A Section 508–conformant HTML version of this article
is available at https://doi.org/10.1289/EHP13393.

# Exposome-Wide Association Study of Body Mass Index Using a Novel Meta-Analytical Approach for Random Forest Models

*Haykanush Ohanyan,[1,2,3,4] Mark van de Wiel,[2,3] Lützen Portengen,[1] Alfred Wagtendonk,[1,2,3] Nicolette R. den Braver,[2,3,4] Trynke R. de Jong,[5] Monique Verschuren,[6,7] Katja van den Hurk,[8,9] Karien Stronks,[9] Eric Moll van Charante,[9] Natasja M. van Schoor,[2,10] Coen D.A. Stehouwer,[11,12] Anke Wesselius,[13,14] Annemarie Koster,[15,16] Margreet ten Have,[17] Brenda W.J.H. Penninx,[18,19] Marieke F. van Wier,[20,21] Irina Motoc,[2,22] Albertine J. Oldehinkel,[23] Gonneke Willemsen,[24] Dorret I. Boomsma,[24] Mariëlle A. Beenackers,[25] Anke Huss,[1] Martin van Boxtel,[26] Gerard Hoek,[1] Joline W.J. Beulens,[2,3,4,5] Roel Vermeulen,[1,5] and Jeroen Lakerveld[1,2,3,4]*

[1]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands
[2]Department of Epidemiology and Data Science, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[3]Health Behaviours and Chronic Diseases, Amsterdam Public Health, Amsterdam, the Netherlands
[4]Upstream Team, Amsterdam UMC, VU University Amsterdam, Amsterdam, the Netherlands
[5]Lifelines Cohort & Biobank, Roden, the Netherlands
[6]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands
[7]National Institute for Public Health and the Environment, Bilthoven, the Netherlands
[8]Donor Medicine Research – Donor Studies, Sanquin Research, Amsterdam, the Netherlands
[9]Department of Public and Occupational Health, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
[10]Aging & Later Life, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
[11]School for Cardiovascular Diseases (CARIM), Maastricht University, Maastricht, the Netherlands
[12]Department of Internal Medicine, Maastricht University Medical Center, Maastricht, the Netherlands
[13]School for Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University, Maastricht, the Netherlands
[14]Department of Epidemiology, Maastricht University, Maastricht, the Netherlands
[15]Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, the Netherlands
[16]Department of Social Medicine, Maastricht University, Maastricht, the Netherlands
[17]Trimbos-Instituut, Netherlands Institute of Mental Health and Addiction, Utrecht, the Netherlands
[18]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[19]Mood, Anxiety, Psychosis, Sleep & Stress Program, Mental Health Program and Amsterdam Neuroscience, Amsterdam Public Health, Amsterdam, the Netherlands
[20]Department of Otolaryngology—Head and Neck Surgery, section Ear and Hearing, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[21]Quality of Care, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
[22]Amsterdam Reproduction & Development Research Institute, Amsterdam, the Netherlands
[23]Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[24]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[25]Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
[26]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience (MHeNs), Maastricht University, Maastricht, the Netherlands

**Background:** Overweight and obesity impose a considerable individual and social burden, and the urban environments might encompass factors that contribute to obesity. Nevertheless, there is a scarcity of research that takes into account the simultaneous interaction of multiple environmental factors.

**Objectives:** Our objective was to perform an exposome-wide association study of body mass index (BMI) in a multicohort setting of 15 studies.

**Methods:** Studies were affiliated with the Dutch Geoscience and Health Cohort Consortium (GECCO), had different population sizes (688–141,825), and covered the entire Netherlands. Ten studies contained general population samples, others focused on specific populations including people with diabetes or impaired hearing. BMI was calculated from self-reported or measured height and weight. Associations with 69 residential neighborhood environmental factors (air pollution, noise, temperature, neighborhood socioeconomic and demographic factors, food environment, drivability, and walkability) were explored. Random forest (RF) regression addressed potential nonlinear and nonadditive associations. In the absence of formal methods for multimodel inference for RF, a rank aggregation-based meta-analytic strategy was used to summarize the results across the studies.

**Results:** Six exposures were associated with BMI: five indicating neighborhood economic or social environments (average home values, percentage of high-income residents, average income, livability score, share of single residents) and one indicating the physical activity environment (walkability in 5-km buffer area). Living in high-income neighborhoods and neighborhoods with higher livability scores was associated with lower BMI. Nonlinear associations were observed with neighborhood home values in all studies. Lower neighborhood home values were associated with higher BMI scores but only for values up to €300,000. The directions of associations were less consistent for walkability and share of single residents.

**Discussion:** Rank aggregation made it possible to flexibly combine the results from various studies, although between-study heterogeneity could not be estimated quantitatively based on RF models. Neighborhood social, economic, and physical environments had the strongest associations with BMI. https://doi.org/10.1289/EHP13393

Address correspondence to Haykanush Ohanyan, Institute for Risk Assessment Sciences, Yalelaan 2, 3584CM Utrecht, Utrecht University, the Netherlands. Telephone: +33 6 52389098. Email: h.ohanyan@uu.nl

Supplemental Material is available online (https://doi.org/10.1289/EHP13393).

The authors declare they have nothing to disclose.



Received 26 May 2023; Revised 4 April 2024; Accepted 20 May 2024; Published 18 June 2024.

## Introduction

Overweight and obesity impose a considerable individual and social burden, with global rates that have nearly tripled since 1975.[1] This striking rise could be attributed to changes in the living environment and behavior occurring due to industrial and technological advances. The combination of these environmental factors is also known as the exposome.[2]

The current evidence base regarding exposome risk factors for overweight and obesity points toward a range of characteristics of the built environment, including urban sprawl and measures of land use.[3,4] Many studies have investigated the associations of the physical activity environment (neighborhood walkability, access to public transport, sport facilities, etc.) and food environment with obesity, but for both of these domains, the results are inconsistent.[4,5] Further inconsistent results were found for green space, air pollution, and road traffic noise.[6–10]

Several factors could have contributed to the heterogeneity in findings: differences in the operationalization of environmental factors, violation of assumptions that relations are linear, or small sample sizes. Moreover, most studies focused on single exposures in isolation, although in real life exposures exist simultaneously and are often correlated due to shared causal factors (e.g., urban planning). To address this issue, some studies used composite index scores. A drawback of this method, however, is that a large volume of data is condensed into a single score, leading to a loss of information, difficulties in interpretation, and measurement error, as the weighting of individual components may not reflect the true importance of factors.

Given the potentially complex interplay between risk factors, it is important to study exposures collectively and to incorporate nonlinear and nonadditive associations in statistical models. More advanced statistical methods, including random forest (RF), are advocated for use in the context of multiple exposures.[11,12] A recent study reported that RF had a good balance between the capacity of addressing complex datasets, the computational burden, and the availability of user-friendly software packages for model training and interpretation.[13]

To accurately assess the associations with environmental risk factors, which tend to have relatively small effect sizes, it is necessary to gather data from large populations. This can be achieved by combining multiple cohorts covering extensive geographical areas. However, it is challenging to combine results from RF or other machine-learning models from different studies into a single, pooled estimate. The reason for this is that the complexity introduced by nonlinear and nonadditive effects in statistical models poses a challenge for estimating precise effect sizes that are required for a meta-analysis.

To quantitatively combine results, meta-analytic methods are necessary, but in their absence, rank-aggregation (RA) methods could serve as a helpful intermediate step for identifying the most important variables across studies and for exploring whether different models consider the same variables as important. Rank aggregation has been applied in genetic studies, which used weighted rank aggregation to determine the most important genes across different studies.[14,15] We apply the RA approach here in the context of an exposome-wide association study comprising 15 Dutch cohorts to identify environmental determinants of body mass index (BMI).

## Methods

### *Study Design and Population*

We conducted cross-sectional analyses in 15 cohort studies affiliated with the Dutch Geoscience and Health Cohort Consortium (GECCO).[16] The studies included cover different target populations across the Netherlands. The studies varied by population size, ranging between 688 and 141,825 participants. Five studies were nationwide, four included participants from more sparsely populated northern regions, three were studies from southern regions, and three were from the center of the country (Table 1). Ten studies included general population samples, and others were conducted in specific populations (hearing impairment, diabetes, women, fetal exposure to famine). Details of the 15 studies can be found in the Supplemental Material ("Description of the included cohort studies and analytical samples").

