

RESEARCH

Open Access



# Comparing different data sources by examining the associations between surrounding greenspace and children's weight status

Yusheng Zhou<sup>1\*</sup> , Thomas von Lengerke<sup>2</sup> and Maren Dreier<sup>1</sup>

## Abstract

**Background:** Studies on the association between surrounding greenspace and being overweight in childhood show inconsistent results, possibly because they differ widely in their definition and measurement of surrounding greenspace. Our aim was to evaluate whether the association of greenspace with being overweight depends on the measurement of greenspace in different data sources.

**Methods:** Based on data from the school entry examinations of 22,678 children in the city of Hannover, Germany, from 2010 to 14, the association between greenspace availability and overweight was examined. Three different sources of greenspace availability were derived for a set of 51 areas of the city: The Normalized Difference Vegetation Index (NDVI), the OpenStreetMap (OSM) dataset, and the European Urban Atlas (UA) dataset. Agreement between the indicators on the quantity of greenspace coverage was compared. The association with children's BMI z-score, including potential interaction terms, was assessed using multilevel regression analysis.

**Results:** Greenspace availability per district area derived by NDVI was on average 42%, by OSM 29% and UA 22%, with OSM and UA being strongly correlated. Only the greenspace availability derived by NDVI showed an association with children's BMI z-score: The higher the greenspace availability was, the lower the BMI. The trend of association was higher for boys and migrant children than for girls and non-migrants and was restricted to the highest levels of greenspace availability.

**Conclusions:** Associations of greenspace with children's weight status depend on the greenspace measurement chosen. Surrounding greenspace was measured more comprehensively by NDVI. Data sources based on land use categories such as UA and OSM may be less suitable to reflect surrounding greenspace relevant for health outcomes. Potential mechanisms warrant further analysis and investigation.

**Keywords:** Greenspace, Children, Pre-schoolers, BMI, Overweight, Spatial source

## Introduction

Childhood is a sensitive developmental period in which children acquire more sophistication in interactions determined by a combination of physical, social, family, and built environmental factors. Scholarly interest in greenspace has increased in response to evidence that environmental benefits on children [1, 2]. It has been

\*Correspondence: Zhou.Yusheng@mh-hannover.de

<sup>1</sup> Hannover Medical School, Institute for Epidemiology, Social Medicine and Health Systems Research, Carl-Neuberg Str. 1, 30625 Hannover, Germany

Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

suggested that exposure to greenspace can contribute to improving public health by multiple pathways, including socio-behaviour (e.g. stress and physical activity) and regulating ecosystem services (e.g. heat reduction) [3, 4]. A shred of growing evidence demonstrates that the rate of overweight is higher among children living in areas lacking sufficient greenspace exposure [5–7]. A study shows that moving to greener neighbourhoods, as measured by density of street trees around home, was associated with lower obesity prevalence [8]. Another study showed that the presence of parks and playgrounds was inversely associated with Body mass index (BMI) [9]. In some studies, using area-level land use data, a null association with BMI was reported [10–12]. The mixed results observed for the association could be due to the differing socioeconomic positions of individuals and the various measurements applied.

As mentioned above, different indices of greenspaces have been used to assess exposures to greenspace in previous epidemiological studies, including the access to greenspace, numbers of greenspace facilities, and the surrounding greenspace (i.e., the presence or absence of greenspace close to participants' residences) [13]. Most studies have introduced a size threshold to represent surrounding greenspace, which refers to the greenspace availability [13, 14]. A UK study assessed the association between physical activity and greenspace availability by measuring the size of the greenspace area [15]. The study emphasized that the size of the greenspace not only matters in terms of the impact on the level of use (quantity) but also because it affects the type of activities (quality) people perform in the area. The larger the greenspace is, the more diverse its flora and the higher its ecological quality are likely to be, which could eventually affect the health and quality of life of populations in the vicinity [16].

Greenspace availability can be formulated differently depending on the data source. Therefore, selecting an appropriate greenspace data source would determine policy intervention, facilitate intercity comparisons, and assess the public health effect. One critical challenge affecting greenspace-related research is the lack of a consistent and strict definition and means of measurement. A range of different geographic datasets are available for assessing greenspace. Some are based on utilization of Landsat satellite images, while others are based on field inventories and mapping. For local-level analysis, such as in one city, local records of greenspace are often preferred because of their high resolution and accuracy [17]. In various European countries, these datasets are prepared by local geographic surveying agencies and constantly updated [18, 19]. However, in many cities, local databases are either not available or not accurate enough.

Researchers, in this case, may be obliged or prefer to use more general databases, such as the Urban Atlas (UA) which provides data for European countries [20] or OpenStreetMap (OSM) which provides global data [21]. Generally, the choice of a greenspace dataset depends on the specific needs of the researcher and the availability of the data set. However, poor data availability or inadequate understanding of the current available data may pose a major challenge to analysing greenspace.

In summary, this study's general aim is to determine the associations between surrounding greenspace and children's weight status by applying three different data sources to assess greenspace availability as a built environment indicator. To address this issue, the present study is guided by the following research questions: (1) To what extent do greenspace availability derived from different data sources agree with respect to the surrounding greenspace in the same area? (2) Does the association between greenspace and children's weight status differ across different data sources? (3) Are the associations modified by other individual or social characteristics?

## Methods

### Study area and participants

A cross-sectional study was conducted to explore associations between the degree of greenspace availability and children's body weight. The study was carried out in 51 areas according to the administrative boundaries of the city of Hannover, the capital of Lower Saxony, Germany. The 51 areas (204 km<sup>2</sup> in total) had a mean area of 4.16 km<sup>2</sup> ranging from 0.73 to 13.89 km<sup>2</sup>. Data on children's body weight were collected from the school entry examinations conducted annually during school entry registration and through the standardized examination programme "SOPHIA" ("Sozialpädiatrisches Programm Hannover—Jugendärztliche Aufgaben"—<http://www.sophia-online.org>). The total dataset encompassed individual data on physical and mental health status, personal characteristics, and family socioeconomic status (SES). In this study, we used data from a five-year period from 2010 to 2014, including information on 22,678 children aged 4 to 8 years old.

### Greenspace indicator and data sources

In this study, the greenspace availability in the locating area was used as the indicator to assess surrounding greenspace. It is measured by the proportion of land dedicated to greenspace per administrative unit area defined by local government. Spatial analyses were performed using the open-source geographic information system QGIS (QGIS Development Team, 2014), version QGIS 3.4.5 LTR.

In this analysis of greenspaces in the city of Hannover, the following datasets were selected, adjusting the greenspace categories to make them comparable:

- Landsat satellite images (2013–2014)
- Open Street Map (OSM) (2017)
- Urban Atlas (UA) (2012)

During the preliminary analysis, a local dataset from the Department of Environment and Urban Greenery of Hannover (Fachbereich Umwelt und Stadtgrün) was omitted. This is because the categories of greenspace were not fully interpreted (state forests, greenspace in residential areas with multi-storey housing, private parks and sports facilities were not recognized), which leads to limited accuracy in terms of the greenspace availability.

First, we used the normalized difference vegetation index (NDVI) to assess the greenspace availability in Hannover city [22]. The NDVI data was derived from Landsat 5 Thematic Mapper data at 30 m × 30 m resolution. Its values range from − 1 to 1, with higher values

indicating a higher proportion of greenness. Two maps of Hannover city were created using cloud-free images, one from 2013, and another from 2014. From these maps, we calculated an average NDVI as a percentage for each area.

