Environment International 188 (2024) 108776

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

A prospective exploration of the urban exposome in relation to headache in the Dutch population-based Occupational and environmental health cohort study (AMIGO)

Eugenio Traini *, Lützen Portengen, Haykanush Ohanyan, Robert van Vorstenbosch, Roel Vermeulen, Anke Huss

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

ARTICLE INFO

Handling Editor: Olga Kalantzi

Keywords:
Urban exposome
Headache
Boruta
Causal forest

ABSTRACT

Objective: Headache is one of the most prevalent and disabling health conditions globally. We prospectively explored the urban exposome in relation to weekly occurrence of headache episodes using data from the Dutch population-based Occupational and Environmental Health Cohort Study (AMIGO).

Material and Methods: Participants (N = 7,339) completed baseline and follow-up questionnaires in 2011 and 2015, reporting headache frequency. Information on the urban exposome covered 80 exposures across 10 domains, such as air pollution, electromagnetic fields, and lifestyle and socio-demographic characteristics. We first identified all relevant exposures using the Boruta algorithm and then, for each exposure separately, we estimated the average treatment effect (ATE) and related standard error (SE) by training causal forests adjusted for age, depression diagnosis, painkiller use, general health indicator, sleep disturbance index and weekly occurrence of headache episodes at baseline.

Results: Occurrence of weekly headache was 12.5 % at baseline and 11.1 % at follow-up. Boruta selected five air pollutants (NO_2, NO_X, PM_{10}, silicon in PM_{10}, iron in $PM_{2.5}$) and one urban temperature measure (heat island effect) as factors contributing to the occurrence of weekly headache episodes at follow-up. The estimated causal effect of each exposure on weekly headache indicated positive associations. NO_2 showed the largest effect (ATE = 0.007 per interquartile range (IQR) increase; SE = 0.004), followed by PM_{10} (ATE = 0.006 per IQR increase; SE = 0.004), heat island effect (ATE = 0.006 per one-degree Celsius increase; SE = 0.007), NO_x (ATE = 0.004 per IQR increase; SE = 0.004), iron in PM2.5 (ATE = 0.003 per IQR increase; SE = 0.004), and silicon in PM_{10} (ATE = 0.003 per IQR increase; SE = 0.004).

Conclusion: Our results suggested that exposure to air pollution and heat island effects contributed to the reporting of weekly headache episodes in the study population.

1. Introduction

Headache disorders, characterized by their diverse intensity and frequency, represent one of the most prevalent and incapacitating health conditions globally (GBD 2016 Headache Collaborators, 2018; Steiner and World Headache Alliance, 2004).

While genetic factors have been acknowledged to play a role in the onset of headaches (Russell et al., 2006; Di Lorenzo et al., 2015), emerging research emphasizes the substantial impact of lifestyle and behavioral characteristics as well as environmental factors on the initiation and persistence of headaches (Molarius et al., 2008; Friedman and De Ver, 2009; Ulrich et al., 2004). As such, no single factor can be considered the sole trigger of headaches in the population; rather, their occurrence is likely the result of a combination of various factors including, among others, stress, lack of sleep, diet, analgesic overuse, environmental stressors, and urban temperature (Lee et al., 2018; Raucci et al., 2021; Winter et al., 2011; Holzhammer and Wöber, 2006; Rains and Poceta, 2012; Nash and Thebarge, 2006; Prince et al., 2004; Nikiforow and Hokkanen, 1978; Ashina et al., 2023). For instance, exposure to high levels of air pollution has been associated with an increased risk of hospitalization for headache (Dales et al., 2009). Similarly, exposure to specific chemicals, such as metals, has been suggested to increase

* Corresponding author at: Utrecht University, Institute for Risk Assessment Sciences, Yalelaan 2, 3584CM Utrecht, The Netherlands.
E-mail address: e.traini@uu.nl (E. Traini).

https://doi.org/10.1016/j.envint.2024.108776
Received 25 March 2024; Received in revised form 3 May 2024; Accepted 24 May 2024
Available online 25 May 2024

headache susceptibility (Donma and Donma, 2002).

As we explore the impact of lifestyle factors, behavioral characteristics, and environmental stressors on headaches, it becomes evident that their intricate dynamics require a comprehensive understanding to inform targeted interventions and personalized approaches to headache management that cannot be achieved by considering each factor separately. The exposome is defined as the totality of exposures that individuals encounter over their lifetimes and the biological reactions that these stressors produce (Vermeulen et al., 2020; Wild, 2005). As such, the urban exposome denotes a complex interplay between the built, social, chemical, food, and lifestyle aspects of the environment where people live. This interaction is characterized by persistent spatial and temporal variations in both quantitative and qualitative measures associated with different aspects of residential surroundings, and, as a consequence, these fluctuations may impact the well-being and health of individuals (Andrianou and Makris, 2018).

To the best of our knowledge, the relationship between urban exposome and headache has not yet been explored. In this study, we aimed to prospectively evaluate factors related to reporting of weekly headache episodes, in a large study population relying on the exposome framework. We conducted an exploratory analysis using data from the urban exposome of the Dutch population-based Occupational and Environmental Health Cohort Study (AMIGO), which represents a rich dataset comprising detailed individual-level information on various determinants (e.g. chemical, biological, physical), lifestyle factors, and health conditions for over 14,000 participants.

2. Methods

2.1. Study design and participants

We used data from the population-based Occupational and Environmental Health Cohort Study (AMIGO) in the Netherlands, established in 2011/2012 to investigate environmental and occupational determinants of diseases and symptoms in the Dutch adult population. The rationale, study design, and participant recruitment in AMIGO were described in detail previously (Slottje et al., 2014). In short, AMIGO participants were recruited from the general population in the Netherlands through the Primary Care Database of the Netherlands Institute for Health Services Research (NIVEL), which consists of routinely recorded data from health care providers to monitor health and utilization of health services in the Dutch population (Nivel Primary Care Database, 2022). The baseline sample includes 14,829 adults (16 % of those invited), aged 31–65 years, who filled in an online questionnaire in 2011/2012 and at follow-up in 2015 (n = 7,905; response rate 54 %). After the exclusion of participants with missing information on headache frequency at baseline and/or follow-up (n = 566), the study population included 7,339 participants who completed baseline and follow-up questionnaires. All cohort members participated voluntarily and gave informed consent prior to their inclusion.

2.2. Exposure factors

The urban exposome of AMIGO was described by Ohanyan et al. previously (Ohanyan et al., 2022). In short, the urban exposome relied on satellite data, monitoring stations, population registry-based data, and geospatial models to estimate participants' exposures at their place of residence at baseline (Martens et al., 2018). In this study, 88 exposures across 10 domains were considered encompassing air pollution (19 factors), quality of drinking water (29 factors), urbanicity and built environment (13 factors), green space density (2 factors), outdoor light at night (1 factor), urban temperature (2 factors), road traffic noise (1 factor), radiofrequency electromagnetic fields (2 factors), sociodemographic characteristics of the neighborhood (17 factors), technology use (2 factors). The list of exposure factors included in this study is provided in Supplementary Table 1.

2.3. Headache

Frequency of headache was self-reported at baseline and follow-up. As primary outcome, we defined the occurrence of weekly headache episodes (referred to as "weekly headache" for brevity, with response categories "yes", "no") according to the question: "How often do you get headache at the moment?" – response categories: "almost every day","5 or 6 days a week", "3–4 days a week", "once or twice a week" "1–2 days per month", "less often".

