HHS Public Access

Author manuscript

Environ Pollut. Author manuscript; available in PMC 2023 December 27.

Published in final edited form as:

Environ Pollut. 2022 February 15; 295: 118658. doi:10.1016/j.envpol.2021.118658.

ALS risk factors: Industrial airborne chemical releases☆

Angeline Andrew[a,*], **Jie Zhou**[b], **Jiang Gui**[b], **Xun Shi**[b], **Meifang Li**[b], **Antoinette Harrison**[c], **Bart Guetti**[a], **Ramaa Nathan**[d], **Tanya Butt**[a], **Daniel Peipert**[a], **Maeve Tischbein**[a], **Erik P. Pioro**[e], **Elijah Stommel**[a], **Walter Bradley**[f]

[a] Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA

[b] Dartmouth College, Hanover, NH, USA

[c] Mitsubishi Tanabe Pharma America, Inc., Jersey City, NJ, USA

[d] Eversana, Wayne, PA, USA

[e] Center for ALS and Related Disorders, Cleveland Clinic, Cleveland, OH, USA

[f] Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA

Abstract

Most amyotrophic lateral sclerosis (ALS) cases are sporadic (~90%) and environmental exposures are implicated in their etiology. Large industrial facilities are permitted the airborne release of certain chemicals with hazardous properties and report the amounts to the US Environmental Protection Agency (EPA) as part of its Toxics Release Inventory (TRI) monitoring program. The objective of this project was to identify industrial chemicals released into the air that may be associated with ALS etiology. We geospatially estimated residential exposure to contaminants using a de-identified medical claims database, the SYMPHONY Integrated Dataverse®, with ~26,000 nationally distributed ALS patients, and non-ALS controls matched for age and gender. We mapped TRI data on industrial releases of 523 airborne contaminants to estimate local residential exposure and used a dynamic categorization algorithm to solve the problem of zero-inflation in the dataset. In an independent validation study, we used residential histories to estimate exposure in each year prior to diagnosis. Air releases with positive associations in both the SYMPHONY analysis and the spatio-temporal validation study included styrene (false discovery rate (FDR) 5.4e-5), chromium (FDR 2.4e-4), nickel (FDR 1.6e-3), and dichloromethane (FDR 4.8e-4). Using a large de-identified healthcare claims dataset, we identified geospatial environmental contaminants associated with ALS. The analytic pipeline used may be applied to

☆This paper has been recommended for acceptance by Wen Chen.

* Corresponding author. Geisel School of Medicine at Dartmouth, One Medical Center Drive, 7936 Rubin Building, Lebanon, NH, 03756, USA. Angeline.Andrew@dartmouth.edu (A. Andrew).

Author contributions

Conceptualization, AH, WB.; Methodology, JZ, XS.; Formal analysis, JZ, ML, AA; Investigation, AA; Resources, ES; Data curation, BG, RN, TB, DP; Writing – original draft Preparation, AA; Writing – review & editing, MT, ES; Visualization, JZ; Supervision, ES, WB; Project administration, MT; Funding acquisition, WB, AA, EP, ES.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2021.118658.

other diseases and identify novel targets for exposure mitigation. Our results support the future evaluation of these environmental chemicals as potential etiologic contributors to sporadic ALS risk.

Keywords

Amyotrophic lateral sclerosis; Risk factors; Airborne; Residential history; Solvents

1. Introduction

In amyotrophic lateral sclerosis (ALS), the progressive loss of both upper and lower motor neurons leads to spasticity, muscle atrophy, frequent respiratory failure and death over a 3–5 year period. Only 10% of ALS cases can be attributed to a familial trait or gene (Mathis et al., 2019). Studies of monozygotic twins discordant for ALS demonstrate that environmental factors play a critical role in the majority of sporadic ALS (sALS) (Meltz Steinberg et al., 2015), perhaps in the setting of also having some genetic predisposition for the disease (Al-Chalabi et al., 2014; Bradley et al., 2018).

Observational studies have revealed that the following environmental exposures are risk factors for ALS: lead and mercury, pesticides, solvents, cyanotoxins, and head trauma (reviewed in Wang et al. (2017)). A case-control study in Pennsylvania (n = 51 cases) showed a 5-fold increased risk of ALS associated with residing in census tracts with exposure to ‘aromatic solvents’ above vs. below the median (Malek et al., 2015). Because Malek et al. analyzed conglomerates of 6–13 chemicals, the identity of the risk-driving chemical(s) is unclear. Together, identifying potential causal exposures could help to prevent ALS and focus interventional studies to block progression.

We and others see each case of ALS and other neurodegenerative diseases as potentially having a variety of different causal environmental factors, centralizing on the ability of a particular stressor to induce a neuronal ‘environmental stress response’ that is not overcome by internal compensatory repair mechanisms. The classical response in this category is the ‘heat shock’ response, involving production of ‘heat shock proteins (HSP)’ that help with stress tolerance and the re-folding and repair of damaged proteins, promoting cell survival (Calabrese et al., 2000; Dattilo et al., 2015). Stressors that necessitate this response range from physical head injuries to chemical exposures (Andrew et al., 2021). Elucidating the relationship between environmental exposures and neurodegenerative disease remains a persistent challenge in the field. Identifying disease-related contaminants likely requires considering complex dose-response relationships in addition to more traditional linear low-dose extrapolation models. These models include non-linear dose-response relationships, such as the biphasic ‘hormetic’ dose-response involving modest adaptive neuro-protective activity at low-doses, yet substantial cell death with exposure to levels of the same stress factor that exceed the limits of adaptation (Dattilo et al., 2015).