The analytical sample included data of adult participants (≥18 years) from each cohort, having available data on geocoded residential addresses (matched with geocoded environmental

**Table 1.** Description of 15 included cohort studies affiliated with the Dutch Geoscience and Health Cohort Consortium (GECCO).

| Cohort study | Participants (*n*) | Year of assessment | Population | Geographic area covered |
|---|---|---|---|---|
| Lifelines[17] | 141,825 | 2006 | General population | Northern region |
| LIFEWORK[18] | 76,567 | 2011–2012 | General population[a] | Nationwide |
| Donor InSight (DIS)[19] | 30,866 | 2007 | Dutch blood donors | Nationwide |
| Healthy Life in an Urban Setting (HELIUS)[20] | 19,054 | 2011 | General population with oversampling of six specific ethnic groups | Amsterdam |
| The Maastricht Study (TMS)[21] | 7,583 | 2010–2020 | Adults 40–75 years old[b] | Maastricht and Heuvelland |
| Netherlands Mental Health Survey and Incidence Study (NEMESIS)[22] | 6,526 | 2007–2009 | General population | Nationwide |
| Netherlands Twin Registry (NTR)[23] | 5,933 | 2004–2008 | Adult twins or their family members | Nationwide |
| Doetinchem Cohort Study (DCS)[24] | 3,983 | 2008–2012 | General population | Doetinchem |
| The Hoorn study (Hoorn)[25] | 2,798 | 2006 | General population | West-Friesland |
| Dutch Cohort Study On Socioeconomic Health Inequalities (GLOBE)[26] | 2,422 | 2014 | General population | Eindhoven and surrounding municipalities |
| Tracking Adolescents' Individual Lives Survey (TRAILS)[27] | 1,832 | 2012–2013 | Adults born in 1989–1990 | Northern region |
| Longitudinal Aging Study Amsterdam (LASA)[28] | 1,411 | 2008–2009 | Adults 55–85 years old | North, south, and secular parts |
| Maastricht Aging Study (MAAS)[29] | 1,106 | 2005 | Adults >50 years old | Province of Limburg |
| Dutch famine birth cohort (DFBC)[30] | 796 | 2002 | Adults with prenatal exposure to Dutch famine | Born in Amsterdam |
| National Longitudinal Study on Hearing (NL-SH)[31,32] | 688 | 2016–2017 | Adults 28–80 years with or without hearing impairment[c] | Nationwide |

Note: Additional information can be found in the Supplemental Material "Description of the included cohort studies and analytical samples."
[a]Combination of Nightingale, EPIC-NL, and AMIGO studies and consists of a large majority of women.
[b]With an oversampling of participants with type 2 diabetes.
[c]With an oversampling of adults with hearing impairment.

exposures) and outcome, i.e., BMI. All participants gave informed consent.

### Study Outcome and Covariates

Self-reported (seven studies) or measured (eight studies) height and weight were used to calculate the outcome, BMI ($kg/m^2$), as a continuous variable. Expert knowledge was used to select the relevant confounders. Several individual sociodemographic characteristics were considered as confounders in this study: age, sex (male/female), ethnicity or country of birth (the Netherlands/other), living situation (with/without a partner), highest degree of education (high/medium or low), employment status (employed/unemployed/student/retired), and smoking (yes/no). For studies where an oversampling of certain groups occurred, the models were further adjusted for relevant group-membership [participants with hearing impairment for the Netherlands Longitudinal Study on Hearing (NL-SH) and those with type 2 diabetes for The Maastricht Study (TMS)]. Ethnicity was not available in the Dutch famine birth cohort (DFBC) and Maastricht Aging Study (MAAS), so models for these studies did not adjust for ethnicity. To ensure the comparability of our models across the studies, we grouped the confounders in similar categories whenever this was possible; however, the confounders were not standardized across the cohorts. Table S1 and S2 give a detailed description of the confounder categories and whether the outcome was measured or self-reported in each cohort.

### Description of the Exposome

Based on the data availability and associations reported in previous studies, we linked 69 exposome factors to the geocoded residential addresses of each participant. The exposure groups included air pollutants (14 factors); traffic noise due to roads, aircraft, or railway (one factor); mean summer temperature (one factor), urbanization degree (one factor); neighborhood built environment (27 factors) including drivability, walkability, green space, food environment, accessibility of public transport and services; neighborhood sociodemographic (15 factors); and economic factors (10 factors). A brief description of the linked data is provided in Table 2.

All environmental data were assessed for three different time points (2000, 2006, 2011) depending on availability and matched with individual-level data as closely as possible to data-collection period of each study. Table S3 provides a detailed description of years of data assessment by study and by exposure.

All environmental data were operationalized and supplied by GECCO. A more comprehensive description of these data is available via GECCO's website.[33]

### Statistical Analysis

***General considerations.*** Statistical analyses were conducted separately for each cohort study because there were obstacles for fitting a single model on merged data, as some cohort data could not be shared and had to be analyzed locally. This was evidently a disadvantage. Nevertheless, an advantage of this approach is that it avoided the imposition of a single model on rather heterogeneous datasets.

To address the issue of possible nonlinearity and varying interactions among studies, the use of RF models was deemed necessary. Consequently, a meta-analytic strategy for synthesizing the results from each individual study was developed.

***Missing data.*** For each study, data quality control was performed prior to imputation. Generally, the missing data percentages were low across the studies and ranged between 0.4 and 5.3% (Table S4). We imputed missing data using the RF algorithm from the Multivariate Imputation by Chained Equations (MICE) package, in R. A single imputation was used. For continuous variables, records with missing values were imputed by random draws from independent normal distributions centered on conditional means predicted using RF. For binary or unordered categorical variables, individual regression trees were fitted to a bootstrap sample of the data, and each missing value was imputed as the prediction of a randomly chosen tree.[40]

***Study-specific analysis using random Forest.*** For each study, exposome-wide associations of BMI were explored using the randomForestSRC package in R. In all models, BMI was the dependent variable, and all exposures and confounders were considered simultaneously, as independent variables. The optimal values of hyperparameters "mtry" and "node size" were decided separately for each study using model-based optimization through the tuneRanger package. To build the trees, we used subsampling without replacement and adhered to the default sample size, which equated to 63.2% of the original dataset size in each study. In all studies, 1,000 trees were used.

We used permutation-based variable importance scores (VIMP) to rank each predictor by its importance. Higher positive values of VIMP scores indicated greater importance. We assessed the stability of rankings by a novel subsampling-based RF approach, implemented in the randomForestSRC package.[41] The idea behind this method is that it takes a small subratio of the data, in a way that subsamples are independent and estimates the variance of VIMP scores across the number of subsamples. We used a constant subratio ($n^{3/4}$) for subsampling and 100 independent subsamples in all studies. The ranking stability of each exposure was assessed by calculating interquartile ranges of the attributed ranks across the subsamples.

To further enhance the interpretation of RF models and gain insights into the nature of the relationship between a specific exposure and BMI (whether it is positive, negative, or has nonlinear associations), we used visualizations of fast approximate Shapley values, calculated using the fastshap package in R. Shapley values are a concept from coalitional game theory, which recently gained popularity for the interpretation of machine learning models, because they indicate the average contribution of a feature value to the prediction when different combinations of features are used to build the model.[42] Shapley values differ across samples, acknowledging the nonlinear and interactive effects. We also assessed the strongest pairwise interactions in each study using the function find.interactions() from the randomForestSRC package. Calculating the interactions was very costly computationally and especially problematic for the two largest studies (LIFEWORK, Lifelines). Hence, pairwise interactions in these studies were assessed using models fitted to a random sample of 10% of the data only.

***Description of meta-analytical approach.*** We employed a RA approach to assess the most important exposures across the 15 study-specific lists containing ordered rankings of VIMPs for each of the 69 exposures that were common to all studies.[14,43] Cross-entropy Monte Carlo iterative algorithm was used to find the "super"-list of exposures, based on Spearman's footrule distance function.[43–45] The latter is a measure of distance between the ranked lists, summing up the absolute value of the element-wise rank-differences between lists. For each element, these values were averaged across the pairwise comparisons of ordered lists.[45] The RA was conducted using the RankAggreg() function from the R package of the same name.

