

Geographic clustering and region-specific determinants of obesity in the Netherlands

Ge Qiu,^{1,2} Xiaojian Liu,² Arsha Yuditha Amiranti,² Mulimba Yasini,² Tong Wu,^{1,3} Sherif Amer,² Peng Jia^{1,2}

¹International Initiative on Spatial Lifecourse Epidemiology (ISLE), Hong Kong, China; ²Faculty of Geo-information Science and Earth Observation, University of Twente, Enschede, The Netherlands; ³Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China

Abstract

As a leading cause of morbidity and premature mortality, obesity has become a major global public health problem. It is therefore important to investigate the spatial variation of obesity prevalence and its associations with environmental and behavioral factors (e.g., food environment, physical activity), to optimize the targeting of scarce public health resources. In this study, the geographic clustering of obesity in the Netherlands was explored by analyzing the local spatial autocorrelation of municipal-level prevalence rates of adulthood obesity (aged ≥ 19 years) in 2016. The potential influential factors that may be associated with obesity prevalence were first selected from five categories of health-related factors through binary and Least Absolute Shrinkage and Selection Operator (LASSO) regressions. Geographically Weighted Regression (GWR) was then used to investigate the spatial variations of the associations between those selected factors and obesity prevalence. The results revealed marked geographic variations in obesity prevalence, with four clusters of high prevalence in the north, south, east, and west, and three clusters of low prevalence in the north and south of the Netherlands. Lack of sports participation, risk of anxiety, falling short of physical activity guidelines, and the number of restaurants around homes were

found to be associated with obesity prevalence across municipalities. Our findings show that effective, region-specific strategies are needed to tackle the increasing obesity in the Netherlands.

Introduction

Obesity is abnormal or excessive fat accumulation that contributes to an increased risk for several chronic diseases, including type 2 diabetes, osteoarthritis, various ischemic heart diseases, and certain types of cancers (Elmokhallalati *et al.*, 2019; Jia *et al.*, 2017; WHO, 2004a). Obesity can severely affect individuals' psychosocial state and quality of life (Pouliou and Elliott, 2009). The Body Mass Index (BMI, kg/m^2), calculated by dividing one's weight (in kg) by the square of one's height (in m), has been widely used as a measure of adulthood overweight ($\text{BMI} \geq 25$) and obesity ($\text{BMI} \geq 30$), according to the standards of the World Health Organization (WHO, 2014b). The global prevalence of adulthood obesity nearly tripled from 4.5% to 12.8% during 1975-2016, accounting for 0.7-2.8% of the total national healthcare expenditures worldwide (Abarca-Gómez *et al.*, 2017; Arroyo-Johnson and Mincey, 2016; FAO *et al.*, 2017; Withrow and Alter, 2011; Zhang *et al.*, 2020b).

In the Netherlands, the mean BMI among adults reached $25.4 \text{ kg}/\text{m}^2$ in 2014, which exceeds the standard for overweight. According to the World Health Organization (WHO, 2014b), the crude adjusted estimates of the prevalence of adulthood obesity in the Netherlands (19.9%) was higher than the average level worldwide (12.8%) in 2014. The obesity prevalence among adults in the Netherlands has been examined in some studies. For example, Visscher *et al.* (2002) reported that the obesity rate of their sample (51,067 people) in the Netherlands has steadily increased from 1976 to 1997 in nearly all age categories. The highest obesity prevalence was found from 1993 to 1997, among men and women with relatively low and high educational levels, respectively. Schokker *et al.* (2007) observed the trend of adulthood obesity from 1981 to 2004, and found that the prevalence of obesity increased from 4% to 12% for males and from 6% to 12% for females. Factors associated with obesity were also studied. For example, based on data from postal and online questionnaires, Bunt *et al.* (2017) analyzed the relationship between BMI and personal characteristics, including age, ethnicity, marital status, education level, and income, among men and women based on multiple logistic regression. However, few studies have investigated the spatial pattern of obesity prevalence. van de Kasstele *et al.*

Correspondence: Peng Jia, International Initiative on Spatial Lifecourse Epidemiology (ISLE), Hong Kong, China.
E-mail: jiapengff@hotmail.com

Key words: Obesity; LASSO; Geographically weighted regression; Netherlands.

Conflict of Interest: The Authors declare no potential conflict of interests.

Received for publication: 29 November 2019.
Accepted for publication: 16 April 2020.

©Copyright: the Author(s), 2020
Licensee PAGEPress, Italy
Geospatial Health 2020; 15:839
doi:10.4081/gh.2020.839

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.



(2017) used a structured additive regression model to estimate the obesity rate and 25 other health-related indicators across neighborhoods, districts, and municipalities based on data collected from 387,195 people in 2012. However, most if not all studies have been conducted based on individual samples. None of existing studies have examined geographic clustering and region-specific determinants of obesity in the Netherlands.

Identifying influential factors and their local associations with obesity prevalence are important to target scarce public health resources as well as suggest region-specific interventions (Bethell *et al.*, 2010; Cobb *et al.*, 2015). The factors that potentially contribute to obesity prevalence are various (Jia *et al.*, 2019b; Jia *et al.*, 2019g; Li *et al.*, 2019; Wang *et al.*, 2020; Xu *et al.*, 2020; Yang *et al.*, 2020; Zhang *et al.*, 2020a). Least Absolute Shrinkage and Selection Operator (LASSO) is an efficient and stable way to identify multicollinearity and select significant influential factors from a set of candidate factors (Tibshirani, 1996). For instance, Guo *et al.* (2015) used a LASSO-based method to select factors relevant to the hepatitis B virus (HBV) infection. Zhang *et al.* (2018) applied the LASSO model to dietary data and identified dietary patterns related to risk factors for cardiovascular diseases. The spatial variation of the relationship between obesity prevalence and associated socio-economic and behavioral factors can be identified by Geographically Weighted Regression (GWR), which is a method extended from the traditional regression framework (Fotheringham *et al.*, 1998). GWR estimates local rather than global parameters (*i.e.*, coefficients) of variables over space (Fraser *et al.*, 2012). Shahid and Bertazzon (2015) studied the spatial relationship between neighborhood walkability and child obesity prevalence in Calgary, Canada. The authors employed GWR to identify the vulnerable neighborhoods where targeted interventions are most needed. With an increasing number of candidate factors that potentially relate to obesity prevalence, machine learning offers effective methods such as LASSO for variable selection. The most relevant influential factors identified using the LASSO model were further combined with GWR to explore their region-specific associations with obesity prevalence.