The objective of the current study was to identify individual airborne contaminants associated with ALS risk using the power of a large, nationally distributed and de-identified healthcare claims dataset. Industrial facilities are permitted airborne release of certain

chemicals with hazardous properties and are required to report these amounts to the United States Environmental Protection Agency (US EPA) as part of its Toxics Release Inventory (TRI) monitoring program. To assess ALS risk in relation to these releases, our study used a two-phase, 'discovery' and 'validation' cohort approach. The SYMPHONY Integrated Dataverse® served as the 'discovery' cohort to identify TRI airborne contaminant releases associated with ALS rates. Our data analysis approach included a categorization algorithm to accommodate the skewed and non-linear distribution of our geospatially estimated contaminant levels. We then performed 'validation' of these top-hit contaminants by calculating exposures prior to ALS onset in a study conducted in the Northern New England and Ohio regions. The advantage of these regional cohorts is that we collected address-level residential histories and could thus examine the relevance of spatial spreading from an industrial site and temporal changes in exposure, focusing on the 5 to 15-year period prior to diagnosis. Using this approach we identified airborne contaminants consistently associated with ALS across the broad SYMPHONY and the detailed regional validation datasets.

2. Materials and methods

Airborne contaminant releases:

We downloaded Toxics Release Inventory (TRI) airborne release data for 523 individual contaminants from the US EPA for the period 1993–2019. For each year, we constructed map based on the mandatory, yearly reporting to US EPA. The centroid of a given site was the geospatial coordinate of each industrial site reporting an airborne contaminant release. We next modeled the spatial spread of reported contaminants from the centroid point assuming a 10k dispersion. Thus, we created annual raster maps with the pixel values reflecting the estimated contaminant concentrations.

Discovery phase:

We used a large de-identified healthcare claims dataset from the SYMPHONY Integrated Dataverse® (herein referred to as SYMPHONY) 2013–2019 as the 'discovery' cohort. ALS patient inclusion criteria were a) minimum of two entries into the SYMPHONY database classified as being due toICD-9/10 codes for ALS at least 3 calendar months apart, b) minimum of 6 months' enrollment in the database prior to first ALS ICD code, c) age at first ALS ICD code ≥18 yr. Supplemental Figure 1 summarizes the analytic pipeline described below.

We randomly selected individuals in the overall SYMPHONY network that were similar to the ALS patient cases in regards to age, gender and length of database history as controls. We excluded patients >80 years old due to low coverage in the network as well as those with ICD-9/10 codes for other neurodegenerative diseases, as there may be shared etiologic factors. We then used the R-package "MatchIt" to perform propensity score matching with a 3:1 ratio to select a subset of controls with the nearest age and same gender as a comparison group (Ho et al., 2007). The selected controls showed similar national distribution to the cases, based on the coverage of the SYMPHONY network.

Within each of the 863 zip3 regions nationwide, we calculated the logit of the proportion of ALS cases to SYMPHONY network controls, r, in each region, i.e., $Y = \log\frac{r}{1-r}$ (mapped in Fig. 1). A histogram of the resulting Y values showed a normal distribution pattern. These Y values were used as the outcome variable for the discovery phase analysis.

Airborne contaminant exposure estimates ––We averaged the annual point-level release amounts (i.e. from industrial facilities) over the 10-years prior to the diagnostic period of the ALS cases 2002–2012. Then we took the average of the values of each of the 523 contaminants released within each zip3 polygon region. The zip3 polygon included all zip-codes that shared the same first three integers. We estimated residential exposure level at the zip3 location, as this was the only spatiotemporal information available for the SYMPHONY cohort.

Because industrial facilities only released each chemical in a small proportion of the zip3 regions, the exposure dataset distribution is skewed and shows a high degree of 'zero-inflation' (Fig. 2). To address this analytical challenge, we applied a categorization procedure. For chemicals with very sparse observations (defined as nonzero observations comprising less than 1% of the zip3 regions), we divided the observations into just two categories (e.g. present/absent). For all of the more frequently released chemicals, we applied the dynamic categorization algorithm that we developed recently to address a similar 'zero-inflation' issue in the context of microbiome data analysis (Zhou et al., 2020). In our optimization of this algorithm, a maximum of 6 binned categories effectively captured the variation in the datasets tested to date. This algorithm efficiently converted the log-transformed continuous concentration of each contaminant into up to $K = 6$ binned categories based on the distribution of that particular contaminant across the 863 zip3 regions. This algorithm runs iteratively with increasing values of K and calculates the within-group variance. The K that minimizes the variance of the contaminant levels within each bin is selected for each contaminant (see Supplement 1 for more details). Fig. 2B shows an example of the continuous levels of a contaminant with a skewed distribution that the algorithm binned into four categories ($K = 4$).

Statistical analysis ––To prioritize the most informative contaminants, we performed a sample-size weighted L1-penalized lasso using the categorized contaminant values as predictors and the logit of the zip3 ALS rate as the outcome. The subset of top-ranking contaminants selected by lasso were subsequently assessed using weighted univariate regression for main-effects on ALS rate, allowing us to assess the direction of the association with ALS and the effect size. We used False Discovery Rate (FDR) correction to account for multiple comparisons, using a significance threshold of <0.05. Weighted pair-wise regression with interaction terms evaluated combinations of contaminants associated with ALS. We also applied nonparametric Bayesian Kernal Machine Regression (BKMR) to assess non-linear relationships (Bobb et al., 2018).

Validation phase:

We used an independent cohort with residential history information for the validation. We obtained mortality records attributed to 'motor neuron disease' using ICD-10 code G12.2

from the following states and years available: New Hampshire (2009–2018 n = 337), Vermont (2008–2016 n = 216), and Ohio (2016–2019 n = 799). From among the same catchment counties as the ALS cases, population controls were identified as residents of New Hampshire (NH)/Vermont (VT) (n = 762), or Ohio (OH) (n = 1336) using the US Postal Service Delivery Sequence file licensed to Marketing Systems Group (Horsham, PA). The sampling algorithm was designed to randomly sample individuals in the population based on the expected demographic distribution of the ALS cases, with over-sampling of 50–75 year-olds and males.

Residential history ––—We obtained the geocodes of addresses held by each validation cohort subject over the 15-year period prior to the index date from a commercial financial marketing database query LexisNexis (Dayton, Ohio). We observed similar results when we tested other periods up to 25-years. We then mapped those addresses into zip3 polygons.