The second source refers to the land use classification by OSM. OSM is a community driven collaborative mapping project that involves contributors from all over the world in the creation of a free, global geospatial database [23]. Countless users are constantly updating open and publicly available resources for the OSM project, although with different degrees of accuracy for different timeframe and places [23]. This fact must be taken into account when processing OSM data together with other spatial data captured. Here, we chose the versions from the end of 2019, which were the most updated versions available when we started our analysis. All categories of the land-use feature classes corresponding to greenspaces were identified from previous studies (Table 1).

For the third source of greenspace, we used the Urban Atlas data recorded in 2012 by the European

**Table 1** Characteristics of the greenspace data sources

	NDVI*	OSM*	UA*
Type of measurements	Biophysical variables in vegetation	Data of categorical land uses	Data of categorical land uses
Type of data	Primary data	Primary data	Secondary data
Key characteristics	Using Landsat image to calculate the visible and near-infrared light reflected by vegetation. The result refers to the NDVI	A project which allows the community of Internet users to continuously update open and publicly available resources	Land cover data based on satellite imagery
Data availability	Worldwide provided	Worldwide provided, although with different degrees of accuracy	Data available for major urban cities in the European area
Time scale	Depends on the Landsat satellites	Continuously update	Data prepared periodically, currently from 2006, 2012 and 2018
Responsible agency	Landsat images available from the United States Geological Survey (USGS)	Database maintained by the OpenStreetMap Foundation	Project coordinated by the European Environment Agency (EEA)
Greenspace categories identified	Greenspace were identified with the use of supervised classification based on the representative samples for the different green space types in the digital image	<ul style="list-style-type: none"> <li>– Allotments</li> <li>– Cemetery</li> <li>– Farmland/farmyard</li> <li>– Forest/wood</li> <li>– Garden</li> <li>– Grassland</li> <li>– Greenfield</li> <li>– Greenhouse horticulture</li> <li>– Meadow</li> <li>– Nature reserve</li> <li>– Orchard</li> <li>– Park</li> <li>– Plant nursery</li> <li>– Scrub</li> <li>– Trees</li> <li>– Village green</li> <li>– Wetland</li> </ul>	<ul style="list-style-type: none"> <li>– Green urban areas (Code 14,100)</li> <li>– Arable land (annual crops) (Code 21,000)</li> <li>– Permanent crops (Code 22,000)</li> <li>– Pastures (Code 23,000)</li> <li>– Complex and mixed cultivation (Code 24,000)</li> <li>– Forests (Code 30,000)</li> <li>– Herbaceous vegetation associations (Code 32,000)</li> <li>– Wetland (Code 40,000)</li> </ul>

\* NDVI: Measurement of calculating the Normalized Difference Vegetation Index (NDVI) by analysing Landsat satellite images; OSM: the OpenStreetMap dataset; UA: the European Urban Atlas dataset

Environment Agency. The Urban Atlas provides detailed vector data on land cover and land use for numerous city regions in Europe [24]. Such data are available at a scale of 1:10 000 for 2006, 2012 and 2018 with a largely standard nomenclature (max. 28 classes). For the latter reference year, the Urban Atlas has been available since 2018 for 788 Functional Urban Areas of Europe as spatial and statistical data [24]. The Urban Atlas is derived from remote sensing (e.g., SPOT satellite images) and contains overlap-free geo-objects (polygons) suitable for GIS processing. Here, we chose UA-2012 to match the year of spatial data with the school examination data (2010–2014). Urban Atlas data on surrounding greenspaces were identified based on the relevant mapping guides' descriptions (Table 1).

### Health outcome and covariates

Objectively standardized measured height and weight were used to calculate the body mass index (BMI) z-score for each subject. The BMI z-score was calculated using the national weight status reference [24]. Weight status was categorized into: normal weight (BMI < 90th percentile), pre-obesity (90th percentile  $\leq$  BMI < 97th percentile), and obesity (BMI  $\geq$  97th percentile). Categorization as overweight in this study refers to the status including both pre-obesity and obesity (BMI  $\geq$  90th percentile).

Multiple factors contribute to childhood overweight disparities including cultural, family and socioeconomic status. In this study, demographic and socioeconomic characteristics of children were collected by structured interviews with children and their parents using a standardized questionnaire. Demographic variables included migration background (self-identified and categorized as migrants vs. non-migrants), family structure (nuclear family vs. others) and number of siblings (one or no siblings vs. two or more siblings). Socioeconomic inequalities among parents may be partly mediated by their health behaviours which eventually reflected among their children. Parents' education is frequently used as an indicator of SES in surveys among children. Also, education may be considered as one of the markers representing childhood social environment, and show differences in awareness of health issues [25]. Here in the school entry examination, parental education status was the only available SES-indicator which was classified into three educational classes (lower, middle, and higher) based on parents' primary qualification and professional education. The birth weight was also included (high: > 4000 g, normal: 2500–4000 g, and low: < 2500 g). In addition, length of child day care participation was included (as at least three years vs. less than three years) to evaluate the influence of preschool childcare services.

On the area level, we selected three sociodemographic characteristics provided by the Statistics Office of State Capital Hannover which publishes the city's structural data annually [26]: the proportion of residents with migration background in each area, the area's unemployment rate, and the residential density (i.e., residents per hectare). Data from 2010 to 2014 were selected to match the school entry examination data.

### Data analysis

To address research question 1, we compared the values of greenspace availability for each area using scatter plots. We explored differences in the mean agreement between the measurements of the greenspace data sources. Correlations between two of the three data sources were determined using the Pearson correlation coefficient  $r$ . The spatial patterns of surrounding greenspace were then mapped separately using GIS software.

To answer research questions 2 and 3, multilevel regression analysis was applied to explore the associations between surrounding greenspace and the BMI z-score of the children. First, an unconditional model with no predictors was estimated to assess the intraclass correlation of the BMI z-score. Then, all individual-level characteristics (i.e., sex and migration background) and area-level sociodemographic factors (i.e., unemployment rates) were added as fixed effects. All the variables were tested for multicollinearity through bivariate correlation and variance inflation factor (VIF). Consequently, one of the area-level indicators (the proportion of residents with migration background in the area) was dropped from the model. To examine the unique contribution of the three greenspace variables (NDVI, UA and OSM), three different models were calculated (Models 1a, 1b, and 1c). To deal with spatial autocorrelation, we applied the Moran's I test by including  $x$ - and  $y$ - coordinates of the unit centroids of each area. As in the analysis, these coordinates did not show to be significant; as a result, they were excluded from the final models. To study whether the changes in children's BMI z-score varied across different sociodemographic subgroups, the data were further analysed as follows: an interaction between greenspace availability and sex, migration background and parental education level was added to the models. A statistically significant interaction term between these factors would suggest that the association between greenspace and children's weight status has a different magnitude in the different categories of that variable (Models 2a, 2b, and 2c).

To study the shape of the association of greenspace availability with being overweight, spline regression models were applied. First, a linear term (1 degree of freedom) was specified. Then, we specified natural splines with 3, 4 and 5 degrees of freedom ( $df=2$  were too small and

therefore excluded from the analysis [27]. To test whether the goodness-of-fit of the spline models was significantly better than that of the linear model, the likelihood ratio test was applied. The plots were presented using the probability of being overweight. Greenspace availability is shown on the x-axis of the plots. To visualize the influence of the interaction terms (sex, migration background and parental education level) on the association, we also performed stratified analyses by these factors in Fig. 3.

P-values < 0.05 were considered statistically significant. All analyses were performed using IBM SPSS Statistics for Windows software, version 25.0 (IBM Corp., Armonk, NY, USA) and R 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

## Results

Table 2 shows the socioeconomic, demographic and built environment characteristics of the complete sample. According to the statistics, approximately half of the children were girls (48.5%) and had a migration background (49.4%), and nearly one in ten was overweight (9.7%).