To the best of author's knowledge, this study for the first time explored geographic clustering of obesity prevalence and their region-specific environmental and behavioral determinants in the Netherlands. The spatial unit of our analysis is the municipality as this level is suitable for both national and local policy-makers to design and implement interventions for controlling and preventing obesity (Jia *et al.*, 2019c; Jia *et al.*, 2019e). This study would contribute to our understanding of the influential factors driving obesity issues in the Netherlands, and provide evidence for region-varying strategies to mitigate obesity prevalence at an actionable level.

Materials and methods

Datasets

The municipality-level obesity rates among the Dutch population aged ≥ 19 years in 2016 were acquired from the Public Health Monitor, a national database designed by Statistics Netherlands (CBS), the Municipal Health Service (MHS), and the National Institute for Public Health and the Environment (RIVM) (National Institute for Public Health and the Environment, 2016). Five categories of factors (details presented below) that are potentially associated with obesity were obtained from CBS and the RIVM, also

for 2016 (National Institute for Public Health and the Environment, 2016; Statistics Netherlands (CBS) 2018).

Variables

Outcome variable

The municipality level obesity rates were calculated on the basis of self-reported height and weight collected from 457,153 people through a nationwide questionnaire survey in 2016. Obesity was defined as $BMI \geq 30 \text{ kg/m}^2$. Detailed information about the method for producing obesity rate data can be found in the study of van de Kasstele *et al.* (2017)

Explanatory variables

Physical Activity (PA) and dietary behavior (*e.g.*, fat intake and alcohol consumption) are decisive to human energy intake and consumption (Arredondo *et al.*, 2019). In this study, PA was approximated by using four indicators: 1) % of sports participation, calculated as the percentage of persons that engage in sports activities at least once a week; 2) % of meeting PA guidelines, measured as the percentage of people who meet the guidelines for physical exercise proposed by the Ministry of Health, Welfare, and Sports of the Netherlands; 3) % of population with one or more chronic conditions (lasting for six months or longer); 4) % of moving restrictions, calculated as the percentage of the population with limitations of movement. As part of dietary behaviors, alcohol consumption is often related to overweight and obesity, since alcohol energy (7kcal/g) is likely to be a contributing factor to weight gain, and unhealthy alcohol consumption habits also have an impact on the development of obesity (Traversy and Chaput, 2015; Yoon *et al.*, 2016). We included alcohol consumption data, which show the percentage of people aged 19 and older who drink zero or a maximum of one glass of alcohol per day as a variable. Socio-demographic factors including age, sex, ethnicity, income, and education level are also related to obesity (Cook *et al.*, 2017). The percentage of different age groups (*i.e.*, 15-24, 25-44, 45-64, and 65 and older (years)), ethnic groups (*e.g.*, Western immigrants (Europe, North America) and immigrants from Morocco, the Antilles and Aruba, Suriname, and Turkey), and the number of men and women were included in this study. Moreover, we employed mental status factors, including the risk of anxiety and loneliness, which were estimated through a questionnaire survey conducted by the Public Health Monitor (National Institute for Public Health and the Environment, 2016).

Environmental factors, including the food environment (*e.g.*, the number of fast food outlets and the diversity of food choices) and the PA environment, are related to the individual weight status. The factors of food environments, such as the number of fast food outlets and the diversity of food choices, indirectly affect energy intake and are related to the obesogenic environment (Fraser *et al.*, 2012). We employed the average road distance from the addresses of residents to the nearest large supermarkets (area larger than 150m^2), small food stores, department stores, fast food restaurants, and restaurants as our food environment factors. The average number of these locations within a road distance of 1, 3, and 5 km (from the addresses of residents) were also used. Factors of the PA environment, like the neighborhood built environment, public transport use, and crime rates, are closely related to walkability and PA (Huang *et al.*, 2015). Also, a high residential density can lead to a vibrant street life, which can make people more physically active (Jia *et al.*, 2019d). In this study, we used the crime rate

(for violent and sexual crimes), the residential density (addresses/km²), and the average number of recreational attractions (e.g., amusement parks, zoos, indoor playgrounds) within road distances of 10, 20, and 50 km. The average road distance from the addresses of residents to the nearest main traffic network (i.e., national and provincial roads), train stations, transfer stations (i.e., railway stations that allow travelers to transfer between different trains), and recreational attractions (e.g., amusement parks, zoos, indoor playgrounds) were also used.

Analysis methods

Identification of spatial clusters

The spatial pattern of obesity prevalence in the Netherlands was explored by analyzing the global Moran's I (which measures the overall spatial clustering of obesity prevalence) and the Local Indicators of Spatial Association (LISA) (which measures the degree of spatial autocorrelation of obesity prevalence at each specific location). The global Moran's I (1) can be used to quantify the overall spatial characteristics (e.g., dispersed, random, and clustered) of a geographic phenomenon (Moran, 1950):

$$I = \frac{N \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2} \quad (\text{Eq. 1})$$

where N represents the number of spatial units indexed by i and j ; x is the variable of interest (obesity rate); \bar{x} is the mean of x ; w_{ij} is a matrix of spatial weight with zeros on the diagonal (i.e., $w_{ii}=0$); and W is the sum of w_{ij} . The result of the global Moran's I ranges from -1 to 1, in which "1" suggests a perfect positive spatial autocorrelation (high and low values clustered, respectively), and "-1" means a perfect negative spatial autocorrelation (perfectly dispersed) (Fu *et al.*, 2014).