Airborne contaminant exposure estimates ––—For the validation cohort, we had residential history information comprised of the geospatial coordinates of each residence for both ALS cases and population controls. To estimate the exposure in each year prior to diagnosis, we read the contaminant amount from the raster maps representing the dispersed contaminant for each case or control residence in each year. We then calculated the median exposure to each contaminant across his/her multiple residences in epochs representing the period prior to the index year (i.e. for the 5-year epoch of a case diagnosed in 2016, we compiled estimated exposures for residences held from 2010 to 2015). We chose to use the median value rather than the cumulative value to avoid introducing bias due to missing residences.

Statistical analysis ––—The subset of contaminants identified by the ‘discovery phase’ entered the ‘validation phase’ analysis, which used the residential history epoch estimates of exposure for the case-control study subjects. The median exposure estimate of each contaminant over the residence prior to diagnosis was binned into categories based on the quartile distribution of each contaminant in the controls. Chi-square tests assessed the univariate difference in proportion of cases and controls by quantile, followed by logistic regression analysis that adjusted for age and gender. These analyses were all performed using R: A Language and Environment for Statistical Computing, version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

Table 1 shows that the age- and gender-distribution of the n =26,199 ALS cases we identified is similar to that of the controls sampled from the SYMPHONY network. The majority (63%) of the cases and controls were 55–75 years of age and approximately 57% were male, as expected based on ALS literature (McCombe and Henderson, 2010). Fig. 1 shows a map of the geospatial variation in the logit of the ALS rate among the 863 zip3 regions.

Fig. 2A shows a map of the average airborne releases of one example contaminant (dichloromethane) across the zip3 regions nationwide. Fig. 2B shows the skewed and

'zero-inflated', non-normal distribution of contaminant levels that led to our dynamic categorization analysis approach. Fig. 3 shows the weighted univariate regression results of the top-ranking contaminants selected by the lasso algorithm, and Supplemental Table 1 contains the full list of coefficients and FDR values, along with the distribution of contaminant levels. Certain identified chemicals had reported releases in many of the 863 counties (e.g. styrene in 578 counties, Supplemental Figure 3), whereas others were rarely reported (e.g. potassium bromate in 2 counties).

Ten contaminants met the FDR cutoff <0.05 from the nationwide SYMPHONY database 'discovery phase' and went into the 'validation phase' of analysis, which used residential history of NH/VT/OH patients and controls. Fig. 4 shows the increased ALS risk for the five contaminants (styrene, nitric acid, nickel, dichloromethane, and chromium) confirmed in the 'validation phase' (Fig. 4). History of residence in areas with styrene release in the 3rd quartile vs. below the median in the 15-years prior to diagnosis was associated increased risk of ALS (1.24-fold). 5-year exposure history in the 50–75th percentile was also associated increased risk of ALS for nitric acid (OR 1.36), nickel (OR 1.27), dichloromethane (OR 1.12), and chromium (OR 1.26) (Fig. 4, Supplemental Table 2). Maps showing the average pounds of dichloromethane and styrene released 2002–2012 across the zip3 regions can be found in Supplemental Figure 3.

We then evaluated the effects of exposure to combinations of the top-ranking contaminants in the SYMPHONY dataset using a pairwise interaction-effects model. As shown in Fig. 5, we did not identify any pairs of contaminants with positive interaction coefficients. Chloromethane paired with either nickel or chromium as the two combinations of contaminants with the lowest interaction p-values corrected for multiple comparisons (FDR) in SYMPHONY (Table 2). Neither pair met our FDR significance threshold of <0.05. Supplemental Table 3 shows the results for the top 20 pairs of contaminants in SYMPHONY.

Supplemental Figure 2 shows graphs of the non-linear and pairwise relationships among contaminants generated by non-parametric regression in SYMPHONY. This BKMR model did not show evidence of positive interactions among pairs of contaminants. Styrene does show a non-linear relationship with ALS, with effects peaking in the middle of the estimated exposure range of the SYMPHONY dataset.

4. Discussion

The role of environmental exposure in the etiology of sporadic ALS remains unexplained in many cases. The objective of this study was to take an unsupervised approach to assess links to airborne releases of contaminants reported to the US EPA in the TRI program. Unique aspects of our study include the use of a large, nationally distributed and de-identified healthcare claims dataset, an analysis pipeline with a dynamic categorization algorithm allowing for non-linear effects, and a validation phase using residential history based exposure assessment. Across these datasets, we observed consistent positive associations for airborne releases of a variety of potential environmental stressors: nitric acid, styrene, chromium, nickel, and dichloromethane.

Airborne styrene was among a group of aromatic solvents implicated in increased risk of ALS in a study of environmental airborne contaminants in the Pittsburgh area (Malek et al., 2015), supporting the styrene association that came up in our nationwide unsupervised analysis of 523 airborne contaminants and ALS also in the residential history validation. Styrene is used in rubber, plastic and fiberglass laminate, and disposable drinking glass manufacturing. Neurotoxicity in styrene-exposed workers has been established with clear symptoms of central nervous system depression, and anesthetic effects that include peripheral nerve conduction velocities and sensorimotor neuropathies (Kovarik et al., 1989; Matikainen et al., 1993). Peripheral markers also document these effects, with significantly lower serum dopamine-beta-hydroxylase activity in exposed vs control workers $p < 0.001$ (Bergamaschi et al., 1996). Ambient airborne styrene levels estimated in the highest quartile using the National Air Toxics Assessment were associated with self-reported central nervous system symptoms among residents of the US Gulf coast (Werder et al., 2018). Styrene oxide is a glutathione-depleting epoxide with cytotoxic effects that are exacerbated in cell lines with abnormal SOD-1 activity (Durham et al., 1995). In our environmental exposure setting, we observed a non-linear relationship with the highest ALS risk associated with doses in the middle of the exposure range. This may be indicative of a hormetic dose-response described by Calabrese et al., where low-levels of exposure activate a protective cellular stress response, inducing production of molecules under control of the “vitagenes” such as HSPs and glutathione, which have antioxidant and anti-apoptotic capabilities (Calabrese et al., 2010). Controlled laboratory experiments are needed to elucidate these mechanisms and clarify the dose response relationship between styrene exposure and levels of cytoprotective proteins such as glutathione.