Within the different greenspace measurements, UA and OSM showed strong positive correlations (Fig. 1). Having moderate positive correlations with UA and OSM, the total greenspace area from NDVI should represent the upper limit of greenspace provision. However, NDVI tended to report high levels of greenspace for some areas in which the other two indicators showed higher levels (Fig. 2).

Figure 2 illustrates spatial differences between the greenspace measurements in the city of Hannover. The maps show a high level of agreement between OSM and UA, but the tendency for UA to indicate lower levels of greenspace is evident. At the same time, the availability index of UA and OSM shows that greenspace availability tends to be high in central areas.

Table 3 presents the results of the multilevel regression analysis. Three models were created to compare the effects of greenspace availability according to different data sources. Across all models, children's sex, migration background, number of siblings in the family, birth weight, and parental educational level were all significantly associated with the children's weight status. For the area-level information, the unemployment rate of the area was significant in all models, while residential density was significant for the two land use driven data source models (UA and OSM datasets). Only greenspace availability as defined by NDVI was significantly associated with BMI, meaning that a lower BMI z-score was associated with higher greenspace in Landsat imagery. The other two data sources showed no significant association.

Table 4 presents the results of the interaction terms between parental educational level, sex and children's migration background and green space availability by different data sources. The association between greenspace and body weight was modified by migration background and sex, restricted to greenspace measured by NDVI. In boys and migrant children, higher greenspace availability was significantly associated with a lower BMI compared to girls and non-migrant children.

To further examine these results, spline regression models stratified by sex and migration background were performed. The association of greenspace availability from NDVI with being overweight was significantly different ( $p < 0.05$ ) from linearity (Fig. 3) in our main model using four degrees of freedom. Similar patterns of the association were found for models with three and five degrees of freedom. From the results of spline models, the exposure–response curves demonstrated patterns with increases and decreases at different levels of greenspace availability (Fig. 3). The downward curve for the exposure–response relationship for higher levels of greenspace is consistent with the results from the BMI z-score interaction term analysis: Greenspace availability tends to show a more negative association with the probability of being overweight for boys and migrant children.

## Discussion

The main goal of this study was to compare results from three different greenspace data sources to determine their relative level of agreement. We used a multilevel modelling approach to analyse associations between greenspace and children's weight status. The degree of surrounding greenspace varied considerably between each of the measurements. The greenspace availability per district area derived by NDVI was on average 42%, by OSM 29% and UA 22%, with OSM and UA being strongly correlated. Only the greenspace availability derived by NDVI was weakly associated with children's weight: The more greenspace availability there was, the lower their BMI. The association was slightly higher for boys and migrant children than for girls and non-migrant children but restricted to the highest levels of greenspace availability.

### Comparison between different data sources

UA and OSM are high-quality and well-validated data sources, while NDVI has been applied in many spatial analyses of greenspaces. Comparisons among these three different data sources clearly showed that NDVI detects surrounding greenspace more comprehensively, presenting the greenspace provision's upper limit among these measurements. The greenspace availability based on UA and OSM is considerably less than that

**Table 2** Characteristics of the study population and the area (data from school entrance examination, city of Hannover, 2010–2014, n = 22,678)

	N (%)	Mean	Standard Deviation	Minimum	Maximum
Dependent variables					
Body mass index percentile	50.1		27.9	0	99.8
Overweight* (> 90th percentile)	9.70%				
Obese > 97th percentile	4.10%				
Individual level factors					
Age			6	0.4	4
Sex					
Boys	51.50%				
Girls	48.50%				
Migration background					
Yes	49.40%				
No	50.60%				
Family structure					
Single parent/ Blended family	2.80%				
Nuclear family	97.20%				
Siblings					
≥ 2 siblings	30.50%				
< 2 siblings	69.50%				
Child day care participation					
< 3 years	18.70%				
≥ 3 years	81.30%				
Parental educational level					
Higher	36.90%				
Middle	26.20%				
Lower	36.90%				
Birth weight					
High (> 4000 g)	10.60%				
Normal (2500–4000 g)	80.10%				
Low (< 2500 g)	6.60%				
Area level factors					
Unemployment rate		8.70%	3.20%	1.90%	16.10%
Proportion of residents with migration background in the area		25.60%	9.50%	6.50%	50.30%
Residential density (residents per hectare)		32.9	32.9	1	168
Greenspace availability indices (Area level)					
Greenspace availability by NDVI		41.70%	20.00%	0.90%	87.10%
Greenspace availability by UA		22.30%	19.10%	0.10%	81.50%
Greenspace availability by OSM		29.80%	16.30%	0.90%	88.90%

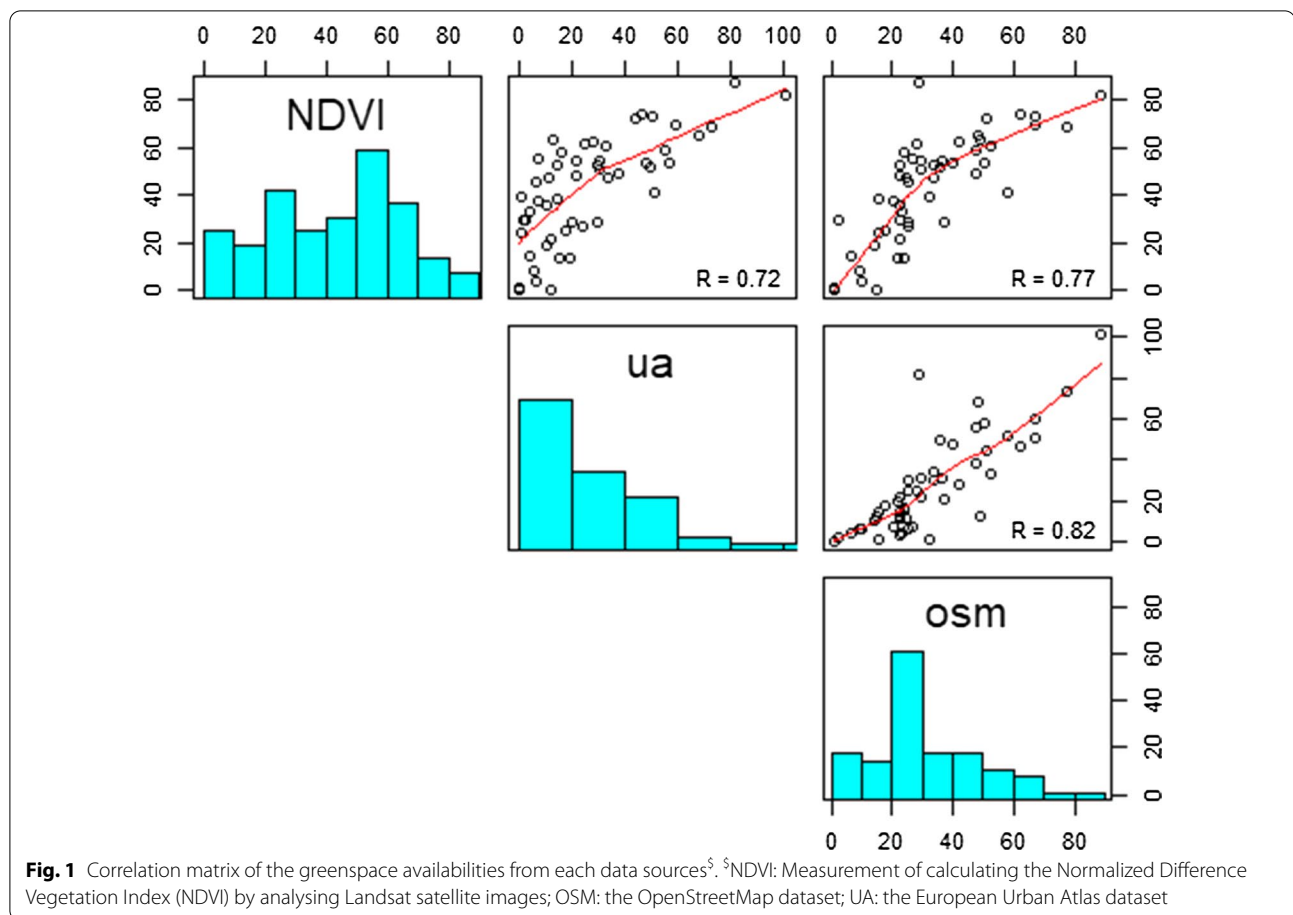
\* Overweight refers to the status including both pre-obesity and obesity