We used the ranking stability (assessed by the interquartile ranges of ranks across the RF subsamples) to weight the exposure ranks (Figure S3). Plus, for each study, we used two importance weighting schemes. First, simply the square root values of the study size. Second, discrete importance weights were assigned: studies with $n < 1,000$ were given a discrete

**Table 2.** Description of exposure data and the sources for each variable from the Dutch Geoscience and Health Cohort Consortium (GECCO). More comprehensive details on the meta-data are available.[33]

| Exposome factor | Description | Map resolution |
|---|---|---|
| Average summer temperature (°C) | Monthly average temperature data[a] (June–September) were interpolated based on 10 automatic monitoring stations.[34,35] | $25 \times 25$ m |
| Traffic noise [dB(A)] | Daily average levels of noise (combining road, rail, and air) were modeled and expressed as Lden (Level day-evening-night) in decibels [dB(A)]. This measure considers a higher penalty for noise exposure during the evening (5 dB) and night (10 dB).[b] | $25 \times 25$ m |
| Urbanization degree | Urbanization level was based on residential density and had the following categories: 1 = very highly urban ≥2,500 addresses per km$^2$; 2 = highly urban 1,500–2,500 addresses; 3 = moderately urban 1,000–1,500 addresses; 4 = less urban 500–1,000 addresses; 5 = nonurban <500 addresses.[36] | Administrative neighborhood |
| Neighborhood drivability[a] | | |
| Accessibility of jobs (two variables) | Average road travel time (in minutes) to access 10,000 and 100,000 jobs. | $100 \times 100$ m |
| Distance to train station (km) | The closest train station. | $100 \times 100$ m |
| Distance to motorway exit (km) | The closest motorway exit. | $100 \times 100$ m |
| Driving destination accessibility index | The index quantifies the ease of reaching different types of destinations (commercial, recreational, and services) by car. It was computed based on a weighting system for areas that are more suitable for active transportation or walking. Index values were normalized to a scale of 0 (low drivability) to 100 (high drivability). | $100 \times 100$ m |
| Paid parking | Percentage of paid parking places in 1-km buffer area. | $100 \times 100$ m |
| Parking pressure | Ratio of registered cars and number of parking places. | $100 \times 100$ m |
| Air pollutants | Annual average concentrations of below-listed air pollutants were modeled by land-use-regression models (data source ESCAPE)[37,38] and a combination of dispersion model calculations and measurements (data source RIVM).[33] | — |
| Benzene ($C_6H_6$) ($\mu g/m^3$) | Dispersion model and measurements. | $1 \times 1$ km |
| Carbon monoxide (CO) ($\mu g/m^3$) | Dispersion model and measurements. | |
| Ammonia ($NH_3$) ($\mu g/m^3$) | Dispersion model and measurements. | |
| Ozone ($O_3$) ($\mu g/m^3$) | Dispersion model and measurements. | |
| Soot (EC) | Dispersion model and measurements. | |
| Sulfur dioxide ($SO_2$) ($\mu g/m^3$) | Dispersion model and measurements. | |
| $NO_2$ ($\mu g/m^3$) | Land-use-regression model. | Point density |
| NOx ($\mu g/m^3$) | Land-use-regression model. | |
| $PM_{10}$ ($\mu g/m^3$) | Land-use-regression model. | |
| $PM_{2.5}$ ($\mu g/m^3$) | Land-use-regression model. | |
| $PM_{2.5}$ absorbance (10-5 m-1) | Land-use-regression model. | |
| PM coarse ($\mu g/m^3$) | Land-use-regression model. | |
| Oxidative potential of $PM_{2.5}$ (two variables) | Oxidative potential was estimated by dithiothreitol or electron spin resonance methods in Land-use regression model.[37] | |
| Neighborhood demographic characteristics[36] | | |
| Neighborhood students | Share of students in neighborhood (%). | Administrative neighborhood |
| Neighborhood employment status | Share of employed residents in neighborhood (%). | |
| Neighborhood age groups (five variables) | Shares of residents 0–14, 15–24, 25–44, 45–65, and 65+ years of age (%). | |
| Neighborhood marital status (four variables) | Shares of single, married, divorced, widowed residents in neighborhood (%). | |
| Neighborhood household sizes (two variables) | Shares of one-person households, households with children (%). | |
| Neighborhood immigration status (two variables) | Shares of immigrants from high-, low-, and middle-income countries (%).[36] | |
| Neighborhood economic characteristics[36] | | |
| Accessibility of housing | People living in private households divided by the number of private households. | Administrative neighborhood |
| Neighborhood home values | Average home values (×1,000 euros). | Administrative neighborhood |
| Neighborhood income (seven variables) | Number of income recipients in neighborhood, mean neighborhood income (×1000 euros), percentages of residents with high (above 80th percentile) or low (below 40th percentile) average income, the average ownership of passenger cars or motorcycles per household, socioeconomic status scores which took into account education, income, and employment status.[36] Neighborhoods were defined by four-digit postal code areas. | Administrative neighborhood |
| Livability score | Neighborhoods were categorized into livability classes ranging from 1 (very insufficient) to 9 (excellent). The score was computed based on factors such as population composition, social cohesion, public space, safety, resources, and housing.[33] | $100 \times 100$ m |

Table 2. (*Continued.*)

| Exposome factor | Description | Map resolution |
|---|---|---|
| Built environment | | |
| Neighborhood walkability (three variables) | Dutch Walkability Index integrated seven components: population density, retail and service destination density, land-use mix, street connectivity, green space, sidewalk presence, and public transport density. The components were summed and normalized to a score 0–100, with higher values indicating higher walkability. It was assessed for 0.5-km, 1-km, and 5-km buffer areas.[39] | $25\times25$ m |
| Green space (three variables) | Green space density was obtained by aggregating *z*-scores of land use data of trees, shrubs, and low vegetation. It was assessed for 0, 0.5-km, 1-km, and 5-km buffer areas.[33] | $25\times25$ m |
| Land use mix (three variables) | Land Use Mix Entropy Index was calculated as the sum of *z*-scores of different land use classes, such as residential, commercial, social-cultural services, offices and public services, green space, and recreation. It was assessed for 0.5-km, 1-km, 5-km buffer areas.[36] | $25\times25$ m |
| Accessibility of neighborhood facilities (three variables) | Average distance to the nearest medical, recreational, or educational facilities in neighbourhood.[36] | Administrative neighborhood |
| Accessibility of sport facilities (two variables) | Density of sport accommodations was assessed for 0.5-km and 1-km buffer areas. Only sports requiring significant physical effort were considered.[33] | $25\times25$ m |
| Accessibility of public transport (three variables) | Density of public transport stops was assessed as sum of the *z*-scores of the entire public transport network in the Netherlands (bus, ferry, metro, taxi, tram). It was assessed for 0.5-km, 1-km, 5-km buffer areas.[33] | $25\times25$ m |
| Neighborhood food environment (three variables) | The Food Environment Healthiness Index was employed to assess food retailer healthfulness. Food retail outlets were assigned values between −5 (very unhealthy) and +5 (very healthy). The score was assessed for 0.5-km, 1-km, 5-km buffer areas.[33] | $25\times25$ m |

Note: —, no data; $PM_{2.5}$, fine particulate matter with aerodynamic diameter ≤2.5 μm; $PM_{10}$, particulate matter with aerodynamic diameter ≤10 μm.
[a]See GECCO (https://www.gecco.nl/).
[b]See Planbureau voor de Leefomgeving (https://www.pbl.nl/en).

importance of one, studies with $1{,}000 < n < 10{,}000$ an importance of two, and the largest studies with $n > 10{,}000$ were allocated an importance of three.

We chose the top 10 exposures as the cut-off point for our RA method. The rationale behind the use of this cut-off is that not all variables within a model have strong predictive capabilities, and we aimed to select only the most important predictors. Consequently, beyond a certain threshold, these variables tend to lose their significance in prediction and become noisy. Determining an entirely objective initial cut-off point is challenging, particularly in the context of multiple studies. Given that the number of influential predictors in our RF models were generally few, we opted for a cut-off of ten to select the top important predictors across the studies. Additionally, we rigorously tested various scenarios of RA with different cut-off values (top 10, top 15, top 20, and top 30). To test if a variable was ranked among the top 10 exposures more often than expected by chance, we used one-sided binomial tests with Bonferroni-corrected *p*-values for all exposures. Thus, the null-hypothesized probability of success was set to 10/69 (for 69 exposures), and the alternative hypothesis was set to greater than the null-hypothesized success.

Rank aggregation enabled the flexible combination of ordered lists regardless of the levels of heterogeneity. However, a downside of this method is that it did not make it possible to estimate the heterogeneity. Instead, we visually assessed the ranking consistency of the most important exposures using a heatmap.

### *Sensitivity Analyses*

Several sensitivity analyses were conducted to test the robustness of our findings: *a*) We assessed the results using only the studies where BMI was objectively measured; *b*) We excluded the two studies with oversampling for particular groups (NL-SH: hearing impairment, TMS: diabetes) and the two studies in which ethnicity was not assessed (DBFC, MAAS); *c*) We performed an additional analysis of nationwide studies [LIFEWORK, Donor InSight (DIS), Netherlands Mental Health Survey and Incidence Study (NEMESIS), Netherlands Twin Registry (NTR), NL-SH], as these have more variability in environmental exposures than localized studies; *d*) We included only the studies with a more homogeneous age range (40–60 years old), because individuals from very different age groups may have different daily routines and dietary habits, hence the Tracking Adolescents' Individual Lives Survey (TRAILS), Longitudinal Aging Study Amsterdam (LASA), and Dutch cohort study on socioeconomic health inequalities (GLOBE) were excluded.