As secondary outcome, we included the occurrence of severe weekly headache episodes (response categories "yes", "no"). The Headache Impact Test (HIT-6) score with a cut-off of 56 points was used to define weekly occurrence of severe headache episodes. The HIT-6 is a tool used to measure the impact headaches have on one's ability to function in various aspects of daily life, including work, school, home, and social contexts. The score, ranging from 36 to 78 points, provides a measure of the degree to which headaches affect daily life and functioning, with higher scores indicating a more significant impact on the participant's overall life (Kosinski et al., 2003).

2.4. Covariates

We assessed the following covariates of the associations between the urban exposome and weekly headache: sex, age, highest level of education attained (elementary, secondary and higher), occupation (employed, unemployed), country of origin (the Netherlands, other), body mass index (BMI) group (normal or underweight, overweight or obese), alcohol consumption (never, former, current), smoking status (never, former, current), sleep disturbance index (Spritzer and Hays, 2003), general health indicator (good, poor), depression diagnosis (yes, no), painkiller use (yes, no).

2.5. Pre-processing of the urban exposome and descriptive statistics

We followed the same approach outlined by Ohanyan et al. to preprocess the urban exposome data in AMIGO (Ohanyan et al., 2022). In short, we excluded exposures that exhibited extremely low variability (9 exposures) or very strong correlations with other exposures (7 exposures). In the latter scenario, where two (or more) exposures showed a Spearman rank correlation coefficient ≥ 0.95, only one of the correlated variables was incorporated into the analysis and treated as a proxy for the other variable(s) (Supplementary Table 2).

Missing values were imputed for exposures and covariates only, and all exposures, covariates, and the study outcome were used as predictors. Thirty imputed datasets were generated through Multivariate Imputation via Chained Equations (MICE) and the imputed values were averaged across the generated datasets, given the considerable computational costs and the absence of a recognized approach to combine results from multiple imputed sets associated with the methods applied in this study.

To mitigate the potential impact of non-normal distribution of the exposures on the imputation process, we applied transformations (logarithmic or square root) to normalize the exposures before incorporating them into MICE, and then back-transformed them after the imputation (van Buuren and Groothuis-Oudshoorn, 2011; White et al., 2011).

Descriptive statistics of the study population were evaluated with regard to the covariates included in the study. We performed a correlation analysis of the urban exposome by visualizing the inter- and intra-group correlations across the 10 domains using the circos and matrix of correlations, respectively (Hernandez-Ferrer et al., 2022).

2.6. Statistical analysis

To evaluate the association between the urban exposome at baseline and weekly occurrence of headache episodes at follow-up, we first performed feature selection using the Boruta algorithm to screen our

dataset and identify relevant exposures for the outcome being investigated (Kursa and Rudnicki, 2010).

Boruta represents a powerful approach for the analysis of high-dimensional datasets that has recently gained popularity particularly in the context of microbiome and omics research (Degenhardt et al., 2017). This method is designed to identify relevant variables and is able to capture interactions and nonlinear associations in complex-dimensional scenarios. Boruta aims to identify all attributes that contribute to some extent to the classification problem based on the so-called *all-relevant* problem approach. This methodology stands in contrast to the *minimal-optimal* problem approach, which focuses on finding the smallest and non-redundant subset of features essential for optimal performance given a specific dataset (Nilsson et al., 2007).

Boruta works as a wrapper algorithm around random forest and operates by comparing the importance of each variable against that of shadow variables, which are randomly permuted versions of the original variables (Breiman, 2001; Liaw and Wiener, 2002). By conducting a series of random forest iterations, Boruta assigns importance scores to each variable, considering both the actual features and their shadow counterparts. Variables that consistently outperform their shadow versions are retained as important, while those that do not are deemed unimportant.

To address class imbalance, we applied the Boruta algorithm repeatedly (250 iterations using 1000 trees at each iteration) and downsampled 85 % of the minority group size without replacement in both groups comprising participants with and without weekly headaches at follow-up in order to obtain 250 different balanced datasets (More and Rana, 2017).

Finally, we retained the features that were labelled as "important" by Boruta in at least 80 % of the 250 iterations, emphasizing their stability in the selection process, and calculated variable importance by averaging the importance of the selected features across iterations.

To evaluate the generalizability of our results, we trained a random forest model on the features selected by Boruta. This evaluation was conducted on an *a priori* sampled independent test set, comprising 20 % of the original dataset. The corresponding Receiver Operating Characteristics (ROC) curve and Area-Under-the-Curve (AUC), along with a 95 % confidence interval (95 % CI), were estimated using 1000 bootstrap replicates.

To visualize the relationship between the response and predictors, and represent the average contribution of a feature value to the prediction (Molnar, 2020), we computed Shapley values by training a random forest model on the features identified by Boruta using the original dataset.

We estimated the causal effect of each exposure identified by Boruta on weekly occurrence of headache episodes by training causal forests. Causal forests represent an extension of random forests to estimate the average treatment effect (ATE) and corresponding standard error (SE) under the assumption of absence of confounding (i.e. the treatment assignment is independent of the potential outcome conditional on the confounders) allowing for covariate adjustment. The ATE represents the average of the difference in potential outcomes in a sample where everyone is treated versus the same sample where everyone is untreated (Jawadekar et al., 2023).

Specifically, when the treatment is continuous, we effectively estimate an average partial effect, which quantifies the change in the expected outcome due to a one-unit change in the treatment, given unconfoundedness. To ensure clarity and consistency in the language used, we will refer to treatment as exposure in the rest of the paper.

Briefly, the algorithm splits the data in order to maximize the difference across splits in the relationship between an outcome and an exposure variable uncovering variations in exposure effects across the sample. Causal forests resemble a randomized controlled trial and estimate exposure propensity weights to create a balanced covariate distribution between the exposed and control groups. It is important to note that, while causal forests identify heterogeneity in causal effects, they do not, per se, establish causation (Athey et al., 2018).

In detail, for each exposure selected by Boruta, we estimated causal forests adjusted for a set of covariates. In estimating causal forests, we used default parameters as they were shown to perform reasonably well with random forests (Athey and Wager, 2019).

To assess the fit of the causal forest, we first examined the distribution of the estimated exposure propensity weights to identify potential extreme values. Second, we explored heterogeneity by grouping observations according to whether their out-of-bag conditional average treatment effect (CATE) estimates (i.e. predictions) were above ("high" region) or below ("low" region) the median CATE estimate. Following this grouping, we calculated the difference in causal effects between regions along with the 95 % confidence interval (95 % CI) to gain insights about the overall strength of heterogeneity in the study population (Athey and Wager, 2019).

As secondary analyses, considering the transient nature of headaches in the population and our predefined interest in assessing whether exposure effects on weekly headaches at follow-up could be mediated by their occurrence at baseline, we trained additional causal forests. Specifically, we estimated the CATE representing the average of the difference in potential outcomes in a specific stratum of the population (here defined by presence/absence of weekly headache at baseline), where everyone in that stratum is exposed versus a scenario where everyone in the same stratum is unexposed (Jawadekar et al., 2023). Finally, to test the null hypothesis of no heterogeneity between the CATEs estimated for the two groups of weekly headache at baseline, we applied Student's *t*-test (Athey and Wager, 2019).

We conducted the following sensitivity analyses: first, we estimated causal forests with mutual adjustment under the assumption that the exposures may act as confounding factors for each other and are therefore not independent entities. This approach involved systematically estimating the causal effect of each exposure separately on weekly headache while simultaneously incorporating the remaining exposures into the adjustment set. This iterative process was repeated for each exposure identified by Boruta.

