollutant-specific associations.

	Results: For a standard deviation (SD) increase in 5-year average concentrations, EC (SD=0.42μg/m 3) was potentially individually associated with an increase in odds (11.5%; 95% credible interval[CrI]:-1.0%,25.6%), with decreases individually for NO x (SD=20μg/m 3) (-4.6%;95%CrI-18.1%,8.9%) and CO (SD=106μg/m 3) (-3.2%;95%CrI-14.4%,10.0%) and a null association of non-EC PM 2 . 5 (SD=2.37μg/m 3) (0.7%;95%CrI-9.2%,12.4%). We found no clear association for joint or overall traffic pollution. There was a 77.8% posterior probability of a positive association between the joint effect of pollutants and ALS diagnosis, 96.3% for EC, 27.8% for NO x and 26.7% for CO. Conclusions: A potential positive association between ALS diagnosis and pollutants, particularly for EC, though results are inconclusive. Further work is needed to understand the role of traffic-related air pollution on ALS pathogenesis.
Additional Information:	
Question	**Response**
Word count, abstract (Enter "0" if not applicable.)	250
Word count, main text. Enter the number of words in the main text. For research articles, main text typically includes the introduction, methods, results, and discussion. Do not include words in title page, bibliography, tables, figure legends, or figures. Do not paste the text from the document itself into this box.	3896
Total word count. Enter the total word count for your submission, excluding the title page and abstract. Include the following in the total word count: (a) main text (for research articles, this typically includes the introduction, methods, results, and discussion), (b) bibliography, (c) tables, (d) figure legends, and (e) figures (calculated as 250 words per figure, each panel in a panel figure should be separately counted). Do not count words in Supplemental Digital Content. Do not paste the text from the document itself into this box.	6383
Funding sources All relevant sources of funding should be included here and on the title page of the manuscript with the heading “Source of Funding.” -Enter “none” if the work was completed without specific funding support. -If the result reported in the submission corresponds directly to the specific aims of a source (or sources) of funding, then describe that source of funding as: “The results reported herein correspond to specific aims of grant (or other source of support) XXX to investigator YYY from ZZZ”, where XXX is a grant or project	Robbie M Parks was supported by the NIEHS K99 ES033742 and 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 AG066793, R21 ES028472, P30 ES009089, and P30 ES000002.

number, YYY is the Principal Investigator of the grant or project, and ZZZ is the funding agency. -Describe all other sources of support as: "This work was (also) supported by grant(s) (or other source of support) XXX from ZZZ", where '(also)' is inserted only if the listed support is in addition to support corresponding directly to a specific aim, XXX is a grant or project number, and ZZZ is a funding agency. Additional sources of support should be added serially (e.g., grants XXX1 from ZZZ1, XXX2 from ZZZ2, and XXX3 from ZZZ3." Sources of support can include general salary support, which may not have a grant or project number. Grant or project numbers should be provided in a format that allows interested parties to find the grant in publicly available databases provided by many funding agencies.	
Describe the process by which someone else could obtain the data and computing code required to replicate the results reported in your submission. If one or the other is not available, please write "The [data, computing code, or both] are not available for replication because" and add an explanation for why one or the other is not available. Authors of submissions without original results, such as commentaries or letters, should enter "Not applicable." Your answer to this question will not affect the editors' decision about the submission's suitability for publication. If your paper is published, your response to this question will be included in the manuscript footnotes.	Danish patient records are available via the Danish National Patient Register (https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-national-patient-register/browse). Danish population records are available via the Danish Civil Registration System (https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-danish-civil-registration-system-cpr/browse). Exposure data are available via the DEHM-UBM-AirGIS website (https://envs.au.dk/en/research-areas/air-pollution-emissions-and-effects/the-monitoring-program/air-pollution-models/airgis/about-airgis/). All code for analysis, results from analysis, and visualization presented in this manuscript will be publicly available via GitHub at https://github.com/rmp15/multipollutants_and_als_code_review.
RETAINED RIGHTS: Except for copyright, other proprietary rights related to the Work (e.g., patent or other rights to any process or procedure) shall be retained by the author. To reproduce any text, figures, tables, or illustrations from this Work in future works of their own, the author must obtain written permission from Wolters Kluwer Health, Inc. ("WKH").	I agree

ORIGINALITY: Each author warrants that his or her submission to the Work is original, does not infringe upon, violate, or misappropriate any copyright or other intellectual property rights, or any other proprietary right, contract or other right or interest of any third party, and that he or she has full power to enter into this agreement. Neither this Work nor a similar work has been published nor shall be submitted for publication elsewhere while under consideration by this Publication.

AUTHORSHIP RESPONSIBILITY: Each author warrants that he or she has participated sufficiently in the intellectual content, the analysis of data, if applicable, and the writing of the Work to take public responsibility for it. Each has reviewed the final version of the Work, believes it represents valid work, and approves it for publication. Moreover, should the editors of the Publication request the data upon which the work is based, they shall produce it.

PREPRINTS: Upon acceptance of the article for publication, each author warrants that he/she will promptly remove any prior versions of this Work (normally a preprint) that may have been posted to an electronic server.

DISCLAIMER: Each author warrants that this Work contains no libelous or unlawful statements and does not infringe or violate the publicity or privacy rights of any third party, libel or slander any third party, contain any scandalous, obscene, or negligently prepared information, or infringe or violate any other personal or proprietary right of others. Each author

warrants that the Work does not contain any fraudulent, plagiarized or incorrectly attributed material. Each author warrants that all statements contained in the Work purporting to be facts are true, and any formula or instruction contained in the Work will not, if followed accurately, cause any injury, illness, or damage to the user. If excerpts (e.g., text, figures, tables, illustrations, or audio/video files) from copyrighted works are included, a written release will be secured by the author prior to submission, and credit to the original publication will be properly acknowledged. Each author further warrants that he or she has obtained, prior to submission, written releases from patients whose names or likenesses are submitted as part of the Work. Should the Editor or WKH request copies of such written releases, the author shall provide them in a timely manner. DISCLOSURES/CONFLICT OF INTEREST Each author must identify any financial interests or affiliations with institutions, organizations, or companies relevant to the manuscript by completing the form below. Additionally, any financial associations involving a spouse, partner or children must be disclosed as well. Note: Some sections below come from the ICMJE Uniform Disclosure Form for Potential Conflicts of Interest at http://www.icmje.org/downloads/coi_disclosure.pdf (dated July 2010).	
Did you or your institution at any time receive payment or support in kind for any aspect of the submitted work (including but not limited to grants, consulting fee or honorarium, support for travel to meetings for the study or other purposes, fees for participation in review activities such as data monitoring boards, statistical analysis, end point committees, and the	No

like, payment for writing or reviewing the manuscript, provision of writing assistance, medicines, equipment, or administrative support, etc...)?	
Other: Did you or your institution at any time receive additional payments or support in kind for any aspect of the submitted work?	No
Please indicate whether you have financial relationships (regardless of amount of compensation) with entities. You should report relationships that were present during the 36 months prior to submission including board membership, consultancy, employment, expert testimony, grants/grants pending, payment for lectures including service on speakers bureaus, payment for manuscript preparation, patents (planned, pending or issued), royalties, payment for development of educational presentations, stock/stock options, travel/accommodations/meeting expenses unrelated to activities listed (for example, if you report a consultancy above there is no need to report travel related to that consultancy), etc.	No
Other (err on the side of full disclosure): Please indicate whether you have any additional financial relationships (regardless of amount of compensation) with entities. You should report relationships that were present during the 36 months prior to submission.	None
Other Relationships Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?	No other relationships/conditions/circumstances that present potential conflict of interest
AUTHOR'S OWN WORK: In consideration of WKH's publication of the Work, the author hereby transfers, assigns, and otherwise conveys all his/her copyright ownership worldwide, in all languages, and in all forms of media now or hereafter known, including electronic media such as CD-ROM, Internet, and Intranet, to WKH. If WKH should decide	I agree

for any reason not to publish the Work, WKH shall give prompt notice of its decision to the corresponding author, this agreement shall terminate, and neither the author nor WKH shall be under any further liability or obligation. Each author grants WKH the rights to use his or her name and biographical data (including professional affiliation) in the Work and in its or the journal's promotion. Notwithstanding the foregoing, this paragraph shall not apply, and any transfer made pursuant to this paragraph shall be null and void if (i) the Work has been accepted by WKH for publication, and (ii) the author chooses to have the Work published by WKH as an open access publication.

WORK MADE FOR HIRE: If this Work or any element thereof has been commissioned by another person or organization, or if it has been written as part of the duties of an employee, an authorized representative of the commissioning organization or employer must also sign this form stating his or her title in the organization.

GOVERNMENT EMPLOYEES: If the Work or a portion of it has been created in the course of any author's employment by the United States Government, check the "Government" box at the end of this form. A work prepared by a government employee as part of his or her official duties is called a "work of the U.S. Government" and is not subject to copyright. If it is not prepared as part of the employee's official duties, it may be subject to copyright.

INSTITUTIONAL REVIEW BOARD/ANIMAL CARE COMMITTEE APPROVAL: Each author warrants that his or her institution has approved the protocol for any investigation involving humans or animals and that all experimentation was conducted in

conformity with ethical and humane principles of research. WARRANTIES: Each author warranty made in this form is for the benefit of WKH and the Editor; each author agrees to defend, indemnify, and hold harmless those parties for any breach of such warranties.	
The journal will permit the author(s) to deposit for display a "final peer-reviewed manuscript" (the final manuscript after peer-review and acceptance for publication but prior to the publisher's copyediting, design, formatting, and other services) 12 months after publication of the final article on the author's personal web site, university's institutional repository or employer's intranet, subject to the following: * You may only deposit the final peer-reviewed manuscript. * You may not update the final peer-reviewed manuscript text or replace it with a proof or with the final published version. * You may not include the final peer-reviewed manuscript or any other version of the article on any commercial site or in any repository owned or operated by any third party. For authors of articles based on research funded by the National Institutes of Health ("NIH"), Wellcome Trust, Howard Hughes Medical Institute ("HHMI"), or other funding agency, see below for the services that WKH will provide on your behalf to comply with "Public Access Policy" guidelines. * You may not display the final peer-reviewed manuscript until twelve months after publication of the final article.	I agree

* You must attach the following notice to the final peer-reviewed manuscript: "This is a non-final version of an article published in final form in (provide complete journal citation)". * You shall provide a link in the final peer-reviewed manuscript to the journal website.	
"Public Access Policy" Funding Disclosure Please disclose below if you have received funding for research on which your article is based from any of the following organizations:	National Institutes of Health (NIH)
Please select:	Author's Own Work
Any additional comments?	None
Compliance with RCUK and Wellcome Trust Open Access Policies Both the Research Councils UK (RCUK) and the Wellcome Trust have adopted policies regarding Open Access to articles that have been funded by grants from the RCUK or the Wellcome Trust. If either “Wellcome Trust” or “Research Councils UK (RCUK)” has been selected above, and the authors of the applicable article choose to have the article published as an open access publication, the following policies will apply: * If the article is to be published pursuant to the “Gold” route of Open Access, both the RCUK and the Wellcome Trust require that WKH make the article freely available immediately pursuant to the Attribution 4.0 Creative Commons License, currently found at http://creativecommons.org/licenses/by/4.0/legalcode (the “CC BY License”). The CC BY License is the most accommodating of the Creative Commons licenses and allows others to distribute, remix,	I agree

tweak, and build upon the article, even commercially, as long as they credit the authors for the original creation.

* If the article is to be published pursuant to the “Green” route of Open Access, both the RCUK and the Wellcome Trust require that WKH make the article freely available within six months pursuant to the Attribution-NonCommerical 4.0 Creative Commons License, currently found at http://creativecommons.org/licenses/by-nc/4.0/legalcode (the “CC BY-NC License”). The CC BY-NC License allows others to remix, tweak, and build upon the article non-commercially, and although their new works must also acknowledge the authors for the original creation and be non-commercial, they don't have to license their derivative works on the same terms.

As a service to our authors, WKH will identify the National Library of Medicine (NLM) articles that require deposit pursuant to the RCUK and Wellcome Trust policies described in this section. This Copyright Transfer Agreement provides the mechanism for identifying such articles.

WKH will transmit the final peer-reviewed manuscript of an article based on research funded in whole or in part by either RCUK or the Wellcome Trust to Pub Med Central.

Upon NIH request, it remains the legal responsibility of the author to confirm with NIH the provenance of his/her manuscript for purposes of deposit. Author will not deposit articles him/herself. Author will not alter the final peer-reviewed manuscript already transmitted to NIH.

With respect to the “Green” route of Open Access, author will not authorize the display of the final peer-reviewed manuscript prior to 6 months following publication of the final article. Authors of articles that have been funded from grants from the RCUK or the Wellcome Trust are required to sign the WKH Open Access License Agreement prior to publication of the applicable article. Please contact the Editorial Office of the applicable journal to receive the Open Access License Agreement that is to be signed in connection with the publication of the article.	
I am the person in question for this submission or otherwise have approval to complete this agreement.	I agree
CME/CE Disclosure Each author must identify and disclose any financial associations involving a spouse, partner or children by completing the Family Disclosure question below, and whether any off-label uses or unapproved drugs or devices are discussed in his/her manuscript by completing the Off-Label Use/Unapproved Drugs or Products question below. In the event that the Work is published as a continuing education or continuing medical education article, this information will be provided to the accrediting body and may be included in the published article. When applicable, articles accepted for publication may need to comply with additional standards related to CME or CE accreditation. Please refer to guidelines for authors for details. WKH and its affiliates reserve the right to publish the manuscript as a continuing education article.	I agree
Family Disclosure	No other relationships/conditions/circumstances that present potential conflict of interest

Do your children or your spouse or partner have financial relationships with entities that have an interest in the content of the submitted work?	
Off-Label Use/Unapproved Drugs or Products If your manuscript discusses an unlabeled use of a commercial product or device or an investigational use of a product or device not yet approved by the FDA for any purpose, you must specifically disclose in the manuscript that the product is not labeled for the use under discussion or that the product is still investigational. Please check the item below that applies to you	I will not discuss unlabeled/investigational uses of any commercial product or device

THE EARTH INSTITUTE
COLUMBIA UNIVERSITY

The Earth Institute
Columbia University
Hogan Hall
MC 3277
2910 Broadway
NY, NY 10025

26 April 2022

Robbie M Parks
Post-doctoral research fellow
Earth Institute, Columbia University
robbie.parks@columbia.edu
www.robbiemparks.github.io

Dr Timothy L. Lash,
Editor-in-Chief, *Epidemiology*

Dear Dr Lash:

Please find enclosed our revised manuscript entitled “Long-term traffic-related air pollutant exposure and amyotrophic lateral sclerosis diagnosis in Denmark: A Bayesian hierarchical analysis” (DE21-0753). We have revised the text according to the comments and the suggestions of the Editors and Reviewers, as outlined in our responses.

The revised manuscript is 3,896 words. The abstract is 250 words. It contains two tables and two figures, with additional information in the Supplementary Digital Content.

We look forward to your response and would be glad to address any further comments you may have on the manuscript.

Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disease, currently without a cure. Approximately half of patients die within three years of symptom onset. Annually, there are nearly 30,000 cases of ALS in Europe and over 200,000 worldwide. Known inherited genetic variants only account for 5–10% of ALS cases. Environmental factors, therefore, are likely important in ALS pathogenesis. Ambient air pollution, especially urban air pollution, is a ubiquitous exposure that has been associated with several other neurodegenerative disorders, and is consistently linked to systemic inflammation, oxidative stress, and neuroinflammation, all of which, in turn, have been reported as key pathways to ALS pathogenesis. Despite the compelling plausibility, few studies to date have evaluated the association between air pollution and ALS.

With this study, *the largest case-control study of ALS and air pollution to date,* we used data on 3,939 ALS cases from the Danish National Patient Register diagnosed between 1989 – 2013 and matched on, sex, year of birth and vital status to 19,298 population-based controls free of ALS at index date. We used predictions of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) from validated spatio-temporal models to assign 1-, 5-, and 10-year average exposures pre-ALS diagnosis at present and historical residential addresses of study participants. Although the last year of data in our analyses is from eight years ago (2013), our findings remain timely due to the uniquely large number of ALS cases we have collected, that air pollution remains so pervasive, and that ALS prevalence is projected to increase nearly 70% by 2040.

We found that an increase in the joint exposure to traffic-related pollutants was associated with an increase in odds of ALS diagnosis, significant for elemental carbon for a 1-year average standard deviation increase, though not significant at the 95% credible interval level for other pollutants. Our results indicate that sources of air pollution with elemental carbon, such as diesel engines and woodburning stoves, might contribute to development of ALS.

This manuscript has not been previously published and is not under review in any other journal. All authors have contributed to the paper, have approved its submission, and take responsibility for its contents. The authors have no actual or potential competing financial interests. There is no closely related paper included with this submission.

Sincerely,

Robbie M. Parks, PhD

Type of Manuscript: Original Research Article

Manuscript Title: Long-term traffic-related air pollutant exposure and amyotrophic lateral sclerosis diagnosis in Denmark: A Bayesian hierarchical analysis

Authors: Robbie M Parks[1,2*], Yanelli Nunez[1], Arin A Balalian[3], Elizabeth A Gibson[1,4], Johnni Hansen[5], Ole Raaschou-Nielsen[5,6], Matthias Ketzel[6,7], Jibran Khan[6], Jørgen Brandt[6,8], Roel Vermeulen[9], Susan Peters[9], Jeff Goldsmith[10], Diane B. Re[1], Marc G. Weisskopf[11], Marianthi-Anna Kioumourtzoglou[1]

[1]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA
[2]The Earth Institute, Columbia University, New York, New York, USA
[3]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA
[4]Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, Massachusetts, USA
[5]Danish Cancer Society Research Center, Copenhagen, Denmark
[6]Department of Environmental Science, Aarhus University, Roskilde, Denmark
[7]Global Centre for Clean Air Research (GCARE), University of Surrey, Guildford, United Kingdom
[8]iClimate – interdisciplinary Center for Climate Change, Aarhus University, Denmark
[9]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands
[10]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York, USA
[11]Departments of Environmental Health and Epidemiology, T. H. Chan School of Public Health, Harvard University, Boston, Massachusetts, 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

Suggestions for running head: Traffic-related air pollutants and ALS

The authors declare they have no actual or potential competing financial interests.

Sources of financial support: Robbie M Parks was supported by the NIEHS K99 ES033742 and 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 AG066793, R21 ES028472, P30 ES009089, and P30 ES000002.

Description of the process by which someone else could obtain the data and computing code required to replicate the results reported in your submission (or explanation why data or code are not available): Danish patient records are available via the Danish National Patient Register (https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-national-patient-register/browse). Danish population records are available via the Danish Civil Registration System (https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-danish-civil-registration-system-cpr/browse). Exposure data are available via the DEHM-UBM-AirGIS website (https://envs.au.dk/en/research-areas/air-pollution-emissions-and-effects/the-monitoring-program/air-pollution-models/airgis/about-airgis/). All code for analysis, results from analysis, and visualization presented in this manuscript will be publicly available via GitHub at https://github.com/rmp15/multipollutants_and_als_code_review.

Abstract

Background: Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease. Limited evidence suggests ALS diagnosis may be associated with air pollution exposure and specifically traffic-related pollutants.

Methods: In this population-based case-control study, we used 3,939 ALS cases from the Danish National Patient Register diagnosed during 1989–2013 and matched on age, sex, year of birth and vital status to 19,298 population-based controls free of ALS at index date. We used validated predictions of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) to assign 1-, 5-, and 10-year average exposures pre-ALS diagnosis at study participants' present and historical residential addresses. We used a Bayesian hierarchical conditional logistic model, adjusting for potential confounders, to estimate overall and joint associations for the traffic-related pollutants (NO_x, CO, EC), and pollutant-specific associations.

Results: For a standard deviation (SD) increase in 5-year average concentrations, EC (SD=0.42μg/m^3) was potentially individually associated with an increase in odds (11.5%; 95% credible interval[CrI]:-1.0%,25.6%), with decreases individually for NO_x (SD=20μg/m^3) (-4.6%;95%CrI-18.1%,8.9%) and CO (SD=106μg/m^3) (-3.2%;95%CrI-14.4%,10.0%) and a null association of non-EC $PM_{2.5}$ (SD=2.37μg/m^3) (0.7%;95%CrI-9.2%,12.4%). We found no clear association for joint or overall traffic pollution. There was a 77.8% posterior probability of a positive association between the joint effect of pollutants and ALS diagnosis, 96.3% for EC, 27.8% for NO_x and 26.7% for CO.

Conclusions: A potential positive association between ALS diagnosis and pollutants, particularly for EC, though results are inconclusive. Further work is needed to understand the role of traffic-related air pollution on ALS pathogenesis.

Abbreviations:

ALS	Amyotrophic lateral sclerosis
BKMR	Bayesian kernel machine regression
CO	Carbon monoxide
CrI	Credible interval
DEHM-UBM-AirGIS	Spatio-temporal air pollution modelling system used in study
EC	Elemental carbon
ICD	International Classification of Diseases
IQR	Interquartile range
IR	Incidence ratio
Non-EC $PM_{2.5}$	Non-elemental carbon fine particles
NO_x	Nitrogen oxides
O_3	Ozone
$PM_{2.5}$	Fine particles
SD	Standard deviation
SES	Socioeconomic status

Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disease,[1] currently without a cure.[2] Approximately half of patients die within three years of symptom onset.[3] Annually, there are nearly 30,000 cases of ALS in Europe and over 200,000 worldwide.[4] Known inherited genetic variants only account for 5–10% of ALS cases.[5,6] Environmental factors, therefore, are likely important in ALS pathogenesis.[7] However, because the disease is relatively rare, it is challenging to conduct large-scale prospective studies. There is a recognized need for more evidence of the environmental contributors of ALS.[5,8]

Although air pollution is commonly studied in association with respiratory- and cardiovascular-related outcomes, e.g., references [9–14], epidemiologic and toxicological studies also support several plausible biological mechanisms in association with the nervous system and neurodegeneration, e.g., references [15–34]. Ambient air pollution, especially urban air pollution, is a ubiquitous exposure that has been associated with several other neurodegenerative disorders, e.g., references [16–21,35,36]. and is consistently linked to systemic inflammation,[22–24] oxidative stress,[25–28] and neuroinflammation,[15,29] all of which, in turn, have been reported as key pathways to ALS pathogenesis, e.g., references [30–34].

Despite the compelling plausibility, few studies to date have evaluated the association between air pollution and ALS.[35,37–39] A study in 2021 found that traffic-related air pollutants may be driving observed associations.[38] No study has hitherto attempted to understand the combined and individual associations of the pollutants in a single model. Air pollutants have been consistently associated with adverse health, primarily in single pollutant analyses.[13,17,40–42] However, they are

highly correlated with one another.[40] It is therefore a mixture modelling challenge to infer the association of multiple air pollutants and health outcomes.[43] Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NO_x), carbon monoxide (CO), and elemental carbon (EC)— we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis, and estimate their (b) joint and (c) overall traffic-related emissions associations.