After confirmation of the existing global spatial autocorrelation, the LISA was used to assess the location, scope, and degree of the spatial autocorrelation of obesity prevalence. In this study, the local Moran's I (2) (Anselin, 1995) was employed to calculate the local spatial autocorrelation:

$$I_i = \frac{x_i - \bar{x}}{\sigma^2} \sum_{j=1, j \neq i}^n [w_{ij} (x_j - \bar{x})] \quad (\text{Eq. 2})$$

where x_i is the value of interest (i.e., obesity rate) at location i ; \bar{x} is the mean of all x ; x_j is another location (where $i \neq j$); σ^2 represents the variance of x ; and w_{ij} is the distance between x_i and x_j . In particular, $I_i \geq E(I_i)$ means that there are units with a similar value surrounding I_i and if $I_i < E(I_i)$, there are negative spatial autocorrelations between the unit I_i and the surrounding units. The LISA uses a Monte Carlo permutation approach to test the local Moran's I result for significance. In the Monte Carlo permutation, the obesity rate in each municipality was randomly assigned. For each permutation, the local Moran's I was calculated. By comparing our observed local Moran's I and the distribution of values from the Monte Carlo permutation, the significance of our observed local Moran's I was tested. The results of the LISA significance map consist of five categories of obesity regions: i) "high-high" regions indicate a high obesity rate clustering; ii) "low-low" regions indicate a low obesity rate clustering; iii) "low-high" and iv) "high-low" regions indicate significant negative spatial autocorrelations;

v) "not significant" regions indicate that no spatial autocorrelation existed. The calculation of the global Moran's I and LISA were processed in GeoDa 1.14.0 (Anselin *et al.*, 2006). The major steps of this study are shown in the flowchart of Figure 1.

Identification of influential factors

The selected explanatory variables were not all equally important for obesity in the Netherlands. To extract the most influential factors, two subsequent analyses were conducted. First, binary regression was performed between each variable and the obesity rate. Variables that were statistically significant under a significance level of 0.05 were selected. After that, LASSO regression (Tibshirani, 1996) was utilized to select important variables. Instead of minimizing the sum of residuals, LASSO regression minimizes the sum of squared residuals subject to a constraint: $\sum |\beta_j| \leq t$, where β_j is the p -vector of parameters (Fu, 1998). Because of the nature of this constraint, according to Tibshirani (1996), it tends to set some coefficients to 0 and therefore removes unimportant variables (variables that cause a large variance which reduce the accuracy of the prediction). LASSO regression can also remove multicollinear variables, which allows LASSO regression to treat data with high dimensions in a time-efficient way. By using LASSO regression, we were able to select the most relevant influential factors and remove multicollinear variables simultaneously. A 20-fold cross-validation was used to tune the regularization parameter of LASSO regression and to save variables with non-zero coefficients. The calculation in this section was processed in R 3.5.1 (R Core Team, 2018).

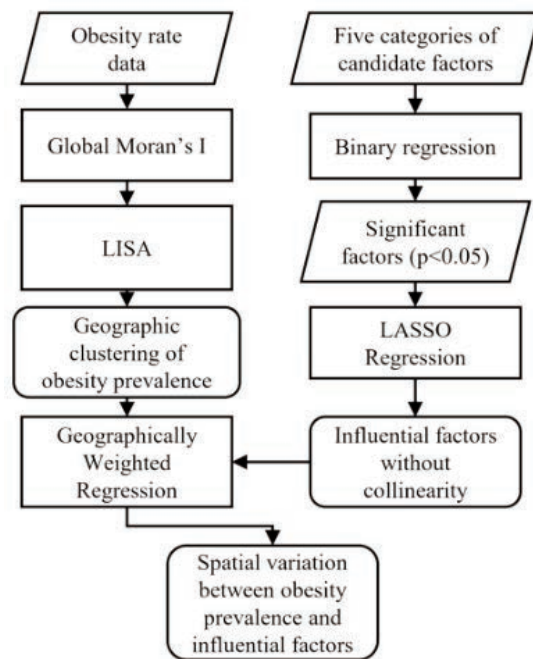


Figure 1. Flowchart of identifying geographic clustering and influential factors of obesity prevalence, and ranking local influential factors by importance. LISA: Local Indicators of Spatial Association; LASSO: Least Absolute Shrinkage and Selection Operator.

Ranking of local influential factors

In this stage, an Ordinary Least Squares (OLS) regression was utilized to explore the global relationship between obesity prevalence and the selected influential factors. GWR was subsequently utilized to examine the variations of the relationship between obesity and the identified covariates across municipalities in the Netherlands. GWR (4) was defined by Fotheringham *et al.* (1998):

$$y_i = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (\text{Eq. 3})$$

where (u_i, v_i) represents the coordinates of the i^{th} point in space (*i.e.*, center of the i^{th} municipality); $a_k(u_i, v_i)$ is the spatial function of the k^{th} parameter of sample i ; x_{ik} is the value of the k^{th} variable at point i ; and ε is the random error, assumed to be normally distributed.