Second, we performed the Boruta feature selection by excluding possible mediators of the association between the urban exposome and weekly headache, namely weekly headache at baseline, general health indicator, sleep disturbance index, and self-reported painkiller use. Third, we replicated the feature selection by adding the perception of environmental factors, such as air pollutants and RF-EMFs, to the list of exposures assessed by Boruta, which may help to disentangle the relationship between actual exposures, their perceptions, and the onset of headache symptoms.

Fourth, we reproduced the Boruta feature selection by excluding participants reporting weekly headache at baseline to assess consistency of determinants of newly reported headaches.

Lastly, based on the results of the sensitivity analyses on Boruta, additional causal forests were trained accordingly.

The analyses were performed with the R statistical software, version 4.0.4, using the packages *mice*, *rexposome*, *Boruta*, *ranger*, and *grf*. Computing code related to all analyses presented is publicly available at https://github.com/eugeniotraini/headache_exposome.

3. Results

In AMIGO, the occurrence of weekly headache episodes at baseline and follow-up showed similar proportions (12.5 % and 11.1 %, respectively). However, out of the 814 participants reporting weekly headache at follow-up, only 55 % reported such headaches at baseline. At the beginning of the study, 5.4 % of participants reported experiencing severe headaches weekly, which decreased to 2 % at follow-up.

Baseline characteristics of the study population are presented in Table 1. Over half of the participants in AMIGO were women and mean age was 52 years old at the time of recruitment. Approximately 44 % of the participants had attained a high level of education, while 70 % were

Table 1
Characteristics of the participants at baseline (N = 7339).

Sex, n (%)	
Male	3499 (47.7)
Female	3840 (52.3)
Age (years)	
Mean (SD)	52.3 (9.00)
Highest level of education attained, n (%)	
High	3215 (43.8)
Low	4122 (56.2)
Missing	2 (0.0)
Occupation, n (%)	
Employed	5161 (70.3)
Unemployed	2178 (29.7)
Country of origin, n (%)	
The Netherlands	7048 (96.0)
Other	291 (4.0)
BMI group, n (%)	
Normal or underweight	3763 (51.3)
Overweight or obese	3576 (48.7)
Alcohol consumption, n (%)	
No	813 (11.1)
Yes	6524 (88.9)
Missing	2 (0.0)
Smoking status, n (%)	
No	6437 (87.7)
Yes	900 (12.3)
Missing	2 (0.0)
Sleep disturbance index	
Mean (SD)	26.5 (18.6)
Missing	23 (0.3 %)
General health indicator, n (%)	
Poor	1098 (15.0)
Good	6239 (85.0)
Missing	2 (0.0)
Depression diagnosis, n (%)	
No	6587 (89.8)
Yes	752 (10.2)
Painkiller use, n (%)	
No	6773 (92.3)
Yes	484 (6.6)
Missing	82 (1.1)
Weekly headache[a], n (%)	
No	6425 (87.5)
Yes	914 (12.5)
Severe weekly headache[a], n (%)	
No	6940 (94.6)
Yes	399 (5.4)

[a] At baseline.

employed. Nearly all participants, specifically 96 %, indicated the Netherlands as their country of origin. Around half of the study population (48.7 %) was classified as overweight or obese, and the proportion of alcohol users and smokers was 88.9 % and 12.3 %, respectively.

Overall, individuals in the AMIGO study reported a good state of health (85 %), low prevalence of painkiller use (6.6 %), and an average sleep disturbance index of 26.5 (on a scale from 0 to 100 with higher scores indicating more sleep disturbances or lower sleep quality). No relevant differences in the distribution of baseline characteristics of the study participants were observed when including those who did not complete the follow-up questionnaire (Supplementary Table 3). The proportion of missing values in the exposures was below 10 % with the highest occurrence observed for the percentage of inhabitants with non-western origins in the neighbourhood (9.7 %) (Supplementary Table 4).

The matrix of correlation plot shows that the strongest intra-group correlations were observed between air pollutants, urbanicity and built environmental variables, RF-EMFs, and socio-demographic area-level factors (Fig. 1). Drinking water components had the lowest intra-group correlations. The circos of correlation plot showed that green space density exhibited a negative inter-family correlation with air pollutants (Supplementary Fig. 1).

Results of the Boruta feature selection showed that five air pollutants (NO_2, NO_X, PM_{10}, Silicon in PM_{10}, Iron in $PM_{2.5}$), one urban temperature measure (heat island effect), five *a priori* defined covariates (age, depression diagnosis, painkiller use, general health indicator, sleep disturbance index), and weekly headache at baseline significantly influenced the reporting of weekly headache at follow-up.

Among those, weekly headache at baseline appeared to be the most important variable, followed by the remaining covariates (with age being the least significant among all selected features). The exposures, listed in descending order of importance, were NO_2, Silicon in PM_{10}, NO_X, Iron in $PM_{2.5}$, PM_{10}, and the heat island effect. Spearman correlation coefficients showed very strong correlations between the exposures selected by Boruta (Fig. 2).

The Shapley plots did not show strong associations between the selected environmental exposures (air pollutants and urban temperature measure) and weekly headache, and, on average, the contributions of individual features to the predicted outcome were modest (Fig. 3).

Being diagnosed with depression, using painkillers, reporting poor general health, experiencing weekly headaches at baseline, being older, and having difficulty sleeping all showed a substantial impact on reporting weekly headache at follow-up (Supplementary Fig. 2).

The ROC analysis of the random forest model including the features selected by Boruta produced an AUC of 0.82 (95 % CI: 0.75–0.88), indicating good discriminatory power in distinguishing individuals with and without weekly headache at follow-up in the independent test set.

Results from causal forests adjusted for age, depression diagnosis, painkiller use, general health indicator, sleep disturbance index, and weekly headache at baseline are presented in Table 2 and showed positive associations between each exposure at baseline and weekly headache at follow-up. In detail, NO_2 showed the largest effect with an ATE of 0.007 (SE = 0.004) per interquartile range (IQR) increase, followed by PM_{10} (ATE = 0.006 (SE = 0.004) per IQR increase), heat island effect (ATE = 0.006 (SE = 0.007) per 1C° increase), NO_x (ATE = 0.004 (SE = 0.004) per IQR increase), Iron in $PM_{2.5}$ (ATE = 0.003 (SE = 0.004) per IQR increase), and Silicon in PM_{10} (ATE = 0.003 (SE = 0.004) per IQR increase). Of note, concerning the highly correlated exposures identified during the pre-processing of the urban exposome, the use of causal forests with Copper in $PM_{2.5}$ serving as a proxy for Iron in $PM_{2.5}$, yielded results consistent with the main findings (Supplementary Table 5).

After conducting a visual inspection of the distribution of estimated propensity weights, no extreme values were identified (Supplementary Fig. 3). The assessment of heterogeneity in causal forests revealed some variation between the regions defined by “high” and “low” CATE estimates, though the strength of heterogeneity appeared to be modest (Supplementary Table 6).

The estimated causal effects conditional on weekly headache at baseline showed distinct patterns between participants who reported symptoms at baseline and those who did not. Specifically, as shown in Table 3, the CATE for participants without weekly headache at baseline was null or negative, whereas the effect observed in those who reported headache at baseline was positive. In detail, the CATE of NO_2 among participants who had weekly headache at baseline was 0.068 (SE = 0.027), whereas for those without symptoms was −0.003 (SE = 0.004). Similar patterns were displayed for all remaining exposures: the heat island effect showed an effect of 0.058 (SE = 0.042) for those who reported weekly headache at baseline, while the effect was null (CATE = 0.000 (SE = 0.006)) for those without weekly headache at baseline. Likewise, PM_{10} showed an effect of 0.054 (SE = 0.027) and 0.000 (SE = 0.004) in those with and without weekly headache at baseline, respectively. Positive yet weaker effects were observed among participants who reported weekly headache at baseline for Silicon in PM_{10} (CATE = 0.043 (SE = 0.027)), NO_x (CATE = 0.035 (SE = 0.022)), and Iron in $PM_{2.5}$ (CATE = 0.013 (SE = 0.026)). We rejected the null hypothesis of no heterogeneity between the CATEs estimated for the two groups of weekly headache at baseline for all exposures (Table 3).