Methods

Study Population and Outcome Assessment

We used data from the Danish National Patient Register during 1989-2013, through which details on demographic characteristics and certain health outcomes of all Danish residents can be linked based on a unique personal identifier.[44] The Register was established in 1977 and is comprehensive, including nationwide clinical and administrative records for all inpatient data, with outpatient data available since 1995.[45]

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, using the date of the first relevant code as the diagnosis date. This was the index date. We only included patients who were at least 20 years old when diagnosed because (i) cases younger than 20 years old were at a greater chance of misclassification, since ALS has been predominantly diagnosed in older adults,[46] and (ii) the very few juvenile ALS cases have been explained to a much larger degree by genetic mutations (~40%).[47] In our validation study, Register data for ALS ascertainment were highly reliable; working with a specialist ALS neurologist to

review medical records and comparing to death certificates and hospital discharges, the Danish National Patient Register was found to have an overall predictive value for ALS of 82%.[48]

We obtained controls through the Danish Civil Registration System, established in 1968 and updated daily, which includes administrative records (e.g., date and place of birth, sex, vital status, and history of civil status and addresses since 1971) on all persons living in Denmark; records are kept even when a person dies or emigrates.[49] We randomly matched five controls per case by age, sex, year of birth, and vital status. Controls were alive and free of diagnosed ALS at the ALS diagnosis date of the matched case (index date). The control-sampling scheme followed a risk-set matching pattern, so cases could have served as controls before diagnosis of ALS.[50]

We obtained all addresses of cases and controls from January 1st 1979 onwards from the Danish Civil Registration System,[49] including the dates of moving to and from each address, before the index date. We then obtained the geographical coordinates at the door of each house of the residential history of the participants, with previous evidence of the high accuracy of this method of geocoding of addresses in Denmark.[17]

This study was approved by the Institutional Review Board Committee at the Columbia University and the Danish Data Protection Agency.

Exposure data

We obtained predictions on monthly concentrations of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) (as well as ozone, O_3, for a sensitivity

analysis, usually negatively correlated with other pollutants due to its chemistry[51]), at residential addresses of study participants from the validated spatio-temporal air pollution modelling system DEHM-UBM-AirGIS that provides full space and time coverage over the study period, described in detail elsewhere.[52–55] In brief, DEHM-UBM-AirGIS is a human exposure modelling system for traffic pollution, developed for application in Danish air pollution epidemiological studies. The modelling system is able to generate street configuration and traffic data based on digital maps and national databases, which enables estimation of air quality levels at a large number of addresses in an automatic and effective way. These predicted pollutant concentrations have been extensively used in previous air pollution epidemiologic studies in Denmark.[17,56–58] The models have good predictive accuracy, with average monthly correlations between measured and modelled results of 0.85 for NO_x, 0.91 for CO, 0.92 for O_3, 0.79 for EC, and 0.83 for annual concentrations of $PM_{2.5}$.[52,55] Because traffic is a major source of $PM_{2.5}$ and EC one of the main $PM_{2.5}$ components in urban environments,[59] we removed the EC concentration from the total $PM_{2.5}$ mass concentration (non-EC $PM_{2.5}$) by subtraction, to avoid overadjustment when including both in the models simultaneously.

Based on the residential history of each case or control, we calculated 1-, 5-, and 10-year average exposure to each pollutant ending at one year before the index date, as diagnosis has been shown previously to occur at a median of 12 months after symptoms onset.[60] Specifically, each case or control average value (1-, 5- or 10-year) was calculated as the mean of all concentrations recorded across time at the recorded addresses within each time window. A small number of Danish residents lack a complete address history (1.7%; lack of house number). To ensure we were including participants with adequately complete exposure records, we set the following minimum

criteria for number of complete exposure record months to include cases and controls: (i) 1-year averages: 9 of 12 months, at least one measurement in each season; (ii) 5-year averages (main exposure): 30 of 60 months; and (iii) 10-year averages: 60 of 120 months.

Covariate data

We included a set of covariates based on as close as possible to index date to account for potential confounding bias, including household socioeconomic status (SES) based on last-reported job title at index date; civil status at index date, last reported place of residence at index date, and place of birth. We used a five-category individual-level SES definition developed by the Danish Institute of Social Sciences, based on job titles from income tax forms, which has been associated with ALS diagnosis in Denmark,[61] as well as how quickly one is identified as having ALS in the Danish Civil Registration System.[62] Group 1 (highest status) includes corporate managers and academics; group 2: proprietors, managers of small businesses and teachers; group 3: technicians and nurses; group 4: skilled workers; and group 5: unspecialized workers, such as entry-level positions within food and retail environments. We also included a group for participants whose job title was unknown (group 9). For each married participant, we used the higher of the couple's individual SES categories, when available. We also used information on civil status (never married, married, divorced, widowed) due to the influence that a spouse may have on visiting a family physician,[63] last reported place of residence from postcode (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland) to account for various local environmental and behavioral stressors,[7] and place of birth (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland, foreign, unknown) to adjust for other potential family-specific, location-specific, and early-life confounders, which may have an impact on the probability of developing ALS.[64] Ultimately, we

were limited by what was available in the Danish Civil Registration System.[62] As part of a sensitivity analysis, we also included parish-level SES, measured by percentage of residents with greater than high-school education, in the model. In Denmark, parishes are administrative units with an average population of ~2,500 residents.

Statistical analysis

We analyzed the association between ALS diagnosis (binary) and exposure to traffic-related pollutants by applying a Bayesian formulation of the conditional logistic model, with Bayesian hierarchy on the traffic-related pollutants (EC, NO_x, CO).[65,66] The conditional approach examines contrasts within matched strata, i.e., groupings of case and matched controls, implicitly adjusting for matching factors (age, sex, year of birth, vital status) within each matched stratum.[65] Matching by finer scale than year of birth was not possible. Bayesian inference allows for full distributional estimation of parameters of interest.[66] We employed a Bayesian hierarchical formulation because it enables estimates of (a) independent pollutant-outcome associations, (b) a joint association of the three pollutants (i.e., total percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO), and (c) an overall average traffic association (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO), while accounting for the variance-covariance structure between the highly-correlated exposures and their coefficients.[66] We included a linear term for each included pollutant and adjusted for individual- and parish-level SES, civil status, last reported place of residence, and place of birth.

Specifically, via a logit function, we modelled the log-odds of ALS diagnosis, as follows:

$$logit[\Pr(Y_{ci} = 1)] = \alpha_c +$$

$$\beta_{\mathrm{NO}_x}\mathrm{NO}_{x\,ci} + \beta_{CO}\mathrm{CO}_{ci} + \beta_{EC}\mathrm{EC}_{ci} +$$

$$\beta_{\mathrm{PM}_{2.5}}(\text{non-EC PM}_{2.5\,ci}) +$$

$$\beta_{\mathrm{SES}}\mathrm{SES}_{ci} + \beta_{\text{Civil Status}}\text{Civil Status}_{ci} + \beta_{\text{Residence}}\text{Residence}_{ci} + \beta_{\text{Birth}}\text{Birth}_{ci},$$

where Y_{ci} denotes whether subject i in matched stratum c was diagnosed with ALS, i.e., c represents a case and its matched controls; α_c the matched stratum-specific intercepts (not estimated in conditional logistic models); $\beta_{\mathrm{NO}_x}, \beta_{CO}, \beta_{EC}, \beta_{\mathrm{PM}_{2.5}}$ the pollutant-specific coefficients (log-odds) per standard deviation (SD) increase in concentration of $\mathrm{NO_x}$, CO, EC, non-EC $\mathrm{PM}_{2.5}$ respectively, scaled by their respective SDs and centered at their means, with each β a pollutant-specific association adjusted by other terms in the model and the rest as coefficients for subject-specific covariates. Interquartile Range (IQR) could equivalently be used to scale pollutant concentrations. If other sources of air pollution are associated with ALS, then including non-EC $\mathrm{PM}_{2.5}$ adjusts for $\mathrm{PM}_{2.5}$ from other sources,[67] as well as indicating whether pollution from other sources not explicitly quantified might also have associations with ALS. Therefore, $\beta_{\mathrm{PM}_{2.5}}$ is interpreted as the association with air pollutants not specifically included in our analysis. In urban European environments, traffic-related pollutants typically represent on-average 14% of $\mathrm{PM}_{2.5}$ concentrations.[68] In a sensitivity analysis, we included O_3 in the model, as O_3 concentrations have been associated with many adverse health outcomes,[69] and were negatively correlated with traffic-related pollutants, and added $ns(\mathrm{SES}_{\mathrm{parish}_{ci}})$, as a natural spline with three degrees of freedom.

In our model, β_{NO_x}, β_{CO}, and β_{EC} represent the independent pollutant-specific associations with ALS diagnosis. In the same model, we estimated the joint association between these three pollutants and ALS diagnosis as:

$$Traffic_{Joint} = \sum_{p=NO_x,CO,EC} \beta_p p.$$

This sum quantifies the association (log-odds) with ALS of a one-SD increase in the three pollutants simultaneously.

Finally, we assumed that the traffic-related pollutant-specific associations arise from a distribution of the overall traffic association with ALS diagnosis. We placed a hierarchy on the traffic-specific pollutant terms in the model to account for the fact that the traffic-related pollutants, EC, NO_x, CO, originate from common sources and primarily traffic in urban environments:

$$\beta_{Traffic} = [\beta_{NO_x}, \beta_{CO}, \beta_{EC}],$$

$$\beta_{Traffic} \sim MVN(\mu, \Sigma),$$

$$\mu \sim N(\lambda, \sigma_\lambda),$$

$$\Sigma = \tau\Omega\tau,$$

where λ denotes the overall average one-SD association of traffic-related pollution with variance σ_λ. Σ, the estimated variance-covariance matrix among pollutant-specific estimates, was expressed as a decomposition into a positive-definite correlation matrix Ω and scale matrix τ.[70]

We used weakly-informative priors so that data drove parameter estimation. Hyper-priors for coefficients on non-EC $PM_{2.5}$ and covariates were N(0,10); for σ_λ and τ we used Half-Cauchy(0,10), as recommended by Gelman, Polson and Scott as a weakly-informative prior;[71,72] Ω was defined by the weakly-informative prior LKJCorr(1).[73] The exception to this was the prior for λ, the average association of traffic-related pollutants, for which estimates became unrealistically high (approaching infinity and not converging with further iterations) with a more

weakly-informative prior. We therefore used a prior of N(0,0.1), which did not affect estimates of other parameters. We conducted sensitivity analyses to understand the influence of priors and the robustness of the results.

We present all results as percentage change in odds of ALS diagnosis per SD increase in pollutant concentration (calculated via e.g., $e^{\beta_{\mathrm{NO}_x}} - 1$, etc. obtained in the modelling process). Due to the risk-set matching pattern of our case-control study, odds ratios are also equivalently incidence ratios (IRs).[65] We ran each model with four chains with a sample size of 1,000 each, after a warm-up of 1,000 samples, for 4,000 total samples. We assessed whether the models converged by checking that the Gelman-Rubin potential scale reduction statistic[74] was below 1.1 for all estimated model parameters. The reported 95% credible intervals (CrI) are the 2.5th to 97.5th percentiles of each parameter's posterior marginal distribution. To calculate the probability that an association estimate was greater than null, we used the 4,000 samples of the posterior distribution and took the proportion of samples which were above the null. A 50% probability means that it is as likely as not that the marginal estimate is null, a probability closer to 100% indicates that the association is more likely to be truly positive, with closer to 0% indicating more likely to be truly negative.

We conducted statistical analyses using the R Statistical Software, version 4.1.1[75] and R-STAN, version 2.21.2.[66] All code for analysis, results from analysis, and visualization presented in this manuscript is publicly available via GitHub at https://github.com/rmp15/traffic_air_pollution_als_denmark_epidemiology.

We assessed the sensitivity of our results to hyper-prior assignment; running more iterations and warm-up per chain; inclusion of O_3; single traffic-related pollutant models adjusting for non-EC $PM_{2.5}$; as well as adjusting by parish-level SES. From the parish-level SES sensitivity analysis we excluded those who lived in areas without parish-level SES data, namely: (i) 819 participants for 1-year average exposure; (ii) 826 participants for 5-year average exposure; and (iii) 838 participants for 10-year average exposure.

Results

After filtering the original 4,011 cases and 20,055 controls based on completeness of exposure records, we used information on 3,934 (98.1% of total) cases and 19,298 (96.2% of total) controls for 5-year average exposure. We also used 3,937 cases,19,333 controls for 1-year average exposure and 3,939 cases, 19,250 controls for 10-year average exposure. Descriptive statistics of included cases and controls for 5-year average exposure can be found in Table 1. Descriptive statistics of controls for 5-year exposure by socioeconomic status, civil status, residence, and place of birth are found in eTables 1-4. For the main results, we present 5-year average exposure associations as a balance between representation of most recent exposure as well as long-term concentration.

The 5-year average traffic-related pollutant concentrations were 27 $\mu g/m^3$ for NO_x (SD=20 $\mu g/m^3$), 238 $\mu g/m^3$ for CO (SD=106 $\mu g/m^3$) and 0.85 $\mu g/m^3$ for EC (SD=0.42 $\mu g/m^3$) (Table 2). Figure 1 shows Spearman correlations between pollutants for 1-, 5-, and 10-year average exposures. Traffic-related pollutants (NO_x, CO, EC) were highly correlated in cases, controls and overall, ranging from correlations of 0.91 to 0.96. Otherwise, non-EC $PM_{2.5}$ was most highly correlated

with CO, ranging from 0.67 to 0.7. O_3 was negatively correlated with other pollutants, ranging from -0.54 to -0.89.

For 5-year average pollutant concentrations, we observed the largest overall association for the individual SD increase in EC (11.5%; 95% CrI: -1.0%, 25.6% per 0.42 $\mu g/m^3$; 96.3% posterior probability of positive association) (Figure 2). SD increases were associated with a decrease in odds of ALS diagnosis in NO_x (-4.6%; 95% CrI: -18.1%, 8.9% per 20 $\mu g/m^3$; 27.8% posterior probability of positive association) and CO (-3.2%; 95% CrI: -14.4%, 10.0% per 106 $\mu g/m^3$; 26.7% posterior probability of positive association). The joint association of traffic-related pollutants (EC, NO_x, CO) was 2.3% (95% CrI: -3.3%, 7.7%), with an 77.8% posterior probability of a positive association. The average overall traffic association was null (-0.1%; 95% CrI: -17.4%, 20.8%). Non-EC $PM_{2.5}$ was not associated with ALS diagnosis (0.7%; 95% CrI: -9.2%, 12.4%). 1-year EC average exposure was associated with an increase in odds of ALS diagnosis (15.4%; 95% CrI: 1.6%, 25.6%) (Figure 2). Compared to the 1- and 5-year results, the 10-year average exposure results were attenuated, as associations tended further to the null. Single-pollutant models for each traffic-related pollutant adjusting for non-EC $PM_{2.5}$ (eFigure 1; single traffic-related pollutant models D, E and F) resulted in positive associations for each of EC, NO_x, CO, with positive associations for non-EC $PM_{2.5}$ in all but the model with EC. The 95% credible interval for EC in the single-pollutant model (eFigure 1; model D) overlapped with the credible intervals of the EC term in the multi-pollutant models (eFigure 1; models B, C, G to P). Results from variations of the main model in the sensitivity analyses were robust to prior choices, inclusion of O_3, and inclusion of parish-level SES (eFigure 1). A map of average concentration of included pollutants (NO_x, EC,

$PM_{2.5}$, CO, O_3) across Denmark for a representative year (2000; middle of study period 1989-2013) is also available in eFigure 2.

Discussion

In the largest case-control study of ALS and traffic-related air pollution to date, we found that a joint increase in average concentrations of traffic-related pollutants was potentially associated with an increase in odds of ALS diagnosis, with the clearest results for EC. We found that EC had the largest-in-magnitude independent association with ALS diagnosis, while associations with NO_x and CO were negative with credible intervals overlapping the null, and smaller in magnitude. Sensitivity analyses demonstrated that for single pollutant models, the association for EC was smaller than for our main multi-pollutant model, which took into account the variance-covariance structure of traffic-related pollutants. Overall conclusions for the association between EC and ALS diagnosis would have been similar from the single- or multi-pollutant models. The inconsistent associations for NO_x and CO in the multi- and single-pollutant models suggest that the model may have had limited success identifying each individual pollutant's association with ALS diagnosis due to the high level of collinearity of traffic-related pollutants. Nevertheless, the consistency of the sign of the central estimate of EC in all models suggests that EC may be a driver of the ALS and traffic-related pollutant association, though further analysis is required. Our results indicate that traffic-related pollutants, hazardous in many ways,[9–21,40–42] may also be associated with ALS diagnosis. Our finding—that increases in EC, are potentially positively associated with ALS diagnosis—is plausible. A recent case-control study in the Netherlands reported that ultrafine particles, another traffic emissions-related surrogate, were associated with ALS diagnosis,[38] while another based in Catalonia, Spain found ALS cases clustered around key road infrastructure.[76]

Although we did not find an association with non-EC $PM_{2.5}$ in our study, our results are not directly comparable to those of the other studies, as our $PM_{2.5}$ effect estimates capture the $PM_{2.5}$ components not accounted for by other pollutants in the analysis.

Our results indicate that EC exposure—a large part of which comes from diesel combustion and small combustion sources (such as wood stoves) in European urban centers, where prevalence of diesel cars is high[77]—has a high probability of a positive association with ALS diagnosis. In our previous study of ALS and occupational exposures in Denmark we found that those working in agriculture and construction, associated with exposure to diesel engine exhausts, were at higher relative risk than those in other employments.[61] Truck drivers, for whom diesel exposure is common, are also at increased risk of sporadic ALS.[78] EC exposure has been associated with inflammation,[79] mitochondrial dysfunction[80] and DNA damage,[80,81] all of which are plausible pathways of neurodegeneration. These factors have also previously been identified as particular pathways to pathogenesis of ALS.[30–34]

We did not find a high probability of a positive association with NO_x in our analyses, in contrast with a previous study, though that study did not include EC.[38] NO_x is also highly correlated with EC (0.95 to 0.96 in our study), which is expected given that they are both combustion products commonly associated with emissions in urban environments. EC exposure was more strongly associated with 1-year than for 5-/10-year average concentrations, which may indicate that the previous year of exposure may be the most relevant exposure window relevant to traffic-related exposures and ALS; this is biologically plausible, as this critical exposure window would be at the pre-symptomatic stage of underlying ALS progression, where traffic-related pollution exposure

may add to the ongoing cellular or molecular process of the disease, to the point where the body can no longer compensate and subsequently enters the clinical phase.[82–84] We do not expect that these results are attributed to reverse causation, as we have lagged these 1-year exposures by one year already prior to diagnosis, and there was likely little substantial residential movement in the year before ALS diagnosis.[85] We do not expect that calendar time was a potential source of confounding, as the controls were matched on age and year of birth. The null joint association, combined with the largest associations from traffic-related pollutant in all models found with EC, further indicates that EC may be driving the association of air pollution with ALS, though further analysis will be necessary to confirm this.

Our study used one the largest number of ALS patients ever included in an environmental health study. Another strength of our study is that we leveraged highly correlated traffic pollutants and Bayesian hierarchical modeling and were able to estimate independent and joint traffic-related pollutant associations, as well as an overall traffic estimate. Although we have adjusted implicitly (by matching; age, sex, year of birth, vital status) and explicitly for many common covariates (SES, civil status, residence, place of birth), we cannot rule out residual confounding. Exposure measurement error is inevitable, as any modelled exposure will be inaccurate to some degree. However, any error is not likely correlated with ALS diagnosis, and therefore any bias would be towards null.[86] While a previous study found that ALS ascertainment from the Danish National Patient Register was highly reliable,[48] outcome misclassification cannot be ruled out, nor can the possibility that date of diagnosis and symptom onset were irregularly aligned. While our analysis adjusted for marital status and household SES, many couples in Denmark cohabitate. This would

not be captured by our analysis, and ALS diagnosis in relation to cohabitation status should be further investigated.[87]

Future research might use larger cohort data to understand the importance of each respective pollutant in a single model. Other mixture model methods, such as Bayesian Kernel Machine Regression (BKMR)[88] might be useful in further exploring the robustness of joint associations in a different framework, though BKMR was not appropriate for our particular research question, since BKMR is not currently appropriate for case-control studies. The timing of exposure will also be an important study route. ALS is projected to increase in prevalence over the next few decades all over the world.[4] Understanding ALS pathogenesis and identifying modifiable risk factors is critical for preventive action.

References

1. Rowland LP, Shneider NA. Amyotrophic lateral sclerosis. *New England Journal of Medicine*. 2001;344(22):1688-1700.

2. Chio A, Logroscino G, Hardiman O, et al. Prognostic factors in ALS: A critical review. *Amyotrophic Lateral Sclerosis*. 2009;10(5-6):310-323.

3. Mitchell JD, Borasio GD. Amyotrophic lateral sclerosis. *The Lancet*. 2007;369(9578):2031-2041.

4. Arthur KC, Calvo A, Price TR, Geiger JT, Chio A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nature Communications*. 2016;7(1):1-6.

5. Al-Chalabi A, Hardiman O. The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology*. 2013;9(11):617-628.

6. Hardiman O, Al-Chalabi A, Chio A, et al. Amyotrophic lateral sclerosis. *Nature reviews Disease primers*. 2017;3(1):1-19.

7. Oskarsson B, Horton DK, Mitsumoto H. Potential environmental factors in amyotrophic lateral sclerosis. *Neurologic Clinics*. 2015;33(4):877-888.

8. Longinetti E, Fang F. Epidemiology of amyotrophic lateral sclerosis: An update of recent literature. *Current Opinion In Neurology*. 2019;32(5):771.

9. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295(10):1127-1134.

10. Bennett JE, Tamura-Wicks H, Parks RM, et al. Particulate matter air pollution and national and county life expectancy loss in the USA: A spatiotemporal analysis. *PLOS Medicine*. 2019;16(7):e1002856. doi:10.1371/journal.pmed.1002856

11. Schwartz J. Particulate air pollution and chronic respiratory disease. *Environmental Research*. 1993;62(1):7-13.

12. Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology*. 2000;11(3):320-326.

13. Brook RD, Rajagopalan S, Pope III CA, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*. 2010;121(21):2331-2378.

14. Dockery DW, Pope CA, Xu X, et al. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*. 1993;329(24):1753-1759. doi:10.1056/NEJM199312093292401

15. Block ML, Elder A, Auten RL, et al. The outdoor air pollution and brain health workshop. *Neurotoxicology*. 2012;33(5):972-984.

16. Zanobetti A, Dominici F, Wang Y, Schwartz JD. A national case-crossover analysis of the short-term effect of PM 2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environmental Health*. 2014;13(1):1-11.

17. Ritz B, Lee PC, 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.

18. Kioumourtzoglou MA, Schwartz JD, Weisskopf MG, et al. Long-term PM2.5 exposure and neurological hospital admissions in the northeastern United States. *Environmental health perspectives*. 2016;124(1):23-29.

19. Levesque S, Surace MJ, McDonald J, Block ML. Air pollution & the brain: Subchronic diesel exhaust exposure causes neuroinflammation and elevates early markers of neurodegenerative disease. *Journal of Neuroinflammation*. 2011;8(1):1-10.

20. Heusinkveld HJ, Wahle T, Campbell A, et al. Neurodegenerative and neurological disorders by small inhaled particles. *Neurotoxicology*. 2016;56:94-106.

21. Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro III A, Schwartz J. Traffic-related air pollution and cognitive function in a cohort of older men. *Environmental Health Perspectives*. 2011;119(5):682-687.

22. Dubowsky SD, Suh H, Schwartz J, Coull BA, Gold DR. Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation. *Environmental Health Perspectives*. 2006;114(7):992-998.

23. Ruckerl R, Ibald-Mulli A, Koenig W, et al. Air pollution and markers of inflammation and coagulation in patients with coronary heart disease. *American Journal of Respiratory and Critical Care Medicine*. 2006;173(4):432-441.

24. Hoffmann B, Moebus S, Dragano N, et al. Chronic residential exposure to particulate matter air pollution and systemic inflammatory markers. *Environmental Health Perspectives*. 2009;117(8):1302-1308.