By introducing geographic coordinates into traditional regression, GWR permits the parameters to differ across regions. Unlike the OLS, which produces coefficients based on the number of variables, GWR generates coefficients across all samples, which means that every municipality in the Netherlands has its own coefficients. The importance of each influential factor in each municipality can be evaluated based on these coefficients. The results of GWR were calculated in the software environment GWR 4.0.8 (Nakaya *et al.*, 2009). A Gaussian GWR model was chosen based on the continuous outcome variable (*i.e.*, municipal level obesity rate). An adaptive kernel width was used, and the optimal bandwidth of 60 was applied on the basis of a minimum corrected Akaike information criterion.

Results

Spatial clusters of obesity prevalence

The result of the global Moran's I was 0.547 with a z-score equal to 16.87 and a p-value nearly equal to 0, which indicated that the spatial pattern of obesity prevalence in the Netherlands was not randomly distributed. In addition, by using the local Moran's I, the LISA demonstrated that four major clusters with a higher obesity prevalence (red color) existed in the north, south, east, and west of the Netherlands (Figure 2). Three clusters with a lower obesity prevalence were also found in the north and south of the Netherlands.

Influential factors of obesity prevalence

Thirty variables were selected by using the binary regression with $p < 0.05$. LASSO regression was used to further select nine variables that were strongly related to obesity prevalence: % of sports participation, % of chronic conditions, % of meeting PA guidelines, % of moving restrictions, average distance from addresses of residents to transfer stations, % of Turkish population, risk of anxiety, average number of restaurants within 1 km (from the addresses of residents), and average number of attractions within 10 km (from the addresses of residents) were strongly associated with the obesity rates of municipalities in the Netherlands.

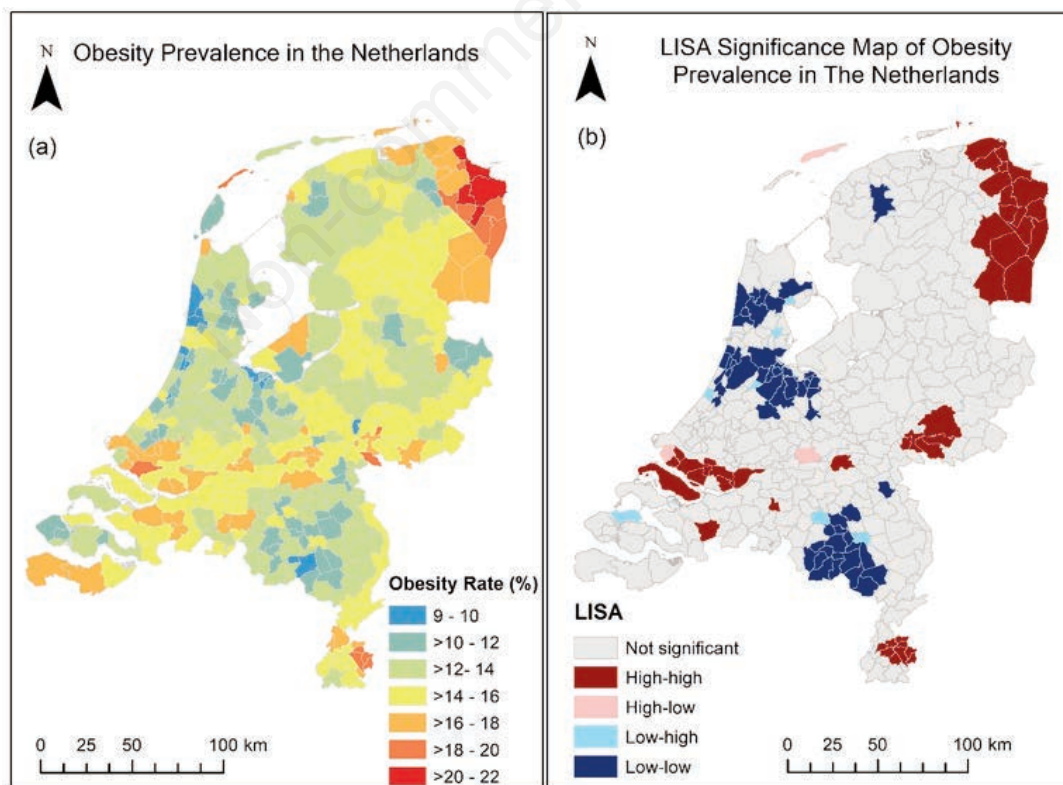


Figure 2. (a) Obesity prevalence map and (b) geographic clustering map of obesity prevalence. LISA: Local Indicators of Spatial Association.

Importance of local influential factors

GWR and the 9 variables listed above were used to measure the spatial variation of the relationship between influential factors and obesity prevalence. Characteristics of these variables and obesity prevalence are shown in Table 1. Results showed that GWR performed better than OLS, given the higher adjusted R-square (0.76) and the lower Akaike Information Criterion (1184.03) (Table 2).

Using the mean coefficient values of the influential factors identified by GWR, the importance of the influential factors in explaining obesity prevalence could be ranked (Table 3). The % of sports participation was the most influential factor with respect to obesity prevalence in the Netherlands. Following that, associations of risk of anxiety, % of meeting PA guidelines, average number of

restaurants within 1 km, and % of chronic conditions with obesity prevalence were also found.

The relative importance of the influential factors was separately ranked for each of the seven identified spatial clusters of high or low obesity prevalence (Figure 3). In the northern clusters of high obesity prevalence, % of sports participation was shown to be the most important factor. In the southern cluster of high obesity prevalence, % of chronic conditions and risk of anxiety were positively associated with obesity prevalence, whereas % of sports participation and % of meeting PA guidelines showed a negative contribution to obesity prevalence. For the eastern cluster of high obesity prevalence, sports participation and anxiety were the most important factors. The % of meeting PA guidelines was the most influential factor with respect to obesity in the western cluster of high obesity prevalence.

Table 1. Statistics of obesity prevalence and the selected influential factors and across municipalities in the Netherlands.