## Results

Overall, we analyzed data of 303,660 participants. The average BMI ranged between 23 and $27\,\mathrm{kg/m^2}$ across the studies. Participant characteristics are listed in Table 3. Exposome factors were mostly similarly distributed across studies, but some differences were related to geographic location and urbanization degree (Table S5). Correlation patterns (based on Spearman's correlation coefficient) between the exposures were also consistent across the studies. The highest correlations were observed among temperature, degree of urbanization, drivability, air pollutants, and socioeconomic and demographic factors. Generally, the correlations were weaker in studies from small towns [Hoorn, MAAS, Doetinchem Cohort Study (DCS)] (Figure S1).

The values of explained variance of BMI from multivariable RF models showed that the total explainability of BMI by the exposome factors was rather low in some of the studies (Table S6). The results from individual RF models showed that the most important exposures were factors of neighborhood social, economic (neighborhood income, livability score), demographic (household types, share of immigrants), built environment (walkability, distance to medical facilities), and urban characteristics (air pollution). In

**Table 3.** Individual characteristics of participants in 15 cohort studies affiliated with the Dutch Geoscience and Health Cohort Consortium (GECCO).

| Cohort[a] | Participants ($n$) | BMI (mean ± SD) | Age (mean ± SD) | Female [$n$ (%)] | Ethnicity: Dutch origin[b] [$n$ (%)] | Civil status: living with a partner [$n$ (%)] | High education [$n$ (%)] | Employed [$n$ (%)] | Currently smoking [$n$ (%)] |
|---|---|---|---|---|---|---|---|---|---|
| Lifelines | 141,825 | 26.1 ± 4.3 | 44.4 ± 12.8 | 88,999 (58.5%) | 138,088 (97.4%) | 113,574 (80.1%) | 42,114 (29.7%) | 110,535 (77.9%) | 30,057 (21.2%) |
| LIFEWORK | 76,567 | 25.3 ± 4.3 | 50.4 ± 12.8 | 68,130 (89.0%) | 73,548 (96.1%) | 62,411 (81.5%) | 32,425 (42.3%) | 55,804 (72.9%) | 9,473 (12.4%) |
| DIS | 30,866 | 25.4 ± 6.3 | 45.6 ± 12.7 | 10,777 (56.6%) | 29,917 (96.9%) | 24,585 (79.7%) | 10,861 (35.2%) | 24,916 (80.7%) | 4,954 (16.1%) |
| HELIUS | 19,054 | 27.1 ± 5.3 | 44.7 ± 13.2 | 10,777 (56.6%) | 4,166 (21.9%) | 9,538 (50.1%) | 4,939 (25.9%) | 11,462 (60.2%) | 8,171 (42.9%) |
| TMS | 7,583 | 27.0 ± 4.5 | 59.8 ± 8.7 | 3,772 (49.7%) | 7,483 (98.7%)[a] | 6,014 (79.3%) | 2,804 (37.0%) | 3,111 (41.0%) | 1,006 (13.3%) |
| NEMESIS | 6,526 | 25.2 ± 4.4 | 44.4 ± 12.5 | 3,600 (55.2%) | 5,630 (86.3%) | 4,446 (68.1%) | 2,305 (35.3%) | 4,538 (69.5%) | 1,936 (29.7%) |
| NTR | 5,933 | 24.7 ± 3.8 | 43.8 ± 14.4 | 3,660 (61.7%) | 5,497 (92.7%) | 4,553 (76.7%) | 2,144 (36.1%) | 3,170 (53.4%) | 1,218 (20.5%) |
| DCS | 3,983 | 27.2 ± 4.3 | 59.9 ± 9.6 | 2,098 (52.7%) | 3,783 (95.0%) | 3,194 (80.2%) | 957 (24.0%) | 1,925 (48.3%) | 674 (16.9%) |
| Hoorn | 2,798 | 26.2 ± 4.0 | 53.4 ± 6.7 | 1,494 (53.4%) | 2,591 (92.6%) | 2,282 (81.6%) | 716 (25.6%) | 1,840 (65.8%) | 590 (21.1%) |
| GLOBE | 2,422 | 26.3 ± 4.9 | 62.9 ± 13.1 | 1,338 (55.2%) | 2,264 (93.5%) | 1,800 (74.3%) | 828 (34.2%) | 895 (35.5%) | 314 (13.0%) |
| TRAILS | 1,832 | 23.3 ± 4.0 | 21.2 ± 1.4 | 939 (51.3%) | 1,676 (91.3%) | 566 (30.9%) | 983 (53.7%) | 1,177 (64.2%) | 686 (37.4%) |
| LASA | 1,411 | 27.6 ± 4.4 | 74.8 ± 7.6 | 776 (55.0%) | 1,395 (98.9%) | 906 (64.2%) | 278 (19.7%) | 233 (16.5%) | 203 (14.4%) |
| MAAS | 1,106 | 26.9 ± 4.4 | 59.4 ± 14.3 | 549 (49.6%) | Not available | 773 (69.9%) | 271 (24.5%) | 498 (45.0%) | 179 (16.1%) |
| DFBC | 796 | 28.6 ± 4.9 | 58.4 ± 1.0 | 434 (54.5%) | Not available | 601 (75.5%) | 135 (17.0%) | 418 (52.5%) | 195 (24.5%) |
| NL-SH | 688 | 26.0 ± 4.5 | 57.6 ± 11.6 | 432 (62.8%) | 658 (95.6%) | 501 (72.8%) | 420 (61.1%) | 361 (52.5%) | 72 (10.5%) |

Note: BMI, body mass index; SD, standard deviation. DCS, Doetinchem Cohort Study; DFBC, Dutch famine birth cohort; DIS, Donor InSight; GLOBE, Dutch cohort study on socioeconomic health inequalities; HELIUS, Healthy Life in an Urban Setting; LASA, Longitudinal Aging Study Amsterdam; MAAS, Maastricht Aging Study; NEMESIS, Netherlands Mental Health Survey and Incidence Study; NL-SH, Netherlands Longitudinal Study on Hearing; NTR, Netherlands Twin Registry; TMS, The Maastricht Study; TRAILS, Tracking Adolescents' Individual Lives Survey.

[a]Complete cohort names are provided in Table 1. For additional information please see the Supplemental Material "Description of the included cohort studies and analytical samples."

[b]For the TMS cohort, the presented percentage of ethnicity refers to European origins rather than native Dutch.

GLOBE, TRAILS, and LASA, the top exposures also included air pollutants [carbon monoxide, soot, benzene, and oxidative potential of fine particulate matter with aerodynamic diameter $\leq 2.5\ \mu m$ ($PM_{2.5}$)]. It is worth noting that in GLOBE, TRAILS, and NL-SH the air pollutants and physical aspects of neighborhoods were more important than socioeconomic factors (Table S7).

Rank aggregation technique identified the 10 most important neighborhood exposures across the studies: average home values, average income, share of high-income residents, livability score, shares of single inhabitants, one-person households, immigrants from low- to middle-income countries (LMICs), density of public transport stops in 5-km buffer area, the accessibility of jobs (10,000), and walkability (5 km) (Table 4). The results of rank aggregation were mostly similar across the application of different weighting scenarios (ranking stability, study size) (Table S8).

The ranking patterns were broadly consistent for the top variables (Figure 1). The frequency of being among the top 10 exposures across the studies were as follows: neighborhood home values (12/15), high-income residents (10/15), livability score (10/15), neighborhood average income (9/15), walkability (5 km) (8/15), share of single residents (7/15), accessibility of jobs (10,000) (6/15), density of public transport stops (5 km) (5/15), shares of LMIC immigrants (4/15), and one-person households in neighborhood (4/15).

**Table 4.** Aggregated top important exposures associated with body mass index and their consistency for being ranked below the 10th rank across the studies based on random forest (RF) models. Results from the one-sample binomial tests.

| Top exposusres[a] | Probability of selection[b] | $p$-Value[c] |
|---|---|---|
| Neighborhood home values | 0.80 | <0.01* |
| High-income residents (%) | 0.67 | <0.01* |
| Livability score | 0.67 | <0.01* |
| Neighborhood average income | 0.60 | <0.01* |
| Walkability (5 km) | 0.53 | 0.03* |
| Single residents (%) | 0.47 | 0.17 |
| Job accessibility (mean travel time to 10,000 jobs) | 0.40 | 0.84 |
| Public transport (5 km) | 0.33 | 1.00 |
| Immigrants from low- or middle-income countries (%) | 0.27 | 1.00 |
| One-person households (%) | 0.27 | 1.00 |

[a]Other exposures are not shown because the corrected $p$-value = 1.