Interestingly, causal forests spent, on average, 23 % of their splits on weekly headache at baseline, making it the most important variable among those included in the algorithm.

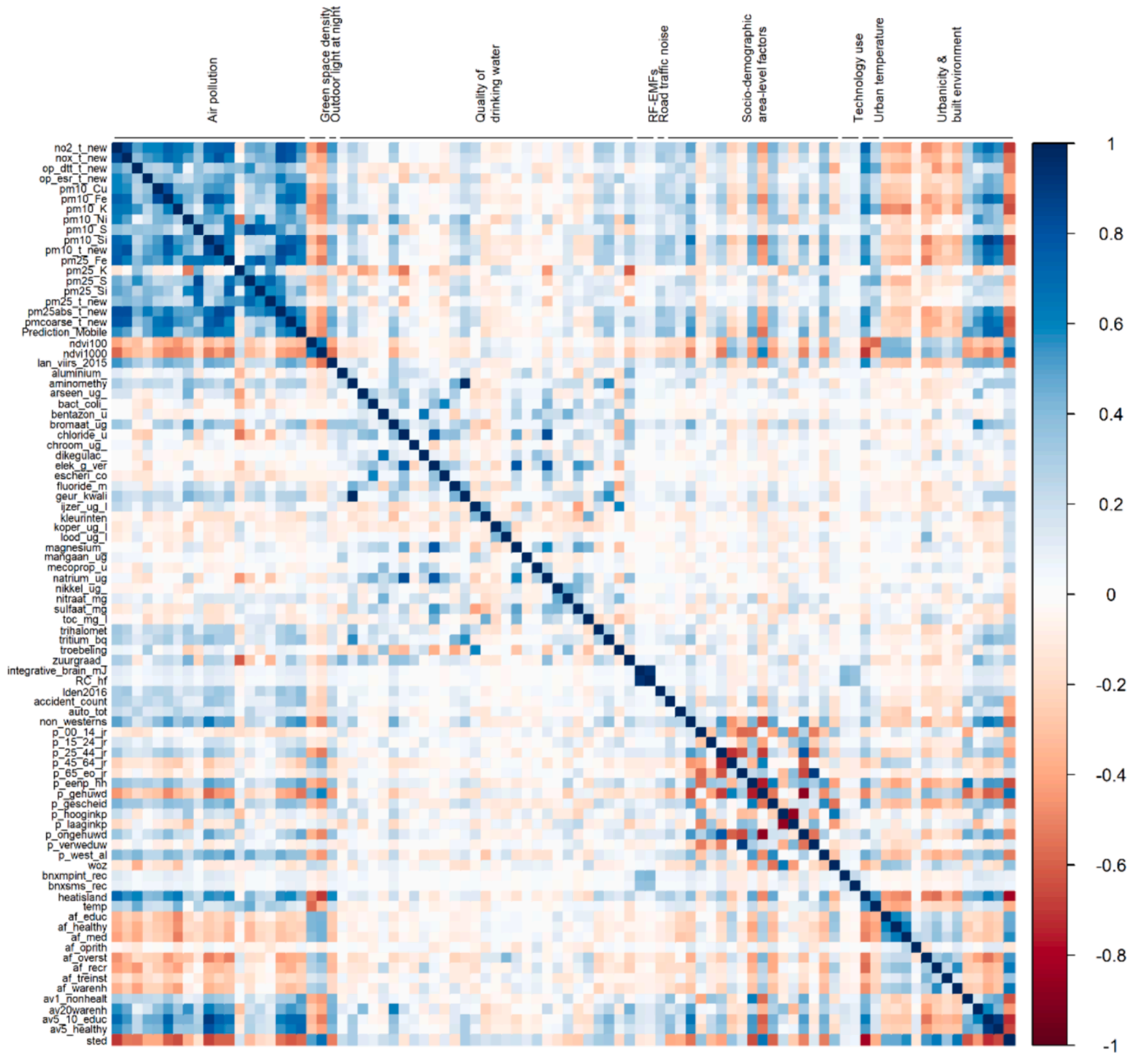

Fig. 1. Matrix of correlation plot showing the intra-family correlations between exposures at baseline.

Supplementary Table 6 displays the results of the mutually adjusted causal forests, where the exposures selected by Boruta were added to the adjustment set.

The mutually adjusted estimates showed reduced precision and some experienced a change in the direction of the effect. In detail, we estimated an ATE of 0.058 (SE = 0.022) for an increase in IQR in NO_2, followed by NO_x (ATE = 0.011 (SE = 0.018)), and PM_{10} (ATE = 0.002 (SE = 0.014)).

In contrast to the main results, the ATEs for Iron in $PM_{2.5}$, Silicon in PM_{10}, and heat island effect were negative (ATE = -0.016 (SE = 0.010); ATE = -0.045 (SE = 0.051); ATE = -0.019 (SE = 0.017), respectively) (Supplementary Table 7).

No extreme values were identified in the distribution of exposure propensity weights (Supplementary Fig. 4), and the comparison between the regions characterized as "high" and "low" CATE estimates aligned with the main results, indicating the presence of some heterogeneity in the dataset (Supplementary Table 8).

Results from the sensitivity analyses using Boruta were in line with the main findings: specifically, by excluding the possible mediators of the association between the urban exposome and weekly headache at follow-up (weekly headache at baseline, general health indicator, sleep disturbance index, and self-reported painkiller use), Boruta retained, in order of decreasing variable importance, depression diagnosis, NO_2, NO_x, Iron in $PM_{2.5}$, and Silicon in PM_{10}. After including the perception of environmental exposures, Boruta selected two air pollutants, namely PM_{coarse} and Potassium in PM_{10}, and road traffic noise, in addition to the features already identified in the main analysis.

Based on these results, we trained additional causal forests including PM_{coarse}, Potassium in PM_{10}, and road traffic noise as exposures, and age, depression diagnosis, and weekly headache at baseline as adjustment factors. The causal effect associated with PM_{coarse}, Potassium in PM_{10}, and road traffic noise on the reporting of weekly headache was 0.005 (SE = 0.004), 0.002 (SE = 0.004), 0.008 (SE = 0.007), respectively (Supplementary Table 9). After excluding participants with

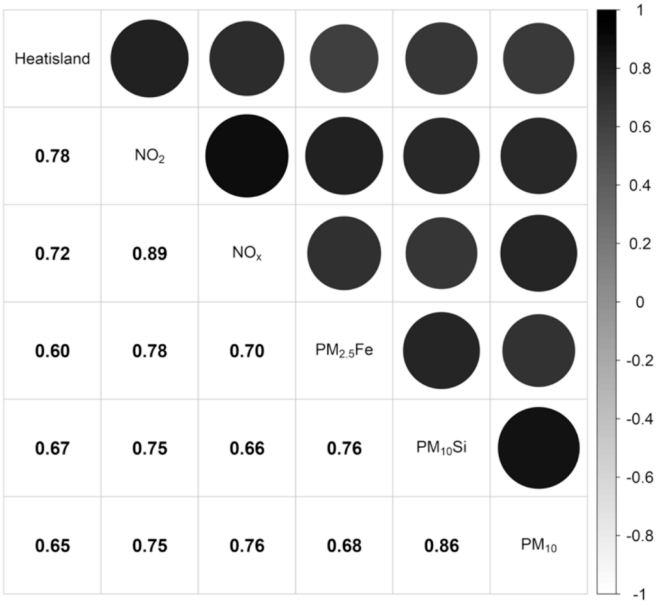

Fig. 2. Spearman rank correlation coefficients and correlation plot of the exposures selected by Boruta. Darker colours and larger circles indicate higher positive correlation levels.

weekly headache at baseline, Boruta only selected depression diagnosis and sleep disturbance index but none of the exposures.