\* NDVI: Measurement of calculating the Normalized Difference Vegetation Index (NDVI) by analysing Landsat satellite images; OSM: the OpenStreetMap dataset; UA: the European Urban Atlas dataset

based on NDVI. This difference derives mainly from the fact that UA and OSM were not able to cover private greenspaces, especially in residential areas. Meanwhile, the NDVI derived from Landsat images includes croplands, forests, lakes and marshes, which are not necessarily recreational friendly and not necessarily

associated with promoting physical activity in children [28].

Theoretically, these two land use data sources may provide more meaningful information to better distinguish the type of greenspace. However, this type of approach has its own limitations for comparative



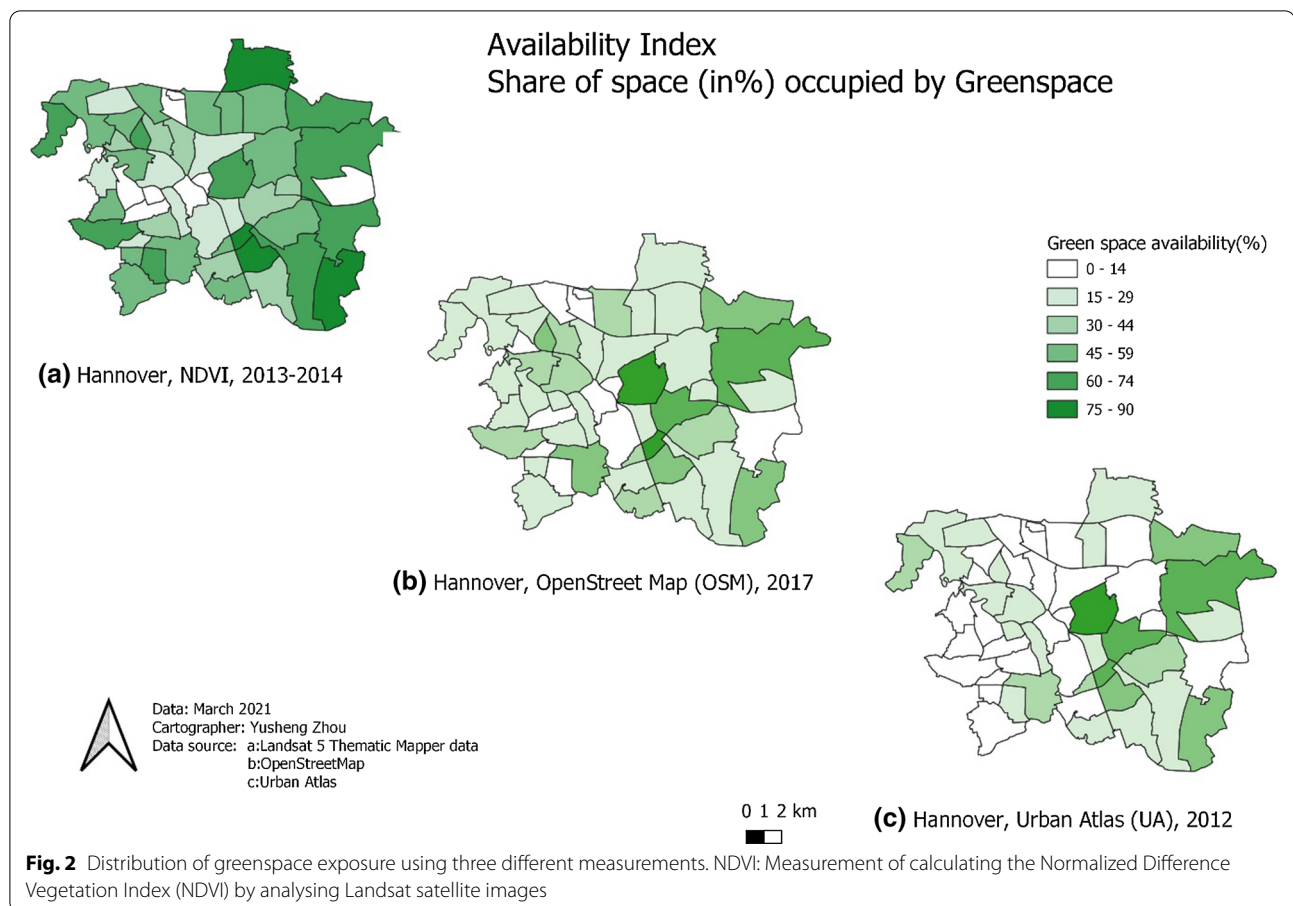
research due to the mean spatial heterogeneity of land cover across cities [29]. The coarseness of some data introduces the problem of derived indicators. For example, indicators based on count metrics may not be completely accurate, because adjacent spaces may appear as separate spaces in the data but exist as one united space [30]. Besides, the edge effect also appears in the area-level analysis, that is, those who live near a vast space may exhibit outliers with an extremely high level of access, especially those located on the edge of town/city or in the rural area [30]. These two land-use data sources have their own issues. According to OSM data, the contributed features may suffer from a lack of consistency such that they have dissimilar geometrical accuracy and frequently overlap with each other [31]. Simultaneously, although the data were generated from a highly reliable approach, UA data maintain issues such as a lack of categories for private greenspaces. To increase the accuracy of the data, researchers have applied OSM data to assist the validation of land use maps from UA data. According to studies of converting OSM data into a map using UA data by Fonte et al. [32, 33], the accuracy indices validated by OSM data were

not significantly different from those obtained by photo interpretation.

For comparison across different data sources, inconsistencies in greenspace definitions are challenging [4, 34]. The greenspace categories in different data sources do not necessarily overlap. This makes it difficult to compare surrounding greenspace obtained from different data sources, not to mention that the detailed categories or types of greenspace are vague in most data sources. For example, none of the included data sources distinguish between the public and private ownership, while some research objectives may be required to distinguish between them. Thus, the study design used here cannot provide a definitive perspective on the choice of data source, but the intriguing result merits further investigation.

#### Association between greenspace and weight status

The built environment in which children live, play and eat provides an important context that may influence overweight risk [35]. Research has shown that publicly provided recreational infrastructure is positively associated with children's physical activity [36, 37]. This indicates



that children and their parents tend to be more physically active in a conducive environment [37]. Mechanistically, exposure to greenspace may promote health-related behaviours, mitigate harmful environmental exposures to air pollution, heat, and noise, and encourage stress alleviation [34, 38]. All these beneficial effects of greenspace are involved with the pathophysiological pathways affecting children's independent mobility and willingness to engage in physical activity behaviour, which eventually lead to changes in weight [34, 39].