[b]Probability of being among the top 10 ranked exposures from the RF model for each study ($n = 15$).

[c]Corrected $p$-values for multiple testing by Bonferroni method (with $n = 69$).

*Asterisks indicate the variables that were ranked below the 10th rank significantly more often than expected by chance.

One-sided binomial tests helped to select the five exposures ranked below the 10th rank significantly more often across studies. Four factors reflected neighborhood social and economic environments (livability score, neighborhood home values, share of inhabitants with high income, neighborhood average income), and one factor reflected the physical activity environment walkability index (5 km) (Table 4). The exploration of pairwise interactions revealed that they were quite heterogeneous across the studies with no clear patterns (Table S9).

Plots based on Shapley values illustrated how changing the exposures (from maximum to minimum values) affected the BMI predictions. Generally, the effect sizes were modest (Figure 2; Figure S2A–S2E). Living in neighborhoods with a higher share of high-income residents or a higher mean income or a better livability score (Figure 2; Figure S2A–S2E) was associated with a lower BMI. Lower neighborhood home values (€150,000–300,000) were associated with higher BMI scores, but no association was found for higher values. In the majority of studies, higher walkability index (5-km buffer) was associated with lower BMI, but no associations were found in NEMESIS, NTR, LASA, and MAAS, perhaps because of lower variability of the index in these studies. The directions of associations were inconsistent across the studies for the share of single residents. Positive associations were observed in large, nationwide studies, and negative associations or no associations were observed in studies covering smaller geographical areas.

Sensitivity analyses limited to studies that measured BMI, only nationwide studies, only studies without additional confounders, studies that adjusted for ethnicity, and studies with more homogeneous average age range (40–60 years old) revealed that the selected exposures mostly remained consistent (Table S10).

## Discussion

In a large, multicohort study, we found that home values, share of high-income residents, livability score, average income, and


**Figure 1.** Distributions of ranks of the six most important exposures (see *y*-axis) associated with body mass index across the 15 cohorts in the Dutch Geoscience and Health Cohort Consortium (GECCO), listed on the *x*-axis. The lower the value of the ranks, the more important the exposure. The columns are ordered from right to left based on the number of observations in each cohort. Thus, Lifelines had the highest number of participants, and NL-SH had the lowest number. Cohort names and participant numbers are provided in Table 1. Note: DCS, Doetinchem Cohort Study; DFBC, Dutch famine birth cohort; DIS, Donor InSight; GLOBE, Dutch cohort study on socioeconomic health inequalities; HELIUS, Healthy Life in an Urban Setting; LASA, Longitudinal Aging Study Amsterdam; MAAS, Maastricht Aging Study; NEMESIS, Netherlands Mental Health Survey and Incidence Study; NL-SH, Netherlands Longitudinal Study on Hearing; NTR, Netherlands Twin Registry; TMS, The Maastricht Study; TRAILS, Tracking Adolescents' Individual Lives Survey.

walkability in residential neighborhoods were the most important aspects of the external exposome associated with adult BMI. Living in high-income neighborhoods was associated with lower BMI. The associations were nonlinear for the neighborhood home values, with an inverse association up to €300,000, after which the association plateaued. Walkability was inversely related to BMI in most studies, but results were heterogeneous between studies. A unique feature of this study was that nonparametric estimates from different RF models were combined in single results, using a ranking based meta-analytical approach. We supplemented the RF analyses with Shapley plots to assess the direction and shape of the association between the most important exposures and BMI.

The same nonlinear shape of association between neighborhood home values and BMI was found in all 15 studies. Living in high-income neighborhoods may be beneficial for health, by providing healthier environments, accessibility of services, better social support, and lower stress levels due to safer surroundings. A recent meta-analysis of 21 observational studies from high-income countries found a strong negative association between the neighborhood socioeconomic position and risk of obesity.[46] The association with BMI was found in studies with cross-sectional and longitudinal design. However, only two of the studies included had a longitudinal design. Another cross-sectional study examined the associations of the neighborhood deprivation and obesity risk in France.[47] The research found similar nonlinear associations, the strength of which varied according to the urbanization level. It was stronger in suburban areas of large cities and weaker in small towns and rural areas. Despite these robust associations, there is a need for caution in the causal interpretations of these findings due to potential residential self-selection bias.

Another important finding was the association of high livability score with lower BMI. Livability is a composite score that was created to reflect the extent to which Dutch residents were satisfied with their living environments. It comprised information on multiple key axes: economic aspects of neighborhood, population composition, housing accessibility, neighborhood social cohesion, and safety. A possible explanation for this association mechanism is that individuals who feel safer and more connected to their community might spend more time outdoors, walking, or exercising, and they might be more motivated to adopt healthy eating.[48–50] Further research could help to explain the mechanisms through which these factors affect BMI.

Some might argue that the exposure domains with a greater number of variables could have been more likely to be selected in our study. The results showed, however, that certain domains (neighborhood economic position) with fewer variables had a higher number of selected variables (three selected out of 10 variables) compared to domains with more variables (i.e., built environment—two selected out of 27, air pollution—nonselected out of 14), highlighting the importance of this domain.

Addressing the challenge of high correlations was important in our study; hence, we undertook some measures to mitigate this issue. We excluded exposure variables with very high levels of correlations ($r_{\text{Spearman}} > 0.95$), as calculated across all the addresses, for the entire Netherlands. We did not exclude pairwise high correlations in each individual study because, despite the similar patterns of correlation structures, there was some degree of variability from one study to another. To ensure the comparability of our models, we chose to maintain a consistent set of exposures across all of our models.

Random forest is commonly applied in the setting of high-dimensional and correlated data in the field of genomics research.[51] Despite this, simulation studies showed that the performance of RF can be impacted when predictors strongly correlate.[52] This is especially problematic when the variable importance score is assessed with the Gini's index, although the permutation importance score can also be subject to this bias.[53]

In our exposome data, the correlations between different exposures ranged from low to moderate, with a few stronger correlations observed mostly within the same domain of exposures (air pollution, neighborhood sociodemographic factors, etc.) (Figure S1). Despite the robustness of our RF models using permutation importance scores, as well as the lack of a large number of strong correlations and large sample sizes in most individual studies, we cannot rule out the possibility of feature importance instability due to correlation in