Finally, regarding severe weekly headaches, it was found that five predetermined covariates (age, depression diagnosis, painkiller use, general health indicator, sleep disturbance index) and the presence of severe weekly headache at baseline influenced reporting at follow-up. However, none of the environmental exposures were identified by Boruta as contributing factors.

4. Discussion

In this study, we prospectively explored the urban exposome with the aim to identify factors associated with reporting of weekly headache episodes by analyzing data from a large cohort of individuals sampled within the Dutch general population.

We applied Boruta, a feature selection algorithm designed to identify relevant variables in complex highly dimensional settings, and causal

Table 2
Average treatment effects (ATEs) and related standard errors (SEs) estimated with causal forests for each exposure separately, adjusted for age, depression diagnosis, painkiller use, general health indicator, sleep disturbance index and weekly headache at baseline.

Exposure	ATE[a] (SE)
NO_2 (μg/m^3)	0.007 (0.004)
PM_{10} (μg/m^3)	0.006 (0.004)
Heat island effect (C◦)	0.006 (0.007)
NO_x (μg/m^3)	0.004 (0.004)
Iron in $PM_{2.5}$ (ng/m^3)	0.003 (0.004)
Silicon in PM_{10} (ng/m^3)	0.003 (0.004)

[a] Results for air pollutants indicate changes per interquartile range (IQR) increase in mean air pollution exposure.

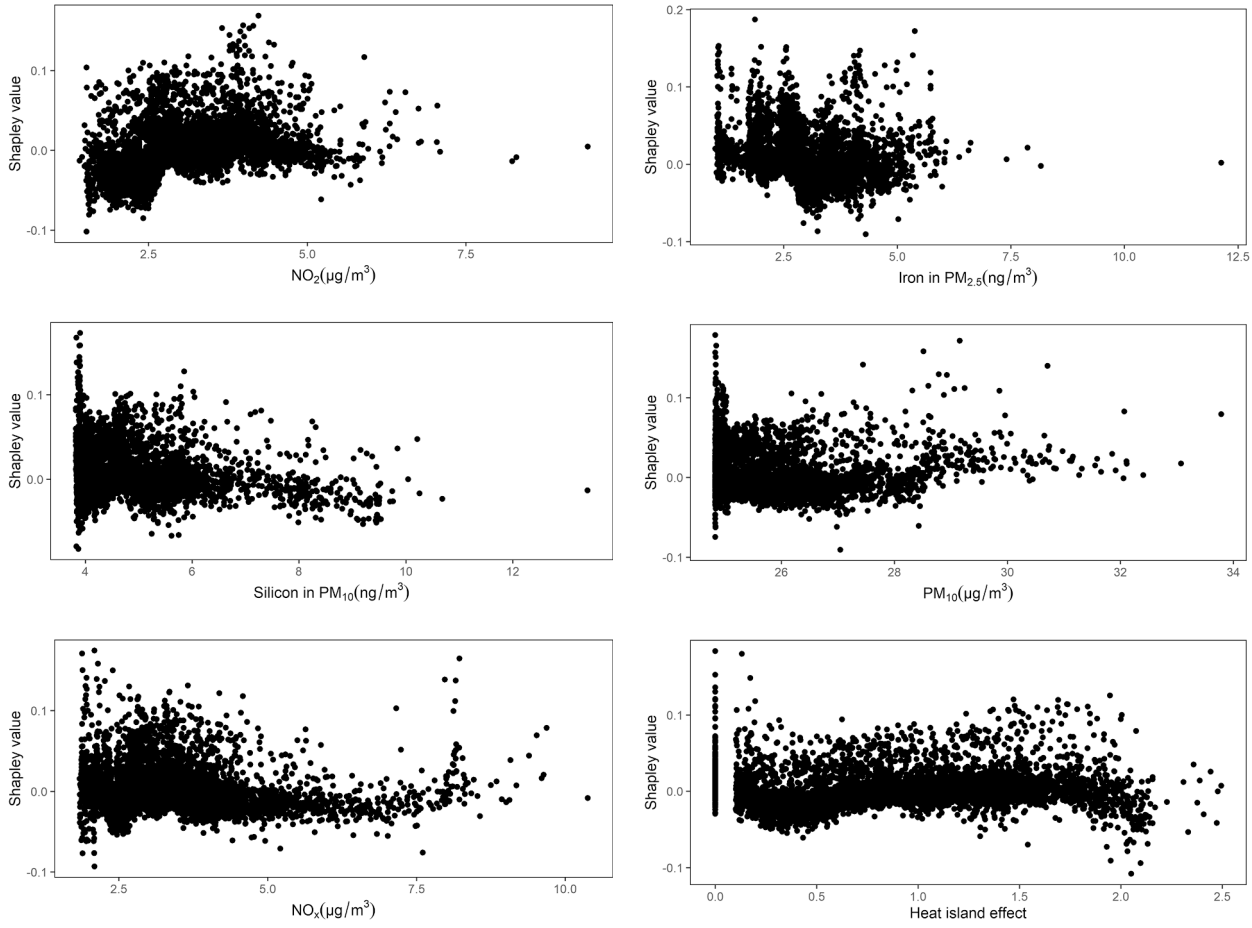

Fig. 3. Shapley plot illustrations of the exposures selected by Boruta.

Table 3
Conditional average treatment effects (CATEs) on weekly headache at baseline and related standard errors (SEs) estimated with causal forests for each exposure separately, adjusted for age, depression diagnosis, painkiller use, general health indicator, and sleep disturbance index.

Exposure	CATE[a] (SE) in the subsample of participants with weekly headache at baseline	CATE[a] (SE) in the subsample of participants without weekly headache at baseline	t-value[b]
NO_2 (μg/m^3)	0.068 (0.027)	−0.003 (0.004)	7.151***
PM_{10} (μg/m^3)	0.054 (0.027)	0.000 (0.004)	6.761***
Heat island effect (C◦)	0.058 (0.042)	0.000 (0.006)	7.645***
NO_x (μg/m^3)	0.035 (0.022)	−0.004 (0.004)	4.471***
Iron in $PM_{2.5}$ (ng/m^3)	0.013 (0.026)	−0.001 (0.004)	6.209***
Silicon in PM_{10} (ng/m^3)	0.043 (0.027)	−0.003 (0.004)	3.916***

[a] Results for air pollutants indicate changes per interquartile range (IQR) increase in mean air pollution exposure.
[b] Null hypothesis of no heterogeneity.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

forests, a statistical method for estimating causal effects of exposures under the assumption of absence of confounding.

Our results consistently showed that, out of 88 factors characterizing the urban exposome, air pollutants and urban temperature measures appeared to contribute most to the reporting of weekly headache at follow-up.