Variable	Min.	Median	Mean	Max.	Standard deviation
Outcome variable					
Obesity rate (%)	9	15	14.54	22	2.19
Influential factors					
% of the Turkish population	0	0	1.05	10	1.72
Average number of restaurants within 1 km	0.1	3.1	4.25	51.7	4.46
Average road distance to transfer stations (in km)	1.3	12.25	14.49	63.1	9.68
Average number of attractions within 10 km	0	1.1	1.50	7.8	1.39
Risk of anxiety (%)	31	42	41.91	56	4.62
% of meeting PA guidelines	35	50.5	49.81	66	4.48
% of moving restrictions	6	10	9.89	17	1.79
% of chronic conditions	27	34	34.13	45	2.94
% of sports participation	35	49	48.95	63	5.48

Table 2. Results of Ordinary Least Squares (OLS) regression and Geographically Weighted Regression (GWR).

Model	Bandwidth	AIC*	R-square	Adjusted R-square	F-test
GWR	60	1,184.03	0.7903	0.7633	3.8078
OLS	N/A	1,262.7143	0.707	0.7	N/A

*AIC: Akaike Information Criterion

Table 3. Summary of the coefficients of the influential factors measured by geographically weighted regression.

Variable	Mean	Standard deviation	Range	Median
% of sports participation	-0.8525	0.3056	1.2285	-0.9062
Risk of anxiety	0.5699	0.1502	0.5706	0.5653
% of meeting PA guidelines	-0.4019	0.3143	1.2602	-0.259
Average number of restaurants within 1 km	-0.3792	0.1007	0.3972	-0.3522
% of chronic conditions	0.3760	0.1566	0.5819	0.399
% of Turkish population	0.2978	0.0889	0.3731	0.3114
Average road distance to transfer stations	0.2047	0.099	0.4174	0.1864
% of moving restrictions	0.1258	0.278	0.9926	0.0214
Average number of attractions within 10 km	-0.1115	0.0532	0.2521	-0.1164



Discussion

Many studies have explored the spatial patterns of obesity prevalence in various countries and regions, including the United States (Michimi and Wimberly, 2010), Canada (Hajizadeh *et al.* 2016), the United Kingdom (Fraser *et al.* 2012), Iran (Farhadian *et al.*, 2013), and mainland China (Cui *et al.*, 2018). However, in the Netherlands, similar efforts have been mostly focused on temporal variations of obesity prevalence or on factors that are associated with obesity at the individual level. In contrast, limited efforts have been focused on the identification of spatial clusters of obesity prevalence at the municipality level, which is more suitable for stakeholders and decision-makers to understand the public health implications of the obesity problem and its potential resolution by implementing relevant policies that focus on the reduction of obesity (which are often targeted at the local government level). To fill this gap, we systematically examined the spatial pattern of obesity prevalence across municipalities in the Netherlands in 2016, as well as the impacts of environmental and social demographic factors on obesity-clustering regions.

Based on the combination of spatial analysis techniques and obesity-related data, this study represents the first effort to map municipality-level obesity spatial clusters in the Netherlands. The results of the global spatial autocorrelation identified by the global Moran's I suggested a significant spatial heterogeneity of obesity prevalence across municipalities in the Netherlands. Such spatial heterogeneity was also found in many other countries and regions at different scales (Michimi and Wimberly, 2010; Poulidou and Elliott, 2009). Additionally, the significant local clusters identified by the LISA demonstrated the spatial connections among high and low obesity rate regions. Therefore, policies should target public health resources more towards these "high-high" regions. In addition,

from the perspective of our GWR results, the relationship between influential factors and obesity prevalence in these "high-high" regions were area-specific. For the western cluster of high obesity prevalence, the most influential factor was % of meeting PA guidelines. By taking the low sports participation rate in this region into account, it is believed that developing campaigns to stimulate physical exercise and increase sport participation in this region could be an effective means to lower the obesity prevalence (Brockmann and Ross, 2020; Keane *et al.*, 2017). Moreover, higher sports participation was the main factor associated with the lower obesity rate in the two northern clusters of low obesity prevalence. Improving the sports participation rate thus appears to be an effective way to mitigate obesity prevalence in the Netherlands, and greater efforts should be undertaken to popularize sport in regions where participation rates are low. In the eastern cluster of high obesity prevalence, the risk of anxiety was the second largest contributor to obesity prevalence, after lack of sports participation. Many studies have revealed the correlation between anxiety/depression and obesity (Baumeister and Harter, 2007; Petry *et al.*, 2008). Here, it is suggested that targeting anxiety – for instance through the improved provisioning of mental health resources (Heuzenroeder *et al.*, 2004), including public education literature and access to medical care – may be useful for reducing the burden of obesity in this region. This could also apply for preventing anxiety, which was another factor contributing to obesity across all regions. For the northern cluster of high obesity prevalence, limitations on movement was a major factor for the higher obesity rates. Region-specific actions that can mitigate the obesity prevalence here could include improved urban planning for pedestrian-friendly surfaces and access to greenspaces and other public venues (Durand *et al.*, 2011).