The epidemiological literature covers several additional contaminants that ranked highly in the SYMPHONY analysis and validated using residential histories. In our assessment of 523 contaminants, dichloromethane exposure was identified as one such risk factor. Interestingly, Dickerson et al. linked dichloromethane (methylene chloride) to ALS in a Danish occupational cohort (Dickerson et al., 2020b). Dichloromethane was used in paint removers, until the EPA banned its use in 2019, and is also an aerosol spray propellant. It acts as a central nervous system toxin causing narcosis, and breaks down into carbon monoxide, impairing oxygen delivery by hemoglobin and myoglobin (Durrani et al., 2020). Hypoxic stress activates a cascade anti-proliferative cellular response initially mediated by hypoxia inducible factor-1α (HIF-1α) induction (Durrani et al., 2020). Yet, carbon monoxide is also considered a “nonconventional neurotransmitter” that can activate mitochondrial redox signaling (Calabrese et al., 2010).

We also found a statistically significant increased risk of ALS associated with chromium estimated over 5-years of prior residences. An analysis of occupational history using the Danish Pension Fund found chromium exposure in the third and fourth quartiles trended towards increased ALS risk, compared to those with no exposure (adjusted ORs = 1.24; 95% CI 0.91, 1.69; 1.19; 95% CI: 0.80, 1.76, respectively) (Dickerson et al., 2020a). Both our geospatial and the Danish occupational studies included a mixture of different forms of these metals, which may obscure the effects. For example, chromium(VI) kills neurons in *Drosophila*, but chromium(III) does not (Singh and Chowdhuri, 2017). A geospatial study in Spain also linked chromium in river water to a 15.7% increased risk of death from

motor neuron disease (Sanchez-Diaz et al., 2018). At the cellular level, chromium(VI) is reduced to chromium (III) producing reactive oxygen species. In a recent review on the neurotoxic effects of chromium, Wise et al. shows substantial evidence of neurodegenerative changes, neuroinflammation, decreased acetylcholinesterase activity, and induction of stress response genes including *Nrf2* and *HO-1* and pro-apoptotic caspases (Wise et al., 2021). While several animal studies show that levels of the antioxidant glutathione decreased after chromium(VI) exposure, others show increases in the cortex, perhaps demonstrating a hormetic adaptive response under certain conditions (Wise et al., 2021).

Likewise, for nickel, we found an adjusted ALS OR 1.27 95%CI 1.06–1.52 associated with 5-year exposure history in the 50th-75th percentile. Among Danish women, occupational nickel exposure was associated with increased ALS risk in the third quartile (adjusted OR = 2.21; 95% CI: 1.14, 4.28), but not the fourth (Dickerson et al., 2020a). Investigation of the neurotoxic effects of nickel is on-going (reviewed (Salimi et al., 2020)). Notably studies of animal models demonstrate that nickel treatment induces oxidative stress and histopathological damage to brain tissue, including mitochondrial damage (Ijomone et al., 2018; Song et al., 2017). We and others have demonstrated that nickel can stabilize HIF-1α, mimicking the hypoxic response that switches cellular metabolism from mitochondrial oxidative phosphorylation to anaerobic glycolysis, greatly reducing the amount of energy produced (Andrew et al., 2001; Song et al., 2017). Nickel exposure also induces production of serum nitric oxide synthase (iNOS) and production of serum nitric oxide concentrations (Cruz et al., 2004; Hattiwale et al., 2013). Nitric oxide and mitochondrial superoxide can combine to form reactive peroxynitrite, which may regulate neurotransmission (Calabrese et al., 2010).

Lastly, we also identified and validated nitric acid as a risk factor. The primary industrial use of nitric acid is in the production of ammonium nitrate for fertilizers, but it is also produced in manufacturing dyes, fungicides, and polymers (National_Research_Council, 2013). The fumes are highly corrosive and damage the respiratory tract, however evidence for neurotoxicity in the literature is lacking (National_Research_Council, 2013). It is possible that nitric acid releases are not causal themselves, but are acting as a surrogate for another chemical produced by the same industrial processes, such as nitrous oxide. Nitrous oxide is an anesthetic that inhibits NMDA-receptor mediated currents and is associated with slowed motor neuron conduction velocities and reduced amplitudes (Li et al., 2016).

Together our analysis identified multiple releases that are associated with ALS and may be relevant to ALS etiology. Neurotoxic effects of solvents occur in workers with high, chronic exposures and include long-term effects on motor neurons. A group of n = 87 workers presented with neurological symptoms and suspected chronic poisoning from either trichloroethylene, perchloroethylene, and/or a mixture of solvents (including butyl acetate). Both at baseline exam, and when re-examined approximately 5 years after exposure cessation, electroneuromyography (EMG) fibrillations indicating denervation of the muscle were observed in 46% of subjects, with loss of motor units documented in 61% of subjects. Among the n = 53 exposed to the solvent mixture, 37% had motor neuropathy, 16% had sensory neuropathy, and 45% had involvement of both (Seppalainen and Antti-Poika, 1983). It is unclear, however, whether the levels of solvents and other airborne contaminants

released from industrial facilities into the general environment can cause neurodegenerative injuries. Our results do support a relationship between residential history in areas with airborne releases of certain contaminants to the environment and increased ALS risk.

As motor neurons transmit signals long distances and have high energy demands, they are particularly vulnerable to the effects of mitochondrial damage, which can be caused by contaminant exposures (Calabrese et al., 2010). Mechanistically, ALS may be the failure of motor neurons to counteract cellular stressors productively. Calabrese et al. suggest the importance of a hormetic response, where low-levels of a stressor may induce a positive, protective response boosting repair factors, whereas high doses of the same chemical cause levels of protein misfolding that overwhelm the capacity of the system (Calabrese et al., 2010). Because of the tradition and ease of linear extrapolation, a hormetic effect may be one of the reasons why identifying neurodegenerative risk factors and therapeutic strategies has been so historically challenging (Calabrese et al., 2010; Martin et al., 2017).