Our previous analysis on the association between area-level greenspace and children's overweight using data from the OSM dataset failed to show a significant association with the BMI z-score [40]. Given the results from this study, this might be due to the data source to measure greenspace exposure. The results from other studies support this hypothesis: a study from New Zealand based on national administrative records discovered that greenspace was not significantly related to children's overweight [12], and a study from Berlin, Germany [41] using a land-use categorical database also showed a nonsignificant association among children. A review of German studies of built environment effects on health

concluded that there was no association between greenspace and obesity [42]. In contrast, studies using NDVI data [43–45] often found that higher surrounding greenspace levels were associated with lower BMI z-scores suggesting that the choice of data source of greenspace has relevant implications on its association with health outcomes.

As NDVI is sensitive to both public and private greenspace, the significant association between children's overweight and greenspace might raise the possibility that it is the private greenspace that holds a specific influence on children's development [46]. Case studies demonstrated that privately owned greenspaces (e.g. backyards or domestic gardens) often acted as a substitute for public greenspaces (e.g. communal gardens or neighbourhood parks) in residential areas [47]. If public greenspace exerted greater influence, we would have expected to see similar associations in UA and OSM. However, the mere presence of greenspace does not necessarily mean its use will provide its effect. The major limitation of applying NDVI in epidemiological studies is its inability to distinguish between different types of greenspaces which may promote physical exercise to varying degrees due to



**Table 3** Estimates from two-level linear modelling predicting pre-school children’s BMI z-score (Data from school entrance examination, city of Hannover, Germany, 2010–2014, n = 22,678)<sup>§</sup>

	Model 1a Availability index by NDVI		Model 1b Availability index by UA		Model 1c Availability index by OSM	
	b (SE)	95% CI	b (SE)	95% CI	b (SE)	95% CI
Individual-level independent factors						
Girls (Ref. Boys)	0.09* (0.01)	(0.06; 0.12)	0.09* (0.01)	(0.06; 0.12)	0.09* (0.01)	(0.06; 0.12)
Children with migration background (Ref. German children)	0.18* (0.02)	(0.15; 0.22)	0.19* (0.02)	(0.16; 0.22)	0.19* (0.02)	(0.15; 0.22)
Single parent/Blended family (Ref. Nuclear family)	0.07 (0.04)	(− 0.01; 0.16)	0.08 (0.04)	(− 0.01; 0.16)	0.08 (0.04)	(− 0.01; 0.16)
Two or more siblings (Ref. one or none siblings)	0.06* (0.02)	(0.02; 0.09)	0.05* (0.02)	(0.02; 0.09)	0.06* (0.02)	(0.02; 0.09)
Childcare less than 3 years (Kita) (Ref. 3 year or longer)	− 0.02 (0.02)	(− 0.06; 0.02)	− 0.02 (0.02)	(− 0.06; 0.02)	− 0.02 (0.02)	(− 0.06; 0.02)
Parental educational level(Ref. Higher education)						
Lower education level	0.27* (0.02)	(0.24; 0.31)	0.27* (0.02)	(0.24; 0.31)	0.27* (0.02)	(0.24; 0.31)
Middle education level	0.15* (0.02)	(0.12; 0.19)	0.15* (0.02)	(0.12; 0.19)	0.15* (0.02)	(0.12; 0.19)
Birth weight (Ref. Normal)						
High (> 4000 g)	0.37* (0.02)	(0.32; 0.42)	0.37* (0.02)	(0.32; 0.41)	0.37* (0.02)	(0.32; 0.42)
Low (< 2500 g)	− 0.25* (0.03)	(− 0.31; − 0.19)	− 0.25* (0.03)	(− 0.31; − 0.19)	− 0.25* (0.03)	(− 0.31; − 0.19)
Area-level socio-demographic factors						
Residential density	0.01 (0.01)	(− 0.02; 0.02)	0.01* (0.01)	(0.001; 0.03)	0.01* (0.01)	(0.001; 0.03)
Unemployment rate (%)	0.04* (0.02)	(0.01; 0.08)	0.05* (0.02)	(0.01; 0.08)	0.04* (0.02)	(0.01; 0.08)
Area-level greenspace indicators						
Greenspace availability index	− 0.01* (0.01)	(− 0.05; − 0.01)	0.01 (0.01)	(− 0.01; 0.03)	0.01 (0.01)	(− 0.01; 0.02)

\* p < 0.05 (highlighted)

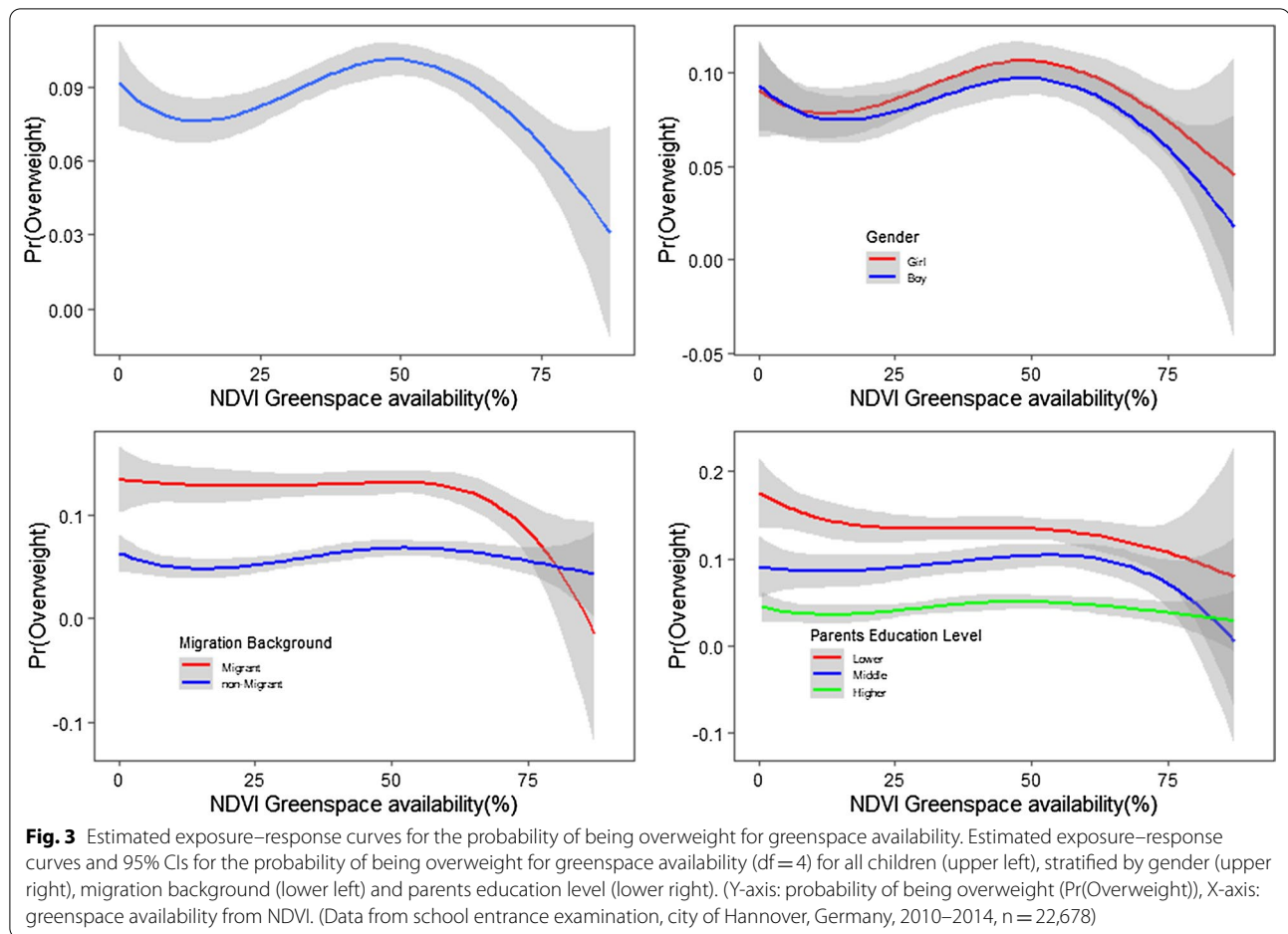
<sup>§</sup> Model 1a: Model on children’s BMI z-score adjusted for individual level factors & other area level factors presented in the table plus greenspace availability by NDVI; Model 1b: Model plus greenspace availability from European Urban Atlas dataset; Model 1c: Model plus greenspace availability from OpenStreetMap dataset