There are several limitations in this study. First, the cross-sectional dataset used in this study cannot provide temporal compar-

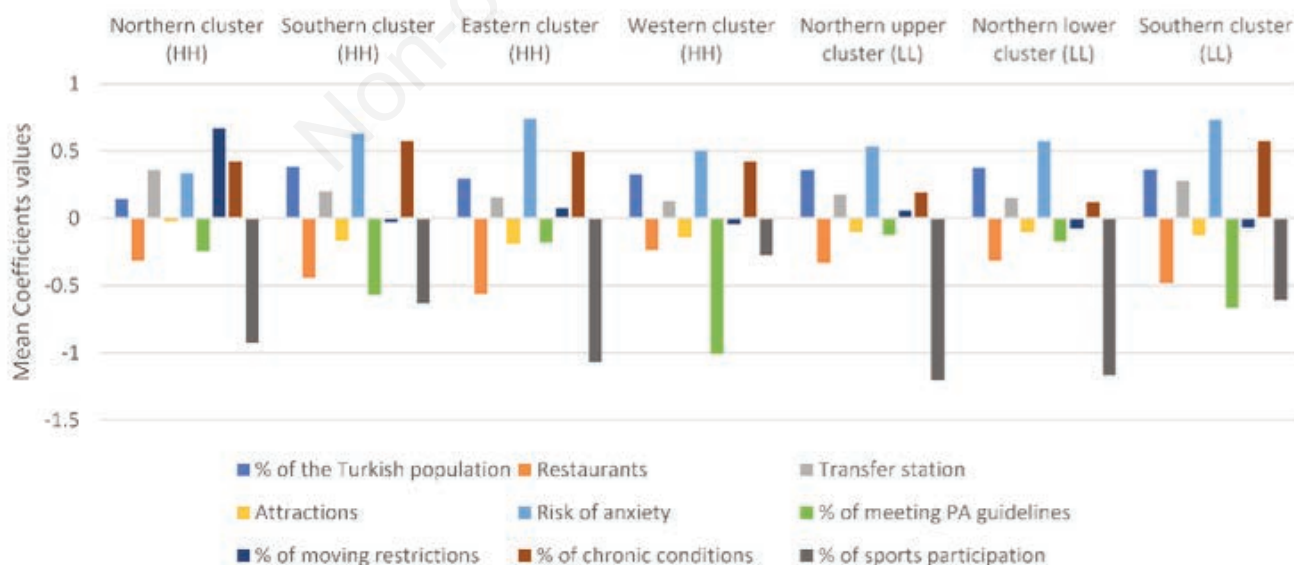


Figure 3. Mean coefficient values of municipalities in seven obesity clusters. Names of clusters: HH = high obesity prevalence clusters; LL = low obesity prevalence clusters. Names of influential factors: Transfer station = Average road distance to transfer stations; Attractions = Average number of attractions within 10 km; Restaurants = Average number of restaurants within 1 km.

isons for result verification. This needs to be further confirmed in future studies with additional datasets featuring relevant time-series information (Jia *et al.*, 2019a; Jia *et al.*, 2019f). Second, although we used binary regressions to confirm the significance of influential factors, we did not conduct significance tests on the coefficient results from GWR, which means that some coefficients at some locations may not be robust. It would be more suitable, whenever possible, to conduct significance tests on the coefficients from GWR based on different methods, such as the Monte Carlo simulation. Third, this study was constrained by a lack of data on income and educational level, which could be significant factors influencing obesity prevalence (Chen *et al.*, 2015; Ogden *et al.*, 2017). Lastly, during the selection of influential factors, the used binary and LASSO regression only took the global relationship between obesity prevalence and explanatory variables into account. It is more suitable to conduct binary and LASSO regression locally (*i.e.*, conduct binary and LASSO regression for each identified cluster), so the spatial variation of the relationship between obesity prevalence and influential factors could be explored more explicitly in the GWR.

Conclusions

The main conclusions of this study are as follows: First, obesity prevalence in the Netherlands exhibits strong spatial patterns. Four areas with significantly higher and three areas with significantly lower obesity rates were identified; Second, the lack of sports participation, risk of anxiety, falling short of PA guidelines, and the number of restaurants around homes were highly related to obesity prevalence across different municipalities. Effective actions, such as strengthening the public education of PA and increasing the public health resources for preventing anxiety, should be adopted; Third, the influential factors driving obesity prevalence in the four high-obesity prevalence clustering areas were not the same. This suggests that specific public health resources and prevention strategies should be targeted to different regions based on the varying combinations of environmental and socioeconomic risk factors that affect obesity.