25. Kelly FJ. Oxidative stress: Its role in air pollution and adverse health effects. *Occupational and Environmental Medicine*. 2003;60(8):612-616.

26. Chuang KJ, Chan CC, Su TC, Lee CT, Tang CS. The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. *American journal of respiratory and critical care medicine*. 2007;176(4):370-376.

27. Li N, Sioutas C, Cho A, et al. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environmental Health Perspectives*. 2003;111(4):455-460.

28. Sørensen M, Daneshvar B, Hansen M, et al. Personal PM2.5 exposure and markers of oxidative stress in blood. *Environmental health perspectives*. 2003;111(2):161-166.

29. Block ML, Calderón-Garcidueñas L. Air pollution: Mechanisms of neuroinflammation and CNS disease. *Trends in neurosciences*. 2009;32(9):506-516.

30. Perry VH, Cunningham C, Holmes C. Systemic infections and inflammation affect chronic neurodegeneration. *Nature Reviews Immunology*. 2007;7(2):161-167.

31. Bergeron C. Oxidative stress: Its role in the pathogenesis of amyotrophic lateral sclerosis. *Journal of the neurological sciences*. 1995;129:81-84.

32. Mhatre M, Floyd RA, Hensley K. Oxidative stress and neuroinflammation in Alzheimer's disease and amyotrophic lateral sclerosis: Common links and potential therapeutic targets. *Journal of Alzheimer's disease*. 2004;6(2):147-157.

33. D'Amico E, Factor-Litvak P, Santella RM, Mitsumoto H. Clinical perspective on oxidative stress in sporadic amyotrophic lateral sclerosis. *Free radical biology and medicine*. 2013;65:509-527.

34. Perry VH, Nicoll JA, Holmes C. Microglia in neurodegenerative disease. *Nature Reviews Neurology*. 2010;6(4):193-201.

35. Nunez Y, Boehme AK, Weisskopf MG, et al. Fine particle exposure and clinical aggravation in neurodegenerative diseases in New York State. *Environmental Health Perspectives*. 2021;129(2):027003.

36. Nunez Y, Boehme AK, Li M, et al. Parkinson's disease aggravation in association with fine particle components in New York State. *Environmental Research*. 2021;201:111554.

37. Malek AM, Barchowsky A, Bowser R, et al. Exposure to hazardous air pollutants and the risk of amyotrophic lateral sclerosis. *Environmental Pollution*. 2015;197:181-186.

38. Yu Z, Peters S, van BL, et al. Long-Term Exposure to Ultrafine Particles and Particulate Matter Constituents and the Risk of Amyotrophic Lateral Sclerosis. *Environmental Health Perspectives*. 2021;129(9):097702. doi:10.1289/EHP9131

39. Seelen M, Toro CRA, Veldink JH, et al. Long-term air pollution exposure and amyotrophic lateral sclerosis in Netherlands: A population-based case–control study. *Environmental Health Perspectives*. 2017;125(9):097023. doi:10.1289/EHP1115

40. Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: Pooled analysis. *BMJ*. 2021;374:n1904. doi:10.1136/bmj.n1904

41. Hamra GB, Laden F, Cohen AJ, Raaschou-Nielsen O, Brauer M, Loomis D. Lung cancer and exposure to nitrogen dioxide and traffic: A systematic review and meta-analysis. *Environmental Health Perspectives*. 2015;123(11):1107-1112.

42. Chen H, Kwong JC, Copes R, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: A population-based cohort study. *The Lancet*. 2017;389(10070):718-726.

43. Gibson EA, Nunez Y, Abuawad A, et al. An overview of methods to address distinct research questions on environmental mixtures: An application to persistent organic pollutants and leukocyte telomere length. *Environmental Health*. 2019;18(1):1-16.

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

45. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: A review of content, data quality, and research potential. *Clinical epidemiology*. 2015;7:449.

46. Trabjerg BB, Garton FC, van Rheenen W, et al. ALS in Danish registries: heritability and links to psychiatric and cardiovascular disorders. *Neurology Genetics*. 2020;6(2).

47. Mathis S, Goizet C, Soulages A, Vallat JM, Le Masson G. Genetics of amyotrophic lateral sclerosis: A review. *Journal of the Neurological Sciences*. 2019;399:217-226.

48. Kioumourtzoglou MA, 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.

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

50. Langholz B, Goldstein L. Risk set sampling in epidemiologic cohort studies. *Statistical Science*. Published online 1996:35-53.

51. Sillman S. The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmospheric Environment*. 1999;33(12):1821-1845.

52. Khan J, Kakosimos K, Raaschou-Nielsen O, et al. Development and performance evaluation of new AirGIS–a GIS based air pollution and human exposure modelling system. *Atmospheric environment*. 2019;198:102-121.

53. Brandt J, Christensen JH, Frohn LM, Palmgren F, Berkowicz R, Zlatev Z. Operational air pollution forecasts from European to local scale. *Atmospheric Environment*. 2001;35:S91-S98.

54. Brandt J, Christensen J, Frohn L, Berkowicz R. Air pollution forecasting from regional to urban street scale—-implementation and validation for two cities in Denmark. *Physics and Chemistry of the Earth, Parts A/B/C*. 2003;28(8):335-344.

55. Frohn LM, Ketzel M, Christensen JH, et al. Modelling ultrafine particle number concentrations at address resolution in Denmark from 1979-2018–Part 1: Regional and urban scale modelling and evaluation. *Atmospheric Environment*. 2021;264:118631.

56. 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.

57. 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.

58. 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.

59. Seinfeld J, Pandis S. Atmospheric chemistry and physics. 1997. *New York*. Published online 2008.

60. Galvin M, Gaffney R, Corr B, Mays I, Hardiman O. From first symptoms to diagnosis of amyotrophic lateral sclerosis: Perspectives of an Irish informal caregiver cohort—a thematic analysis. *BMJ Open*. 2017;7(3). doi:10.1136/bmjopen-2016-014985

61. Dickerson AS, Hansen J, Kioumourtzoglou MA, Specht AJ, Gredal O, Weisskopf MG. Study of occupation and amyotrophic lateral sclerosis in a Danish cohort. *Occup Environ Med*. 2018;75(9):630-638. doi:10.1136/oemed-2018-105110

62. Roberts AL, Johnson NJ, Chen JT, Cudkowicz ME, Weisskopf MG. Race/ethnicity, socioeconomic status, and ALS mortality in the United States. *Neurology*. 2016;87(22):2300-2308.

63. Bucher BT, Shi J, Pettit RJ, Ferraro J, Chapman WW, Gundlapalli A. Determination of marital status of patients from structured and unstructured electronic healthcare data. In: *AMIA Annual Symposium Proceedings*. Vol 2019. American Medical Informatics Association; 2019:267.

64. Norman RE, Carpenter DO, Scott J, Brune MN, Sly PD. Environmental exposures: an underrecognized contribution to noncommunicable diseases. *Reviews on environmental health*. 2013;28(1):59-65.

65. Rothman KJ, Greenland S, Lash TL, others. *Modern Epidemiology*. Vol 3. Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia; 2008.

66. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis, Third Edition*. CRC Press; 2013.

67. Mostofsky E, Schwartz J, Coull BA, et al. Modeling the association between particle constituents of air pollution and health outcomes. *American journal of epidemiology*. 2012;176(4):317-326.

68. Thunis P, Degraeuwe B, Pisoni E, et al. PM2.5 source allocation in European cities: A SHERPA modelling study. *Atmospheric Environment*. 2018;187:93-106.

69. Nuvolone D, Petri D, Voller F. The effects of ozone on human health. *Environmental Science and Pollution Research*. 2018;25(9):8074-8088.

70. Martin R, Peters G, Wilkinson J. Symmetric decomposition of a positive definite matrix. *Numerische Mathematik*. 1965;7(5):362-383.

71. Polson NG, Scott JG. On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis*. 2012;7(4):887-902.

72. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal*. 2006;1(3):515-534. doi:10.1214/06-BA117A

73. Lewandowski D, Kurowicka D, Joe H. Generating random correlation matrices based on vines and extended onion method. *Journal of multivariate analysis*. 2009;100(9):1989-2001.

74. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical science*. 1992;7(4):457-472.

75. R Core Team. R: A language and environment for statistical computing. Published online 2013.

76. Povedano M, Saez M, Martinez-Matos JA, Barceló MA. Spatial assessment of the association between long-term exposure to environmental factors and the occurrence of amyotrophic lateral sclerosis in Catalonia, Spain: A population-based nested case-control study. *Neuroepidemiology*. 2018;51(1-2):33-49.

77. von Schneidemesser E, Mar KA, Saar D. Black carbon in Europe: Targeting an air Pollutant and climate forcer. Published online 2017.

78. Pamphlett R, Rikard-Bell A. Different occupations associated with amyotrophic lateral sclerosis: Is diesel exhaust the link? *PloS One*. 2013;8(11):e80993.

79. Zhang R, Dai Y, Zhang X, et al. Reduced pulmonary function and increased pro-inflammatory cytokines in nanoscale carbon black-exposed workers. *Part Fibre Toxicol*. 2014;11:73. doi:10.1186/s12989-014-0073-1

80. Gao X, Xu H, Shang J, et al. Ozonized carbon black induces mitochondrial dysfunction and DNA damage. *Environ Toxicol*. 2017;32(3):944-955. doi:10.1002/tox.22295

81. Kyjovska ZO, Jacobsen NR, Saber AT, et al. DNA damage following pulmonary exposure by instillation to low doses of carbon black (Printex 90) nanoparticles in mice. *Environ Mol Mutagen*. 2015;56(1):41-49. doi:10.1002/em.21888

82. Benatar M, Turner MR, Wuu J. Defining pre-symptomatic amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2019;20(5-6):303-309.

83. Benatar M, Wuu J, McHutchison C, et al. Preventing amyotrophic lateral sclerosis: insights from pre-symptomatic neurodegenerative diseases. Published online 2021.

84. Eisen A, Kiernan M, Mitsumoto H, Swash M. Amyotrophic lateral sclerosis: A long preclinical period? *Journal of Neurology, Neurosurgery & Psychiatry*. 2014;85(11):1232-1238.

85. Weisskopf MG, Webster TF. Trade-offs of personal vs. more proxy exposure measures in environmental epidemiology. *Epidemiology (Cambridge, Mass)*. 2017;28(5):635.

86. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. CRC press; 2006.

87. Frisch M, Simonsen J. Marriage, cohabitation and mortality in Denmark: national cohort study of 6.5 million persons followed for up to three decades (1982–2011). *International Journal of Epidemiology*. 2013;42(2):559-578.

88. Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*. 2015;16(3):493-508.

Table 1. Demographic characteristics of cases and controls for 5-year average exposure group.

Characteristic	Overall, N = 23,232[a]	Case, N = 3,934[a]	Control, N = 19,298[a]
Average age (years)	66 (12)	66 (12)	66 (12)
Sex			
Female	10,973 (47%)	1,854 (47%)	9,119 (47%)
Male	12,259 (53%)	2,080 (53%)	10,179 (53%)
Socioeconomic status (SES)			
Group 1 (Highest)	2,337 (10%)	451 (11%)	1,886 (9.8%)
Group 2	2,839 (12%)	499 (13%)	2,340 (12%)
Group 3	4,360 (19%)	785 (20%)	3,575 (19%)
Group 4	6,598 (28%)	1,076 (27%)	5,522 (29%)
Group 5 (Lowest)	4,419 (19%)	717 (18%)	3,702 (19%)
Group 9 (Unknown)	2,679 (12%)	406 (10%)	2,273 (12%)
Place of birth			
Greater Copenhagen	4,858 (21%)	831 (21%)	4,027 (21%)
Big cities of Denmark	7,923 (34%)	1,357 (34%)	6,566 (34%)
Rest of Denmark	9,009 (39%)	1,548 (39%)	7,461 (39%)
Greenland	243 (1.0%)	53 (1.3%)	190 (1.0%)
Foreign	1,065 (4.6%)	122 (3.1%)	943 (4.9%)
Unknown	134 (0.6%)	23 (0.6%)	111 (0.6%)
Civil status			
Married	14,158 (61%)	2,411 (61%)	11,747 (61%)
Divorced	2,703 (12%)	433 (11%)	2,270 (12%)
Widowed	4,224 (18%)	726 (18%)	3,498 (18%)
Never married	2,147 (9.2%)	364 (9.3%)	1,783 (9.2%)
Last reported place of residence			
Greater Copenhagen	1,887 (8.1%)	335 (8.5%)	1,552 (8.0%)
Big cities of Denmark	9,385 (40%)	1,590 (40%)	7,795 (40%)
Rest of Denmark	11,954 (51%)	2,008 (51%)	9,946 (52%)
Greenland	6 (<0.1%)	1 (<0.1%)	5 (<0.1%)

[a]Mean (SD); n (%)

Table 2. Summary of 5-year average pollutant concentrations (all in $\mu g/m^3$).

Pollutant	Overall, N = 23,232[a]	Case, N = 3,934[a]	Control, N = 19,298[a]
NO_x	27 (20)	28 (21)	27 (20)
CO	238 (106)	239 (112)	237 (105)
EC	0.85 (0.42)	0.86 (0.45)	0.85 (0.42)
non-EC $PM_{2.5}$	11.76 (2.37)	11.78 (2.41)	11.76 (2.37)
O_3	51.9 (6.0)	51.9 (6.1)	52.0 (6.0)

[a]Mean (SD)

Figure Captions

Figure 1. Spearman correlation of 1,- 5-, and 10-year average pollutant concentrations.

Figure 2. Percentage change in odds of ALS diagnosis per 1-, 5- and 10-year average standard deviation (SD) increase for each pollutant. Results are from the Bayesian hierarchical model including each of EC, NO_x, CO, and non-EC $PM_{2.5}$ together, and were additionally adjusted by age, sex, year of birth, vital status, socioeconomic status, civil status, last reported place of residence, and place of birth.

Figure 1

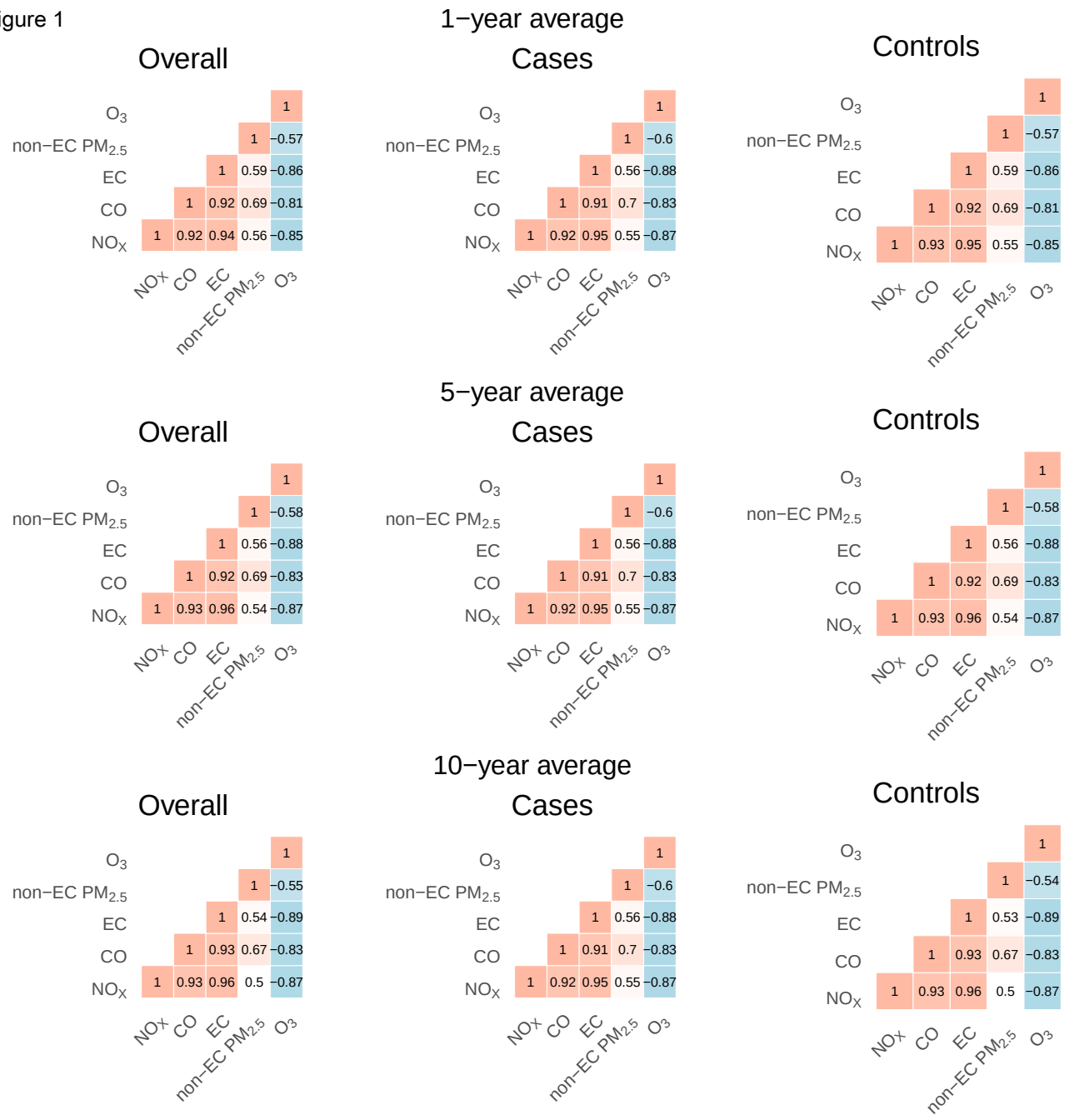
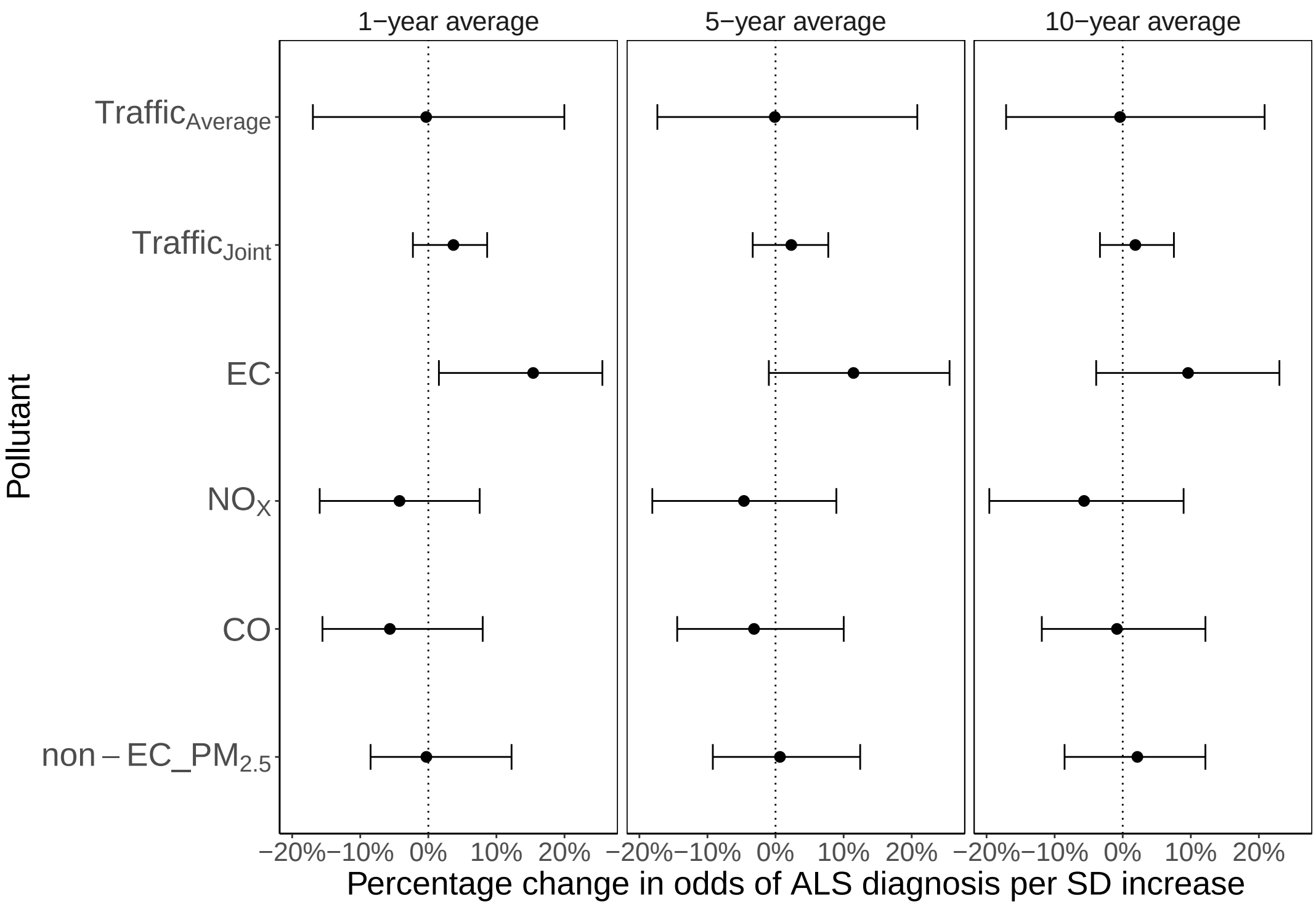

1−year average
Overall
Cases
Controls
5−year average
Overall
Cases
Controls
10−year average
Overall
Cases
Controls
O3
non−EC PM2.5
EC
CO
NOX
1
0.92
0.94
0.56
−0.85
0.59
−0.86
0.69
−0.81
−0.57
−0.6
0.91
0.7
−0.83
0.95
0.55
−0.87
0.93
0.96
0.54
−0.58
0.56
−0.88
0.69
−0.83
−0.87
0.67
−0.83
0.5
−0.87
−0.55
0.54
−0.89
0.53
−0.89
−0.54

Figure 2


1−year average
5−year average
10−year average
Pollutant
Traffic Average
Traffic Joint
EC
NOX
CO
non − EC_PM2.5
−20%
−10%
0%
10%
20%
Percentage change in odds of ALS diagnosis per SD increase

Click here to access/download

Supplemental Digital Content

supplemental digital content.pdf

Comments from the editors

We thank the Editors and Reviewers for their thoughtful and constructive suggestions. We have revised the manuscript in response to their comments, as detailed below.

All page/line/reference numbers refer to the clean revised manuscript.

Please take particular note of reviewer comments below on justifying and describing the use of the Bayesian hierarchical model, on determination and clarification of confounder control, and the lack of control for potentially important factors such as BMI, smoking, and indoor air quality.

We have addressed this, as detailed throughout relevant responses to Reviewers below. In particular, please see responses to Major Comments 4, 6, 10, 15 and Minor Comments 3, 5, 8 for Reviewer 1, and responses to Major Comments 1, 2, 4, 6, 7, 8, 9, 10, 12 and Minor Comments 14, 18, 19, 21, 22 for Reviewer 2.

Second, please ensure that the summary of main findings are fully supported by the results, e.g., take into account and explain sensitivity analyses in eFigure 1 that show weaker single-pollutant estimates for EC than those from the main models.

Our main model demonstrated that there was a potentially positive association between EC and ALS diagnosis, which contained traffic-related pollutants (EC, NO_x, CO) and non-EC $PM_{2.5}$, as well as adjusting implicitly by matching cases and controls on age, sex, year of birth, vital status, and explicitly by including terms in the model for SES, civil status, residence, place of birth.