**Table 4** Associations between children’s BMI z-score and interaction terms of greenspace availability (Data from school entrance examination, city of Hannover, 2010–2014, n = 22,678)<sup>§</sup>

	Model 2a		Model 2b		Model 2c	
	(Availability index by -NDVI)		(Availability index by UA)		(Availability index by OSM)	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Sex * Greenspace Availability	0.03* (0.01)	(0.01, 0.06)	− 0.02 (0.01)	(− 0.04, 0.01)	− 0.01 (0.01)	(− 0.03, 0.02)
Migration background* Greenspace Availability	− 0.001	(− 0.09, − 0.03)	0.01 (0.02)	(− 0.02, 0.04)	− 0.01 (0.02)	(− 0.03, 0.03)
Lower education level * Greenspace Availability	− 0.02 (0.02)	(− 0.06, 0.01)	− 0.02 (0.02)	(− 0.06, 0.13)	− 0.03 (0.02)	(− 0.06, 0.01)
Middle education level * Greenspace Availability	0.01 (0.02)	(− 0.03, 0.04)	− 0.01 (0.02)	(− 0.04, 0.03)	− 0.01 (0.02)	(− 0.04, 0.03)

<sup>§</sup> Adjusted for sex, migration background, family structure, number of siblings, length of child day care participation, parental educational level, birth weight, residential density, unemployment rate

\* p < 0.05 (highlighted)



characteristics such as size and available facilities [48]. Simultaneously, the greenest areas in an urban setting (e.g. urban forests) might relate to having fewer transport destinations within walking distance, which leads to greater car dependency [49]. Therefore, further study should consider that the different types of greenspace may have different effects on promoting physical activity and hence different protective influences on children's weight gain.

#### Effect of greenspace may differ in subgroups

Greenspace enables outdoor physical activity which may be an important preventive factor for children's overweight. However, research evidence suggests that the influence of individual characteristics is greater than the influence of built environmental determinants on the level of physical activity [50]. Individual socioeconomic status may modify correlations between greenspace availability and the willingness to engage in physical activity [2, 14]. Correspondingly, Lovasi et al. found a weak association between greenspace and the

physical activity rate, especially in low-income communities [8]. Another reason might be safety aspects, as research from North America points out that areas with high-density trees may have higher risk of crime, which leads to a lower willingness to engage in physical activity due to safety considerations [51].

Our study noted that the association between greenspace and body weight could be modified by sex and migration background, which is consistent with several greenspace-related research findings in adulthood. The evidence for sex is mixed and restricted to adults. Some studies indicate a tendency for a more substantial effect in women [1, 52–54], while others have shown no associations [55]. Several studies considered race/ethnicity effects for pregnant woman [56, 57] and school children [28]. They note that the potential built environment reflecting cultural preferences has plausible implications for health behaviour and outcomes [58]. Therefore, further research on children should include more comprehensive and multidimensional metrics.

### Limitations

Greenspace availability is not the only indicator to measure greenspace. Generally, access to greenspace is often used in epidemiological studies as well [13, 14, 59]. It can be estimated by multiple indicators, including the Euclidean distance to a park, the network distance to a greenspace area, and the presence of park within a given area around the home address [14]. Furthermore, some studies explored the perception of greenspace on willingness to use greenspace facilities and actual use by questionnaires [37, 39, 60]. The absence of a standard indicator representing greenspace has become an important barrier to the generalization of the study results.

This study was limited to analysing greenspace at the area level. This is partly due to the unavailability of individual home address information. Additional dimensions of indicators on a better scale (i.e., zip code and home address) would greatly enhance the precision of exposure measurement [14]. When performing spatial analysis, defining built environment values is undermined by being dependent on spatial scales and spatial units. This phenomenon is described by the modifiable areal unit problem (MAUP) [61]. The MAUP is an issue in spatial analysis for studies in geography. It describes that built environment values may vary with the spatial scale for which data are available and the boundaries between spatial units [62]. For now, the MAUP is still rarely addressed in practice. A systematic review in 2014 reports that MAUP was recognized in only 1% of papers using spatially aggregated data [61]. A potential way to mitigate the impact of the MAUP is to create a geographical structure with high heterogeneity zoning units. Other studies proposed to conduct a sensitivity analysis [63]. Therefore, increasing awareness of this methodologically relevant issue is critically needed in health geography analysis.

This study has other limitations in the method design. First, focusing on a sample of one city was informative with respect to the accuracy of greenspace measurements achieved, but had reduced generalizability. Greenspaces in our target, the city of Hannover, constitute 11.4 percent of the total surface area, earning it the title of Germany's greenest city. This study may not cover the entire situation of the built environment in German cities, and the results might be different for other cities and countries. Second, the OSM data used were from a slightly later period (2017) than the other two data sources (approximately 2012–2013); it is possible that differences in the greenspace detected were due to changes over time. As the presence of greenspace in Germany is relatively persistent over these ten years, the bias resulting from this measurement error is likely to be minor [29]. Third, our results cannot prove any causal relationship between greenspace exposure and children's overweight according

to the cross-sectional study. Associations may be prone to residential self-selection, as families are likely to select their neighbourhood according to their personal preferences, lifestyle, and culture, and consequently are likely to prefer healthier, greener neighbourhoods [40, 64, 65]. Residual confounding cannot be excluded, as the range of confounders adjusted for in this study was limited and did not include potential risk factors for being overweight such as eating habits and actual physical activity. Regarding the outcome measure, BMI percentiles and z-scores have the potential to misclassify health-related outcomes in children. Previous studies have evaluated the validity of BMI z-score predicting overweight in children and conclude that it is a weak predictor of the changes in total body fat [66]. However, due to ease of acquisition, BMI standardized for age and sex is still the most widely used clinical outcome variable. The addition of other anthropometric measurements such as waist circumference and skinfolds measurements to BMI z-score assessment may help to identify those children with excess body fat.

### Conclusion

Surrounding greenspace measured by NDVI reaches nearly twice the amount of the greenspace proportion of OSM and UA, identifying NDVI as the more comprehensive data source. Negative associations with BMI were only detected for NDVI-based greenspace availability, suggesting that data sources based on land use categories, such as UA and OSM, are less suitable to reflect the surrounding greenspace that is relevant for health outcomes. Ultimately, achieving accurate and standardized surrounding greenspace measurements can be difficult, especially considering individual characteristics and data availability. It is beyond the scope of this article to recommend the most appropriate greenspace data source. However, researchers should consider and transparently report the impact of different greenspace data sources on the relationship between surrounding greenspace and health outcomes.

### Abbreviations

NDVI: The Normalized Difference Vegetation Index; OSM: The OpenStreetMap dataset; UA: The European Urban Atlas dataset; BMI: The body mass index; SES: Socioeconomic status.