References

- Abarca-Gómez L, Abdeen ZA, Hamid ZA, Abu-Rmeileh NM, Acosta-Cazares B, Acuin C, Adams RJ, Aekplakorn W, Afsana K, Aguilar-Salinas CA, 2017. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128· 9 million children, adolescents, and adults. *Lancet* 390:2627-42.
- Anselin L, 1995. Local indicators of spatial association—LISA. *Geogr Anal* 27:93-115.
- Anselin L, Syabri I, Kho Y, 2006. GeoDa: An Introduction to Spatial Data Analysis. *Geogr Anal* 38:5-22.
- Arredondo A, Torres C, Orozco E, Pacheco S, Huang FY, Zambrano E, Bolanos-Jimenez F, 2019. Socio-economic indicators, dietary patterns, and physical activity as determinants of maternal obesity in middle-income countries: Evidences from a cohort study in Mexico. *Int J Health Plann Manage* 34:E713-25.
- Arroyo-Johnson C, Mincey KD, 2016. Obesity epidemiology worldwide. *Gastroenterol Clin N* 45:571-9.
- Baumeister H, Harter M, 2007. Mental disorders in patients with obesity in comparison with healthy probands. *Int J Obes* 31:1155-64.
- Bethell C, Simpson L, Stumbo S, Carle AC, Gombojav N, 2010. National, State, And Local Disparities In Childhood Obesity. *Health Aff (Millwood)* 29:347-56.
- Brockmann AN, Ross KM, 2020. Bidirectional association between stress and physical activity in adults with overweight and obesity. *J Behav Med* 43:246-53.
- Bunt S, Mérelle S, Steenhuis I, Kroeze W, 2017. Predictors of need for help with weight loss among overweight and obese men and women in the Netherlands: a cross-sectional study. *BMC Health Serv Res* 17:819.
- Chen Y, Rennie DC, Karunanayake CP, Janzen B, Hagel L, Pickett W, Dyck R, Lawson J, Dosman JA, Pahwa P, Saskatchewan Assoc HLth Study G, 2015. Income adequacy and education associated with the prevalence of obesity in rural Saskatchewan, Canada. *BMC Public Health* 15:700.
- Cobb LK, Appel LJ, Franco M, Jones-Smith JC, Nur A, Anderson CAM, 2015. The relationship of the local food environment with obesity: A systematic review of methods, study quality, and results. *Obesity* 23:1331-44.
- Cook WK, Tseng W, Tam C, John I, Lui C, 2017. Ethnic-group socioeconomic status as an indicator of community-level disadvantage: A study of overweight/obesity in Asian American adolescents. *Soc Sci Med* 184:15-22.
- Cui J, Sun XF, Li XJ, Ke M, Sun JP, Yasmeeen N, Khan JM, Xin HL, Xue SY, Baloch Z, 2018. Association Between Different Indicators of Obesity and Depression in Adults in Qingdao, China: A Cross-Sectional Study. *Front Endocrinol (Lausanne)* 9:549
- Durand CP, Andalib M, Dunton GF, Wolch J, Pentz MA, 2011. A systematic review of built environment factors related to physical activity and obesity risk: implications for smart growth urban planning. *Obes Rev* 12:e173-82.
- Elmokhallalati Y, FarajAllah H, Albarqouni L, 2019. Socio-demographic and economic determinants of overweight and obesity in preschool children in Palestine: analysis of data from the Palestinian Multiple Indicator Cluster Survey. *Lancet* 393:S22.
- FAO, IFAD, UNICEF, WFP and WHO, 2017. The State of Food Security and Nutrition in the World 2017. Building resilience for peace and food security. Rome, FAO. Available from: <https://www.who.int/nutrition/publications/foodsecurity/state-food-security-nutrition-2017/en/>.
- Farhadian M, Moghimbeigi A, Aliabadi M, 2013. Mapping the Obesity in Iran by Bayesian Spatial Model. *Iran J Public Health* 42:581-7.
- Fotheringham AS, Charlton ME, Brunsdon C, 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A* 30:1905-27.
- Fraser LK, Clarke GP, Cade JE, Edwards KL, 2012. Fast Food and Obesity A Spatial Analysis in a Large United Kingdom Population of Children Aged 13-15. *Am J Prev Med* 42: E77-E85.
- Fu WJ, Jiang PK, Zhou GM, Zhao KL, 2014. Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeosciences* 11:2401-9.