Because the main models contain more terms which are highly correlated, it was not surprising that the estimated standard errors would be larger than single-pollutant models. Furthermore, because of the high correlations among traffic-related pollutants, the point estimates are also expected to somewhat vary between analyses. However, because of the consistency of the sign of the EC association throughout (i.e., whether in single-pollutant or multi-pollutant models), we have suggested that EC may be a driver of the relationship in the revised manuscript (P. 14, Lines 298-309):

We found that EC had the largest-in-magnitude independent association with ALS diagnosis, while associations with NO_x and CO were negative with credible intervals overlapping the null, and smaller in magnitude. Sensitivity analyses demonstrated that for single pollutant models, the association for EC was smaller than for our main multi-pollutant model, which took into account the variance-covariance structure of traffic-related pollutants. Overall conclusions for the association between EC and ALS diagnosis would have been similar from the single- or multi-pollutant models. The inconsistent associations for NO_x and CO in the multi- and single-pollutant models suggest that the model may have had limited success identifying each individual pollutant's association with ALS diagnosis due to the high level of collinearity of traffic-related pollutants. Nevertheless, the consistency of the sign of the central estimate of EC in all models suggests that EC may be a driver of the ALS and traffic-related pollutant association, though further analysis is required.

Finally, please ensure consistency in results presentation throughout the abstract and main text; e.g., the abstract presents posterior probabilities that are not included in the results section.

We have ensured that there is consistency in results presentation throughout the abstract and main text, e.g., (P. 13, Lines 273-278):

For 5-year average pollutant concentrations, we observed the largest overall association for the individual SD increase in EC (11.5%; 95% CrI: -1.0%, 25.6% per 0.42 μg/m^3; 96.3% posterior probability of positive association) (Figure 2). SD increases were associated with a decrease in odds of ALS diagnosis in NO_x (-4.6%; 95% CrI: -18.1%, 8.9% per 20 μg/m^3; 27.8% posterior probability of positive association) and CO (-3.2%; 95% CrI: -14.4%, 10.0% per 106 μg/m^3; 26.7% posterior probability of positive association).

Comments from reviewers

Reviewer #1: In this manuscript, the authors investigate the association between long-term traffic related air pollution and amyotrophic lateral sclerosis (ALS) in Denmark. The authors used Bayesian hierarchical modeling to estimate joint and individual effects of traffic related air pollutants (NOx, CO, EC) on odds of developing ALS. The major finding of this paper is that 5-year average EC concentration was individually associated with ALS. The study found no overall or joint association of traffic related pollution with ALS. This is an important topic and using methods to estimate joint effects of correlated pollutants is a necessary next step in air pollution studies. This paper was well done, though authors should explain certain analysis choices more clearly and provide more detailed results and discussion.

We thank the Reviewer for the thoughtful and constructive suggestions. We have responded point-by-point to the Reviewer's questions and comments below.

Major Comments

1) Methods, page 5, line 46: Why did authors only include patients that were at least 20 years old at diagnosis? This choice needs to be explained/motivated.

Cases in younger patients (i.e., younger than 20 years old) would have been at a much greater chance of misclassification, with a very high likelihood that a case identified in such a young person is an error in diagnosis coding (Trabjerg et al. 2020). Further, juvenile ALS cases have been explained to much larger degree by genetic mutations (Mathis et al. 2019). We have added to the explanation of the 20-year-old limit in the revised manuscript (P. 4, Lines 85-88):

We only included patients who were at least 20 years old when diagnosed because (i) cases younger than 20 years old were at a greater chance of misclassification, since ALS has been predominantly diagnosed in older adults,[46] and (ii) the very few juvenile ALS cases have been explained to a much larger degree by genetic mutations (~40%).[47]

2) Methods, page 6, line 51: Please include more detail about the spatio-temporal air pollution modeling system.

We have now added more detail about the spatio-temporal air pollution modelling system in revised manuscript (PP. 5-6, Lines 111-120):

We obtained predictions on monthly concentrations of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) (as well as ozone, O_3, for a sensitivity analysis, usually negatively correlated with other pollutants due to its chemistry[51]), at residential addresses of study participants from the validated spatio-temporal air pollution modelling system DEHM-UBM-AirGIS that provides full space and time coverage over the study period, described in detail elsewhere.[52–55] In brief, DEHM-UBM-AirGIS is a human exposure modelling system for traffic pollution, developed for application in Danish air pollution epidemiological studies. The modelling system is able to generate street configuration and traffic data based on digital maps and national databases, which enables estimation of air quality levels at a large number of addresses in an automatic and effective way.

We have additionally provided metrics of predictive accuracy for each of the used pollutants in the revised manuscript (P. 6, Lines 121-124):

The models have good predictive accuracy, with average monthly correlations between measured and modelled results of 0.85 for NO_x, 0.91 for CO, 0.92 for O_3, 0.79 for EC, and 0.83 for annual concentrations of $PM_{2.5}$.[52,55]

3) Methods, page 7, line 17: Please provide more details on how the 1-, 5-, and 10-year averages were created for air pollution exposures. Was a weighted average created based on how long the participant lived in one location?

We have given more details on how the 1-, 5-, and 10-year averages were created for the air pollution exposures in the revised manuscript (P. 6, Lines 129-133):

Based on the residential history of each case or control, we calculated 1-, 5-, and 10-year average exposure to each pollutant ending at one year before the index date, as diagnosis has been shown previously to occur at a median of 12 months after symptoms onset.[60] *Specifically, each case or control average value (1-, 5- or 10-year) was calculated as the mean of all concentrations recorded across time at the recorded addresses within each time window.*

4) Methods, page 7, line 41: How did the authors determine confounding variables?

We matched cases and controls on age, sex, year of birth, and vital status, as ALS prevalence varies according to these characteristics. We also accounted for socioeconomic status (SES), civil status, last reported place of residence, and place of birth. SES influences many lifestyle factors, such as obesity, and is associated with ALS diagnosis in Denmark (Dickerson et al. 2018). Civil status was included due to the influence that a spouse has on visiting a family physician (Bucher et al. 2019). Last reported place of residence was included to account for various local environmental and behavioral stressors, such heavy metals, which may have an influence on ALS prevalence (Oskarsson, Horton, and Mitsumoto 2015). Place of birth was included to account for the variety of childhood exposures, which vary by location, which may have an impact on the probability of developing ALS (Norman et al. 2013). Ultimately, we were limited by what was available in the Danish Civil Registration System.

We have added these details to the revised manuscript (PP. 7-8, Lines 141-159):

We included a set of covariates based on as close as possible to index date to account for potential confounding bias, including household socioeconomic status (SES) based on last-reported job title at index date; civil status at index date, last reported place of residence at index date, and place of birth. We used a five-category individual-level SES definition developed by the Danish Institute of Social Sciences, based on job titles from income tax forms, which has been associated with ALS diagnosis in Denmark,[61] *as well as how quickly one is identified as having ALS in the Danish Civil Registration System.*[62] *Group 1 (highest status) includes corporate managers and academics; group 2: proprietors, managers of small businesses and teachers; group 3: technicians and nurses; group 4: skilled workers; and group 5: unspecialized workers, such as entry-level positions within food and retail environments. We also included a group for participants whose job title was unknown (group 9). For each married participant, we used the higher of the couple's individual SES categories, when available. We also used information on civil status (never married, married, divorced, widowed) due to the influence that a spouse may have on visiting a family physician,*[63] *last reported place of residence from postcode (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland) to account for various local environmental and behavioral stressors,*[7]

and place of birth (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland, foreign, unknown) to adjust for other potential family-specific, location-specific, and early-life confounders, which may have an impact on the probability of developing ALS.[64] Ultimately, we were limited by what was available in the Danish Civil Registration System.[62]

5) Methods, page 8, line 50: Please include more motivation and reasoning for ozone sensitivity analysis.

There is evidence from other studies that ozone concentrations are associated with many different adverse health outcomes (Nuvolone, Petri, and Voller 2018). However, ozone is highly negatively correlated with traffic-related air pollutants, which is also the case in our study, as can be seen in Figure 1 in the main manuscript. Adding an extra pollutant which is highly-correlated with the others may have posed challenges to the statistical inference of the model, which is why we included it in the sensitivity analysis rather than the main analysis.

Nevertheless, our sensitivity analyses demonstrated that inclusion of ozone did not noticeably change our results (eFigure 1). Our conclusions from the main analysis, thus, would not have changed had we included ozone.

We have clarified why we included ozone in the sensitivity analysis in the revised manuscript (P. 9, Lines 197-199):

In a sensitivity analysis, we included O_3 in the model, as O_3 concentrations have been associated with many adverse health outcomes,[69] and were negatively correlated with traffic-related pollutants […]

6) Methods, page 9-10: When discussing priors used for the Bayesian model why are weakly-informative priors given to non-EC PM2.5, but non-informative priors are given to other parameters? Please give more detail when justifying use of priors.

We have edited the description of priors used in the Bayesian model to clarify that weakly-informative priors were given to all of the pollutant coefficients, with the other pollutant coefficients (i.e., not that of non-EC $PM_{2.5}$) given the same prior via the hierarchical structure of the traffic-related pollutant concentrations. The revised description can be found in the revised manuscript (P. 10, Lines 220-223):

We used weakly-informative priors so that data drove parameter estimation. Hyper-priors for coefficients on non-EC $PM_{2.5}$ and covariates were N(0,10); for σ_λ and τ we used Half-Cauchy(0,10), as recommended by Gelman, Polson and Scott as a weakly-informative prior;[71,72] Ω was defined by the weakly-informative prior LKJCorr(1).[73]

7) Results, page 11: Why are only 5-year average exposures presented in the results section? Please provide more justification or present the 1- and 10-year data as well in the main tables/figures.

We provide the 5-year average exposures as the main results here because a 5-year average exposure is a balance between most recent exposure as well as long-term concentration. We have clarified this by adding more justification in the revised manuscript (P. 12, Lines 262-263):

For the main results, we present 5-year average exposure associations as a balance between representation of most recent exposure as well as long-term concentration.

8) Results, page 12: Please provide more discussion of the protective effect of NOX and CO. How does this effect the null joint effect of NOX, CO, and EC?

While the point estimates of the association of ALS and NO_x and CO are less than 0% change, the credible intervals widely overlap with the null. Nevertheless, we were also surprised by this result. We discuss further why this might be in the Discussion of the revised manuscript (P. 14, Lines 303-309):

Overall conclusions for the association between EC and ALS diagnosis would have been similar from the single- or multi-pollutant models. The inconsistent associations for NO_x and CO in the multi- and single-pollutant models suggest that the model may have had limited success identifying each individual pollutant's association with ALS diagnosis due to the high level of collinearity of traffic-related pollutants. Nevertheless, the consistency of the sign of the central estimate of EC in all models suggests that EC may be a driver of the ALS and traffic-related pollutant association, though further analysis is required.

9) Results, page 12, line 46: Please provide correlation coefficient for ozone and other pollutants.

We now provide the correlation coefficient for ozone and the other pollutants in the revised manuscript (P. 13, Lines 270-271):

O_3 was negatively correlated with other pollutants, ranging from -0.54 to -0.89.

10) Results, page 13, line 9: Can the authors also provide posterior probability for the null?

The posterior probability is the amount of the marginal posterior distribution estimate of the coefficient of interest which is above the null. Therefore a 50% probability means it is as likely as not that the marginal estimate is null. We clarify this in the text with a further exposition of what various posterior probabilities mean (P. 11, Lines 237-241):

To calculate the probability that an association estimate was greater than null, we used the 4,000 samples of the posterior distribution and took the proportion of samples which were above the null. A 50% probability means that it is as likely as not that the marginal estimate is null, a probability closer to 100% indicates that the association is more likely to be truly positive, with closer to 0% indicating more likely to be truly negative.

11) Results page 13, line 22: "(eFigure 1) resulted in positive associations for each of EC, NOx, CO, with positive associations for non-EC PM2.5 in all but the model with EC." Is eFigure 1 the correct figure? eFigure 1 shows protective effects for all but EC.

We have clarified that we are only referring to the single-pollutant models D, E and F in eFigure 1 in the revised manuscript (P. 13, Lines 284-287):

Single-pollutant models for each traffic-related pollutant adjusting for non-EC $PM_{2.5}$ (eFigure 1; single traffic-related pollutant models D, E and F) resulted in positive associations for each of EC, NO_x, CO, with positive associations for non-EC $PM_{2.5}$ in all but the model with EC.

12) Discussion, page 13, line 38: Authors state that they found an average increase in concentration of traffic-related pollutants was associated with and increase in odds of ALS. Though only EC showed a positive association and joint effect was null?

We have clarified that we are referring to the joint association here (P. 14, Lines 296-298):

In the largest case-control study of ALS and traffic-related air pollution to date, we found that a joint increase in average concentrations of traffic-related pollutants was potentially associated with an increase in odds of ALS diagnosis, with the clearest results for EC.

13) Discussion, page 14, line 48: If EC and NOX are so highly correlated why are their associations with ALS so different?

While EC and NO_x are highly-correlated (0.94-0.96), they are not perfectly correlated. Our Bayesian hierarchical model structure allowed us to incorporate their high correlation while also leveraging the differences to make inferences about their single-pollutant associations with ALS diagnosis. With this high collinearity in mind, however, it is reassuring that single-pollutant model associations for EC and NO_x (eFigure 1; Models D and F) give similar results. We discuss that while the EC results were consistent, the NO_x results were not in the revised manuscript, indicating that while this would indicate that EC is potentially driving the association, further work is necessary (P. 14, Lines 304-309):

The inconsistent associations for NO_x and CO in the multi- and single-pollutant models suggest that the model may have had limited success identifying each individual pollutant's association with ALS diagnosis due to the high level of collinearity of traffic-related pollutants. Nevertheless, the consistency of the sign of the central estimate of EC in all models suggests that EC may be a driver of the ALS and traffic-related pollutant association, though further analysis is required.

14) Discussion, page 14, line 58: the 1-year estimate may be the most robust to exposure misclassification, provide more justification for why it may be the most relevant exposure window.

We understand why the Reviewer suggests that the 1-year estimate may be the most robust to exposure misclassification. However, for every case and control, we have addresses geocoded over time for each year during our study period. This would minimize the potential for exposure misclassification for earlier years in the study period compared with later years. Therefore, we do not think that the 1-year estimate may necessarily be the most robust to exposure misclassification.

We have added more justification for why the 1-year average concentrations may be the most relevant exposure window in the revised manuscript (PP. 15-16, Lines 333-339):

EC exposure was more strongly associated with 1-year than for 5-/10-year average concentrations, which may indicate that the previous year of exposure may be the most relevant exposure window relevant to traffic-related exposures and ALS; this is biologically plausible,

as this critical exposure window would be at the pre-symptomatic stage of underlying ALS progression, where traffic-related pollution exposure may add to the ongoing cellular or molecular process of the disease, to the point where the body can no longer compensate and subsequently enters the clinical phase.[82–84]

15) Please include justification for use of the Bayesian hierarchical model, as opposed to other mixture modeling methods (Bayesian kernel machine regression, etc.) that are more established.

Bayesian Kernel Machine Regression is not currently appropriate for matched case-control studies. Further, we required a Bayesian model to be able to take into account variance-covariance structure of the traffic-related pollutants, as well as being able to optimally capture uncertainty. We have clarified this in the future research paragraph in the revised manuscript (P. 17, Lines 364-367):

Other mixture model methods, such as Bayesian Kernel Machine Regression (BKMR)[88] *might be useful in further exploring the robustness of joint associations in a different framework, though BKMR was not appropriate for our particular research question, since BKMR is not currently appropriate for case-control studies.*

16) Figure 1. Please provide figure for 1-, and 10-year exposure estimates.

On Figure 1, we now provide Spearman correlations for 1-, 5-, and 10-year pollutant concentrations, as seen below:

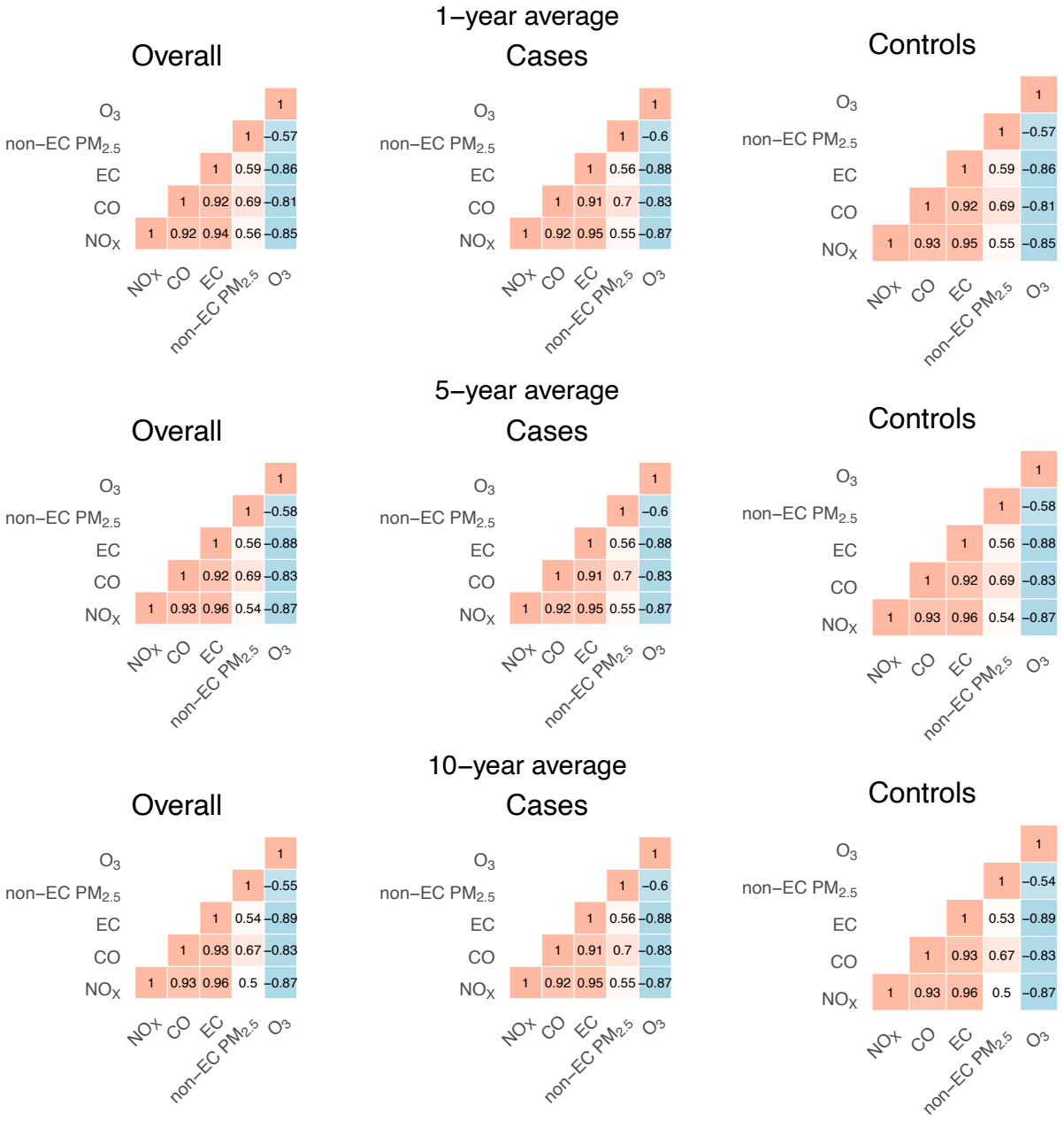


Figure 1. Spearman correlation of 1,- 5-, and 10-year average pollutant concentrations.

17) Table 1. SES group 9, why were unemployed and unclassified grouped together?

This should have been stated as 'unknown' here and we have corrected this in the revised manuscript (P. 7, Lines 150-151):

We also included a group for participants whose job title was unknown (group 9).

Minor Comments

1) Throughout the paper there is inconsistent use of the abbreviation SD for standard deviation.

We have now introduced the abbreviation for the term standard deviation (SD), which is also in the list of Abbreviations. We now consistently use this throughout the revised manuscript e.g., (P. 13, Lines 273-275):

For 5-year average pollutant concentrations, we observed the largest overall association for the individual SD increase in EC (11.5%; 95% CrI: -1.0%, 25.6% per 0.42 μg/m^3; 96.3% posterior probability of positive association) (Figure 2).

2) Introduction, page 4, line 26: Unclear why this sentence is a contradiction.

We have added the word 'also' in the subclause of the sentence to demonstrate that not only should air pollution be studied in relation to respiratory- and cardiovascular-related outcomes, but also with nervous system-related outcomes and neurodegeneration in the revised manuscript (P. 3, Lines 53-56):

Although air pollution is commonly studied in association with respiratory- and cardiovascular-related outcomes, e.g., references [9–14], epidemiologic and toxicological studies also support several plausible biological mechanisms in association with the nervous system and neurodegeneration, e.g., references [15–34].

3) Methods, page 8 line 42: Authors say non-EC PM2.5 adjust for other air pollutants from other sources. This would only adjust for PM2.5 from other sources.

This Reviewer is correct. We have also clarified that $PM_{2.5}$ is also an indicator of overall air pollution mixture, and that adjusting by $PM_{2.5}$ will also indicate whether pollution from other sources not explicitly quantified might also have associations with ALS. We have clarified this in the revised manuscript (P. 9, Lines 192-194):

If other sources of air pollution are associated with ALS, then including non-EC $PM_{2.5}$ adjusts for $PM_{2.5}$ from other sources,[67] as well as indicating whether pollution from other sources not explicitly quantified might also have associations with ALS.

4) Methods, page 9, line 16: Model quantifies log-odds of one standard deviation increase. Please add explanation of why you chose one standard deviation increase instead of interquartile range.

Presenting effect estimates for a standardized increase in pollutant concentrations is preferrable when examining multiple pollutants and interested in the overall effect, because we combined concentration associations via the traffic terms in the model. Both standard deviation (SD) and interquartile range (IQR) are measures of the spread of values, which can be equivalently used to present effect estimates for an increase in standardized pollutant concentrations. There is no inherent benefit to picking one or the other in this case, as the role of dividing by both measures of spread is to normalize concentrations. We have added that both are equivalent ways of standardizing in the revised manuscript (P. 9, Lines 187-192):

β_{NO_x}, β_{CO}, β_{EC}, $\beta_{PM_{2.5}}$ the pollutant-specific coefficients (log-odds) per standard deviation (SD) increase in concentration of NO_x, CO, EC, non-EC $PM_{2.5}$ respectively, scaled by their respective SDs and centered at their means, with each β a pollutant-specific association adjusted by other terms in the model and the rest as coefficients for subject-specific covariates. Interquartile Range (IQR) could equivalently be used to scale pollutant concentrations.

5) Results, page 12, line 4: Joint association of which pollutants?

We have clarified which pollutants in the revised manuscript (P. 13, Lines 278-280):

The joint association of traffic-related pollutants (EC, NO_x, CO) was 2.3% (95% CrI: -3.3%, 7.7%), with an 77.8% posterior probability of a positive association.

6) Results, page 13, line 17: "10-year average exposure results were attenuated versions of the 1- and 5-year results." Wording could be improved.

We have expanded upon this phrase to be more descriptive in the revised manuscript (P. 13, Lines 283-284):

Compared to the 1- and 5-year results, the 10-year average exposure results were attenuated, as associations tended further to the null.

7) Results overall: eFigure 2 is not mentioned.

We now mention eFigure 2 in the revised manuscript (PP. 13-14, Lines 291-293):

A map of average concentration of included pollutants (NO_x, EC, $PM_{2.5}$, CO, O_3) across Denmark for a representative year (2000; middle of study period 1989-2013) is also available in eFigure 2.

8) Discussion, page 14, lines 29-39: If BMI is not a confounder this is unnecessary to include.

We have deleted these lines in the revised manuscript.