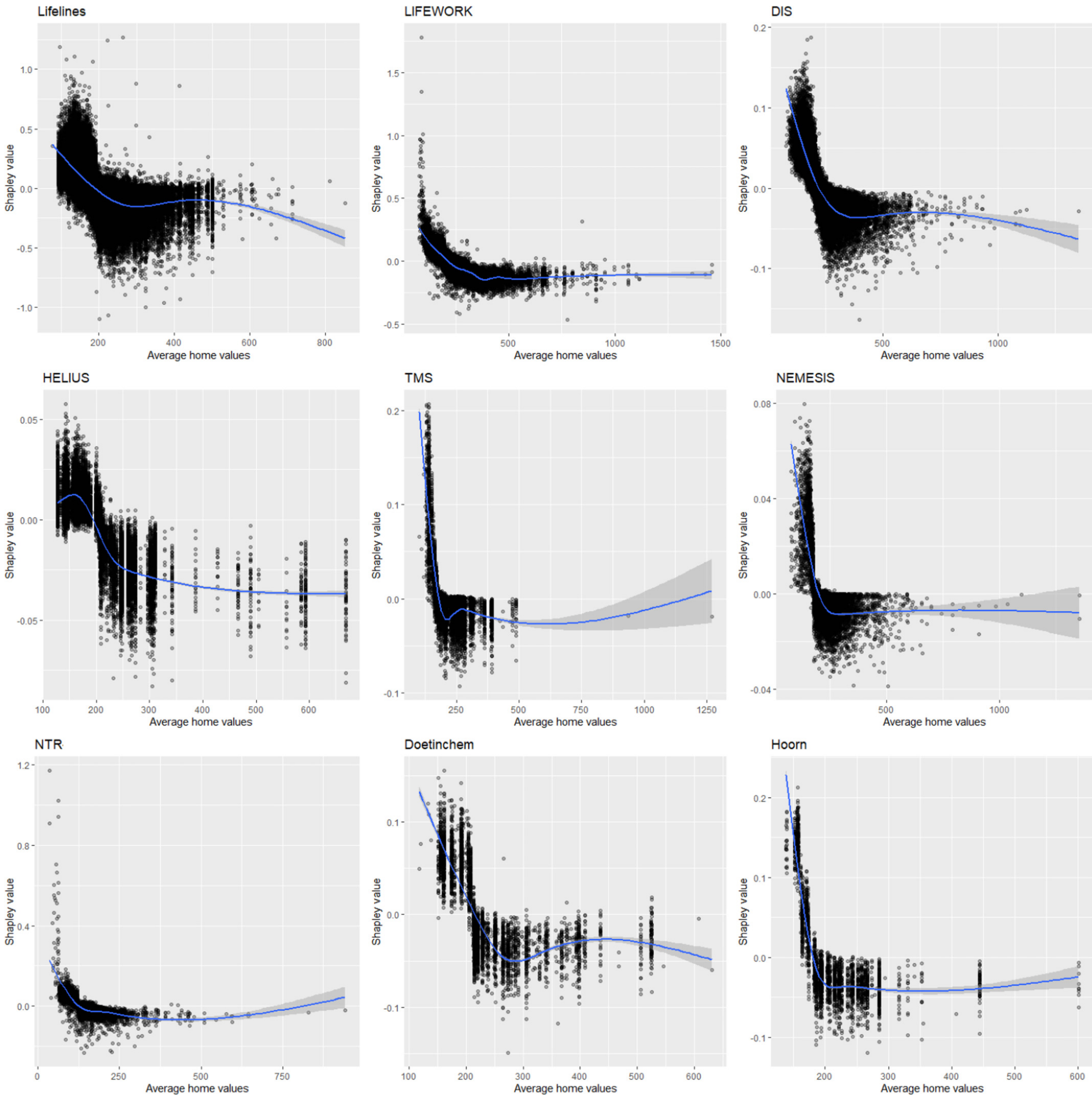


**Figure 2.** Shapley plots of neighborhood home values across the studies for each of the cohorts included in the Dutch Geoscience and Health Cohort Consortium (GECCO). Shapley values represent the difference between a prediction and the average prediction of BMI ($kg/m^2$). The size of Shapley values vary; hence, the *y*-axes are on different scales depending on each cohort. The *x*-axis represents average administrative neighborhood-level home values, in 1,000s of euros. The solid line represents the smoothed relationship between the observed neighborhood home values and Shapley values for neighborhood home values. Complete cohort names and numbers of participants per cohort are provided in Table 1. Note: BMI, body mass index; DCS, Doetinchem Cohort Study; DFBC, Dutch famine birth cohort; DIS, Donor InSight; GLOBE, Dutch cohort study on socioeconomic health inequalities; HELIUS, Healthy Life in an Urban Setting; LASA, Longitudinal Aging Study Amsterdam; MAAS, Maastricht Aging Study; NEMESIS, Netherlands Mental Health Survey and Incidence Study; NL-SH, National Longitudinal Study on Hearing; NTR, Netherlands Twin Registry; TMS, The Maastricht Study; TRAILS, Tracking Adolescents' Individual Lives Survey.

individual studies. The current literature presents conflicting information regarding the impact of correlated variables on permutation importance scores. One simulation study suggested that highly correlated variables might lead to an underestimation of their importance,[54] while another study indicated that correlated variables could receive higher importance values in random forest models.[55]

Random forest enabled us to study the combination of a large set of exposures to reveal the important exposure–outcome associations, including nonlinear associations and interactions. However, an important challenge in this study was how to combine the results from different RF models, as there are no established meta-analytical methods for combining estimates of associations. We

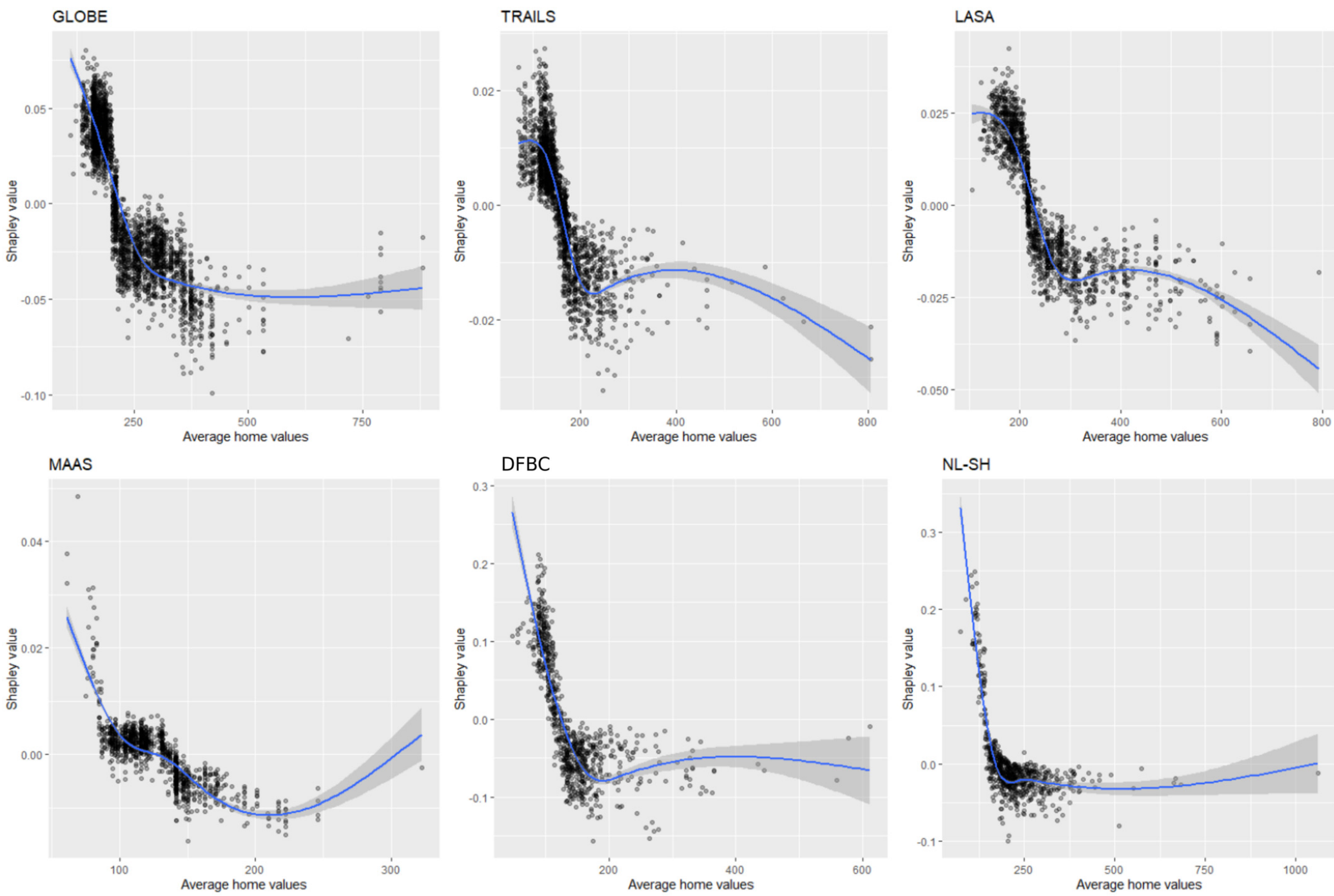


**Figure 2.** (*Continued.*)

applied rank aggregation because it is an interpretable method that has several interesting properties. It enabled us to control the aggregation process by using weights on the ranks based on their stability, as well as importance weights on the ordered lists based on cohort size. We found that rank aggregation with cross-entropy Monte Carlo had a good stability for identifying the top variables, but the order in which aggregated variables appeared in the top list was less stable. However, by using one-sided binomial tests to select the factors consistently ranked higher than 10th throughout the cohorts, we circumvented this instability. Moreover, using rank aggregation rather than a meta-analysis of regression coefficients provided the possibility of ultimate standardization, meaning that the effect sizes were standardized to ranks, allowing us to compare the relative importance of variables across studies, which is particularly useful when dealing with heterogeneous populations with disparate models or covariates.

This study has two main strengths. First, it disentangles the relations of various interrelated aspects of the urban exposome and BMI in a large population. The second strength of this study is that it provides a first example of the use of a meta-analytical approach for combining the results from a machine-learning approach within the context of exposome research, thus pushing the boundaries of the interpretability of such models and facilitating their further application in this field. This study also has some limitations. First, compared to the standard meta-analysis, rank aggregation does not allow for obtaining pooled estimates for effect sizes. Despite this, in each individual study, we were able to calculate Shapley values, which indicated by how much/to what extent a factor was contributing to the predicted outcome (Figure S2A–S2E). Another limitation of the RA approach is the difficulty of estimating the between-study heterogeneity because the rank aggregation does not account for the precision of these measured ranks. Despite this, we evaluated the consistency of ranks across the studies by means of visual assessment (Figure 1). Third, we analyzed the data from each study cross-sectionally, meaning that our findings could be subject to residential self-selection bias, as the temporal link between the associations cannot be established. Furthermore, some exposures were correlated and could potentially serve as a proxy for other aspects of living environment. Therefore, these findings cannot be interpreted from a causal perspective. Fourth, the environmental data might be prone to measurement error, as a mixture of modeled and measured exposures was used, and also exposures and outcome data were not always perfectly matched with the year. Measurement error would under most scenarios lead to attenuation of the effects. Therefore, more mismeasured exposures would have a lower probability of selection.