- Fu WJ, 1998. Penalized regressions: The bridge versus the lasso. *J Comput Graph Stat* 7:397-416.
- Guo P, Zeng F, Hu X, Zhang D, Zhu S, Deng Y, Hao Y, 2015. Improved Variable Selection Algorithm Using a LASSO-Type Penalty, with an Application to Assessing Hepatitis B Infection Relevant Factors in Community Residents. *PLoS One* 10:e0134151
- Hajizadeh M, Campbell MK, Sarma S, 2016. A Spatial Econometric Analysis of Adult Obesity: Evidence from Canada. *Appl Spat Anal Polic* 9:329-63.
- Heuzenroeder L, Donnelly M, Haby MM, Mihalopoulos C, Rossell R, Carter R, Andrews G, Vos T, 2004. Cost-effectiveness of psychological and pharmacological interventions for generalized anxiety disorder and panic disorder. *Aust N Z J Psychiatry* 38:602-12.
- Huang R, Moudon AV, Cook AJ, Drewnowski A, 2015. The spatial clustering of obesity: does the built environment matter? *J Hum Nutr Diet* 28:604-12.
- Jia P, Cheng X, Xue H, Wang Y, 2017. Applications of geographic information systems (GIS) data and methods in obesity-related research. *Obes Rev* 18:400-11.
- Jia P, Lakerveld J, Wu JG, Stein A, Root ED, Sabel CE, Vermeulen R, Remais JV, Chen X, Brownson RC, Amer S, Xiao Q, Wang LM, Verschuren WMM, Wu T, Wang YF, James P, 2019a. Top 10 Research Priorities in Spatial Lifecourse Epidemiology. *Environ Health Perspect* 127:074501.
- Jia P, Luo MY, Li YM, Zheng JS, Xiao Q, Luo JY, 2019b. Fast-food restaurant, unhealthy eating, and childhood obesity: A systematic review and meta-analysis. *Obes Rev*: 1-27.
- Jia P, Ma S, Qi X, Wang YF, 2019c. Spatial and Temporal Changes in Prevalence of Obesity Among Chinese Children and Adolescents, 1985-2005. *Prev Chronic Dis* 16:E160
- Jia P, Xue H, Cheng X, Wang YG, Wang YF, 2019d. Association of neighborhood built environments with childhood obesity: Evidence from a 9-year longitudinal, nationally representative survey in the US. *Environ Int* 128:158-64.
- Jia P, Xue H, Yin L, Stein A, Wang MQ, Wang YF, 2019e. Spatial Technologies in Obesity Research: Current Applications and Future Promise. *Trends Endocrinol Metab* 30:211-23.
- Jia P, Yu C, Remais JV, Stein A, Liu Y, Brownson RC, Lakerveld J, Wu T, Yang L, Smith M, Amer S, Pearce J, Kestens Y, Kwan M-P, Lai S, Xu F, Chen X, Rundle A, Xiao Q, Xue H, Luo M, Zhao L, Cheng G, Yang S, Zhou X, Li Y, Panter J, Kingham S, Jones A, Johnson BT, Shi X, Zhang L, Wang L, Wu J, Mavoa S, Toivonen T, Mwenda KM, Wang Y, Verschuren WMM, Vermeulen R, James P, 2019f. Spatial Lifecourse Epidemiology Reporting Standards (ISLE-ReSt) statement. *Health Place* 61:102243. doi: 10.1016/j.healthplace. 2019. 102243.
- Jia P, Zou Y, Wu Z, Zhang D, Wu T, Smith M, Xiao Q, 2019g. Street connectivity, physical activity, and childhood obesity: A systematic review and meta-analysis. *Obes Rev* 1-11. Available from: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/obr.12943>.
- Keane E, Li X, Harrington JM, Fitzgerald AP, Perry IJ, Kearney PM, 2017. Physical Activity, Sedentary Behaviour and the Risk of Overweight and Obesity in School Aged Children. *Pediatr Exerc Sci* 29:408-18.
- Schokker DF, Visscher TLS, Nooyens ACJ, van Baak MA, Seidell JC, 2007. Prevalence of overweight and obesity in the Netherlands. *Obes Rev* 8:101-7.
- Li YM, Luo MY, Wu XY, Xiao Q, Luo JY, Jia P, 2019. Grocery store access and childhood obesity: A systematic review and meta-analysis. *Obes Rev*. doi: 10.1111/obr.12945.
- Michimi A, Wimberly MC, 2010. Spatial Patterns of Obesity and Associated Risk Factors in the Conterminous U.S. *Am J Prev Med* 39:E1-12.
- Moran PA, 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.
- Nakaya T, Fotheringham S, Charlton M, Brunsdon C, 2009. Semiparametric geographically weighted generalised linear modelling in GWR 4.0. In Lees (Ed.), 10th International Conference on GeoComputation (pp. 1-5). UNSW, Sydney
- National Institute for Public Health and the Environment, 2016. [Gezondheid per buurt, wijk en gemeente.] Available from: <https://www.rivm.nl/media/smapp/>. [Website in Dutch].
- Ogden CL, Fakhouri TH, Carroll MD, Hales CM, Fryar CD, Li XF, Freedman DS, 2017. Prevalence of Obesity Among Adults, by Household Income and Education - United States, 2011-2014. *Mmwr-Morbidity and Mortality Weekly Report* 66:1369-73.
- Petry NM, Barry D, Pietrzak RH, Wagner JA, 2008. Overweight and obesity are associated with psychiatric disorders: Results from the national epidemiologic survey on alcohol and related conditions. *Psychosom Med* 70:288-97.
- Pouliou T, Elliott SJ, 2009. An exploratory spatial analysis of overweight and obesity in Canada. *Prev Med* 48:362-7.
- R Core Team, 2018. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing, Vienna, Austria. Available from: <https://www.r-project.org/>
- Shahid R, Bertazzon S, 2015. Local Spatial Analysis and Dynamic Simulation of Childhood Obesity and Neighbourhood Walkability in a Major Canadian City. *Aims Public Health* 2:616-37.
- Statistics Netherlands (CBS), 2018. Toelichting wijk en buurtkaart 2016-2017-2018
- Tibshirani R, 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B* 58:267-88.
- Traversy G, Chaput JP, 2015. Alcohol Consumption and Obesity: An Update. *Current Obesity Reports* 4:122-30.
- van de Kasstele J, Zwakhals L, Breugelmans O, Ameling C, van den Brink C, 2017. Estimating the prevalence of 26 health-related indicators at neighbourhood level in the Netherlands using structured additive regression. *Int J Health Geogr* 16:23.
- Visscher T, Kromhout D, Seidell J, 2002. Long-term and recent time trends in the prevalence of obesity among Dutch men and women. *Int J Obes* 26: 1218.
- Wang Z, Zhao L, Huang Q, Hong A, Yu C, Xiao Q, Zou B, Ji S, Zhang L, Zou K, Ning Y, Zhang J, Jia P, 2020. Traffic-related environmental factors and childhood obesity: A systematic review and meta-analysis. *Obes Rev*: 1-15. 10.1111/obr.12995.
- Withrow D, Alter DA, 2011. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. *Obes Rev* 12:131-41.
- WHO, 2014. Global status report on noncommunicable diseases 2014, World Health Organization. Available from: <https://www.who.int/nmh/publications/ncd-status-report-2014/en/>.
- Xu F, Jin LL, Qin ZZ, Chen X, Xu Z, He J, Wang ZY, Ji W, Ren F, Du QY, Xiong YQ, Jia P, 2020. Access to public transport and childhood obesity: A systematic review. *Obes Rev*. doi: 10.1111/obr.12987.

- Yang S, Zhang X, Feng P, Wu T, Tian R, Zhang D, Zhao L, Xiao C, Zhou Z, He F, Cheng G, Jia P, 2020. Access to fruit and vegetable markets and childhood obesity: A systematic review. *Obes Rev*: 1-12. doi: 10.1111/obr.12980.
- Yoon SJ, Kim HJ, Doo M, 2016. Association between perceived stress, alcohol consumption levels and obesity in Koreans. *Asia Pac J Clin Nutr* 25: 316-25.
- Zhang F, Tapera TM, Gou J, 2018. Application of a new dietary pattern analysis method in nutritional epidemiology. *BMC Med Res Methodol* 18:119.
- Zhang X, Zhang M, Zhao Z, Huang Z, Deng Q, Li Y, Pan A, Li C, Chen Z, Zhou M, Yu C, Stein A, Jia P, Wang L, 2020a. Obesogenic environmental factors of adult obesity in China: a nationally representative cross-sectional study. *Environ Res Lett* 15:044009.
- Zhang X, Zhang M, Zhao ZP, Huang ZJ, Deng Q, Li YC, Pan A, Li C, Chen ZH, Zhou MG, Yu C, Stein A, Jia P, Wang LM, 2020b. Geographic Variation in Prevalence of Adult Obesity in China: Results From the 2013-2014 National Chronic Disease and Risk Factor Surveillance. *Ann Intern Med* 172:291-3.

Non-commercial use only