9) Figure 2. Please make the point estimate dots bigger.

We have made these points bigger in Figure 2 in the revised manuscript, copied below for convenience:

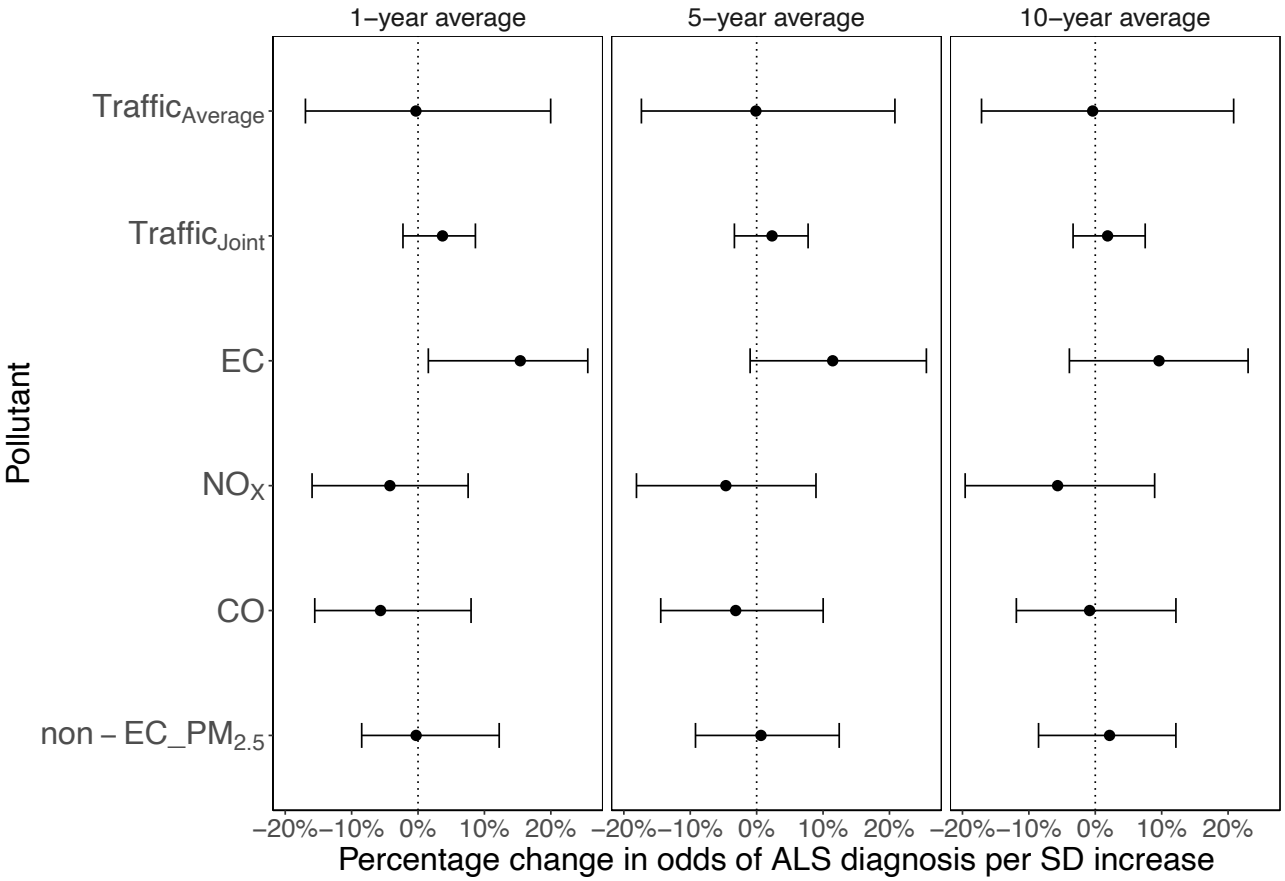


Figure 2. Percentage change in odds of ALS diagnosis per 1-, 5- and 10-year average standard deviation (SD) increase for each pollutant. Results are from the Bayesian hierarchical model including each of EC, NO_x, CO, and non-EC $PM_{2.5}$ together, and were additionally adjusted by age, sex, year of birth, vital status, socioeconomic status, civil status, last reported place of residence, and place of birth.

Reviewer #2: Comments pasted below. The uploaded review contains a figure.

This paper describes a study of the effect of traffic-related air pollution exposures on amyotrophic lateral sclerosis (ALS). The study, set in a very large healthcare administrative database in Denmark, has several strengths, including exposure assessments that span decades and an attempt to account for a latent period between disease onset and clinical diagnosis. The report lacks clarity and detail on important aspects of the investigation, encompassing the target estimands, validity of the ALS measurement, and sources of confounding. My concerns follow:

We thank the Reviewer for the thoughtful and constructive suggestions. We have responded point-by-point to the Reviewer's questions and comments below.

MAJOR

1. The interpretive distinction between the 3 estimands pursued in this study was unclear. E.g., from the abstract, "… the overall and joint association for the three traffic-related pollutants (NOx, CO, and EC), as well as pollutant-specific associations." What does "overall" mean exactly, and how does "overall" differ from "joint"? Does "joint" include interactions? Does "pollutant-specific" reflect adjustment for other pollutants? For example, how would you express the parameter estimate from each of these in words that are true to the underlying mathematics?

We distinguished between the "joint" association of the three pollutants (i.e., total percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO), and "overall" association of the three pollutants (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO). We have stated this in the revised manuscript (P. 8, Lines 171-176):

We employed a Bayesian hierarchical formulation because it enables estimates of (a) independent pollutant-outcome associations, (b) a joint association of the three pollutants (i.e., total percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO), and (c) an overall average traffic association (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO), while accounting for the variance-covariance structure between the highly-correlated exposures and their coefficients.[66]

The Reviewer is correct that the "pollutant-specific" associations reflect adjustment not only for other pollutants, but also for other covariates to account for potential confounding bias, described in more detail in the revised manuscript (P. 9, Lines 187-191):

β_{NO_x}, β_{CO}, β_{EC}, $\beta_{PM_{2.5}}$ *the pollutant-specific coefficients (log-odds) per standard deviation (SD) increase in concentration of* NO_x, CO, EC, *non-EC* $PM_{2.5}$ *respectively, scaled by their respective SDs and centered at their means, with each* β *a pollutant-specific association adjusted by other terms in the model and the rest as coefficients for subject-specific covariates. […]*

2. Related to #1, the joint association is described in these terms: "This sum quantifies the association (log-odds) with ALS of a one-SD increase in the three pollutants simultaneously." Although it is mathematically possible to compute this, how well does an increment of SD in all 3 pollutants match up with the joint distribution of these

pollutants in the population? Do the concentrations vary at about the same pace? i.e., can you identify locations (or location-periods) that are 1-SD apart (or 0.5 or 0.1 SD apart, etc.) on all 3 pollutants?

As described in the revised manuscript, the pollutants are highly correlated (PP. 12-13, Lines 265-271):

The 5-year average traffic-related pollutant concentrations were 27 µg/m^3 for NO_x (SD=20 µg/m^3), 238 µg/m^3 for CO (SD=106 µg/m^3) and 0.85 µg/m^3 for EC (SD=0.42 µg/m^3) (Table 2). Figure 1 shows Spearman correlations between pollutants for 1-, 5-, and 10-year average exposures. Traffic-related pollutants (NO_x, CO, EC) were highly correlated in cases, controls and overall, ranging from correlations of 0.91 to 0.96. Otherwise, non-EC $PM_{2.5}$ was most highly correlated with CO, ranging from 0.67 to 0.7. O_3 was negatively correlated with other pollutants, ranging from -0.54 to -0.89.

This would imply that the increase in one pollutant by 1-SD would result in a similar relative increase in the other related pollutants. Therefore, each of the three traffic-related pollutants each increasing simultaneously by 1-SD is a physically-plausible scenario.

3. Introduction: the literature review seemed cursory (e.g., "… epidemiologic and toxicological studies support several plausible biological mechanisms in association with the nervous system and neurodegeneration.15-34"). I recommend citing systematic reviews (e.g., Integrated Science Assessments from the US EPA) or using "e.g." before some of the citations.

As per the Reviewer's suggestion we have added "e.g." before some of the example citations in the Introduction here (P. 3, Lines 53-60):

Although air pollution is commonly studied in association with respiratory- and cardiovascular-related outcomes, e.g., references [9–14], epidemiologic and toxicological studies also support several plausible biological mechanisms in association with the nervous system and neurodegeneration, e.g., references [15–34]. Ambient air pollution, especially urban air pollution, is a ubiquitous exposure that has been associated with several other neurodegenerative disorders, e.g., references [16–21,35,36]. and is consistently linked to systemic inflammation,[22–24] oxidative stress,[25–28] and neuroinflammation,[15,29] all of which, in turn, have been reported as key pathways to ALS pathogenesis, e.g., references [30–34].

4. Introduction/Methods: the use of ozone in this investigation was confusing. The Introduction states, "Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NOx), carbon monoxide (CO), and elemental carbon (EC)— as well as fine particles (PM2.5) and ozone (O3), we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis …," which loosely implied that ozone was a traffic-related pollutant of interest. Later, in the Methods, O3 is described as being part of "a sensitivity analysis, usually negatively correlated with other pollutants due to its chemistry." This requires more explanation. Was exposure to O3 not expected to be relevant to ALS? Are the predicted O3 concentrations inaccurate?

Our study aim was to investigate whether each traffic-related air pollutant (EC, NO_x, CO), individually, jointly and overall, was associated with ALS diagnosis, while also fully adjusting

for other pollutants ($PM_{2.5}$ from other sources and O_3) and other relevant covariates. To clarify that our main focus was the traffic-related pollutants, we have edited the last paragraph of the Introduction in the revised manuscript (P. 4, Lines 68-72):

Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NO_x), carbon monoxide (CO), and elemental carbon (EC)— we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis, and estimate their (b) joint and (c) overall traffic-related emissions associations.

While there is no existing evidence, to our knowledge, to assess whether O_3 concentrations may impact ALS diagnosis, we included a sensitivity analysis to examine whether including it made a difference to the main set of results, which it did not. We explain this in the revised manuscript (P. 9, Lines 197-198):

In a sensitivity analysis, we included O_3 in the model, as O_3 concentrations have been associated with many adverse health outcomes,[69] […]

The models we have used for air pollution prediction, DEHM-UBM-AirGIS, have good predictive accuracy for all pollutants, including predicted concentrations of O_3, with average monthly correlations between measured and modelled results quoted in the revised manuscript (P. 13, Lines 289-291):

Results from variations of the main model in the sensitivity analyses were robust to prior choices, inclusion of O_3, and inclusion of parish-level SES (eFigure 1).

5. Methods: The validity of the registry for identifying ALS cases requires more detail, as it is a fundamental aspect of this investigation. Was the validation against an in-person assessment? a neurologist's in-depth review of medical records? Did the validation compare date of diagnosis with the date of symptom onset? What were the quantified indices of accuracy? As appropriate, it could be useful to mention the potential influence of outcome misclassification (or lack thereof) on the findings and mention determinants of misclassification if known.

We had claimed the validity of obtaining ALS diagnosis data from the Danish National Patient Register based on previous work from some of our co-authors, including the senior author of this current study (Kioumourtzoglou et al. 2015). In that work, a specialist ALS neurologist reviewed medical records of registry and mortality data for factors that may have been slightly related to agreement of ALS diagnosis. This previous study found that the use of hospital discharge and death certificate records are a valuable and highly reliable tool for ALS epidemiologic studies such as our current one. We have added detail to the revised manuscript (PP. 4-5, Lines 88-91):

In our validation study, Register data for ALS ascertainment were highly reliable; working with a specialist ALS neurologist to review medical records and comparing to death certificates and hospital discharges, the Danish National Patient Register was found to have an overall predictive value for ALS of 82%.[48]

We have added to the Limitations in the Discussion that while the Danish National Patient Register was highly reliable for ALS ascertainment, outcome misclassification may still have been possible (P. 16, Lines 356-357):

While a previous study found that ALS ascertainment from the Danish National Patient Register was highly reliable,[48] outcome misclassification cannot be ruled out, […]

While the studied quoted above did not examine date of system onset compared with date of diagnosis, another study in Ireland found that diagnosis of ALS occurs at a median of 12 months after symptoms onset (Galvin et al. 2017), which we have stipulated in the revised manuscript (P. 6, Lines 129-131):

Based on the residential history of each case or control, we calculated 1-, 5-, and 10-year average exposure to each pollutant ending at one year before the index date, as diagnosis has been shown previously to occur at a median of 12 months after symptoms onset.[60]

Nevertheless, we have added that this may also be a limitation of the work in the Discussion of the revised manuscript (P. 16, Lines 356-358):

While a previous study found that ALS ascertainment from the Danish National Patient Register was highly reliable,[48] outcome misclassification cannot be ruled out, nor can the possibility that date of diagnosis and symptom onset were irregularly aligned.

6. Methods: "… we removed the EC concentration from the total PM2.5 mass concentration …." How did you "remove" EC? Subtraction? Using residuals from a regression model?

Our pollutant model for $PM_{2.5}$, DEHM-UBM-AirGIS (Khan et al. 2019; Jørgen Brandt et al. 2001; J Brandt et al. 2003; Frohn et al. 2021), constructed $PM_{2.5}$ concentrations by adding from specific species of pollutants, one of which was EC. We were therefore able to subtract the EC concentration from the $PM_{2.5}$ concentration to obtain non-EC $PM_{2.5}$ concentrations. We have added a clarification to the revised manuscript (P. 6, Lines 124-127):

Because traffic is a major source of $PM_{2.5}$ and EC one of the main $PM_{2.5}$ components in urban environments,[59] we removed the EC concentration from the total $PM_{2.5}$ mass concentration (non-EC $PM_{2.5}$) by subtraction, to avoid overadjustment when including both in the models simultaneously.

7. Methods, adjustment for SES/occupation: was the goal to adjust for individual-level education and/or occupation (and its attendant exposures) or to adjust for overall SES in the household? E.g., "For each married participant, we used the higher of the couple's individual SES categories." This seems to be getting at household SES. Co-habitation among unmarried couples is common in Denmark. How was this addressed? https://academic.oup.com/ije/article/42/2/559/737789 Was it important to capture information differently from previously married people? This should come down to the construct you are trying to measure. (The SES-ALS paper cited was about individual-level occupation.

Our goal in adjusting for SES was to adjust for occupational exposure as well as household SES, as this would be potentially related to how quickly one is identified as having ALS in the

registry system and to the predicted air pollution concentrations at their residence. Related to this, when compared with previous marriage information, current household SES may have had more influence on the speed at which ALS diagnosis could have occurred. In any case, we have a category of 'divorced' people in our civil status adjustment (P. 7, Lines 151-153):

For each married participant, we used the higher of the couple's individual SES categories, when available. We also used information on civil status (never married, married, divorced, widowed) […]

The Reviewer is correct to point out that co-habitation is common among couples in Denmark, which would not be captured by our analysis. We have no detail in our dataset which would allow this to be adjusted for. We have added this as a Limitation in the Discussion of the revised manuscript (PP. 16-17, Lines 358-361):

While our analysis adjusted for marital status and household SES, many couples in Denmark cohabitate. This would not be captured by our analysis, and ALS diagnosis in relation to cohabitation status should be further investigated.[87]

8. Methods, timing of covariates: for covariates whose values could vary over time (e.g., occupation, civil status, parish-level SES), what was the timing of these covariates relative to the exposure and outcome periods?

All covariates for cases and controls were constructed similarly, using exact dates and matching to index date where the data allowed. Specifically, civil status was assessed at index date; last reported place of residence was the last residence the case or control had reported at the index date; parish-level SES is based on last reported address at index date; employment was based on last self-reported job title; SES was obtained based on the last self-reported job title, which for older ages, during which most the ALS cases in our study occurred, was not usually obtained at the index date, which we have clarified in the revised manuscript (P. 7, Lines 141-144):

We included a set of covariates based on as close as possible to index date to account for potential confounding bias, including household socioeconomic status (SES) based on last-reported job title at index date; civil status at index date, last reported place of residence at index date, and place of birth.

9. Methods: calendar time as a course of confounding? Given that air pollutant concentrations and other determinants of ALS may have changed over time, might calendar time be a source of confounding?

In a cohort study, this could certainty be the case. However, ours was a case-control study. Given that our cases and controls were matched on age and year of birth (among other characteristics) the controls would have been identified at a point very close in time to the associated controls. Therefore, we do not hold that calendar time could have been a source of confounding. We have clarified this in the revised manuscript (P. 16, Lines 342-343):

We do not expect that calendar time was a potential source of confounding, as the controls were matched on age and year of birth.

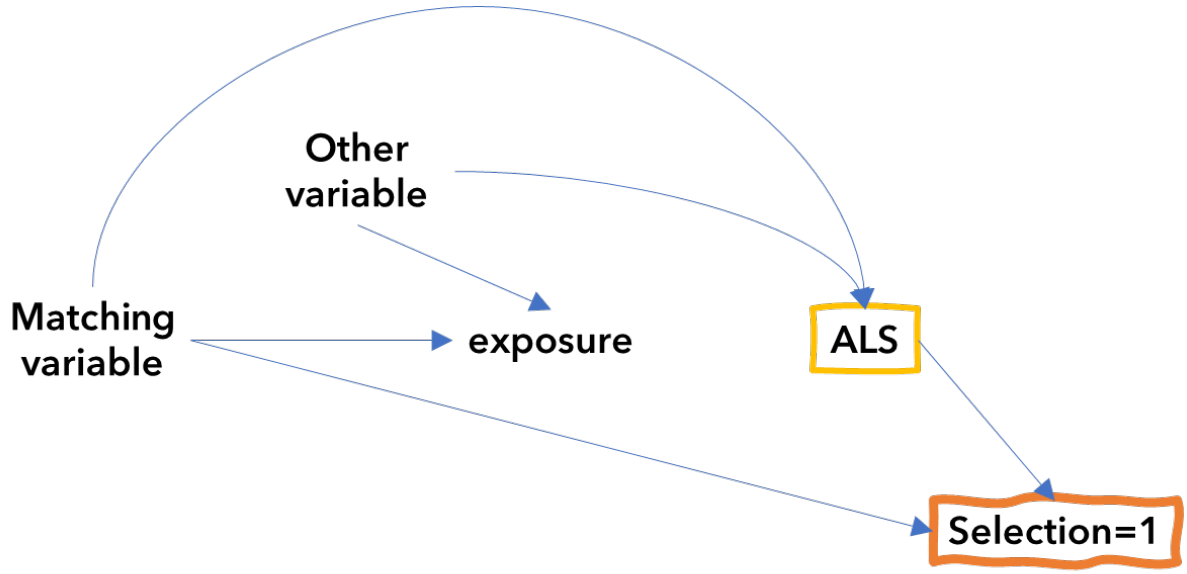
10. Discussion: potential that smoking and/or BMI could be sources of confounding. The

paper states, "… to induce confounding bias, any unaccounted-for variable would have to influence both ALS diagnosis and air pollution. BMI, previously associated with ALS, would not confound the association between traffic-related air pollution and ALS, as pollutant concentrations are derived independently from BMI distribution. Any BMI-air pollution association in our study, thus, would be via SES." Through complex social and economic mechanisms, the association of air pollution exposure with any given factor (e.g., BMI) can vary across populations. For example, in some study populations, areas where smoking is more common have higher concentrations of traffic-related air pollutants; in other settings, the pattern is reversed; and still others, there is little association. How is it known that BMI is not associated with exposure in this study population? Furthermore, how is it known that any such association in this study population would operate through SES? Analogous questions could be posed about smoking, as well. In the absence of firm answers to these questions, it could be useful to conduct a quantitative bias analysis, particularly as the estimated effect sizes are small.

We agree with the Reviewer that there are complex social and economic mechanisms through which BMI may well be related to pollution levels in Denmark. However, to be a potential source of confounding, BMI would have to be causing the variation in pollution levels. There is no evidence that we know of that would suggest that BMI drives variation in pollution levels, and the way that the exposures were predicted did not in any way take into account individual-level BMI data. Rather, SES predicts where one lives, and also one's BMI, and where one lives is a driver of air pollution levels, which warranted adjusting for in our analysis. We have blocked the path via SES if there were such an association.

Because Reviewer 1 asked us to remove most of this section, we have largely deleted and summarized the original statements.

11. Table 1/Table 2: Distribution of covariates by exposure level. To provide more information on correlates of exposure in the underlying population, conditional on the matching factors, it would be helpful to provide a table showing the co-distribution of key covariates and air pollutant exposure among the controls. This could be a challenging proposition, though. The matching scheme means that the associations of matching factors with exposure may be distorted so long as there are other common causes of exposure and ALS. Thus, it is not clear whether, without extensive additional exploration, it would be possible to show meaningful co-distributions of exposure with the matching variables. However, conditional on the matching variables, it may be informative to show associations of other covariates with exposure. See example DAG below.



The case-control study is, by definition, restricted to Selection=1. Showing data additionally restricted to controls from this study also conditions on ALS by restricting to ALS=0. Thus, in this subset …

- "matching variable" is associated with exposure through a non-causal path involving the collider at ALS.
- "other variable" is associated with exposure through a non-causal path involving the collider at ALS. Conditioning on the matching factor will block this.

More background on structure by Mansournia et al 2013:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3733703/pdf/dyt083.pdf

We thank the Reviewer for this comment. We now provide summaries of 5-year average pollutant concentrations of controls by socioeconomic status, civil status, last reported place of residence, and place of birth in eTables 1-4, with eTable 1 below as an example:

eTable 1. Summary of 5-year average pollutant concentrations of controls by socioeconomic status (all in µg/m3).

Pollutant	Overall, N = 19,298[1]	Group 1 (Highest), N = 1,886[1]	Group 2, N = 2,340[1]	Group 3, N = 3,575[1]	Group 4, N = 5,522[1]	Group 5 (Lowest), N = 3,702[1]	Group 9 (Unknown), N = 2,273[1]
NOX	27 (20)	29 (20)	27 (20)	25 (17)	27 (20)	27 (19)	30 (23)
CO	237 (105)	244 (103)	233 (104)	225 (89)	237 (104)	237 (102)	258 (130)
EC	0.85 (0.42)	0.89 (0.42)	0.84 (0.42)	0.79 (0.37)	0.85 (0.42)	0.84 (0.41)	0.92 (0.48)
non-EC PM2.5	11.76 (2.37)	11.75 (2.21)	11.54 (2.24)	11.58 (2.30)	11.69 (2.34)	11.93 (2.43)	12.13 (2.62)
O3	52.0 (6.0)	51.1 (5.9)	52.0 (5.9)	53.0 (5.6)	51.9 (5.9)	52.2 (5.9)	50.7 (6.4)

Pollutant	Overall, N = 19,298[1]	Group 1 (Highest), N = 1,886[1]	Group 2, N = 2,340[1]	Group 3, N = 3,575[1]	Group 4, N = 5,522[1]	Group 5 (Lowest), N = 3,702[1]	Group 9 (Unknown), N = 2,273[1]

[1]Mean (SD)

12. Methods/Bayesian hierarchical approach: I am not particularly fluent in Bayesian methods, so will leave to others to evaluate the particulars of this approach. Nonetheless, if possible but without offering an entire course in Bayesian methods, it would be helpful for readers like me to see a clearer justification for choosing this method over, say, conventional conditional logistic regression (some of this clarification might tie into clarifying the target estimands mentioned in #1), and motivation or intuition for some of the specific steps and interpretation of posterior probability. For example, in writing, "We placed a hierarchy on the traffic-specific pollutant terms in the model," does this mean that you have assumed that one pollutant emanates from another? Also, does it make sense to compute credible intervals for the posterior probabilities?