To the best of our knowledge, this is the first study to use a meta-analytical approach that combines the results from nonparametric RF models in an exposome context. Despite its exploratory nature, this study offers some important insight into the relation of the neighborhood environment and high BMI risk across various population groups. Neighborhood-level socioeconomic factors, livability score, and walkability were the most important factors of the urban exposome, associated with BMI. The importance of these factors was largely sustained across different studies. Future studies should employ longitudinal or experimental designs to investigate causal pathways underlying the associations observed in this study and to help address issues related to residential self-selection bias.

## Acknowledgments

The authors are grateful to Susan Picavet and Petra Vissink for their contribution to the Doetinchem cohort data; to Femmeke Prinsze for her contribution to the data linkage processes of Donor InSight data; to Henrike Galenkamp for her contribution to the Healthy Life in an Urban Setting (HELIUS) data; to Martijn Huisman for his contribution to the data of the LASA study; to Nicolaas C. Schaper, Pieter C. Dagnelie, Carla J. H. van der Kallen, Miranda T. Schram, Marleen M. J. van Greevenbroek, and Ronald M. A. Henry for their contribution to The Maastricht study data; to Sophia E. Kramer for her contribution to the NL-SH study data; to Inka Pieterson for her contribution to LIFEWORK study data; to Sebastian Köhler for his contribution to MAAS study data; and to Matti Rookus for her contribution for the EPIC-NL and LIFEWORK data.

Geo-data were collected as part of the Geoscience and Health Cohort Consortium (GECCO), which was financially supported by the Netherlands Organisation for Scientific Research (NWO), the Netherlands Organisation for Health Research and Development (ZonMw; project number 91118017), and Amsterdam UMC. This work was also supported by EXPOSOME-NL, 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).

All included studies were approved by an ethical committee. Detailed descriptions of these procedures and the specificities for each study is given elsewhere (Table 1).

Since the data underlying this article contain privacy-sensitive data, access is restricted by the ethical approvals and the legislation of the European Union.

## References

1. WHO (World Health Organization). 2020. *Obesity and Overweight*. Geneva, Switzerland: WHO. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 20 January 2021].
2. Wild CP. 2012. The exposome: from concept to utility. Int J Epidemiol 41(1):24–32, PMID: 22296988, https://doi.org/10.1093/ije/dyr236.
3. Chandrabose M, Rachele JN, Gunn L, Kavanagh A, Owen N, Turrell G, et al. 2019. Built environment and cardio-metabolic health: systematic review and meta-analysis of longitudinal studies. Obes Rev 20(1):41–54, PMID: 30253075, https://doi.org/10.1111/obr.12759.
4. Lam TM, Vaartjes I, Grobbee DE, Karssenberg D, Lakerveld J. 2021. Associations between the built environment and obesity: an umbrella review. Int J Health Geogr 20(1):7, PMID: 33526041, https://doi.org/10.1186/s12942-021-00260-6.
5. Jia P, Luo M, Li Y, Zheng JS, Xiao Q, Luo J. 2021. Fast-food restaurant, unhealthy eating, and childhood obesity: a systematic review and meta-analysis. Obes Rev 22(suppl 1):e12944, PMID: 31507064, https://doi.org/10.1111/obr.12944.
6. Lin L, Li T, Sun M, Liang Q, Ma Y, Wang F, et al. 2022. Global association between atmospheric particulate matter and obesity: a systematic review and meta-analysis. Environ Res 209:112785, PMID: 35077718, https://doi.org/10.1016/j.envres.2022.112785.
7. Luo Y-N, Huang W-Z, Liu X-X, Markevych I, Bloom MS, Zhao T, et al. 2020. Greenspace with overweight and obesity: a systematic review and meta-analysis of epidemiological studies up to 2020. Obes Rev 21(11):e13078, https://doi.org/10.1111/obr.13078.
8. Huang W-Z, Yang B-Y, Yu H-Y, Bloom MS, Markevych I, Heinrich J, et al. 2020. Association between community greenness and obesity in urban-dwelling Chinese adults. Sci Total Environ 702:135040, PMID: 31726339, https://doi.org/10.1016/j.scitotenv.2019.135040.
9. O'Callaghan-Gordo C, Espinosa A, Valentin A, Tonne C, Pérez-Gómez B, Castaño-Vinyals G, et al. 2020. Green spaces, excess weight and obesity in Spain. Int J Hyg Environ Health 223(1):45–55, PMID: 31679857, https://doi.org/10.1016/j.ijheh.2019.10.007.
10. Gui S-Y, Wu K-J, Sun Y, Chen Y-N, Liang H-R, Liu W, et al. 2022. Traffic noise and adiposity: a systematic review and meta-analysis of epidemiological studies. Environ Sci Pollut Res Int 29(37):55707–55727, PMID: 35320480, https://doi.org/10.1007/s11356-022-19056-7.
11. Choirat C, Braun D, Kioumourtzoglou M-A. 2019. Data science in environmental health research. Curr Epidemiol Rep 6(3):291–299, PMID: 31723546, https://doi.org/10.1007/s40471-019-00205-5.
12. Stafoggia M, Breitner S, Hampel R, Basagaña X. 2017. Statistical approaches to address multi-pollutant mixtures and multiple exposures: the state of the science. Curr Environ Health Rep 4(4):481–490, PMID: 28988291, https://doi.org/10.1007/s40572-017-0162-z.
13. Ohanyan H, Portengen L, Huss A, Traini E, Beulens JWJ, Hoek G, et al. 2022. Machine learning approaches to characterize the obesogenic urban exposome. Environ Int 158:107015, PMID: 34991269, https://doi.org/10.1016/j.envint.2021.107015.
14. Pihur V, Datta S, Datta S. 2009. RankAggreg, an R package for weighted rank aggregation. BMC Bioinformatics 10(1):62, https://doi.org/10.1186/1471-2105-10-62.
15. Kaur P, Singh M, Singh Josan G. 2017. Comparative analysis of rank aggregation techniques for metasearch using genetic algorithm. Educ Inf Technol 22(3):965–983, https://doi.org/10.1007/s10639-016-9467-z.
16. Lakerveld J, Wagtendonk A, Vaartjes I, Karssenberg D, GECCO Consortium. 2020. Deep phenotyping meets big data: the geoscience and hEalth cohort COnsortium (GECCO) data to enable exposome studies in the Netherlands. Int J Health Geogr 19(1):49, PMID: 33187515, https://doi.org/10.1186/s12942-020-00235-z.
17. Stolk RP, Rosmalen JGM, Postma DS, de Boer RA, Navis G, Slaets JPJ, et al. 2008. Universal risk factors for multifactorial diseases: lifelines: a three-generation population-based study. Eur J Epidemiol 23(1):67–74, PMID: 18075776, https://doi.org/10.1007/s10654-007-9204-4.
18. Reedijk M, Lenters V, Slottje P, Pijpe A, Peeters PH, Korevaar JC, et al. 2018. Cohort profile: LIFEWORK, a prospective cohort study on occupational and environmental risk factors and health in the Netherlands. BMJ Open 8(2):e018504, PMID: 29431129, https://doi.org/10.1136/bmjopen-2017-018504.
19. Timmer TC, de Groot R, Habets K, Merz E-M, Prinsze FJ, Atsma F, et al. 2019. Donor InSight: characteristics and representativeness of a Dutch cohort study on blood and plasma donors. Vox Sang 114(2):117–128, PMID: 30590867, https://doi.org/10.1111/vox.12731.
20. Stronks K, Snijder MB, Peters RJ, Prins M, Schene AH, Zwinderman AH. 2013. Unravelling the impact of ethnicity on health in Europe: the HELIUS study. BMC Public Health 13(1):402, PMID: 23621920, https://doi.org/10.1186/1471-2458-13-402.
21. Schram MT, Sep SJS, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. 2014. The Maastricht study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol 29(6):439–451, PMID: 24756374, https://doi.org/10.1007/s10654-014-9889-0.
22. De Graaf R, Ten Have M, Van Gool C, Van Dorsselaer S. 2012. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. Soc Psychiatry Psychiatr Epidemiol 47(2):203–213, PMID: 21197531, https://doi.org/10.1007/s00127-010-0334-8.
23. Ligthart L, van Beijsterveldt CEM, Kevenaar ST, de Zeeuw E, van Bergen E, Bruins S, et al. 2019. The Netherlands twin register: longitudinal research based on twin and twin-family designs. Twin Res Hum Genet 22(6):623–636, PMID: 31666148, https://doi.org/10.1017/thg.2019.93.
24. Verschuren WMM, Blokstra A, Picavet HSJ, Smit HA. 2008. Cohort profile: the Doetinchem Cohort Study. Int J Epidemiol 37(6):1236–1241, PMID: 18238821, https://doi.org/10.1093/ije/dym292.
25. Rutters F, Nijpels G, Elders P, Stehouwer CDA, van der Heijden AA, Groeneveld L, et al. 2018. Cohort profile: the Hoorn studies. Int J Epidemiol 47(2):396–396J, PMID: 29244153, https://doi.org/10.1093/ije/dyx227.
26. van Lenthe FJ, Kamphuis CBM, Beenackers MA, Jansen T, Looman CWN, Nusselder WJ, et al. 2014. Cohort profile: understanding socioeconomic inequalities in health and health behaviours: the GLOBE study. Int J Epidemiol 43(3):721–730, PMID: 23716619, https://doi.org/10.1093/ije/dyt040.
27. Huisman M, Oldehinkel AJ, de Winter A, Minderaa RB, de Bildt A, Huizink AC, et al. 2008. Cohort profile: the Dutch "TRacking adolescents" individual lives' survey'; TRAILS. Int J Epidemiol 37(6):1227–1235, PMID: 18263649, https://doi.org/10.1093/ije/dym273.
28. Huisman M, Poppelaars J, van der Horst M, Beekman ATF, Brug J, van Tilburg TG, et al. 2011. Cohort profile: the longitudinal aging study Amsterdam. Int J Epidemiol 40(4):868–876, PMID: 21216744, https://doi.org/10.1093/ije/dyq219.
29. Jolles J, Boxtel MV, Ponds RWHM, Metsemakers JFM, Houx PJ. 1998. Het longitudinaal perspectief van cognitieve veroudering [The Maastricht aging study (MAAS). The longitudinal perspective of cognitive aging]. Tijdschr Gerontol Geriatr 29(3):120–129 [in Dutch], PMID: 9675779.
30. Bleker LS, De Rooij SR, Painter RC, Ravelli ACJ, Roseboom TJ. 2021. Cohort profile: the Dutch famine birth cohort (DFBC)—a prospective birth cohort study in the Netherlands. BMJ Open 11(3):e042078, PMID: 33664071, https://doi.org/10.1136/bmjopen-2020-042078.
31. Nachtegaal J, Kuik DJ, Anema JR, Goverts ST, Festen JM, Kramer SE. 2009. Hearing status, need for recovery after work, and psychosocial work characteristics: results from an internet-based national survey on hearing. Int J Audiol 48(10):684–691, PMID: 19863354, https://doi.org/10.1080/14992020902962421.
32. van Wier MF, Jansen LA, Goderie T, Stam M, Nachtegaal J, van Beek JHM, et al. 2023. Cohort profile: Netherlands longitudinal study on hearing (NL-SH). BMJ Open 13(4):e070180, PMID: 37068904, https://doi.org/10.1136/bmjopen-2022-070180.