Advantages of our study include a large de-identified database with over 26,000 ALS patients occurring in the nationally distributed SYMPHONY healthcare network based population. Our analytic approach used a dynamic categorization algorithm to accommodate the zero-inflated, skewed, and varied distributions of the levels of 523 contaminants. We used a sample-size weighting during our contaminant selection to minimize chance findings associated with small numbers. However, the geospatial data for the SYMPHONY dataset was limited to the zip3 location at diagnosis. Some of the contaminants selected in the nationally distributed SYMPHONY analysis were released at only a small number of sites, particularly potassium bromate and catechol. Thus, we were unable to validate these findings using our regional datasets.

In our regional validation case-control studies, we used commercial financial credit and marketing databases to ascertain exact location coordinates with residential histories for the ~15-year pre-diagnostic period. The address information from the credit reporting company, LexisNexis (Dayton, Ohio), compared well to self-reported address information. Among 1099 Michigan bladder cancer study participants with both types of data collected over similar periods of time, 96.8% of the LexisNexis three most recent addresses were concordant with those reported in the survey (Jacquez et al., 2011). Our validation study used ALS mortality to identify patients, which had the advantage of including all patients who died in each state, independent of diagnosing facility, although it may contain some misclassification of ‘motor neuron disease’ as the cause of death.

5. Conclusions

We demonstrated an analytic pipeline using a dynamic categorization algorithm to identify contaminants geospatially associated with risk of ALS in a large healthcare claims network, and validation using residential histories. This analytic approach is applicable to other contaminant datasets and additional disease outcomes in the future. Our analysis identified industrial releases of nitric acid, styrene, chromium, nickel, and dichloromethane (methylene chloride) as ALS risk factors that particularly warrant further investigation in the laboratory

to assess mechanisms, as potential etiologic contributors to sporadic ALS risk, and as targets for exposure mitigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the research staff who made this work possible. The New Hampshire, Vermont, and Ohio Departments of Health compiled the mortality data.

Funding

This research was funded by MTPA. AH is an employee of MTPA. The CDC/ATSDR grant R01TS000288 supported the validation studies.

References

Al-Chalabi A, Calvo A, Chio A, Colville S, Ellis CM, Hardiman O, Heverin M, Howard RS, Huisman MH, Keren N, Leigh PN, Mazzini L, Mora G, Orrell RW, Rooney J, Scott KM, Scotton WJ, Seelen M, Shaw CE, Sidle KS, Swingler R, Tsuda M, Veldink JH, Visser AE, van den Berg LH, Pearce N, 2014. Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. Lancet Neurol. 13, 1108–1113. [PubMed: 25300936]

Andrew AS, Bradley WG, Peipert D, Butt T, Amoako K, Pioro EP, Tandan R, Novak J, Quick A, Pugar KD, Sawlani K, Katirji B, Hayes TA, Cazzolli P, Gui J, Mehta P, Horton DK, Stommel EW, 2021. Risk factors for amyotrophic lateral sclerosis: a regional United States case-control study. Muscle Nerve 63, 52–59. [PubMed: 33006184]

Andrew AS, Klei LR, Barchowsky A, 2001. Nickel requires hypoxia-inducible factor-1 alpha, not redox signaling, to induce plasminogen activator inhibitor-1. Am. J. Physiol. Lung Cell Mol. Physiol. 281, L607–L615. [PubMed: 11504687]

Bergamaschi E, Mutti A, Cavazzini S, Vettori MV, Renzulli FS, Franchini I, 1996. Peripheral markers of neurochemical effects among styrene-exposed workers. Neurotoxicology 17, 753–759. [PubMed: 9086498]

Bobb JF, Claus Henn B, Valeri L, Coull BA, 2018. Statistical software for analyzing the health effects of multiple concurrent exposures via Bayesian kernel machine regression. Environ. Health 17, 67. [PubMed: 30126431]

Bradley WG, Miller RX, Levine TD, Stommel EW, Cox PA, 2018. Studies of environmental risk factors in amyotrophic lateral sclerosis (ALS) and a phase I clinical trial of L-serine. Neurotox. Res. 33, 192–198. [PubMed: 28527102]

Calabrese V, Copani A, Testa D, Ravagna A, Spadaro F, Tendi E, Nicoletti VG, Giuffrida Stella AM, 2000. Nitric oxide synthase induction in astroglial cell cultures: effect on heat shock protein 70 synthesis and oxidant/antioxidant balance. J. Neurosci. Res. 60, 613–622. [PubMed: 10820432]

Calabrese V, Cornelius C, Dinkova-Kostova AT, Calabrese EJ, Mattson MP, 2010. Cellular stress responses, the hormesis paradigm, and vitagenes: novel targets for therapeutic intervention in neurodegenerative disorders. Antioxidants Redox Signal. 13, 1763–1811.

Cruz MT, Goncalo M, Figueiredo A, Carvalho AP, Duarte CB, Lopes MC, 2004. Contact sensitizer nickel sulfate activates the transcription factors NF-kB and AP-1 and increases the expression of nitric oxide synthase in a skin dendritic cell line. Exp. Dermatol. 13, 18–26. [PubMed: 15009112]

Dattilo S, Mancuso C, Koverech G, Di Mauro P, Ontario ML, Petralia CC, Petralia A, Maiolino L, Serra A, Calabrese EJ, Calabrese V, 2015. Heat shock proteins and hormesis in the diagnosis and treatment of neurodegenerative diseases. Immun. Ageing 12, 20. [PubMed: 26543490]

Dickerson AS, Hansen J, Gredal O, Weisskopf MG, 2020a. Study of occupational chromium, iron, and nickel exposure and amyotrophic lateral sclerosis in Denmark. Int. J. Environ. Res. Publ. Health 17.

Dickerson AS, Hansen J, Thompson S, Gredal O, Weisskopf MG, 2020b. A mixtures approach to solvent exposures and amyotrophic lateral sclerosis: a population-based study in Denmark. Eur. J. Epidemiol.