The main advantage of the Bayesian hierarchical structure we have utilized in our analysis is that the variance-covariance structure of the traffic-related pollutants (EC, NO_x, CO) can be incorporated into the model, enabling an estimate of each individual pollutant's association with ALS diagnosis, as well as joint (i.e., percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO) and overall (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO) associations. The Reviewer is correct that this enables the model to account for the fact that the traffic-related pollutants originate from common sources. We have added this description in the revised manuscript (P. 10, Lines 209-211):

We placed a hierarchy on the traffic-specific pollutant terms in the model to account for the fact that the traffic-related pollutants, EC, NO_x, CO, originate from common sources and primarily traffic in urban environments:

The posterior probabilities are generated from the full posterior probability distributions of the marginal estimates for each pollutant-specific, overall, and joint association. To calculate the posterior probability that an effect estimate was greater than null, we took a large amount of draws from these full distributions (4,000 in our case) and estimated the proportion of samples which were above zero (null association). To clarify how to interpret this value, for which there is no credible interval, we have added a description to the revised manuscript (P. 11, Lines 237-241):

To calculate the probability that an association estimate was greater than null, we used the 4,000 samples of the posterior distribution and took the proportion of samples which were above the null. A 50% probability means that it is as likely as not that the marginal estimate is null, a probability closer to 100% indicates that the association is more likely to be truly positive, with closer to 0% indicating more likely to be truly negative.

MINOR

13. Abstract: "For a standard deviation (SD) increase in 5-year average…." For more context, please provide the SD for each pollutant.

We now provide the SD values for each pollutant in the Abstract of the revised manuscript (P. 1, Lines 15-19):

For a standard deviation (SD) increase in 5-year average concentrations, EC (SD=0.42μg/m^3) was potentially individually associated with an increase in odds (11.5%; 95% credible interval[CrI]:-1.0%,25.6%), with decreases individually for NO_x (SD=20μg/m^3) (-4.6%;95%CrI-18.1%,8.9%) and CO (SD=106μg/m^3) (-3.2%;95%CrI-14.4%,10.0%) and a null association of non-EC $PM_{2.5}$ (SD=2.37μg/m^3) (0.7%;95%CrI-9.2%,12.4%).

14. How PM2.5 was used in this investigation was presented with some ambiguity. The introduction states, "Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NOx), carbon monoxide (CO), and elemental carbon (EC)— as well as fine particles (PM2.5) and ozone (O3), we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis," suggesting that PM2.5 was being considered as a traffic-related air pollutant (similar to the aforementioned situation for ozone). PM2.5 is not necessarily traffic-related, as the authors later state, but here it appears to be one of the primary exposures of interest.

To avoid ambiguity, which the Reviewer has correctly pointed out, we have removed reference to $PM_{2.5}$ and ozone at this point in the Introduction in the revised manuscript (P. 4, Lines 68-72):

Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NO_x), carbon monoxide (CO), and elemental carbon (EC)— we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis, and estimate their (b) joint and (c) overall traffic-related emissions associations.

15. A curiosity: the relative difference in odds (percentage difference in odds) is effectively an arithmetic variation on the odds ratio. Was there a particular reason the authors opted for the percentage difference expression?

The Reviewer is correct, and we opted for percentage difference purely for wider interpretability, i.e., so a non-specialist epidemiologist might be able to read through the results and understand the numerical output.

16. Abstract: Given the results, it was a surprise to see this conclusion in the abstract "Our results indicate a potential positive association between ALS diagnosis and pollutants, particularly for EC." Perhaps this ties into clarifying the contribution of the Bayesian approach?

We attempted to make our Conclusions in the original Abstract indicative but not conclusive, which is the reason behind using the phrase 'potential association'. We have further clarified with the clear statement that our results are inconclusive in the revised manuscript (P. 2, Lines 24-26):

Conclusions: *A potential positive association between ALS diagnosis and pollutants, particularly for EC, though results are inconclusive. Further work is needed to understand the role of traffic-related air pollution on ALS pathogenesis.*

17. Methods/Index date: Please state earlier that the date of diagnosis as indicated in the database is the index date. I.e., "We identified ALS cases based on their International Classification of Diseases (ICD) discharge diagnoses, …, using the date of the first relevant code as the diagnosis date. This was the index date."

We have done this (P. 4, Lines 82-84):

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, using the date of the first relevant code as the diagnosis date.

18. Methods/matching scheme: what was the degree of match sought for age and year of birth (within months, years?).

The matching was made by age, sex, single year of birth, and vital status. We have clarified that matching by finer scale than year of birth was not possible in the revised manuscript (P. 8, Lines 169-170):

Matching by finer scale than year of birth was not possible.

19. Methods/study design: The control-sampling scheme seems to follow a risk-set matching pattern, so cases could serve as controls. If that is correct, could state that. It also means that computed ORs are estimates of IRs.

The Reviewer is correct. We have stated this in the revised manuscript (P. 5, Lines 97-99):

Controls were alive and free of diagnosed ALS at the ALS diagnosis date of the matched case (index date). The control-sampling scheme followed a risk-set matching pattern, so cases could have served as controls before diagnosis of ALS.[50]

We have also added clarification that (pg. 112) to be non-cases at the time of occurrences, odds ratios are IRs in the revised manuscript (P. 11, Lines 230-233):

We present all results as percentage change in odds of ALS diagnosis per SD increase in pollutant concentration (calculated via e.g., $e^{\beta_{NO_x}} - 1$*, etc. obtained in the modelling process). Due to the risk-set matching pattern of our case-control study, odds ratios are also equivalently incidence ratios (IRs).*[65]

20. Methods/occupational classes: these are likely official terms of the DK government, but they are not very descriptive and "unskilled" is somewhat derogatory. Although extensive detail is not needed, a little more would be informative.

We accept the Reviewer's point that the 'unskilled' by itself as a term is derogatory, so have replaced with 'unspecialized', and have provided some context in the revised manuscript (P. 7, Lines 147-150):

Group 1 (highest status) includes corporate managers and academics; group 2: proprietors, managers of small businesses and teachers; group 3: technicians and nurses; group 4: skilled

workers; and group 5: unspecialized workers, such as entry-level positions within food and retail environments.

21. Discussion: "If other sources of air pollution are associated with ALS, then including non-EC PM2.5 adjusts for other air pollutants from other sources." Is it known that air pollutants that fall outside of PM2.5 (most obviously, anything in the coarse fraction of PM10) are not related to ALS risk?

We have corrected this sentence to indicate that we are talking about other sources of $PM_{2.5}$ in the revised manuscript (P. 9, Lines 192-194):

If other sources of air pollution are associated with ALS, then including non-EC $PM_{2.5}$ adjusts for $PM_{2.5}$ from other sources,[67] as well as indicating whether pollution from other sources not explicitly quantified might also have associations with ALS.

22. This phrasing was unexpected: "The conditional approach automatically accounts for matching factors (age, sex, year of birth, vital status) …." What is meant by "automatically accounts for"?

We have clarified the language in the revised manuscript (P. 8, Lines 167-169):

The conditional approach examines contrasts within matched strata, i.e., groupings of case and matched controls, implicitly adjusting for matching factors (age, sex, year of birth, vital status) within each matched stratum.[65]

23. Discussion: Please take care to avoid relying on null hypothesis significance testing to interpret the findings. See the journal's guidance here: https://edmgr.ovid.com/epid/accounts/ifauth.htm/ In addition, the American Statistical Association issued a strong critique of significance testing (https://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf), and additional cogent arguments along these lines have been issued elsewhere, including Nature. (https://www.nature.com/articles/d41586-019-00857-9)

We agree wholeheartedly with the Reviewer that testing significance as the sole mechanism to deciding whether results are non-null should be avoided, as it focuses purely on the extreme part of the distribution and is prone to dismissing potentially important conclusions from results.

One of the main benefits of our Bayesian approach is that each marginal estimate's distribution probability mass is described fully, and therefore one is able to examine how much of the probability mass is above a null association, which in our view, and many statisticians' views, is a more 'natural' way of describing the confidence that a result is non-null.

We have avoided any reference to p-values, as they do not have a place in Bayesian analysis, and have now avoided any use of words related to the significance of a result in the revised manuscript, e.g., (P. 14, Lines 296-300):

In the largest case-control study of ALS and traffic-related air pollution to date, we found that a joint increase in average concentrations of traffic-related pollutants was potentially associated with an increase in odds of ALS diagnosis, with the clearest results for EC. We

found that EC had the largest-in-magnitude independent association with ALS diagnosis, while associations with NO_x and CO were negative with credible intervals overlapping the null, and smaller in magnitude.

* * * * *

Preparing a revision

1. For estimates of causal effects, we strongly discourage the use of categorized P-values and language referring to statistical significance, including whether a confidence interval covers the null. We prefer instead interval estimation, which conveys the precision of the estimate with respect to sampling variability. We are more open to testing with respect to modeling decisions, such as for tests of interaction and for tests for trend.

We have avoided p-values throughout.

2. We do not permit acronyms unless they are generally recognized by epidemiologists (e.g. HIV is okay, but LVA is not). When in doubt, we recommend that you spell out.

We have been careful to introduce acronyms where used.

3. Please do not include uninformative precision (excessive decimal places). For example, percents should be rounded to nn%, n.n%, or 0.0n% and risk ratios should be rounded to nn, n.n, or 0.nn unless clarity of the presentation and the sample size justify more significant digits.

We have done this.

4. Please be sure to include explicit information about approval of human subjects research by an independent review board. If no such review was required, include an explicit statement about why the requirement for review was waived.

We have done this in the manuscript (P. 5, Lines 107-108):

This study was approved by the Institutional Review Board Committee at the Columbia University and the Danish Data Protection Agency.

5. Do not include public health policy recommendations in Brief Reports or Original Articles that present new research findings.

We have not included any public health policy recommendations.

6. Data appearing in the abstract must also be cited in the main text, not just in tables or figures.

We have done this.

7. Resubmissions must adhere to word limits. The word limits for main text (generally the introduction, methods, results, and discussion) are 1500 words for Brief Reports (plus 150 words for its abstract), 4000 words for Original Articles (plus 250 words for its abstract), 5000 words for reviews (plus 250 words for its abstract), 2000 words for Commentaries (no abstract), 600 words for Research Letters (no abstract), and 400 words for Letters to the Editor (no abstract).

We have done this, with an Abstract of 250 words and an Original Article of 3,896 words in the revised manuscript.

8. We advise that total word counts for Original Articles should not exceed 7500 words and for Brief Reports should not exceed 3500 words. The total word count includes main text (introduction, methods, results, and discussion), bibliography, figure legends, tables, and figures (250 words per figure, including each figure in a panel). The title page, abstract, acknowledgments, and funding information do not count in the total word count.

We have adhered to this, with a total word count of 6,383 words in the revised manuscript.

9. Figure labels: Make font size as large as possible, so as to be legible when figures are reduced for publication (typically one column [8.5cm] in width).

We have made the Figure labels large and legible.

10. Footnotes to tables and figures should use superscript lowercase letters to link content to the footnote, not symbols or numerals.

The footnote in Table 1 uses a superscript lowercase letter.

11. Do not use parenthetical phrases like "(data not shown), (results not shown), or (available from the authors upon request)." In these circumstances, the data or results should be provided in Supplementary Digital Content.

We have avoided any use of these phrases.

12. Additional details regarding submission requirements can be found in the Instructions for Authors, which are posted at http://edmgr.ovid.com/epid/accounts/ifauth.htm .

Preparing for resubmission

13. Prepare a response document for the Editor that responds point-by-point to the reviewers' comments (presenting each comment followed by your response). Give the page number where revised text can be found and, where practical, paste revised text directly into the reply document.

We have done this.

14. Submit versions of the manuscript with and without your changes displayed.

We have submitted clean and tracked versions of the revised manuscript.

15. Supplementary Digital Content should be submitted as a single PDF file, and you should use our convention - e.g. eFigure 1, eAppendix 2 - to label and refer to online content.

We have done this.

16. Authors should submit copies of any closely related manuscripts (published, in press, or under review).

We do not have similar manuscripts to submit at this time.

17. Please revisit information about page charges and color printing charges available in the Instructions for Authors, which are posted at http://edmgr.ovid.com/epid/accounts/ifauth.htm.

We acknowledge the charges on the link provided.

18. We request that the complete revised manuscript (with all tables and figures) be completed by 05 May 2022. If you are not able to meet this deadline, please notify the editorial office.

We have submitted before 5th May 2022.

Resubmitting via Editorial Manager

19. Log-in to Editorial Manager as an author using the credentials above.
20. Click on the "Submissions Needing Revision" link.
21. To view the previous decision letter and reviewer comments, please click the blue decision term listed under the View Decision menu.
22. If you would like to download the previous manuscript to make revisions, click on "Download Files" under the Action menu.
23. To begin the resubmission: Click "Submit Revision" under the Action menu.
24. Proof each screen to ensure the information is still correct (the Title, Authors, etc.), then click Next at the bottom of each page.
25. On the Attach Files screen, select each previous submission item that you would like to carry forward to the resubmission.
26. Upload the revised versions of the main text (with and without tracked changes), and order them with the highlighted version first.
27. Upload the point-by-point reply to review.
28. When you are finished uploading, please click Next.
29. Click "Build PDF for My Approval."
30. Click "Go to Submissions Waiting for Author's Approval."
31. Wait for the PDF to build. When it has been built, you will see the link "View Submission" in the Action menu. Click "View Submission," and open the manuscript to proof your work.
32. If you find problems with the manuscript, click "Edit Submission" from the Action menu. Make the required changes, and begin again at the file uploads.
33. Once the submission is complete and acceptable, click "Approve Submission" from the Action menu.
34. If you have difficulty with these procedures, you may send questions to timothy.lash@epidemiology-journal.com.

Thank you for the resubmission instructions. We have followed them.

References

Brandt, J, JH Christensen, LM Frohn, and R Berkowicz. 2003. 'Air Pollution Forecasting from Regional to Urban Street Scale—-Implementation and Validation for Two Cities in Denmark'. *Physics and Chemistry of the Earth, Parts A/B/C* 28 (8): 335–44.

Brandt, Jørgen, Jesper H Christensen, Lise M Frohn, Finn Palmgren, Ruwim Berkowicz, and Zahari Zlatev. 2001. 'Operational Air Pollution Forecasts from European to Local Scale'. *Atmospheric Environment* 35: S91–98.

Bucher, Brian T, Jianlin Shi, Robert John Pettit, Jeffrey Ferraro, Wendy W Chapman, and Adi Gundlapalli. 2019. 'Determination of Marital Status of Patients from Structured and Unstructured Electronic Healthcare Data'. In *AMIA Annual Symposium Proceedings*, 2019:267. American Medical Informatics Association.

Dickerson, Aisha S, Johnni Hansen, Marianthi-Anna Kioumourtzoglou, Aaron J Specht, Ole Gredal, and Marc G Weisskopf. 2018. 'Study of Occupation and Amyotrophic Lateral Sclerosis in a Danish Cohort'. *Occupational and Environmental Medicine* 75 (9): 630–38. https://doi.org/10.1136/oemed-2018-105110.

Frohn, Lise Marie, Matthias Ketzel, Jesper Heile Christensen, Jørgen Brandt, Ulas Im, Andreas Massling, Christopher Andersen, et al. 2021. 'Modelling Ultrafine Particle Number Concentrations at Address Resolution in Denmark from 1979-2018–Part 1: Regional and Urban Scale Modelling and Evaluation'. *Atmospheric Environment* 264: 118631.

Galvin, Miriam, Rebecca Gaffney, Bernie Corr, Iain Mays, and Orla Hardiman. 2017. 'From First Symptoms to Diagnosis of Amyotrophic Lateral Sclerosis: Perspectives of an Irish Informal Caregiver Cohort—a Thematic Analysis'. *BMJ Open* 7 (3). https://doi.org/10.1136/bmjopen-2016-014985.

Khan, Jibran, Konstantinos Kakosimos, Ole Raaschou-Nielsen, Jørgen Brandt, Steen Solvang Jensen, Thomas Ellermann, and Matthias Ketzel. 2019. 'Development and Performance Evaluation of New AirGIS–a GIS Based Air Pollution and Human Exposure Modelling System'. *Atmospheric Environment* 198: 102–21.

Kioumourtzoglou, Marianthi-Anna, Ryan M Seals, Liselotte Himmerslev, Ole Gredal, Johnni Hansen, and Marc G Weisskopf. 2015. '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* 16 (3–4): 224–29.

Mathis, Stéphane, Cyril Goizet, Antoine Soulages, Jean-Michel Vallat, and Gwendal Le Masson. 2019. 'Genetics of Amyotrophic Lateral Sclerosis: A Review'. *Journal of the Neurological Sciences* 399: 217–26.

Norman, Rosana E, David O Carpenter, James Scott, Marie Noel Brune, and Peter D Sly. 2013. 'Environmental Exposures: An Underrecognized Contribution to Noncommunicable Diseases'. *Reviews on Environmental Health* 28 (1): 59–65.

Nuvolone, Daniela, Davide Petri, and Fabio Voller. 2018. 'The Effects of Ozone on Human Health'. *Environmental Science and Pollution Research* 25 (9): 8074–88.

Oskarsson, Björn, D Kevin Horton, and Hiroshi Mitsumoto. 2015. 'Potential Environmental Factors in Amyotrophic Lateral Sclerosis'. *Neurologic Clinics* 33 (4): 877–88.

Trabjerg, Betina B, Fleur C Garton, Wouter van Rheenen, Fang Fang, Robert D Henderson, Preben Bo Mortensen, Esben Agerbo, and Naomi R Wray. 2020. 'ALS in Danish Registries: Heritability and Links to Psychiatric and Cardiovascular Disorders'. *Neurology Genetics* 6 (2).

Abstract

Background: Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease. Limited evidence suggests ~~that~~ ALS ~~symptoms onset is~~diagnosis may be associated with air pollution exposure and specifically ~~to~~ traffic-related pollutants.

Methods: In this population-based case-control study, we used ~~data on~~ 3,939 ALS cases from the Danish National Patient Register diagnosed ~~between~~during 1989~~–~~–2013 and matched on age, sex, year of birth and vital status to 19,298 population-based controls free of ALS at index date. We used validated predictions of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) ~~from validated spatio-temporal models~~ to assign 1-, 5-, and 10-year average exposures pre-ALS diagnosis at study participants' present and historical residential addresses ~~of study participants.~~. We used a Bayesian hierarchical conditional logistic model ~~and adjusted~~, adjusting for potential confounders, to estimate ~~the~~ overall and joint ~~association~~associations for the ~~three~~ traffic-related pollutants (NO_x, CO, ~~and~~ EC), ~~as well as~~and pollutant-specific associations.

Results: For a standard deviation (SD) increase in 5-year average concentrations, EC ~~was~~(SD=0.42µg/m^3) was potentially individually associated with an increase in odds (11.5%; 95% credible interval-[CrI~~]:~~–]:-1.0%,-25.6%), with decreases individually for NO_x (SD=20µg/m^3) (-4.6%;95%CrI-–18.1%,8.9%) and CO (SD=106µg/m^3) (-3.2%;-95%CrI-–14.4%,-10.0%) and a null ~~effect~~association of non-EC $PM_{2.5}$ (SD=2.37µg/m^3) (0.7%;95%CrI-–9.2%,12.4%). We found no clear association for joint or overall traffic pollution. There was a 77.8% posterior probability of a

Formatted: Justified

Formatted: Numbering: Continuous

Formatted: Font color: Auto

positive association between the joint effect of pollutants and ALS diagnosis, 96.3% for EC, 27.8% for NO_x and 26.7% for CO.

Conclusions: ~~Our results indicate a~~A potential positive association between ALS diagnosis and pollutants, particularly for EC~~.~~, though results are inconclusive. Further work is needed to understand the role of traffic-related air pollution on ALS pathogenesis ~~and timing of onset~~.

Abbreviations:

ALS	Amyotrophic lateral sclerosis
~~BMI~~	~~Body mass index~~
BKMR	Bayesian kernel machine regression
CO	Carbon monoxide
CrI	Credible interval
DEHM-UBM-AirGIS	Spatio-temporal air pollution modelling system used in study
EC	Elemental carbon
ICD	International Classification of Diseases
IQR	Interquartile range
IR	Incidence ratio
Non-EC $PM_{2.5}$	Non-elemental carbon fine particles
NO_x	Nitrogen oxides
O_3	Ozone
$PM_{2.5}$	Fine particles
SD	Standard deviation

Formatted: Font color: Text 1

46	SES	Socioeconomic status

Formatted: Font: Not Bold, Font color: Text 1

Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating and fatal neurodegenerative disease,[1] currently without a cure.[2] Approximately half of patients die within three years of symptom onset.[3] Annually, there are nearly 30,000 cases of ALS in Europe and over 200,000 worldwide.[4] Known inherited genetic variants only account for 5–10% of ALS cases.[5,6] Environmental factors, therefore, are likely important in ALS pathogenesis.[7] However, because the disease is relatively rare, it is challenging to conduct large scale prospective studies. There is a recognized need for more evidence of the environmental contributors of ALS.[5,8]

Although air pollution is commonly studied in association with respiratory- and cardiovascular-related outcomes,[9–14] epidemiologic and toxicological studies support several plausible biological mechanisms in association with the nervous system and neurodegeneration.[15–34] Ambient air pollution, especially urban air pollution, is a ubiquitous exposure that has been associated with several other neurodegenerative disorders,[16–21,35,36] and is consistently linked to systemic inflammation,[22–24] oxidative stress,[25–28] and neuroinflammation,[15,29] all of which, in turn, have been reported as key pathways to ALS pathogenesis.[30–34]

Despite the compelling plausibility, few studies to date have evaluated the association between air pollution and ALS.[35,37–39] A recent study found that traffic-related air pollutants may be driving observed associations.[38] No study has hitherto attempted to understand the combined and individual associations of the pollutants in a single model. Air pollutants have been consistently associated with adverse health, primarily in single pollutant analyses.[13,17,40–42] However, they are highly correlated with one another.[40] It is therefore a mixture modelling challenge to infer the

Formatted: Justified

~~association of multiple air pollutants and health outcomes.[43] Using three air pollutants commonly used in health studies as traffic-related emissions tracers – nitrogen oxides (NO_x), carbon monoxide (CO), and elemental carbon (EC) – as well as fine particles ($PM_{2.5}$) and ozone (O_3), we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis, and estimate their (b) joint and (c) overall traffic-related emissions associations.~~[1] currently without a cure.[2] Approximately half of patients die within three years of symptom onset.[3] Annually, there are nearly 30,000 cases of ALS in Europe and over 200,000 worldwide.[4] Known inherited genetic variants only account for 5–10% of ALS cases.[5,6] Environmental factors, therefore, are likely important in ALS pathogenesis.[7] However, because the disease is relatively rare, it is challenging to conduct large-scale prospective studies. There is a recognized need for more evidence of the environmental contributors of ALS.[5,8]

Although air pollution is commonly studied in association with respiratory- and cardiovascular-related outcomes, e.g., references [9–14], epidemiologic and toxicological studies also support several plausible biological mechanisms in association with the nervous system and neurodegeneration, e.g., references [15–34]. Ambient air pollution, especially urban air pollution, is a ubiquitous exposure that has been associated with several other neurodegenerative disorders, e.g., references [16–21,35,36]. and is consistently linked to systemic inflammation,[22–24] oxidative stress,[25–28] and neuroinflammation,[15,29] all of which, in turn, have been reported as key pathways to ALS pathogenesis, e.g., references [30–34].