33. GECCO (Geoscience and Health Cohort Consortium). 2022. *Available Exposure Data*. https://www.gecco.nl/exposure-data-1/ [accessed 28 November 2022].
34. KNMI (The Royal Netherlands Meteorological Institute). 2024. *KNMI Data Platform: Temperature*. https://dataplatform.knmi.nl/group/temperature [accessed 28 March 2024].
35. RIVM (Dutch National Institute for Public Health and the Environment). 2020. *Stedelijk hitte-eiland effect (UHI) in Nederland*. https://data.rivm.nl/meta/srv/api/records/c9aa9109-3f32-4f65-84e5-bb1c9ebdfbec [accessed 3 January 2023].
36. Statistics Netherlands. 2012. *District and Neighborhood Map*. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2011 [accessed 18 December 2020].
37. Yang A, Wang M, Eeftens M, Beelen R, Dons E, Leseman DLAC, et al. 2015. Spatial variation and land use regression modeling of the oxidative potential of fine particles. Environ Health Perspect 123(11):1187–1192, PMID: 25840153, https://doi.org/10.1289/ehp.1408916.
38. Ohanyan H, Portengen L, Kaplani O, Huss A, Hoek G, Beulens JWJ, 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 170:107592, PMID: 36306550, https://doi.org/10.1016/j.envint.2022.107592.
39. Lam TM, Wang Z, Vaartjes I, Karssenberg D, Ettema D, Helbich M, et al. 2022. Development of an objectively measured walkability index for the Netherlands. Int J Behav Nutr Phys Act 19(1):50, https://doi.org/10.1186/s12966-022-01270-8.
40. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. 2014. Comparison of random Forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. Am J Epidemiol 179(6):764–774, PMID: 24589914, https://doi.org/10.1093/aje/kwt312.
41. Ishwaran H, Lu M. 2019. Standard errors and confidence intervals for variable importance in random Forest regression, classification, and survival. Stat Med 38(4):558–582, PMID: 29869423, https://doi.org/10.1002/sim.7803.
42. Molnar C. 2020. *Shapley Values | Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/shapley.html [accessed 17 March 2021].
43. Pihur V, Datta S, Datta S. 2008. Finding common genes in multiple cancer types through meta-analysis of microarray experiments: a rank aggregation approach. Genomics 92(6):400–403, PMID: 18565726, https://doi.org/10.1016/j.ygeno.2008.05.003.
44. Pihur V, Datta S, Datta S. 2007. Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach. Bioinformatics 23(13):1607–1615, PMID: 17483500, https://doi.org/10.1093/bioinformatics/btm158.
45. Li X, Wang X, Xiao G. 2019. A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. Brief Bioinform 20(1):178–189, PMID: 28968705, https://doi.org/10.1093/bib/bbx101.
46. Mohammed SH, Habtewold TD, Birhanu MM, Sissay TA, Tegegne BS, Abuzerr S, et al. 2019. Neighbourhood socioeconomic status and overweight/obesity: a systematic review and meta-analysis of epidemiological studies. BMJ Open 9(11):e028238, https://doi.org/10.1136/bmjopen-2018-028238.
47. Feuillet T, Valette JF, Charreire H, Kesse-Guyot E, Julia C, Vernez-Moudon A, et al. 2020. Influence of the urban context on the relationship between neighbourhood deprivation and obesity. Soc Sci Med 265:113537, PMID: 33250318, https://doi.org/10.1016/j.socscimed.2020.113537.
48. Cuevas AG, Kawachi I, Ortiz K, Pena M, Reitzel LR, McNeill LH. 2020. Greater social cohesion is associated with lower body mass index among African American adults. Prev Med Rep 18:101098, PMID: 32382494, https://doi.org/10.1016/j.pmedr.2020.101098.
49. Fields ND, Whitcomb BW, Bertone-Johnson ER, Martínez AD, VanKim NA. 2023. Race-specific associations between psychological distress and obesity: the role of social cohesion. Ethn Health 28(3):446–457, PMID: 35289677, https://doi.org/10.1080/13557858.2022.2052713.
50. Daniels KM, Schinasi LH, Auchincloss AH, Forrest CB, Diez Roux AV. 2021. The built and social neighborhood environment and child obesity: a systematic review of longitudinal studies. Prev Med 153:106790, PMID: 34506813, https://doi.org/10.1016/j.ypmed.2021.106790.
51. Gregorutti B, Michel B, Saint-Pierre P. 2017. Correlation and variable importance in random forests. Stat Comput 27(3):659–678, https://doi.org/10.1007/s11222-016-9646-1.
52. Darst BF, Malecki KC, Engelman CD. 2018. Using recursive feature elimination in random Forest to account for correlated variables in high dimensional data. BMC Genet 19:65, https://doi.org/10.1186/s12863-018-0633-8.
53. Nicodemus KK, Malley JD. 2009. Predictor correlation impacts machine learning algorithms: implications for genomic studies. Bioinformatics 25(15):1884–1890, PMID: 19460890, https://doi.org/10.1093/bioinformatics/btp331.
54. Tolosi L, Lengauer T. 2011. Classification with correlated features: unreliability of feature ranking and solutions. Bioinformatics 27(14):1986–1994, PMID: 21576180, https://doi.org/10.1093/bioinformatics/btr300.
55. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. 2008. Conditional variable importance for random forests. BMC Bioinformatics 9(1):307, https://doi.org/10.1186/1471-2105-9-307.