### Acknowledgements

The authors gratefully acknowledge the hard work of the team of the Hannover Region Social Pediatrics and Youth Medicine Office and the Statistics Office of State Capital Hannover for providing the valuable data, respectively. We also acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Fund of Hannover Medical School (MHH).

### Authors' contributions

YZ led the work and was responsible for the conceptualization of the project, data processing, formulation of the methodology, statistical analysis, modelling and writing the original draft and interpreting the results. MD guided

the data analysis, formulating some of the concepts and interpretation of the results, edited the manuscript. TvL supervised statistical analysis, edited and improved the manuscript. All authors read and approved the final manuscript.

#### Funding

Open Access funding enabled and organized by Projekt DEAL. Open Access funding provided by the German Research Foundation (DFG) and the Open Access Fund of Hannover Medical School (MHH).

#### Availability of data and materials

All data generated or analyzed during this study are available from the corresponding author on reasonable request.

#### Declarations

#### Ethics approval and consent to participate

Not applicable. Consent to publish. Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Hannover Medical School, Institute for Epidemiology, Social Medicine and Health Systems Research, Carl-Neuberg Str. 1, 30625 Hannover, Germany.

<sup>2</sup>Department of Medical Psychology, Hannover Medical School, Centre for Public Health and Healthcare, Carl-Neuberg Str. 1, 30625 Hannover, Germany.

Received: 12 January 2021 Accepted: 13 May 2021

Published online: 28 May 2021

#### References

- Astell-Burt T, Feng X, Kolt GS. Greener neighborhoods, slimmer people? Evidence from 246,920 Australians. *Int J Obes (Lond)*. 2014;38(1):156–9.
- Markevych I, Smith MP, Jochner S, Standl M, Brüske I, von Berg A, et al. Neighbourhood and physical activity in German adolescents: GINIplus and LISAPlus. *Environ Res*. 2016;147:284–93.
- van den Bosch M, Ode SÅ. Urban natural environments as nature-based solutions for improved public health—a systematic review of reviews. *Environ Res*. 2017;158:373–84.
- Lee ACK, Maheswaran R. The health benefits of urban green spaces: a review of the evidence. *J Public Health (Oxf)*. 2011;33(2):212–22.
- Lakes T, Burkart K. Childhood overweight in Berlin: intra-urban differences and underlying influencing factors. *Int J Health Geogr*. 2016;15:12.
- Müller G, Harhoff R, Rahe C, Berger K. Inner-city green space and its association with body mass index and prevalent type 2 diabetes: a cross-sectional study in an urban German city. *BMJ Open*. 2018;8(1):e019062.
- Schüle SA, Fromme H, Bolte G. Built and socioeconomic neighbourhood environments and overweight in preschool aged children. A multilevel study to disentangle individual and contextual relationships. *Environ Res*. 2016;150:328–36.
- Lovasi GS, Schwartz-Soicher O, Quinn JW, Berger DK, Neckerman KM, Jaslow R, et al. Neighborhood safety and green space as predictors of obesity among preschool children from low-income families in New York City. *Prev Med*. 2013;57(3):189–93.
- Nesbit KC, Kolobe TA, Arnold SH, Sisson SB, Anderson MP. Proximal and distal environmental correlates of adolescent obesity. *J Phys Act Health*. 2014;11(6):1179–86.
- Richardson EA, Pearce J, Mitchell R, Kingham S. Role of physical activity in the relationship between urban green space and health. *Public Health*. 2013;127(4):318–24.
- Picavet HSJ, Milder I, Kruize H, de Vries S, Hermans T, Wendel-Vos W. Greener living environment healthier people? Exploring green space, physical activity and health in the Doetinchem Cohort Study. *Prev Med*. 2016;89:7–14.
- Jenkin GL, Pearson AL, Bentham G, Day P, Kingham S. Neighbourhood influences on children's weight-related behaviours and body mass index. *AIMS Public Health*. 2015;2(3):501–15.
- Nieuwenhuijsen MJ, Kruize H, Gidlow C, Andrusaityte S, Antó JM, Basagaña X, et al. Positive health effects of the natural outdoor environment in typical populations in different regions in Europe (PHENOTYPE): A study programme protocol. *BMJ Open*. 2014;4(4):e004951.
- Jia P, Cao X, Yang H, Dai S, He P, Huang G, et al. Green space access in the neighbourhood and childhood obesity. *Obes Rev*. 2020. <https://doi.org/10.1111/obr.13100>.
- Coombes E, Jones AP, Hillsdon M. The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Soc Sci Med*. 2010;70(6):816–22.
- James P, Banay RF, Hart JE, Laden F. A review of the health benefits of greenness. *Curr Epidemiol Rep*. 2015;2(2):131–42.
- Kabisch N, Haase D. Green justice or just green? Provision of urban green spaces in Berlin, Germany. *Landsc Urban Plan*. 2014;122:129–39.
- Burckhardt M. Erstellung eines differenzierten Gebäudebestandes auf Grundlage des ATKIS-Basis-DLM und Daten örtlicher Bauleitplanung. *J Cartogr Geogr Inf*. 2015;65(4):194–201.
- Bielecka E, Medyńska-Gulij B. Zur Geodateninfrastruktur in Polen. *J Cartogr Geogr Inf*. 2015;65(4):201–8. <https://doi.org/10.1007/BF03545142>.
- Vilcea C, Şoşea C. A GIS-based analysis of the urban green space accessibility in Craiova city, Romania. *Geografisk Tidsskrift-Danish J Geogr*. 2020;120(1):19–34.
- Goodchild MF. Citizens as sensors: the world of volunteered geography. *Geojournal*. 2007;69(4):211–21.
- Gascon M, Cirach M, Martínez D, Dadvand P, Valentín A, Plasència A, et al. Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: the case of Barcelona city. *Urban For Urban Green*. 2016;19:88–94.
- Viana CM, Encalada L, Rocha J. The value of OpenStreetMap historical contributions as a source of sampling data for multi-temporal land use/cover maps. *IJGI*. 2019;8(3):116.
- Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, et al. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd*. 2001;149(8):807–18.
- Geyer S, Hemström O, Peter R, Vågerö D. Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *J Epidemiol Community Health*. 2006;60(9):804–10.
- Landeshauptstadt Hannover. Strukturdaten der Stadtteile und Stadtbezirke; 2014 [cited 2020 Apr 1]. <https://www.hannover.de/Leben-in-der-Region-Hannover/Politik/Wahlen-Statistik/Statistikstellen-von-Stadt-und-Region/Statistikstelle-der-Landeshauptstadt-Hannover/Strukturdaten-der-Stadtteile-und-Stadtbezirke>.
- Eisen EA, Agalliu I, Thurston SW, Coull BA, Checkoway H. Smoothing in occupational cohort studies: an illustration based on penalised splines. *Occup Environ Med*. 2004;61(10):854–60.
- Wilhelmsen CK, Skalleberg K, Raanaas RK, Tveite H, Aamodt G. Associations between green area in school neighbourhoods and overweight and obesity among Norwegian adolescents. *Prev Med Rep*. 2017;7:99–105.
- Kremer P, Hamstead Z, Haase D, McPhearson T, Frantzeskaki N, Andersson E et al. Key insights for the future of urban ecosystem services research. *Ecol Soc* 2016; 21(2).
- Smith G, Cirach M, Swart W, Dédélé A, Gidlow C, van Kempen E, et al. Characterisation of the natural environment: quantitative indicators across Europe. *Int J Health Geogr*. 2017;16(1):16.
- Ehrig-Page JC. Evaluating methods for downloading OpenStreetMap data. *Cartogr Perspect*. 2020;95:42–9.
- Fonte C, Minghini M, Patriarca J, Antoniou V, See L, Skopeliti A. Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *IJGI*. 2017;6(4):125.
- Fonte CC, Martinho N. Assessing the applicability of OpenStreetMap data to assist the validation of land use/land cover maps. *Int J Geogr Inf Sci*. 2017;31(12):2382–400.
- Hartig T, Mitchell R, de Vries S, Frumkin H. Nature and health. *Annu Rev Public Health*. 2014;35:207–28.
- Hillsdon M, Panter J, Foster C, Jones A. The relationship between access and quality of urban green space with population physical activity. *Public Health*. 2006;120(12):1127–32.