In particular, exposure to NO_2 at baseline was the most important environmental factor identified by Boruta in relation to reporting of weekly headache at follow-up, followed by Silicon in PM_{10}, NO_X, Iron in $PM_{2.5}$, PM_{10}, and the heat island effect. Finally, causal forests estimated the largest effect on the reporting of weekly headache at follow-up for NO_2, PM_{10}, and the heat island effect. While the magnitude of these estimates may be modest, even minor associations can carry important implications when considering the widespread exposure to air pollution and the higher temperatures in urban areas on population health.

Previous research showed that exposure to air pollution may act as trigger in the onset of headaches in the population (Nattero and Enrico, 1996; Szyszkowicz, 2008; Vodonos et al., 2015). The mechanism by which this occurs, however, is not fully understood. Air pollutants can impact the nervous system by entering through the olfactory and lower respiratory tracts. This process involves the direct initiation of inflammatory processes and the release of cytokines, allowing them to reach the central nervous system, triggering symptoms (Block and Calderón-Garcidueñas, 2009). Previous studies conducted in large urban areas in Canada and China suggested that particularly short-term exposure to NO_2 was associated to an increase in the number of emergency department visits for headaches (Szyszkowicz, 2008; Xu et al., 2023). A study conducted in the urban area of Turin in Italy, exploring the relationship between exposure to air pollutants and meteorological factors in relation to headaches, found that simultaneous exposure to carbon monoxide (CO) and NO_2 increased incidence of headache attacks along with wind velocity which was linked to frequency and severity of episodes (Nattero and Enrico, 1996).

It is noteworthy that, despite extensive research indicating positive associations between migraines and air pollution exposure, and particularly NO_2 (Portt et al., 2023; Elser et al., 2021; Lee et al., 2018), exploration of headaches is still limited. In this regard, our study highlights the link between exposure to air pollution and the occurrence of headaches, suggesting that measures aimed at decreasing emissions could be beneficial to reduce the impact of pollutants on symptoms. Furthermore, Iron in $PM_{2.5}$ and Silicon in PM_{10} were linked to reporting of headaches in our study population, although their impact appeared to be less prominent compared to NO_2, PM_{10}, and NO_x. To the best of our knowledge, this is the first study to identify specific fine particulate components in relation to headache, and future research should investigate these associations further to elucidate the contribution of individual components, both independently and in combination, to headache occurrence.

Our results showed that the heat island effect, that is the temperature difference between an urban area and the rural surrounding, was causally related with more frequent weekly headache at follow-up. Previously, the increase in temperature, particularly in densely urban areas and especially during summer heat waves, has been linked to immediate body reactions such as heavy sweating, dehydration, skin rashes, and headaches, among others (O'Malley et al., 2015; Tong et al., 2021; Arifwidodo and Chandrasiri, 2020; Aghamohammadi et al., 2021).

In the urban exposome, we used satellite pictures to estimate the surface temperature on a hot day as urban heat island effect is best assessed during heatwaves. Therefore, the effect that we observed in AMIGO could be partially explained by some residual urbanicity effect, which may include air pollutants and green space density.

Given the strong interplay between air pollutants and other environmental determinants assessed in the urban exposome, such as road traffic noise and urban temperature, the effect that we found may have, independently or in conjunction with air pollution exposure, exacerbated the reporting of headache.

In our cohort, about half of the participants who indicated to suffer from headache at baseline did not report the same at follow-up, meaning that headache represents a transient condition in the study population. To explore how the exposures may affect different subgroups of participants, specifically those with and without symptoms at the baseline, and therefore improve our understanding of the potential underlying mechanisms that triggered the symptoms, we estimated causal effects conditional on weekly headache at baseline. Interestingly, the estimated effects appeared to be mediated by weekly headache symptoms at baseline. Moreover, baseline weekly headache emerged as the most important variable in the causal forests. These suggest the potential existence of a vulnerable subpopulation, represented by those reporting symptoms at baseline, that is more susceptible and therefore at a higher risk of adverse health outcomes if exposed to air pollution and heat island effect.

In our study, we identified a subset of exposures from the urban exposome which contributed to the occurrence of headache in the population, and estimated the magnitude of their effect under the assumption of absence of confounding using a combination of state-of-the-art statistical methods that, in part, were previously identified as valid tools to address the complexity of the exposome (Ohanyan et al., 2022; Maitre et al., 2022).

Results from the main analyses indicated that each exposure at baseline identified by Boruta was positively associated with reporting of headache at follow-up. In the mutually adjusted models, the estimates of the causal effects showed some increase for NO_2 and NO_x but with reduced precision, and the direction of the effects was not always consistent with the main results. Spearman correlation coefficients showed very strong correlations (ranging from 0.60 to 0.89) between the exposures selected by Boruta and included in the causal forests.

In methods that rely on propensity scores to balance covariates between exposed and unexposed, such as causal forests, many issues that arise with traditional regression modelling, such as multicollinearity, should no longer be a threat to validity (Arbour et al., 2014; McMurry et al., 2015).

Based on our results, multicollinearity clearly affected the precision of effect estimates produced by causal forests, given the larger standard

error associated with the estimates in the mutually adjusted models.

Furthermore, the balancing property of propensity scores which assumes that, conditional on the propensity score, the distribution of observed covariates is expected to be similar between the treated and untreated groups, is only true if the propensity scores are relatively well-behaved and no extreme values are present (Lee et al., 2011). However, following an inspection of the distribution of exposure propensities, this did not appear to be the situation in our analysis.

Our study has strengths: first, to our knowledge, this is the first study conducted within the exposome framework to explore the association between the urban exposome and headache. Given the high prevalence of individuals reporting recurrent or chronic headaches in the population, our study provides important insights into the relationship between environmental stressors that are ubiquitous in urban areas and the occurrence of headache symptoms. Our results aim to support the formulation of more tailored public health interventions targeting air quality improvement and a healthier urban environment in order to reduce the burden of headache in the population.

Second, we used data from a large prospective cohort of Dutch individuals, and detailed information about individual-level exposures, including perceived exposures, and neighborhood characteristics, all elements that strengthen the robustness and facilitate the causal interpretation of our results. With regard to generalizability of the data, compared to the general Dutch population, AMIGO participants consisted of more females and older subjects, although no indications of systematic health-related participation bias based on morbidity and associated lifestyle information such as smoking and medication use was found (Slottje et al., 2014).

Third, we used a combination of cutting-edge statistical techniques, that is Boruta and causal forests, to explore the urban exposome in relation to headache. In particular, causal forests and random forest, upon which Boruta is built, were previously identified as valuable tools to study the complexity characterizing exposome research and show good interpretability of the results (Maitre et al., 2022). Furthermore, training Boruta in iterations, despite being time consuming and computationally intensive, helped mitigate the effects of class imbalance present in our dataset, and ensured stability as well as generalizability of our results. In conclusion, we showed that the combination of statistical methods used in this study represents a robust approach to identify influential predictors, particularly in highly dimensional settings, and generate accurate machine learning models to estimate causal effects (Kursa and Rudnicki, 2010; Degenhardt et al., 2017; Athey and Wager, 2019).

Our study has some limitations: first, weekly headache was self-reported by the participants, which may be prone to recall bias or over/under reporting of symptoms. However, we assessed weekly headache using the validated HIT-6 questionnaire, which is considered a reliable and valid tool for discriminating headache impact in daily life, and it is employed as a screening tool in clinical practice (Kosinski et al., 2003). Furthermore, we evaluated severity of weekly headaches as secondary outcome to further strengthen our findings, although results of this analysis did not lead to the identification of any specific environmental exposures associated to the outcome, likely due to diminished statistical power and even more problematic class imbalance compared to the primary endpoint.