Durham HD, O'Brien C, Nalbantoglu J, Figlewicz DA, 1995. Use of tissue culture models to study environmental-genetic interactions relevant to neurodegenerative diseases. Clin. Exp. Pharmacol. Physiol. 22, 366–367. [PubMed: 7554432]

Durrani T, Clapp R, Harrison R, Shusterman D, 2020. Solvent-based paint and varnish removers: a focused toxicologic review of existing and alternative constituents. J. Appl. Toxicol. 40, 1325–1341. [PubMed: 32342556]

Hattiwale SH, Saha S, Yendigeri SM, Jargar JG, Dhundasi SA, Das KK, 2013. Protective effect of L-ascorbic acid on nickel induced pulmonary nitrosative stress in male albino rats. Biometals 26, 329–336. [PubMed: 23463385]

Ho D, Imai K, King G, Stuart E, 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit. Anal. 15, 199–236.

Ijomone OM, Olatunji SY, Owolabi JO, Naicker T, Aschner M, 2018. Nickel-induced neurodegeneration in the hippocampus, striatum and cortex; an ultrastructural insight, and the role of caspase-3 and alpha-synuclein. J. Trace Elem. Med. Biol. 50, 16–23. [PubMed: 30262274]

Jacquez GM, Slotnick MJ, Meliker JR, AvRuskin G, Copeland G, Nriagu J, 2011. Accuracy of commercially available residential histories for epidemiologic studies. Am. J. Epidemiol. 173, 236–243. [PubMed: 21084554]

Kovarik J, Pithartova R, Ehler E, Salandova J, Kuzelova M, Popler A, Sirl J, Ciharova M, 1989. [Health status of workers exposed to styrene used in the manufacture of glass laminates]. Sb Ved Pr Lek Fak Karlovy Univerzity Hradci Kralove Suppl 32, 617–624. [PubMed: 2638043]

Li HT, Chu CC, Chang KH, Liao MF, Chang HS, Kuo HC, Lyu RK, 2016. Clinical and electrodiagnostic characteristics of nitrous oxide-induced neuropathy in Taiwan. Clin. Neurophysiol. 127, 3288–3293. [PubMed: 27567448]

Malek AM, Barchowsky A, Bowser R, Heiman-Patterson T, Lacomis D, Rana S, Ada Y, Talbott EO, 2015. Exposure to hazardous air pollutants and the risk of amyotrophic lateral sclerosis. Environ. Pollut. 197, 181–186. 10.1016/j.envpol.2014.1012.1010. Epub 2014 Dec 1023. [PubMed: 25544309]

Martin S, Khaleifat AA, Al-Chalabi A, 2017. What causes amyotrophic lateral sclerosis? F1000Research, 6.

Mathis S, Goizet C, Soulages A, Vallat JM, Masson GL, 2019. Genetics of amyotrophic lateral sclerosis: a review. J. Neurol. Sci. 399, 217–226. [PubMed: 30870681]

Matikainen E, Forsman-Gronholm L, Pfaffli P, Juntunen J, 1993. Neurotoxicity in Workers Exposed to Styrene. IARC Sci Publ, pp. 153–161.

McCombe PA, Henderson RD, 2010. Effects of gender in amyotrophic lateral sclerosis. Gend. Med. 7, 557–570. [PubMed: 21195356]

Meltz Steinberg K, Nicholas TJ, Koboldt DC, Yu B, Mardis E, Pamphlett R, 2015. Whole genome analyses reveal no pathogenetic single nucleotide or structural differences between monozygotic twins discordant for amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 16, 385–392. [PubMed: 25960086]

National_Research_Council, 2013. Nitric acid. In: Levels C.o.A.E.G. (Ed.), Acute Exposure Guideline Levels for Selected Airborne Chemicals. National Academies Press, Washington (DC).

Salimi A, Jamali Z, Atashbar S, Khezri S, Ghorbanpour AM, Etefaghi N, 2020. Pathogenic mechanisms and therapeutic implication in nickel-induced cell damage. Endocr. Metab. Immune Disord. - Drug Targets 20, 968–984. [PubMed: 32056533]

Sanchez-Diaz G, Escobar F, Badland H, Arias-Merino G, Posada de la Paz M, Alonso-Ferreira V, 2018. Geographic analysis of motor neuron disease mortality and heavy metals released to rivers in Spain. Int. J. Environ. Res. Publ. Health 15.

Seppalainen AM, Antti-Poika M, 1983. Time course of electrophysiological findings for patients with solvent poisoning. A descriptive study. Scand. J. Work. Environ. Health 9, 15–24.

Singh P, Chowdhuri DK, 2017. Environmental presence of hexavalent but not trivalent chromium causes neurotoxicity in exposed Drosophila melanogaster. Mol. Neurobiol. 54, 3368–3387. [PubMed: 27167131]

Song X, Fiati Kenston SS, Kong L, Zhao J, 2017. Molecular mechanisms of nickel induced neurotoxicity and chemoprevention. Toxicology 392, 47–54. [PubMed: 29032222]

Wang MD, Little J, Gomes J, Cashman NR, Krewski D, 2017. Identification of risk factors associated with onset and progression of amyotrophic lateral sclerosis using systematic review and meta-analysis. Neurotoxicology 61, 101–130. [PubMed: 27377857]

Werder EJ, Engel LS, Richardson DB, Emch ME, Gerr FE, Kwok RK, Sandler DP, 2018. Environmental styrene exposure and neurologic symptoms in U. S. Gulf coast residents. Environ. Int. 121, 480–490. [PubMed: 30278311]

Wise JP Jr., Young JL, Cai J, Cai L, 2021. Current understanding of hexavalent chromium [Cr(VI)] neurotoxicity and new perspectives. Environ. Int. 158, 106877. [PubMed: 34547640]

Zhou J, Viles WD, Lu B, Li Z, Madan JC, Karagas MR, Gui J, Hoen AG, 2020. Identification of microbial interaction network: zero-inflated latent Ising model based approach. BioData Min. 13, 16. [PubMed: 33042226]