Despite the compelling plausibility, few studies to date have evaluated the association between air pollution and ALS.[35,37–39] A study in 2021 found that traffic-related air pollutants may be driving

Formatted: Font: Not Bold, Italic

observed associations.[38] No study has hitherto attempted to understand the combined and individual associations of the pollutants in a single model. Air pollutants have been consistently associated with adverse health, primarily in single pollutant analyses.[13,17,40–42] However, they are highly correlated with one another.[40] It is therefore a mixture modelling challenge to infer the association of multiple air pollutants and health outcomes.[43] Using three air pollutants commonly used in health studies as traffic-related emissions tracers—nitrogen oxides (NO_x), carbon monoxide (CO), and elemental carbon (EC)— we aimed to assess whether exposure to (a) each individual air pollutant is independently associated with ALS diagnosis, and estimate their (b) joint and (c) overall traffic-related emissions associations.

Formatted: Justified

Methods

Study Population and Outcome Assessment

We used data from the Danish National Patient Register during 1989-2013, through which details on demographic characteristics and certain health outcomes of all Danish residents can be linked based on a unique personal identifier.[44~~44~~44] The Register was established in 1977 and is comprehensive, including nationwide clinical and administrative records for all inpatient data, with outpatient data available since 1995.[45~~45~~45]

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, using the date of the first relevant code as the diagnosis date. ~~We only included patients who were at least 20 years old when diagnosed. In our validation study, Register data for ALS ascertainment were highly reliable.[46]~~This was the index date. We only included patients who were

at least 20 years old when diagnosed because (i) cases younger than 20 years old were at a greater chance of misclassification, since ALS has been predominantly diagnosed in older adults,[46] and (ii) the very few juvenile ALS cases have been explained to a much larger degree by genetic mutations (~40%).[47] In our validation study, Register data for ALS ascertainment were highly reliable; working with a specialist ALS neurologist to review medical records and comparing to death certificates and hospital discharges, the Danish National Patient Register was found to have an overall predictive value for ALS of 82%.[48]

We obtained controls through the Danish Civil Registration System, established in 1968 and updated daily, which includes administrative records (e.g., date and place of birth, sex, vital status, and history of civil status and addresses since 1971) on all persons living in Denmark; records are kept even when a person dies or emigrates.[47][49] We randomly matched five controls per case by age, sex, year of birth, and vital status. Controls were alive and free of diagnosed ALS at the ALS diagnosis date of the matched case (index date). The control-sampling scheme followed a risk-set matching pattern, so cases could have served as controls before diagnosis of ALS.[50]

We obtained all addresses of cases and controls from January 1st 1979 onwards from the Danish Civil Registration System,[47] ~~including the dates of moving to and from each address, prior to the index date.~~[49] including the dates of moving to and from each address, before the index date. We then obtained the geographical coordinates at the door of each house of the residential history of the participants, with previous evidence of the high accuracy of this method of geocoding of addresses in Denmark.[17][17]

Formatted: Font color: Auto

Formatted: Font: Italic, Font color: Auto

This study was approved by the Institutional Review Board Committee at the Columbia University and the Danish Data Protection Agency.

Exposure data

~~We obtained predictions on monthly concentrations of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) (as well as ozone (O_3) for a sensitivity analysis, usually negatively correlated with other pollutants due to its chemistry[48]), at residential addresses of study participants from the validated spatio-temporal air pollution modelling system (DEHM-UBM-AirGIS) with full space and time coverage over our study period, described in detail elsewhere.[49–52] These predicted pollutant concentrations have been extensively used in previous air pollution epidemiologic studies in Denmark.[17,53–55] The models have good predictive accuracy, with average monthly correlations between measured and modelled results of 0.85 for NO_x, 0.91 for CO, 0.92 for O_3, 0.79 for EC, and 0.83 for annual concentrations of $PM_{2.5}$.[49,52] Because traffic is a major source of $PM_{2.5}$ and EC one of the main $PM_{2.5}$ components in urban environments,[56] we removed the EC concentration from the total $PM_{2.5}$ mass concentration (non-EC $PM_{2.5}$), to avoid overadjustment when including both in the models simultaneously.~~

We obtained predictions on monthly concentrations of nitrogen oxides (NO_x), carbon monoxide (CO), elemental carbon (EC), and fine particles ($PM_{2.5}$) (as well as ozone, O_3, for a sensitivity analysis, usually negatively correlated with other pollutants due to its chemistry[51]), at residential addresses of study participants from the validated spatio-temporal air pollution modelling system DEHM-UBM-AirGIS that provides full space and time coverage over the study period, described in detail elsewhere.[52–55] In brief, DEHM-UBM-AirGIS is a human exposure modelling system for traffic pollution, developed for application in Danish air pollution epidemiological studies. The

modelling system is able to generate street configuration and traffic data based on digital maps and national databases, which enables estimation of air quality levels at a large number of addresses in an automatic and effective way. These predicted pollutant concentrations have been extensively used in previous air pollution epidemiologic studies in Denmark.[17,56–58] The models have good predictive accuracy, with average monthly correlations between measured and modelled results of 0.85 for NO_x, 0.91 for CO, 0.92 for O_3, 0.79 for EC, and 0.83 for annual concentrations of $PM_{2.5}$.[52,55] Because traffic is a major source of $PM_{2.5}$ and EC one of the main $PM_{2.5}$ components in urban environments,[59] we removed the EC concentration from the total $PM_{2.5}$ mass concentration (non-EC $PM_{2.5}$) by subtraction, to avoid overadjustment when including both in the models simultaneously.

Based on the residential history of each case or control, we calculated 1-, 5-, and 10-year average exposure to each pollutant ending at one year before the index date, as diagnosis has been shown previously to occur at a median of 12 months after symptoms onset.[57~~60~~] Specifically, each case or control average value (1-, 5- or 10-year) was calculated as the mean of all concentrations recorded across time at the recorded addresses within each time window. A small number of Danish residents lack a complete address history (1.7%; lack of house number). To ensure we were including participants with adequately complete exposure records, we set the following minimum criteria for number of complete exposure record months to include cases and controls: (i) 1-year averages: 9 of 12 months, at least one measurement in each season; (ii) 5-year averages (main exposure): 30 of 60 months; and (iii) 10-year averages: 60 of 120 months.

Covariate data

Formatted: Justified

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Italic, Font color: Auto

We included a set of covariates to account for potential confounding bias. We used a five-category individual-level socioeconomic status (SES) definition developed by the Danish Institute of Social Sciences, based on job titles from income tax forms, which we have shown as having an association with ALS diagnosis in Denmark,[58]. Group 1 (highest status) includes corporate managers and academics; group 2: proprietors, managers of small businesses and teachers; group 3: technicians and nurses; group 4: skilled workers; and group 5: unskilled workers. We included a group for participants who were unemployed or unclassified (group 9). For each married participant, we used the higher of the couple's individual SES categories, where available. We also used information on civil status (never married, married, divorced, widowed), last reported place of residence from postcode (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 potential family-specific, location-specific, and early-life confounders. As part of a sensitivity analysis, we also included parish-level SES, measured by percentage of residents with greater than high-school education, in the model. In Denmark, parishes are administrative units with an average population of ~2,500 residents.

Statistical analysis

We analyzed the association between ALS diagnosis (binary) and exposure to traffic-related pollutants by applying a Bayesian formulation of the conditional logistic model, with Bayesian hierarchy on the traffic-related pollutants (EC, NO_x, CO).[59,60] The conditional approach automatically accounts for matching factors (age, sex, year of birth, vital status) between cases and controls within each matched stratum, i.e., groupings of case and matched controls.[59]

~~Bayesian inference allows for full distributional estimation of parameters of interest.[60] We employed a Bayesian hierarchical formulation because it enables estimates of (a) independent pollutant-outcome associations, (b) a joint association of the three pollutants (i.e., percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO), and (c) an overall average traffic association (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO), while accounting for the variance-covariance structure between the highly-correlated exposures and their coefficients.[60] We included a linear term for each included pollutant and adjusted for individual- and parish-level SES, civil status, last reported place of residence, and place of birth.~~

We included a set of covariates based on as close as possible to index date to account for potential confounding bias, including household socioeconomic status (SES) based on last-reported job title at index date; civil status at index date, last reported place of residence at index date, and place of birth. We used a five-category individual-level SES definition developed by the Danish Institute of Social Sciences, based on job titles from income tax forms, which has been associated with ALS diagnosis in Denmark,[61] as well as how quickly one is identified as having ALS in the Danish Civil Registration System.[62] Group 1 (highest status) includes corporate managers and academics; group 2: proprietors, managers of small businesses and teachers; group 3: technicians and nurses; group 4: skilled workers; and group 5: unspecialized workers, such as entry-level positions within food and retail environments. We also included a group for participants whose job title was unknown (group 9). For each married participant, we used the higher of the couple's individual SES categories, when available. We also used information on civil status (never married, married, divorced, widowed) due to the influence that a spouse may have on visiting a family physician,[63] last reported place of residence from postcode (Greater Copenhagen, big cities of Denmark, rest

of Denmark, Greenland) to account for various local environmental and behavioral stressors,[7] and place of birth (Greater Copenhagen, big cities of Denmark, rest of Denmark, Greenland, foreign, unknown) to adjust for other potential family-specific, location-specific, and early-life confounders, which may have an impact on the probability of developing ALS.[64] Ultimately, we were limited by what was available in the Danish Civil Registration System.[62] As part of a sensitivity analysis, we also included parish-level SES, measured by percentage of residents with greater than high-school education, in the model. In Denmark, parishes are administrative units with an average population of ~2,500 residents.

Statistical analysis

We analyzed the association between ALS diagnosis (binary) and exposure to traffic-related pollutants by applying a Bayesian formulation of the conditional logistic model, with Bayesian hierarchy on the traffic-related pollutants (EC, NO_x, CO).[65,66] The conditional approach examines contrasts within matched strata, i.e., groupings of case and matched controls, implicitly adjusting for matching factors (age, sex, year of birth, vital status) within each matched stratum.[65] Matching by finer scale than year of birth was not possible. Bayesian inference allows for full distributional estimation of parameters of interest.[66] We employed a Bayesian hierarchical formulation because it enables estimates of (a) independent pollutant-outcome associations, (b) a joint association of the three pollutants (i.e., total percentage change in odds of ALS diagnosis with increase in each of EC, NO_x, CO), and (c) an overall average traffic association (i.e., average percentage change in odds of ALS diagnosis from each of EC, NO_x, CO), while accounting for the variance-covariance structure between the highly-correlated exposures and their coefficients.[66] We included a linear

term for each included pollutant and adjusted for individual- and parish-level SES, civil status, last reported place of residence, and place of birth.

Specifically, via a logit function, we modelled the log-odds of ALS diagnosis, as follows:

$$logit[\Pr(Y_{ci} = 1)] = \alpha_c +$$
$$\beta_{NO_x} NO_{x\,ci} + \beta_{CO} CO_{ci} + \beta_{EC} EC_{ci} +$$
$$\beta_{PM_{2.5}} (\text{non-EC } PM_{2.5\,ci}) +$$
$$\beta_{SES} SES_{ci} + \beta_{\text{Civil Status}} \text{Civil Status}_{ci} + \beta_{\text{Residence}} \text{Residence}_{ci} + \beta_{\text{Birth}} \text{Birth}_{ci},$$

where Y_{ci} denotes whether subject i in matched stratum c was diagnosed with ALS, i.e., c represents a case and its matched controls; α_c the matched stratum-specific intercepts (not estimated in conditional logistic models); $\beta_{NO_x}, \beta_{CO}, \beta_{EC}, \beta_{PM_{2.5}}$ the pollutant-specific coefficients (log-odds) per standard deviation (SD) increase in concentration of ~~NO_x~~ NO_x, CO, EC, non-EC $PM_{2.5}$ respectively, scaled by their respective ~~standard deviations~~SDs and centered at their means~~;~~, with each β a pollutant-specific association adjusted by other terms in the model and the rest as coefficients for subject-specific covariates. Interquartile Range (IQR) could equivalently be used to scale pollutant concentrations. If other sources of air pollution are associated with ALS, then including non-EC $PM_{2.5}$ adjusts for ~~other air pollutants from other sources.~~[61] $PM_{2.5}$ from other sources,[67] as well as indicating whether pollution from other sources not explicitly quantified might also have associations with ALS. Therefore, $\beta_{PM_{2.5}}$ is interpreted as the association with air pollutants not specifically included in our analysis. In urban European environments, traffic-related pollutants typically represent on-average 14% of $PM_{2.5}$ concentrations.[62] ~~In a sensitivity analysis, we included β_{O_3} to account for O_3 exposures in the model~~[68] In a sensitivity analysis, we

Formatted: Justified

Formatted: Justified

Formatted: Font: Times New Roman, Not Italic

Formatted: Font: Times New Roman

Formatted: Font: Italic

included O_3 in the model, as O_3 concentrations have been associated with many adverse health outcomes,[69] and were negatively correlated with traffic-related pollutants, and added $ns(\text{SES}_{\text{parish}_{ci}})$, as a natural spline with three degrees of freedom.

In our model, β_{NO_x}, β_{CO}, and β_{EC} represent the independent pollutant-specific associations with ALS diagnosis. In the same model, we estimated the joint association between these three pollutants and ALS diagnosis as:

$$Traffic_{Joint} = \sum_{p=NO_x,CO,EC} \beta_p p.$$

This sum quantifies the association (log-odds) with ALS of a one-SD increase in the three pollutants simultaneously.

Finally, we assumed that the traffic-related pollutant-specific associations arise from a distribution of the overall traffic association with ALS diagnosis. We placed a hierarchy on the traffic-specific pollutant terms in the model to account for the fact that the traffic-related pollutants, EC, NO_x, CO, originate from common sources and primarily traffic in urban environments:

$$\beta_{Traffic} = [\beta_{\text{NO}_x}, \beta_{CO}, \beta_{EC}],$$

$$\beta_{Traffic} \sim MVN(\mu, \Sigma),$$

$$\mu \sim N(\lambda, \sigma_\lambda),$$

$$\Sigma = \tau\Omega\tau,$$

where λ denotes the overall average one-SD association of traffic-related pollution with variance σ_λ. Σ, the estimated variance-covariance matrix among pollutant-specific estimates, was expressed as a decomposition into a positive-definite correlation matrix Ω and scale matrix τ.[63~~70~~]

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Justified

Field Code Changed

We used weakly-informative priors so that data drove parameter estimation. Hyper-priors for coefficients on non-EC $PM_{2.5}$ and covariates were N(0,10); for ~~σ_λ~~σ_λ and τ we used Half-Cauchy(0,10), as recommended by Gelman, Polson and Scott as a weakly-informative prior;[~~64,65~~] ~~and Ω was defined by~~[~~71,72~~] Ω was defined by the weakly-informative prior LKJCorr(1).[~~6673~~] The exception to this was the prior for λ, the average association of traffic-related pollutants, for which estimates became unrealistically high (approaching infinity and not converging with further iterations) with a ~~non~~more weakly-informative prior. We therefore used a prior of N(0,0.1), which did not affect estimates of other parameters. We conducted sensitivity analyses to understand the influence of priors and the robustness of the results.

We present all results as percentage change in odds of ALS diagnosis per ~~standard deviation (SD)~~SD increase in pollutant concentration (calculated via e.g., $e^{\beta_{NO_x}} - 1$, etc. obtained in the modelling process). Due to the risk-set matching pattern of our case-control study, odds ratios are also equivalently incidence ratios (IRs).[65] We ran each model with four chains with a sample size of 1,000 each, after a warm-up of 1,000 samples, for 4,000 total samples. We assessed whether the models converged by checking that the Gelman-Rubin potential scale reduction statistic[~~6774~~] was below 1.1 for all estimated model parameters. The reported 95% credible intervals (CrI) are the 2.5th to 97.5th percentiles of each parameter's posterior marginal distribution. To calculate the probability that an association estimate was greater than null, we used the 4,000 samples of the posterior distribution and took the proportion of samples which were above ~~a~~the null. A 50% probability means that it is as likely as not that the marginal estimate is null, a probability closer

Formatted: Font: Italic

to 100% indicates that the association. is more likely to be truly positive, with closer to 0% indicating more likely to be truly negative.

~~We conducted statistical analyses using the R Statistical Software, version 4.1.1[68] and R-STAN, version 2.21.2.[60] All code for analysis, results from analysis, and visualization presented in this manuscript will be publicly available via GitHub at https://github.com/rmp15/multipollutants_and_als_code_review.~~

We conducted statistical analyses using the R Statistical Software, version 4.1.1[75] and R-STAN, version 2.21.2.[66] All code for analysis, results from analysis, and visualization presented in this manuscript is publicly available via GitHub at https://github.com/rmp15/traffic_air_pollution_als_denmark_epidemiology.

Formatted: Justified

We assessed the sensitivity of our results to hyper-prior assignment; running more iterations and warm-up per chain; inclusion of O_3; single traffic-related pollutant models adjusting for non-EC $PM_{2.5}$; as well as adjusting by parish-level SES. From the parish-level SES sensitivity analysis we excluded those who lived in areas without parish-level SES data, namely: (i) 819 participants for 1-year average exposure; (ii) 826 participants for 5-year average exposure; and (iii) 838 participants for 10-year average exposure.

Results

After filtering the original 4,011 cases and 20,055 controls based on completeness of exposure records, we used information on 3,934 (98.1% of total) cases and 19,298 (96.2% of total) controls for 5-year average exposure. We also used 3,937 cases,19,333 controls for 1-year average exposure

and 3,939 cases, 19,250 controls for 10-year average exposure. Descriptive statistics of included cases and controls for 5-year average exposure can be found in Table 1. Descriptive statistics of controls for 5-year exposure by socioeconomic status, civil status, residence, and place of birth are found in eTables 1-4. For the main results, we present 5-year average exposure associations as a balance between representation of most recent exposure as well as long-term concentration.

The 5-year average traffic-related pollutant concentrations were 27 µg/m^3 for NO_x (SD=20 µg/m^3), 238 µg/m^3 for CO (SD=106 µg/m^3) and 0.85 µg/m^3 for EC (SD=0.42 µg/m^3) (Table 2). Figure 1 shows Spearman correlations between pollutants for ~~5~~1-, 5-, and 10-year average exposures. Traffic-related pollutants (NO_x, CO, EC) were highly correlated in cases, controls and overall, ranging from correlations of 0.91 to 0.96. Otherwise, non-EC $PM_{2.5}$ was most highly correlated with CO~~.~~, ranging from 0.67 to 0.7. O_3 was negatively correlated with other pollutants, ranging from -0.54 to -0.89.

For 5-year average pollutant concentrations, we observed the largest overall association for the individual ~~standard deviation~~SD increase in EC (11.5%; 95% CrI: -1.0%, 25.6~~%;~~% per 0.42 µg/m^3; 96.3% posterior probability of positive association ~~per 0.42 µg/m^3~~) (Figure 2). ~~Standard deviation~~SD increases were associated with a decrease in odds of ALS diagnosis in NO_x (-4.6%; 95% CrI: -18.1%, 8.9% per 20 µg/m^3; 27.8% posterior probability of positive association) and CO (-3.2%; 95% CrI: -14.4%, 10.0% per 106 µg/m^3~~).~~; 26.7% posterior probability of positive association). The joint association of traffic-related pollutants (EC, NO_x, CO) was 2.3% (95% CrI: -3.3%, 7.7%), with an 77.8% posterior probability of a positive association. ~~Finally, the~~The average overall traffic association was null (-0.1%; 95% CrI: -17.4%, 20.8%). Non-EC $PM_{2.5}$ was

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Italic, Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto, English (United States)

Formatted: Font color: Auto, English (United States)

Formatted: Font color: Auto, English (United States)

Formatted: Font: Italic, Font color: Auto, English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

not associated with ALS diagnosis (0.7%; 95% CrI: -9.2%, 12.4%). 1-year EC average exposure was associated with ~~a significant~~an increase in odds of ALS diagnosis (15.4%; 95% CrI: 1.6%, 25.6%) (Figure 2). Compared to the 1- and 5-year results, the 10-year average exposure results were attenuated ~~versions of~~, as associations tended further to the ~~1- and 5-year results~~null. Single-pollutant models for each traffic-related pollutant adjusting for non-EC $PM_{2.5}$ (eFigure 1; single traffic-related pollutant models D, E and F) resulted in positive associations for each of EC, NO_x, CO, with positive associations for non-EC $PM_{2.5}$ in all but the model with EC. The 95% credible interval for EC in the single-pollutant model (eFigure 1; model D) overlapped with the credible intervals of the EC term in the multi-pollutant models (eFigure 1; models B, C, G to P). Results from variations of the main model in the sensitivity analyses were robust to prior choices, inclusion of O_3, and inclusion of parish-level SES (eFigure 1). A map of average concentration of included pollutants (NO_x, EC, $PM_{2.5}$, CO, O_3) across Denmark for a representative year (2000; middle of study period 1989-2013) is also available in eFigure 2.

Formatted: Font color: Auto, English (United States)

Formatted: Font color: Auto, English (United States)

Formatted: Font color: Auto, English (United States)

Formatted: Font: Italic, Font color: Auto, English (United States)

Formatted: Font: Italic, Font color: Auto, English (United States)

Discussion

~~In the largest case-control study of ALS and traffic-related air pollution to date, we found that an increase in average concentrations of traffic-related pollutants was associated with an increase in odds of ALS diagnosis, though not significant at the 95% credible interval level, apart from EC for 1-year average SD increase. We found that EC had the largest in magnitude independent association with ALS diagnosis, while the non-significant associations with NOx and CO were negative and smaller in magnitude.~~

Our results indicate that traffic-related pollutants, hazardous in many ways,[9–21,40–42] may also be associated with ALS diagnosis. Our finding – that increases in EC, are potentially positively associated with ALS diagnosis – is plausible. A recent case-control study in the Netherlands reported that ultrafine particles, another traffic emissions-related surrogate, were associated with ALS diagnosis,[38] while another based in Catalonia, Spain found ALS cases clustered around key road infrastructure.[69] Although we did not find an association with non-EC $PM_{2.5}$ in our study, our results are not directly comparable to those of the other studies, as our $PM_{2.5}$ effect estimates capture the $PM_{2.5}$ components not accounted for by other pollutants in the analysis.

Our results indicate that EC exposure – a large part of which comes from diesel combustion and small combustion sources (such as wood stoves) in European urban centers, where prevalence of diesel cars is high[70] – has a high probability of a positive association with ALS diagnosis. In our previous study of ALS and occupational exposures in Denmark we found that those working in agriculture and construction, associated with exposure to diesel engine exhausts, were at higher relative risk than those in other employments.[58] Truck drivers, for whom diesel exposure is common, are also at increased risk of sporadic ALS.[71] EC exposure has been associated with inflammation,[72] mitochondrial dysfunction[73] and DNA damage,[73,74] all of which are plausible pathways of neurodegeneration. These factors have also previously been identified as particular pathways to pathogenesis of ALS.[30–34]

We did not find a high probability of a positive association with NO_x in our analyses, in contrast with a previous study, though that study did not include EC.[38] NO_x is also highly correlated with EC (0.95 to 0.96 in our study), which is expected given that they are both combustion products

~~commonly associated with emissions in urban environments. EC exposure was more strongly associated with 1-year than for 5-/10-year average concentrations, which may indicate that the previous year may be the most relevant exposure window. We do not expect that these results are attributed to reverse causation, as we have lagged these 1-year exposures by one-year already prior to diagnosis, and there was likely little substantial residential movement in the year before ALS diagnosis.[75]~~

In the largest case-control study of ALS and traffic-related air pollution to date, we found that a joint increase in average concentrations of traffic-related pollutants was potentially associated with an increase in odds of ALS diagnosis, with the clearest results for EC. We found that EC had the largest-in-magnitude independent association with ALS diagnosis, while associations with NO_x and CO were negative with credible intervals overlapping the null, and smaller in magnitude. Sensitivity analyses demonstrated that for single pollutant models, the association for EC was smaller than for our main multi-pollutant model, which took into account the variance-covariance structure of traffic-related pollutants. Overall conclusions for the association between EC and ALS diagnosis would have been similar from the single- or multi-pollutant models. The inconsistent associations for NO_x and CO in the multi- and single-pollutant models suggest that the model may have had limited success identifying each individual pollutant's association with ALS diagnosis due to the high level of collinearity of traffic-related pollutants. Nevertheless, the consistency of the sign of the central estimate of EC in all models suggests that EC may be a driver of the ALS and traffic-related pollutant association, though further analysis is required. Our results indicate that traffic-related pollutants, hazardous in many ways,[9–21,40–42] may also be associated with ALS diagnosis. Our finding—that increases in EC, are potentially positively associated with ALS diagnosis—is plausible. A recent case-control study in the Netherlands reported that ultrafine

particles, another traffic emissions-related surrogate, were associated with ALS diagnosis,[38] while another based in Catalonia, Spain found ALS cases clustered around key road infrastructure.[76] Although we did not find an association with non-EC $PM_{2.5}$ in our study, our results are not directly comparable to those of the other studies, as our $PM_{2.5}$ effect estimates capture the $PM_{2.5}$ components not accounted for by other pollutants in the analysis.