Second, information on weekly headache was available at baseline and follow-up, and no information was available in between. As a result, the outcome assessment may not precisely capture symptoms occurring between these two time points, especially for a transient condition like headache. In future studies, it would be beneficial to confirm the associations that we found between air pollutants and the urban temperature and headache exploring the dynamics of these associations over time.

Third, the exposures included in the urban exposome of AMIGO, such as the air pollutants, were modeled at the home address of the participants. As a consequence, it was impossible to quantify the exposure levels in places where participants could have spent some of their time during the day or when, for example, commuting between work and home. In fact, in cohorts such as AMIGO, directly measuring exposure levels for the single participant proves impractical due to the large sample size and the high costs associated. As a result, it is common to rely on exposure modeling, such as land-use regression models to estimate air pollution levels, which might introduce additional complexity due to the use of shared predictors that may lead to stronger correlations between exposures than those existing in the real world (Szpiro and Paciorek, 2013). In addition, in our study we did not directly evaluate residential self-selection bias, where the decision to relocate is influenced by various factors such as age, ethnicity, professional or life choices, and socioeconomic status. This dynamic may ultimately result in changes in environmental exposures across different life stages (Saucy et al., 2023). Given the complex interplay of these factors with the exposures assessed in AMIGO, we cannot rule out the possibility of residual bias in our dataset originating from residential self-selection.

Finally, despite recent developments in causal inference methods for multiple exposures (Williams and Crespi, 2020), we acknowledge a substantial gap in statistical methods for estimating the effect of multiple exposures, particularly in situations where these are represented by a combination of continuous and categorical exposures, as is common in the context of the urban exposome, and high correlation levels between exposures are present. Nevertheless, the approach followed in this study allowed us to identify a group of exposures involved in the exposome-outcome association and estimate the direct effect of single exposures on reporting of headache controlling for potential confounding variables to obtain more accurate estimates of causal effects.

5. Conclusions

Our study indicated that the exposure to environmental stressors, in particular air pollutants and urban heat island effects, contributed to reporting of weekly headache episodes in our population. Given the high global burden associated with headache, understanding the role of environmental factors becomes imperative not only for advancing our comprehension of the mechanisms generating symptoms but also for formulating effective preventing strategies.

Source of funding

We acknowledge financial support from the EXPANSE (EC H2020, grant agreement No 874627) and EXPOSOME-NL. EXPOSOME-NL is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant number 024.004.017).

We further acknowledge financial support by The Netherlands Organization for Health Research (ZonMw) within the program on Epidemiological Perception Research in the field of Electromagnetic Fields and Health under grant numbers 8520000, 85200002 and 85200003.

CRediT authorship contribution statement

Eugenio Traini: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lützen Portengen:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Haykanush Ohanyan:** Writing – review & editing, Methodology, Data curation, Conceptualization. **Robert van Vorstenbosch:** Writing – review & editing, Software, Methodology, Conceptualization. **Roel Vermeulen:** Writing – review & editing, Supervision, Resources, Conceptualization. **Anke Huss:** Writing – review & editing, Supervision, Resources, Conceptualization.

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.

Data availability

Data will be made available on request.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2024.108776.

References

Aghamohammadi, N., Fong, C.S., Idrus, M.H.M., Ramakreshnan, L., Sulaiman, N.M., 2021. Environmental heat-related health symptoms among community in a tropical city. Sci. Total Environ. 15 (782), 146611.

Andrianou, X.D., Makris, K.C., 2018. The framework of urban exposome: Application of the exposome concept in urban health studies. Sci. Total Environ. 15 (636), 963–967.

Arbour, D., Marazopoulou, K., Garant, D., Jensen, D., 2014. Propensity score matching for causal inference with relational data. CEUR Workshop Proceedings. 1 (1274), 25–34.

Arifwidodo, S.D., Chandrasiri, O., 2020. Urban heat stress and human health in Bangkok. Thailand. Environ Res. 185, 109398.

Ashina, S., Terwindt, G.M., Steiner, T.J., Lee, M.J., Porreca, F., Tassorelli, C., et al., 2023. Medication overuse headache. Nat. Rev. Dis. Primers 9 (1), 5.

Athey, S., Wager, S., 2019. Estimating Treatment Effects with Causal Forests: An Application [Internet]. arXiv [cited 2024 Jan 11]. Available from: http://arxiv.org/abs/1902.07409.

Athey, S., Tibshirani, J., Wager, S., 2018. Generalized Random Forests [Internet]. arXiv; [cited 2024 Jan 11]. Available from: http://arxiv.org/abs/1610.01271.

Block, M.L., Calderón-Garcidueñas, L., 2009. Air Pollution: Mechanisms of Neuroinflammation & CNS Disease. Trends Neurosci. 32 (9), 506–516.

Breiman, L., 2001. Random Forests. Mach. Learn. 45 (1), 5–32.

Dales, R.E., Cakmak, S., Vidal, C.B., 2009. Air Pollution and Hospitalization for Headache in Chile. Am. J. Epidemiol. 170 (8), 1057–1066.

Degenhardt, F., Seifert, S., Szymczak, S., 2017. Evaluation of variable selection methods for random forests and omics data sets. Brief. Bioinform. 20 (2), 492–503.

Di Lorenzo, C., Santorelli, F.M., van den Maagdenberg, A.M.J.M., 2015. Genetics of Headache. In: Ashina, M., Geppetti, P. (Eds.), Pathophysiology of Headaches: From Molecule to Man [Internet]. Springer International Publishing, Cham, pp. 83–99. https://doi.org/10.1007/978-3-319-15621-7_4.

Donma, O., Donma, M.M., 2002. Association of headaches and the metals. Biol. Trace Elem. Res. 90 (1), 1–14.

Elser, H., Morello-Frosch, R., Jacobson, A., Pressman, A., Kioumourtzoglou, M.A., Reimer, R., et al., 2021. Air pollution, methane super-emitters, and oil and gas wells in Northern California: the relationship with migraine headache prevalence and exacerbation. Environ. Health 20 (1), 45.

Friedman, D.I., De Ver, D.T., 2009. Migraine and the Environment. Headache: J. Head Face Pain 49 (6), 941–952.

GBD 2016 Headache Collaborators, 2018. Global, regional, and national burden of migraine and tension-type headache, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 17 (11), 954–976.

Hernandez-Ferrer, C., Gonzalez, J., 2022. Escribà-Montagut X. rexposome: Exposome exploration and outcome data analysis. R Package Version 1.18.2.

Holzhammer, J., Wöber, C., 2006. Alimentary trigger factors that provoke migraine and tension-type headache. Schmerz 20 (2), 151–159.

Jawadekar, N., Kezios, K., Odden, M.C., Stingone, J.A., Calonico, S., Rudolph, K., et al., 2023. Practical Guide to Honest Causal Forests for Identifying Heterogeneous Treatment Effects. Am. J. Epidemiol. 192 (7), 1155–1165.

Kosinski, M., Bayliss, M.S., Bjorner, J.B., Ware, J.E., Garber, W.H., Batenhorst, A., et al., 2003. A six-item short-form survey for measuring headache impact: the HIT-6. Qual. Life Res. 12 (8), 963–974.

Kursa, M.B., Rudnicki, W.R., 2010. Feature Selection with the Boruta Package. J. Stat. Softw. 16 (36), 1–13.

Lee, B.K., Lessler, J., Stuart, E.A., 2011. Weight trimming and propensity score weighting. PLoS One 6 (3), e18174.