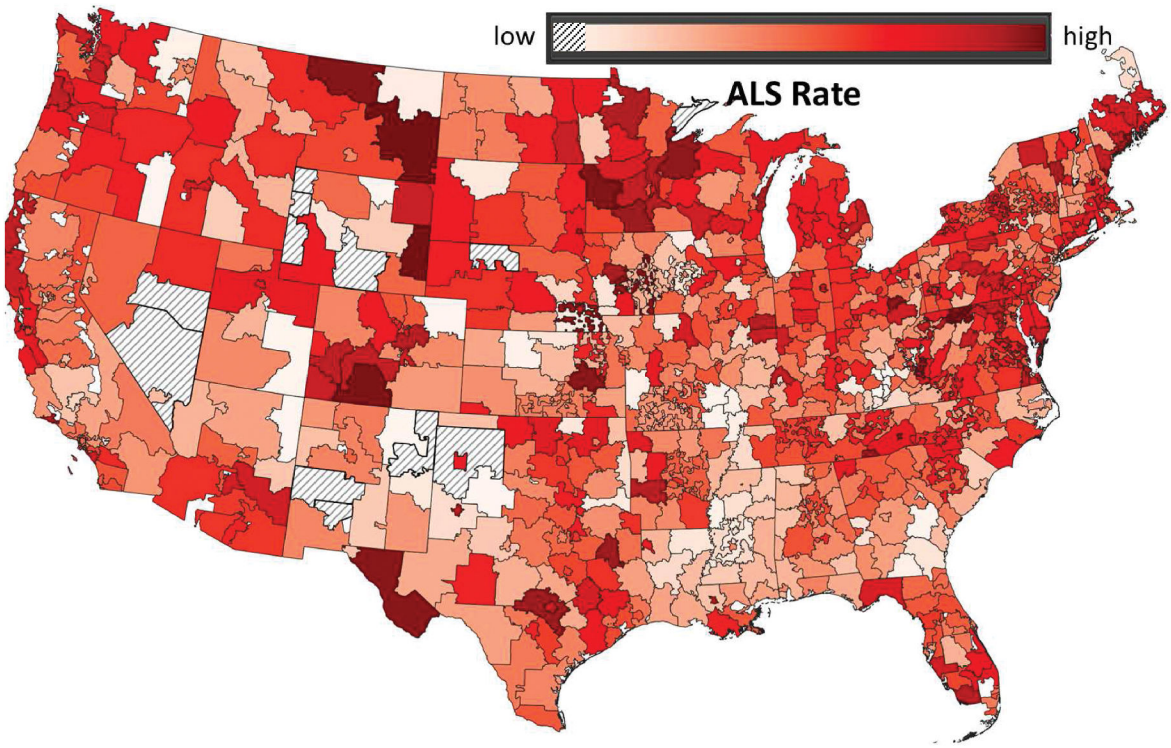

Fig. 1.

ALS rate in 863 zip3 polygons calculated in the SYMPHONY dataset. Within each of the 863 zip3 regions nationwide, we calculated the logit of the proportion of ALS cases to SYMPHONY network controls, r, in each region, i.e., $Y = \log \frac{r}{1-r}$.

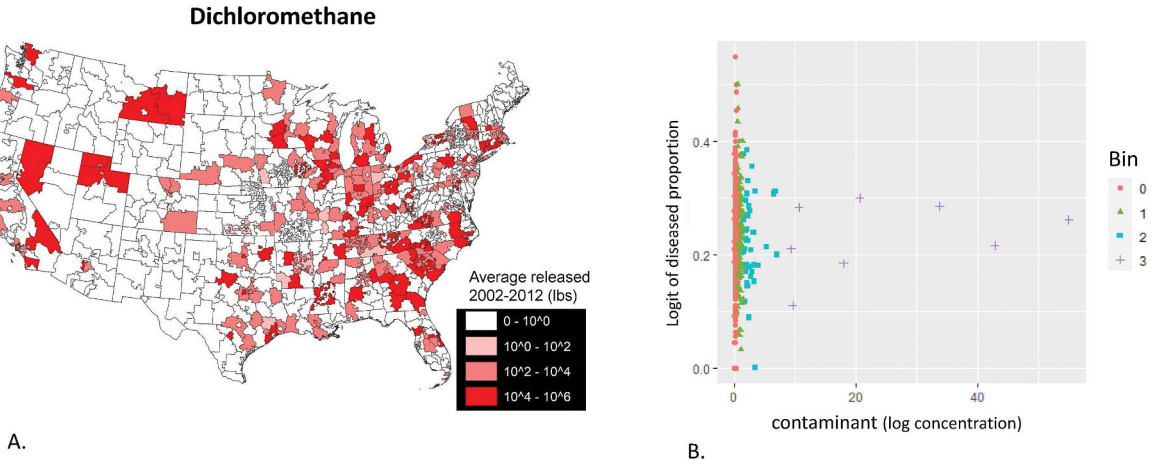

Fig. 2.

US EPA TRI reported contaminant levels are zero-inflated and right-skewed. A) Average dichloromethane release concentrations by zip3 region over the period 2002–2012, as an example contaminant. B) Example showing that the distribution of contaminant levels is skewed and shows a high degree of ‘zero-inflation’. We used an algorithm to bin the continuous values into four categories, as shown by the colored symbols. The purple ‘+’ represents the bin with the highest contaminant level.

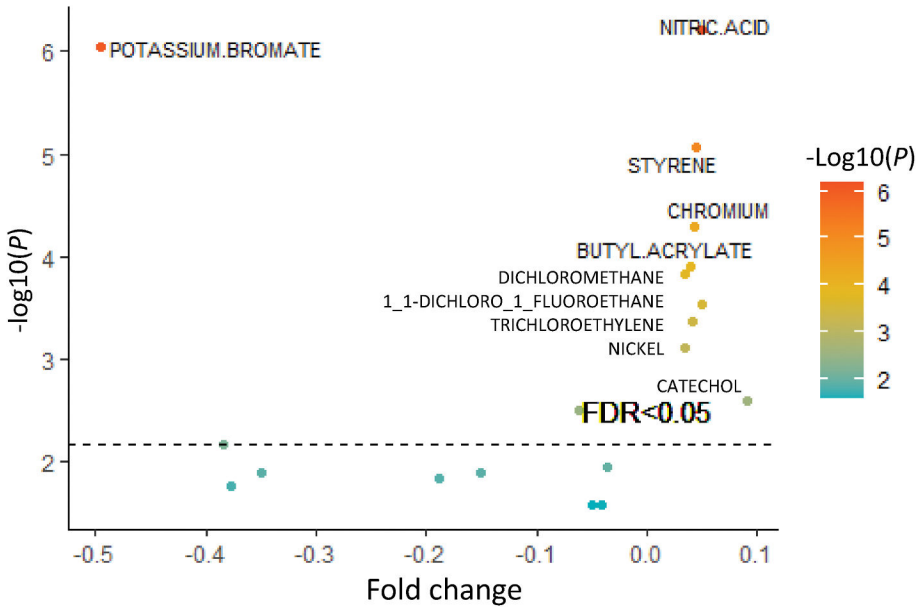

Fig. 3.