36. Kaczynski AT, Potwarka LR, Saelens BE. Association of park size, distance, and features with physical activity in neighborhood parks. *Am J Public Health*. 2008;98(8):1451–6.
37. Schalkwijk AAH, van der Zwaard BC, Nijpels G, Elders PJM, Platt L. The impact of greenspace and condition of the neighbourhood on child overweight. *Eur J Public Health*. 2018;28(1):88–94.
38. Rojas-Rueda D, Nieuwenhuijsen MJ, Gascon M, Perez-Leon D, Mudu P. Green spaces and mortality: a systematic review and meta-analysis of cohort studies. *Lancet Planet Health*. 2019;3(11):e469–77.
39. Dzhambov A, Hartig T, Markevych I, Tilov B, Dimitrova D. Urban residential greenspace and mental health in youth: different approaches to testing multiple pathways yield different conclusions. *Environ Res*. 2018;160:47–59.
40. Zhou Y, Buck C, Maier W, von Lengerke T, Walter U, Dreier M. Built environment and childhood weight status: a multi-level study using population-based data in the city of Hannover, Germany. *Int J Environ Res Public Health*. 2020;17(8):2694.
41. Kabisch N, Strohbach M, Haase D, Kronenberg J. Urban green space availability in European cities. *Ecol Ind*. 2016;70:586–96.
42. Schulz M, Romppel M, Grande G. Built environment and health: a systematic review of studies in Germany. *J Public Health (Oxf)*. 2018;40(1):8–15.
43. van der Zwaard BC, Schalkwijk AAH, Elders PJM, Platt L, Nijpels G. Does environment influence childhood BMI? A longitudinal analysis of children aged 3–11. *J Epidemiol Community Health*. 2018;72(12):1110–6.
44. Petravičienė I, Grazulevičienė R, Andrusaitė S, Dedele A, Nieuwenhuijsen MJ. Impact of the social and natural environment on preschool-age children weight. *Int J Environ Res Public Health*. 2018;15(3):449.
45. Dadvand P, Villanueva CM, Font-Ribera L, Martínez D, Basagaña X, Belmonte J, et al. Risks and benefits of green spaces for children: a cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy. *Environ Health Perspect*. 2014;122(12):1329–35.
46. Richardson EA, Pearce J, Shortt NK, Mitchell R. The role of public and private natural space in children's social, emotional and behavioural development in Scotland: a longitudinal study. *Environ Res*. 2017;158:729–36.
47. Rupprecht CDD, Byrne JA. Informal urban green-space: Comparison of quantity and characteristics in Brisbane, Australia and Sapporo, Japan. *PLoS ONE*. 2014;9(6):e99784.
48. Lachowycz K, Jones AP, Page AS, Wheeler BW, Cooper AR. What can global positioning systems tell us about the contribution of different types of urban greenspace to children's physical activity? *Health Place*. 2012;18(3):586–94.
49. Richardson EA, Mitchell R, Hartig T, de Vries S, Astell-Burt T, Frumkin H. Green cities and health: a question of scale? *J Epidemiol Community Health*. 2012;66(2):160–5.
50. de Vries S, van Dillen SME, Groenewegen PP, Spreeuwenberg P. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Soc Sci Med*. 2013;94:26–33.
51. Mackenbach JD, Rutter H, Compernelle S, Glonti K, Oppert J-M, Charreire H, et al. Obesogenic environments: a systematic review of the association between the physical environment and adult weight status, the SPOT-LIGHT project. *BMC Public Health*. 2014;14:233.
52. Reklaitienė R, Grazulevičienė R, Dedele A, Virvičiūtė D, Vensloviene J, Tamosiunas A, et al. The relationship of green space, depressive symptoms and perceived general health in urban population. *Scand J Public Health*. 2014;42(7):669–76.
53. Sander HA, Ghosh D, Hodson CB. Varying age-gender associations between body mass index and urban greenspace. *Urban For Urban Green*. 2017;26:1–10.
54. Markevych I, Fuertes E, Tiesler CMT, Birk M, Bauer C-P, Koletzko S, et al. Surrounding greenness and birth weight: results from the GINIplus and LISAplus birth cohorts in Munich. *Health Place*. 2014;26:39–46.
55. van den Berg M, van Poppel M, van Kamp I, Andrusaityte S, Balseviciene B, Cirach M, et al. Visiting green space is associated with mental health and vitality: a cross-sectional study in four European cities. *Health Place*. 2016;38:8–15.
56. Dadvand P, Wright J, Martinez D, Basagaña X, McEachan RRC, Cirach M, et al. Inequality, green spaces, and pregnant women: roles of ethnicity and individual and neighbourhood socioeconomic status. *Environ Int*. 2014;71:101–8.
57. McEachan RRC, Prady SL, Smith G, Fairley L, Cabieses B, Gidlow C, et al. The association between green space and depressive symptoms in pregnant women: moderating roles of socioeconomic status and physical activity. *J Epidemiol Community Health*. 2016;70(3):253–9.
58. Luo Y-N, Huang W-Z, Liu X-X, Markevych I, Bloom MS, Zhao T, et al. Greenspace with overweight and obesity: a systematic review and meta-analysis of epidemiological studies up to 2020. *Obes Rev*. 2020;21(11):e13078.
59. Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Fornis J, Plasència A, et al. Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *Int J Environ Res Public Health*. 2015;12(4):4354–79.
60. Alexander DS, Huber LRB, Piper CR, Tanner AE. The association between recreational parks, facilities and childhood obesity: a cross-sectional study of the 2007 National Survey of Children's Health. *J Epidemiol Community Health*. 2013;67(5):427–31.
61. Manley D. Scale, aggregation, and the modifiable areal unit problem. In: Fischer MM, Nijkamp P, editors. *Handbook of regional science*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p. 1157–71. [https://doi.org/10.1007/978-3-642-23430-9\\_69](https://doi.org/10.1007/978-3-642-23430-9_69).
62. Tuson M, Yap M, Kok MR, Boruff B, Murray K, Vickery A, et al. Overcoming inefficiencies arising due to the impact of the modifiable areal unit problem on single-aggregation disease maps. *Int J Health Geogr*. 2020;19(1):40.
63. Duque JC, Laniado H, Polo A. S-maup: Statistical test to measure the sensitivity to the modifiable areal unit problem. *PLoS ONE*. 2018;13(11):e0207377.
64. Baar J, Romppel M, Igel U, Brähler E, Grande G. The independent relations of both residential self-selection and the environment to physical activity. *Int J Environ Health Res*. 2015;25(3):288–98.
65. Boone-Heinonen J, Guilkey DK, Evenson KR, Gordon-Larsen P. Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *Int J Behav Nutr Phys Act*. 2010;7:70.
66. Wilkes M, Thornton J, Horlick M, Sopher A, Wang J, Widen EM, et al. Relationship of BMI z score to fat percent and fat mass in multiethnic prepubertal children. *Pediatr Obes*. 2019;14(1):e12463.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