Lee, H., Myung, W., Cheong, H.K., Yi, S.M., Hong, Y.C., Cho, S.I., et al., 2018. Ambient air pollution exposure and risk of migraine: Synergistic effect with high temperature. Environ. Int. 1 (121), 383–391.

Lee, M., Ohde, S., Urayama, K.Y., Takahashi, O., Fukui, T., 2018. Weather and Health Symptoms. Int. J. Environ. Res. Public Health 15 (8), 1670.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News. 2 (3), 18–22.

Maitre, L., Guimbaud, J.B., Warembourg, C., Güil-Oumrait, N., Petrone, P.M., Chadeau-Hyam, M., et al., 2022. State-of-the-art methods for exposure-health studies: Results from the exposome data challenge event. Environ. Int. 1 (168), 107422.

Martens, A.L., Reedijk, M., Smid, T., Huss, A., Timmermans, D., Strak, M., et al., 2018. Modeled and perceived RF-EMF, noise and air pollution and symptoms in a population cohort. Is perception key in predicting symptoms? Sci. Total Environ. 15 (639), 75–83.

McMurry, T.L., Hu, Y., Blackstone, E.H., Kozower, B.D., 2015. Propensity scores: Methods, considerations, and applications in the *Journal of Thoracic and Cardiovascular Surgery*. J. Thorac. Cardiovasc. Surg. 150 (1), 14–19.

Molarius, A., Tegelberg, A., Ohrvik, J., 2008. Socio-economic factors, lifestyle, and headache disorders - a population-based study in Sweden. Headache 48 (10), 1426–1437.

Molnar, C., 2020. Shapley Values | Interpretable Machine Learning. Interpretable [Internet]. 2020. Available from: Mach. Learn. 5.9-5.10 https://christophm.github.io/interpretable-ml-book/shapley.html.

More, A.S., Rana, D.P., 2017. Review of random forest classification techniques to resolve data imbalance. In: 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM) [Internet]. [cited 2024 Mar 7]. p. 72–78. Available from: https://ieeexplore.ieee.org/document/8122151.

Nash, J.M., Thebarge, R.W., 2006. Understanding psychological stress, its biological processes, and impact on primary headache. Headache 46 (9), 1377–1386.

Nattero, G., Enrico, A., 1996. Outdoor pollution and headache. Headache 36 (4), 243–245.

Nikiforow, R., Hokkanen, E., 1978. An Epidemiological Study of Headache In an Urban and a Rural Population In Northern Finland. Headache: J. Head Face Pain 18 (3), 137–145.

Nilsson, R., Peña, J.M., Björkegren, J., Tegnér, J., 2007. Consistent Feature Selection for Pattern Recognition in Polynomial Time. J. Mach. Learn. Res. 8 (21), 589–612.

Nivel Primary Care Database, Nivel [Internet]. 2022 [cited 2022 Aug 5]. Available from: https://www.nivel.nl/en/nivel-zorgregistraties-eerste-lijn/nivel-primary-care-database.

O'Malley, C., Piroozfar, P., Farr, E.R.P., Pomponi, F., 2015. Urban Heat Island (UHI) mitigating strategies: A case-based comparative analysis. Sustain. Cities Soc. 1 (19), 222–235.

Ohanyan, H., Portengen, L., Huss, A., Traini, E., Beulens, J.W.J., Hoek, G., et al., 2022. Machine learning approaches to characterize the obesogenic urban exposome. Environ. Int. 158, 107015.

Ohanyan, H., Portengen, L., Kaplani, O., Huss, A., Hoek, G., Beulens, J.W.J., et al., 2022. Associations between the urban exposome and type 2 diabetes: Results from penalised regression by least absolute shrinkage and selection operator and random forest models. Environ. Int. 1 (170), 107592.

Portt, A.E., Orchard, C., Chen, H., Ge, E., Lay, C., Smith, P.M., 2023. Migraine and air pollution: A systematic review. Headache: J. Head Face Pain 63 (9), 1203–1219.

Prince, P.B., Rapoport, A.M., Sheftell, F.D., Tepper, S.J., Bigal, M.E., 2004. The effect of weather on headache. Headache 44 (6), 596–602.

Rains, J.C., Poceta, J.S., 2012. Sleep-Related Headaches. Neurol. Clin. 30 (4), 1285–1298.

Raucci, U., Boni, A., Evangelisti, M., Della Vecchia, N., Velardi, M., Ursitti, F., et al., 2021. Lifestyle Modifications to Help Prevent Headache at a Developmental Age [cited 2024 Jan 8];11. Available from: Front. Neurol. [Internet] https://www.frontiersin.org/articles/10.3389/fneur.2020.618375.

Russell, M.B., Šaltytš-Benth, J., Levi, N., 2006. Are infrequent episodic, frequent episodic and chronic tension-type headache inherited? A population-based study of 11 199 twin pairs. J. Headache Pain 7 (3), 119–126.

Saucy, A., Gehring, U., Olmos, S., Delpierre, C., de Bont, J., Gruzieva, O., et al., 2023. Effect of residential relocation on environmental exposures in European cohorts: An exposome-wide approach. Environ. Int. 173, 107849.

Slottje, P., Yzermans, C.J., Korevaar, J.C., Hooiveld, M., Vermeulen, R.C.H., 2014. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in The Netherlands. BMJ Open 4 (11), e005858.

Spritzer, K., Hays, R., 2003. MOS Sleep Scale: A Manual for Use and Scoring. Version 1.0. RAND Corporation, Los Angeles, CA.

Steiner, T.J., World Headache Alliance, 2004. Lifting the burden: The global campaign against headache. Lancet Neurol. 3 (4), 204–205.

Szpiro, A.A., Paciorek, C.J., 2013. Measurement error in two-stage analyses, with application to air pollution epidemiology. Environmetrics 24 (8), 501–517.

Szyszkowicz, M., 2008. Air pollution and daily emergency department visits for headache in Montreal. Canada. Headache. 48 (3), 417–423.

Tong, S., Prior, J., McGregor, G., Shi, X., Kinney, P., 2021. Urban heat: an increasing threat to global health. BMJ 25 (375), n2467.

Ulrich, V., Gervil, M., Olesen, J., 2004. The relative influence of environment and genes in episodic tension-type headache. Neurology 62 (11), 2065–2069.

van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. J. Stat. Softw. 12 (45), 1–67.

Vermeulen, R., Schymanski, E.L., Barabási, A.L., Miller, G.W., 2020. The exposome and health: where chemistry meets biology. Science 367 (6476), 392–396.

Vodonos, A., Novack, V., Zlotnik, Y., Ifergane, G., 2015. Ambient air pollution, weather and daily emergency department visits for headache. Cephalalgia 35 (12), 1085–1091.

White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: Issues and guidance for practice. Stat. Med. 30 (4), 377–399.

Wild, C.P., 2005. Complementing the genome with an 'exposome': the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol. Biomark. Prev. 14 (8), 1847–1850.

Williams, J.R., Crespi, C.M., 2020. Causal inference for multiple continuous exposures via the multivariate generalized propensity score. arXiv preprint arXiv:2008.13767.

Winter, A.C., Hoffmann, W., Meisinger, C., Evers, S., Vennemann, M., Pfaffenrath, V., et al., 2011. Association between lifestyle factors and headache. J. Headache Pain 12 (2), 147–155.

Xu, H., Xu, M., Jc, Z., Ye, F., Liu, X., Liu, Y., et al., 2023. Short-term environmental nitrogen dioxide exposure and neurology clinic visits for headaches, a time-series study in Wuhan, China. BMC Public Health 23 (1), 828.