Volcano plot showing top-ranking contaminants selected from SYMPHONY by the lasso algorithm. We performed weighted univariate regression of airborne contaminants using the logit of the ALS rate in the SYMPHONY dataset as the outcome. The y-axis shows increasing statistical significance, while the x-axis reflects the size of the effect. Contaminants with increasing ALS risk are shown in the top right portion of the plot.

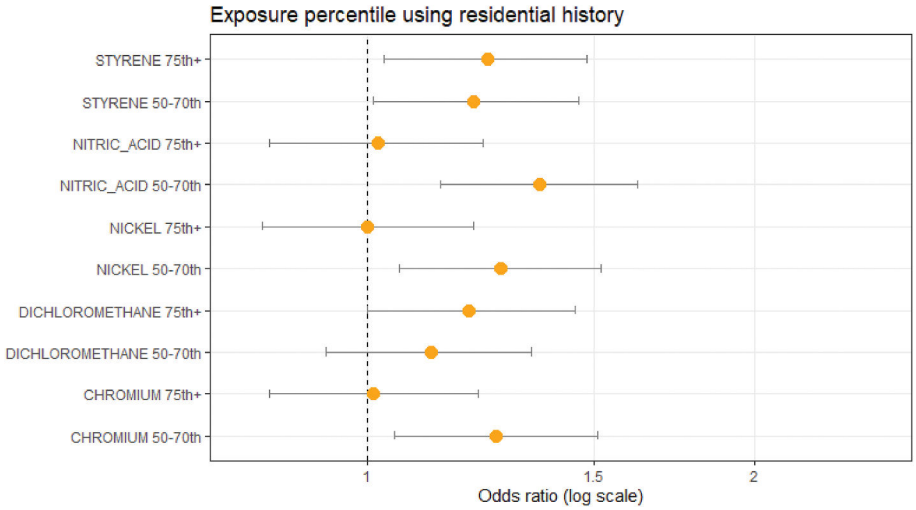

Fig. 4.

Top-ranking contaminants validated in NH/VT/OH residential history studies of ALS. We used residential history data for the epochs prior to the index date to estimate exposure at the geospatial coordinates of each residence. Bars depict the Odds Ratio (OR) and 95% confidence interval by quartile, using exposure below the median as the reference. We show the epochs with the largest magnitude effect size, which was 15-years for styrene, and 5-years for the other contaminants.

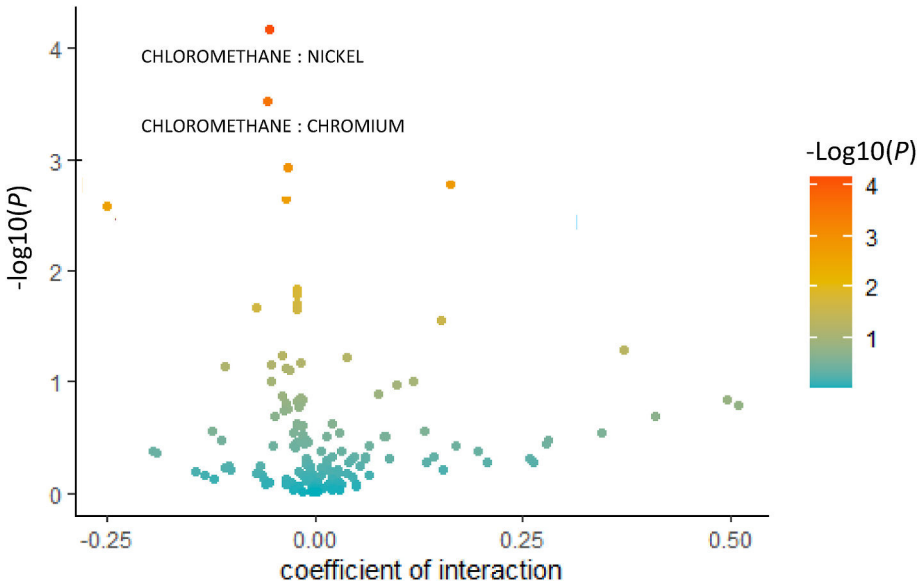

Fig. 5.

Volcano plot showing combinations of contaminants in the SYMPHONY dataset. We assessed a pair-wise interaction-effects model of airborne contaminants using the logit of the ALS rate as the outcome. Higher vertical points depict stronger statistical significance, however none of the pairs of contaminants met our FDR significance threshold. Positive interaction coefficients indicating synergistic combinations are shown to the right of the plot.

Table 1

SYMPHONY population characteristics.

		Controls	ALS patients	*p*-Value
		N = 78,597 (%)	N = 26,199 (%)	
Age	<45	5823 (7.4)	1941 (7.4)	1
	45–55	12,816 (16.3)	4272 (16.3)	
	55–65	24,561 (31.2)	8187 (31.2)	
	65–75	25,269 (32.2)	8423 (32.2)	
	75 +	10,128 (12.9)	3376 (12.9)	
Sex	Female	33,264 (42.3)	11,286 (43.1)	0.033
	Male	45,328 (57.7)	14,912 (56.9)	

Table 2

Pairwise interaction-effects model of airborne contaminants vs. ALS in the SYMPHONY dataset.

		Coefficients		Coefficient of	FDR
Contaminant_1	**Contaminant_2**	**main_1**	**main_2**	**interaction_1*2**	
CHLOROMETHANE	NICKEL	0.057	0.051	− 0.055	0.062
CHLOROMETHANE	CHROMIUM	0.060	0.057	− 0.057	0.13