Our results indicate that EC exposure—a large part of which comes from diesel combustion and small combustion sources (such as wood stoves) in European urban centers, where prevalence of diesel cars is high[77]—has a high probability of a positive association with ALS diagnosis. In our previous study of ALS and occupational exposures in Denmark we found that those working in agriculture and construction, associated with exposure to diesel engine exhausts, were at higher relative risk than those in other employments.[61] Truck drivers, for whom diesel exposure is common, are also at increased risk of sporadic ALS.[78] EC exposure has been associated with inflammation,[79] mitochondrial dysfunction[80] and DNA damage,[80,81] all of which are plausible pathways of neurodegeneration. These factors have also previously been identified as particular pathways to pathogenesis of ALS.[30–34]

We did not find a high probability of a positive association with NO_x in our analyses, in contrast with a previous study, though that study did not include EC.[38] NO_x is also highly correlated with EC (0.95 to 0.96 in our study), which is expected given that they are both combustion products commonly associated with emissions in urban environments. EC exposure was more strongly associated with 1-year than for 5-/10-year average concentrations, which may indicate that the previous year of exposure may be the most relevant exposure window relevant to traffic-related

exposures and ALS; this is biologically plausible, as this critical exposure window would be at the pre-symptomatic stage of underlying ALS progression, where traffic-related pollution exposure may add to the ongoing cellular or molecular process of the disease, to the point where the body can no longer compensate and subsequently enters the clinical phase.[82–84] We do not expect that these results are attributed to reverse causation, as we have lagged these 1-year exposures by one year already prior to diagnosis, and there was likely little substantial residential movement in the year before ALS diagnosis.[85] We do not expect that calendar time was a potential source of confounding, as the controls were matched on age and year of birth. The null joint association, combined with the largest associations from traffic-related pollutant in all models found with EC, further indicates that EC may be driving the association of air pollution with ALS, though further analysis will be necessary to confirm this.

Formatted: Justified

Our study used one the largest number of ALS patients ever included in an environmental health study. Another strength of our study is that we leveraged highly correlated traffic pollutants and Bayesian hierarchical modeling and were able to estimate independent and joint traffic-related pollutant associations, as well as an overall traffic estimate. Although we have adjusted implicitly (by matching; age, sex, year of birth, vital status) and explicitly for many common covariates (SES, civil status, residence, place of birth), we cannot rule out residual confounding ~~(e.g., from smoking or body mass index (BMI)). However, to induce confounding bias, any unaccounted for variable would have to influence both ALS diagnosis and air pollution. BMI, previously associated with ALS,[76,77] would not confound the association between traffic-related air pollution and ALS,[75] as pollutant concentrations are derived independently from BMI distribution. Any BMI-air pollution association in our study, thus, would be via SES, for which we adjusted at both the individual and~~

~~parish level.~~. Exposure measurement error is inevitable, as any modelled exposure will be inaccurate to some degree. However, any error is not likely correlated with ALS diagnosis, and therefore any bias would be towards null.[~~7886~~] While a previous study found that ALS ascertainment from the Danish National Patient Register was highly reliable,[48] outcome misclassification cannot be ruled out, nor can the possibility that date of diagnosis and symptom onset were irregularly aligned. While our analysis adjusted for marital status and household SES, many couples in Denmark cohabitate. This would not be captured by our analysis, and ALS diagnosis in relation to cohabitation status should be further investigated.[87]

Future research might use larger cohort data to understand the importance of each respective pollutant in a single model. Other mixture model methods, such as Bayesian Kernel Machine Regression (BKMR)[88] might be useful in further exploring the robustness of joint associations in a different framework, though BKMR was not appropriate for our particular research question, since BKMR is not currently appropriate for case-control studies. The timing of exposure will also be an important study route. ALS is projected to increase in prevalence over the next few decades all over the world.[44] Understanding ALS pathogenesis and identifying modifiable risk factors is critical for preventive action.

Formatted: Font: Italic, Font color: Auto

Formatted: Font: Not Bold, Italic

References

1. Rowland LP, Shneider NA. Amyotrophic lateral sclerosis. *New England Journal of Medicine*. 2001;344(22):1688-1700.

2. Chio A, Logroscino G, Hardiman O, et al. Prognostic factors in ALS: A critical review. *Amyotrophic Lateral Sclerosis*. 2009;10(5-6):310-323.

3. Mitchell JD, Borasio GD. Amyotrophic lateral sclerosis. *The Lancet*. 2007;369(9578):2031-2041.

4. Arthur KC, Calvo A, Price TR, Geiger JT, Chio A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nature Communications*. 2016;7(1):1-6.

5. Al-Chalabi A, Hardiman O. The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology*. 2013;9(11):617-628.

6. Hardiman O, Al-Chalabi A, Chio A, et al. Amyotrophic lateral sclerosis. *Nature reviews Disease primers*. 2017;3(1):1-19.

7. Oskarsson B, Horton DK, Mitsumoto H. Potential environmental factors in amyotrophic lateral sclerosis. *Neurologic Clinics*. 2015;33(4):877-888.

8. Longinetti E, Fang F. Epidemiology of amyotrophic lateral sclerosis: An update of recent literature. *Current Opinion In Neurology*. 2019;32(5):771.

9. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295(10):1127-1134.

10. Bennett JE, Tamura-Wicks H, Parks RM, et al. Particulate matter air pollution and national and county life expectancy loss in the USA: A spatiotemporal analysis. *PLOS Medicine*. 2019;16(7):e1002856. doi:10.1371/journal.pmed.1002856

11. Schwartz J. Particulate air pollution and chronic respiratory disease. *Environmental Research*. 1993;62(1):7-13.

12. Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology*. 2000;11(3):320-326.

13. Brook RD, Rajagopalan S, Pope III CA, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*. 2010;121(21):2331-2378.

14. Dockery DW, Pope CA, Xu X, et al. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*. 1993;329(24):1753-1759. doi:10.1056/NEJM199312093292401

Formatted: Font: Not Bold

15. Block ML, Elder A, Auten RL, et al. The outdoor air pollution and brain health workshop. *Neurotoxicology*. 2012;33(5):972-984.

16. Zanobetti A, Dominici F, Wang Y, Schwartz JD. A national case-crossover analysis of the short-term effect of PM 2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environmental Health*. 2014;13(1):1-11.

17. Ritz B, Lee PC, 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.

18. Kioumourtzoglou MA, Schwartz JD, Weisskopf MG, et al. Long-term PM2.5 exposure and neurological hospital admissions in the northeastern United States. *Environmental health perspectives*. 2016;124(1):23-29.

19. Levesque S, Surace MJ, McDonald J, Block ML. Air pollution & the brain: Subchronic diesel exhaust exposure causes neuroinflammation and elevates early markers of neurodegenerative disease. *Journal of Neuroinflammation*. 2011;8(1):1-10.

20. Heusinkveld HJ, Wahle T, Campbell A, et al. Neurodegenerative and neurological disorders by small inhaled particles. *Neurotoxicology*. 2016;56:94-106.

21. Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro III A, Schwartz J. Traffic-related air pollution and cognitive function in a cohort of older men. *Environmental Health Perspectives*. 2011;119(5):682-687.

22. Dubowsky SD, Suh H, Schwartz J, Coull BA, Gold DR. Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation. *Environmental Health Perspectives*. 2006;114(7):992-998.

23. Ruckerl R, Ibald-Mulli A, Koenig W, et al. Air pollution and markers of inflammation and coagulation in patients with coronary heart disease. *American Journal of Respiratory and Critical Care Medicine*. 2006;173(4):432-441.

24. Hoffmann B, Moebus S, Dragano N, et al. Chronic residential exposure to particulate matter air pollution and systemic inflammatory markers. *Environmental Health Perspectives*. 2009;117(8):1302-1308.

25. Kelly FJ. Oxidative stress: Its role in air pollution and adverse health effects. *Occupational and Environmental Medicine*. 2003;60(8):612-616.

26. Chuang KJ, Chan CC, Su TC, Lee CT, Tang CS. The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. *American journal of respiratory and critical care medicine*. 2007;176(4):370-376.

27. Li N, Sioutas C, Cho A, et al. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environmental Health Perspectives*. 2003;111(4):455-460.

28. Sørensen M, Daneshvar B, Hansen M, et al. Personal PM2.5 exposure and markers of oxidative stress in blood. *Environmental health perspectives*. 2003;111(2):161-166.

29. Block ML, Calderón-Garcidueñas L. Air pollution: Mechanisms of neuroinflammation and CNS disease. *Trends in neurosciences*. 2009;32(9):506-516.

30. Perry VH, Cunningham C, Holmes C. Systemic infections and inflammation affect chronic neurodegeneration. *Nature Reviews Immunology*. 2007;7(2):161-167.

31. Bergeron C. Oxidative stress: Its role in the pathogenesis of amyotrophic lateral sclerosis. *Journal of the neurological sciences*. 1995;129:81-84.

32. Mhatre M, Floyd RA, Hensley K. Oxidative stress and neuroinflammation in Alzheimer's disease and amyotrophic lateral sclerosis: Common links and potential therapeutic targets. *Journal of Alzheimer's disease*. 2004;6(2):147-157.

33. D'Amico E, Factor-Litvak P, Santella RM, Mitsumoto H. Clinical perspective on oxidative stress in sporadic amyotrophic lateral sclerosis. *Free radical biology and medicine*. 2013;65:509-527.

34. Perry VH, Nicoll JA, Holmes C. Microglia in neurodegenerative disease. *Nature Reviews Neurology*. 2010;6(4):193-201.

35. Nunez Y, Boehme AK, Weisskopf MG, et al. Fine particle exposure and clinical aggravation in neurodegenerative diseases in New York State. *Environmental Health Perspectives*. 2021;129(2):027003.

36. Nunez Y, Boehme AK, Li M, et al. Parkinson's disease aggravation in association with fine particle components in New York State. *Environmental Research*. 2021;201:111554.

37. Malek AM, Barchowsky A, Bowser R, et al. Exposure to hazardous air pollutants and the risk of amyotrophic lateral sclerosis. *Environmental Pollution*. 2015;197:181-186.

38. Yu Z, Peters S, van BL, et al. Long-Term Exposure to Ultrafine Particles and Particulate Matter Constituents and the Risk of Amyotrophic Lateral Sclerosis. *Environmental Health Perspectives*. 2021;129(9):097702. doi:10.1289/EHP9131

39. Seelen M, Toro CRA, Veldink JH, et al. Long-term air pollution exposure and amyotrophic lateral sclerosis in Netherlands: A population-based case–control study. *Environmental Health Perspectives*. 2017;125(9):097023. doi:10.1289/EHP1115

40. Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: Pooled analysis. *BMJ*. 2021;374:n1904. doi:10.1136/bmj.n1904

41. Hamra GB, Laden F, Cohen AJ, Raaschou-Nielsen O, Brauer M, Loomis D. Lung cancer and exposure to nitrogen dioxide and traffic: A systematic review and meta-analysis. *Environmental Health Perspectives*. 2015;123(11):1107-1112.

42. Chen H, Kwong JC, Copes R, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: A population-based cohort study. *The Lancet*. 2017;389(10070):718-726.

43. Gibson EA, Nunez Y, Abuawad A, et al. An overview of methods to address distinct research questions on environmental mixtures: An application to persistent organic pollutants and leukocyte telomere length. *Environmental Health*. 2019;18(1):1-16.

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

45. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: A review of content, data quality, and research potential. *Clinical epidemiology*. 2015;7:449.

~~46~~46. Trabjerg BB, Garton FC, van Rheenen W, et al. ALS in Danish registries: heritability and links to psychiatric and cardiovascular disorders. *Neurology Genetics*. 2020;6(2).

47. Mathis S, Goizet C, Soulages A, Vallat JM, Le Masson G. Genetics of amyotrophic lateral sclerosis: A review. *Journal of the Neurological Sciences*. 2019;399:217-226.

48. Kioumourtzoglou MA, 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.

~~47~~49. Pedersen CB. The Danish civil registration system. *Scandinavian journal of public health*. 2011;39(7_suppl):22-25.

~~48~~50. Langholz B, Goldstein L. Risk set sampling in epidemiologic cohort studies. *Statistical Science*. Published online 1996:35-53.

51. Sillman S. The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmospheric Environment*. 1999;33(12):1821-1845.

~~49~~52. Khan J, Kakosimos K, Raaschou-Nielsen O, et al. Development and performance evaluation of new AirGIS–a GIS based air pollution and human exposure modelling system. *Atmospheric environment*. 2019;198:102-121.

~~50~~53. Brandt J, Christensen JH, Frohn LM, Palmgren F, Berkowicz R, Zlatev Z. Operational air pollution forecasts from European to local scale. *Atmospheric Environment*. 2001;35:S91-S98.

~~51~~54. Brandt J, Christensen J, Frohn L, Berkowicz R. Air pollution forecasting from regional to urban street scale—-implementation and validation for two cities in Denmark. *Physics and Chemistry of the Earth, Parts A/B/C*. 2003;28(8):335-344.

~~52~~55. Frohn LM, Ketzel M, Christensen JH, et al. Modelling ultrafine particle number concentrations at address resolution in Denmark from 1979-2018–Part 1: Regional and urban scale modelling and evaluation. *Atmospheric Environment*. 2021;264:118631.

~~53~~56. 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.

~~54~~57. 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.

~~55~~58. 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.

~~56~~59. Seinfeld J, Pandis S. Atmospheric chemistry and physics. 1997. *New York*. Published online 2008.

~~57~~60. Galvin M, Gaffney R, Corr B, Mays I, Hardiman O. From first symptoms to diagnosis of amyotrophic lateral sclerosis: Perspectives of an Irish informal caregiver cohort—a thematic analysis. *BMJ Open*. 2017;7(3). doi:10.1136/bmjopen-2016-014985

~~58~~61. Dickerson AS, Hansen J, Kioumourtzoglou MA, Specht AJ, Gredal O, Weisskopf MG. Study of occupation and amyotrophic lateral sclerosis in a Danish cohort. *Occup Environ Med*. 2018;75(9):630-638. doi:10.1136/oemed-2018-105110

~~59~~62. Roberts AL, Johnson NJ, Chen JT, Cudkowicz ME, Weisskopf MG. Race/ethnicity, socioeconomic status, and ALS mortality in the United States. *Neurology*. 2016;87(22):2300-2308.

63. Bucher BT, Shi J, Pettit RJ, Ferraro J, Chapman WW, Gundlapalli A. Determination of marital status of patients from structured and unstructured electronic healthcare data. In: *AMIA Annual Symposium Proceedings*. Vol 2019. American Medical Informatics Association; 2019:267.

64. Norman RE, Carpenter DO, Scott J, Brune MN, Sly PD. Environmental exposures: an underrecognized contribution to noncommunicable diseases. *Reviews on environmental health*. 2013;28(1):59-65.

65. Rothman KJ, Greenland S, Lash TL, others. *Modern Epidemiology*. Vol 3. Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia; 2008.

~~60~~66. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis, Third Edition*. CRC Press; 2013.

~~61~~67. Mostofsky E, Schwartz J, Coull BA, et al. Modeling the association between particle constituents of air pollution and health outcomes. *American journal of epidemiology*. 2012;176(4):317-326.

~~62~~68. Thunis P, Degraeuwe B, Pisoni E, et al. PM2.5 source allocation in European cities: A SHERPA modelling study. *Atmospheric Environment*. 2018;187:93-106.

~~63~~69. Nuvolone D, Petri D, Voller F. The effects of ozone on human health. *Environmental Science and Pollution Research*. 2018;25(9):8074-8088.

70. Martin R, Peters G, Wilkinson J. Symmetric decomposition of a positive definite matrix. *Numerische Mathematik*. 1965;7(5):362-383.

~~64~~71. Polson NG, Scott JG. On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis*. 2012;7(4):887-902.

~~65~~72. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal*. 2006;1(3):515-534. doi:10.1214/06-BA117A

~~66~~73. Lewandowski D, Kurowicka D, Joe H. Generating random correlation matrices based on vines and extended onion method. *Journal of multivariate analysis*. 2009;100(9):1989-2001.

~~67~~74. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical science*. 1992;7(4):457-472.

~~68~~75. R Core Team. R: A language and environment for statistical computing. Published online 2013.

~~69~~76. Povedano M, Saez M, Martinez-Matos JA, Barceló MA. Spatial assessment of the association between long-term exposure to environmental factors and the occurrence of amyotrophic lateral sclerosis in Catalonia, Spain: A population-based nested case-control study. *Neuroepidemiology*. 2018;51(1-2):33-49.

~~70~~77. von Schneidemesser E, Mar KA, Saar D. Black carbon in Europe: Targeting an air Pollutant and climate forcer. Published online 2017.

~~71~~78. Pamphlett R, Rikard-Bell A. Different occupations associated with amyotrophic lateral sclerosis: Is diesel exhaust the link? *PloS One*. 2013;8(11):e80993.

~~72~~79. Zhang R, Dai Y, Zhang X, et al. Reduced pulmonary function and increased pro-inflammatory cytokines in nanoscale carbon black-exposed workers. *Part Fibre Toxicol*. 2014;11:73. doi:10.1186/s12989-014-0073-1

~~73~~80. Gao X, Xu H, Shang J, et al. Ozonized carbon black induces mitochondrial dysfunction and DNA damage. *Environ Toxicol*. 2017;32(3):944-955. doi:10.1002/tox.22295

~~74~~81. Kyjovska ZO, Jacobsen NR, Saber AT, et al. DNA damage following pulmonary exposure by instillation to low doses of carbon black (Printex 90) nanoparticles in mice. *Environ Mol Mutagen*. 2015;56(1):41-49. doi:10.1002/em.21888

~~75~~82. Benatar M, Turner MR, Wuu J. Defining pre-symptomatic amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2019;20(5-6):303-309.

83. Benatar M, Wuu J, McHutchison C, et al. Preventing amyotrophic lateral sclerosis: insights from pre-symptomatic neurodegenerative diseases. Published online 2021.

84. Eisen A, Kiernan M, Mitsumoto H, Swash M. Amyotrophic lateral sclerosis: A long preclinical period? *Journal of Neurology, Neurosurgery & Psychiatry*. 2014;85(11):1232-1238.

85. Weisskopf MG, Webster TF. Trade-offs of personal vs. more proxy exposure measures in environmental epidemiology. *Epidemiology (Cambridge, Mass)*. 2017;28(5):635.

~~76. Nakken O, Meyer HE, Stigum H, Holmøy T. High BMI is associated with low ALS risk: A population-based study. *Neurology*. 2019;93(5):e424-e432.~~

~~77. Jawaid A, Murthy SB, Wilson AM, et al. A decrease in body mass index is associated with faster progression of motor symptoms and shorter survival in ALS. *Amyotrophic Lateral Sclerosis*. 2010;11(6):542-548.~~

~~78~~86. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. CRC press; 2006.

87. Frisch M, Simonsen J. Marriage, cohabitation and mortality in Denmark: national cohort study of 6.5 million persons followed for up to three decades (1982–2011). *International Journal of Epidemiology*. 2013;42(2):559-578.

88. Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*. 2015;16(3):493-508.

Table 1. Demographic characteristics of cases and controls for 5-year average exposure group.

Characteristic	Overall, N = 23,232[a]	Case, N = 3,934[a]	Control, N = 19,298[a]
Average age (years)	66 (12)	66 (12)	66 (12)
Sex			
Female	10,973 (47%)	1,854 (47%)	9,119 (47%)
Male	12,259 (53%)	2,080 (53%)	10,179 (53%)
Socioeconomic status (SES)			
Group 1 (Highest)	2,337 (10%)	451 (11%)	1,886 (9.8%)
Group 2	2,839 (12%)	499 (13%)	2,340 (12%)
Group 3	4,360 (19%)	785 (20%)	3,575 (19%)
Group 4	6,598 (28%)	1,076 (27%)	5,522 (29%)
Group 5 (Lowest)	4,419 (19%)	717 (18%)	3,702 (19%)
Group 9 (~~Unemployed or unclassified~~Unknown)	2,679 (12%)	406 (10%)	2,273 (12%)
Place of birth			
Greater Copenhagen	4,858 (21%)	831 (21%)	4,027 (21%)
Big cities of Denmark	7,923 (34%)	1,357 (34%)	6,566 (34%)
Rest of Denmark	9,009 (39%)	1,548 (39%)	7,461 (39%)
Greenland	243 (1.0%)	53 (1.3%)	190 (1.0%)
Foreign	1,065 (4.6%)	122 (3.1%)	943 (4.9%)
Unknown	134 (0.6%)	23 (0.6%)	111 (0.6%)
Civil status			
Married	14,158 (61%)	2,411 (61%)	11,747 (61%)
Divorced	2,703 (12%)	433 (11%)	2,270 (12%)
Widowed	4,224 (18%)	726 (18%)	3,498 (18%)
Never married	2,147 (9.2%)	364 (9.3%)	1,783 (9.2%)
Last reported place of residence			
Greater Copenhagen	1,887 (8.1%)	335 (8.5%)	1,552 (8.0%)
Big cities of Denmark	9,385 (40%)	1,590 (40%)	7,795 (40%)
Rest of Denmark	11,954 (51%)	2,008 (51%)	9,946 (52%)
Greenland	6 (<0.1%)	1 (<0.1%)	5 (<0.1%)

[a]Mean (SD); n (%)

725 **Table 2.** Summary of 5-year average pollutant concentrations (all in μg/m^3).

Pollutant	Overall, N = 23,232[a]	Case, N = 3,934[a]	Control, N = 19,298[a]
NO_X	27 (20)	28 (21)	27 (20)
CO	238 (106)	239 (112)	237 (105)
EC	0.85 (0.42)	0.86 (0.45)	0.85 (0.42)
non-EC $PM_{2.5}$	11.76 (2.37)	11.78 (2.41)	11.76 (2.37)
O_3	51.9 (6.0)	51.9 (6.1)	52.0 (6.0)

[a]Mean (SD)

726

Figure Captions

Figure 1. Spearman correlation of ~~5~~1,- 5-, and 10-year average pollutant concentrations.

Figure 2. Percentage change in odds of ALS diagnosis per 1-, 5- and 10-year average standard deviation (SD) increase for each pollutant. Results are from the Bayesian hierarchical model including each of EC, NO_x, CO, and non-EC $PM_{2.5}$ together, and were additionally adjusted by age, sex, year of birth, vital status, socioeconomic status, civil status, last reported place of residence, and place of birth.

Formatted: Justified