Spatial Characteristics of Urban Green Spaces and Human Health: An Exploratory Analysis of Canonical Correlation

Chia-Tsung Yeh [1],*, Ya-Yun Cheng [2] and Tsai-Yun Liu [2]

[1] Graduate Institute of Urban Planning, National Taipei University, No. 151, University Rd., Sanxia Dist., New Taipei City 237303, Taiwan

[2] Urban and Rural Development Bureau, New Taipei City Government, No. 161, Sec. 1, Zhongshan Rd., Banqiao Dist., New Taipei City 22001, Taiwan; 0910129@gmail.com (Y.-Y.C.); AM5180@ms.ntpc.gov.tw (T.-Y.L.)

* Correspondence: alexyeh@mail.ntpu.edu.tw; Tel.: +886-2-8674-1111 (ext. 67351)

check for
updates

Abstract: In highly urbanized areas, urban green spaces (UGSs) are important natural and cultural entities. Previous studies have shown some evidence of positive relationships between UGSs and human health. Most of these studies relied on self-reported health data and often used institutional quantitative measures of UGSs instead of the spatial characteristics of UGSs. This study analyzed the relationships between the spatial characteristics of UGSs and morbidity of diseases, which were considered variables of human health in the Taipei Metro. The Longitudinal Health Insurance Database of Taiwan was applied as the source of morbidity of diseases. A canonical correlation analysis was performed by using the six variables of spatial characteristics of UGSs as predictors and three variables of morbidity as criterion variables to evaluate the multivariate shared relationships between the two variable sets. The results found a strong canonical correlation between the spatial characteristics of UGSs and human health. Furthermore, the results revealed that living in districts with a higher area percentage of green spaces and denser vegetation cover, as well as exposure to more aggregative and irregular-shape green spaces, can reduce the morbidity of diseases.

Keywords: urban green spaces; spatial characteristics; human health; canonical correlation analysis

1. Introduction

Green space can be defined as lands that consist predominantly of unsealed, pervious, soft surfaces, such as soil, grass, shrubs, and trees, with few structures present [1]. In highly urbanized areas sealed by artificial constructions, green spaces are important natural and cultural entities for urban residents. Green spaces in urban areas provide important ecosystem services, such as air quality improvement, atmospheric carbon dioxide reduction, and recreational and cultural values [2–4]. However, urbanization, environmental degradation, and lifestyle changes are diminishing opportunities for human contact with nature [5].

Amenities in green spaces encourage urban residents to increase physical activities, to pursue more leisure and recreation, and to have contact with the natural environment. Increasing evidence indicates that the presence of green spaces can enhance the quality of life [6,7]. For example, easy access to green spaces increases physical activity [5,8,9] and promotes physical and mental health of urban residents [6,10–13]. More greenness in urban residential areas is associated with more physical activity and lower risk of disease mortality [14]. Moreover, urban green spaces (UGSs) can facilitate social interactions, social networking, social support, social inclusion, and improve the mental health of residents [15–17].

More people live in cities now than at any other point in history. Urban growth is presenting numerous challenges for the maintenance of urban green spaces, and, consequently, for human health. An emerging body of studies has linked exposures to UGSs with improving human health, through increasing physical activity, reducing levels of noise and pollution, relieving physiological and psychological stress, and enhancing social contacts [18–25]. Epidemiological studies have focused on the association between UGSs and the risks of specific diseases, such as obesity (e.g., [26]), cardiovascular disease (e.g., [21,27–29]), and type 2 diabetes (e.g., [30,31]). However, few studies have simultaneously investigated the relationships between the spatial characteristics of UGSs and the risks of two or more diseases [11].

Previous studies heavily relied on self-reported data to investigate the association between exposures to UGSs and health benefits to participants (e.g., [20,24,25,32,33]). Questionnaire surveys and interviews are frequently used to obtain the self-reported data [17,34,35]. The criterion variables used in the analyses of associations were self-reported health indicators, e.g., number of symptoms experienced in the last 14 days, perceived general health, and score on the general health questionnaire [19,24]. These studies had relatively small study population sizes or were cross-sectional in design, limiting the generalizability of the results [29]. However, rare studies conducted their studies based on the large-scale datasets of human health insurance and analyzed the association between UGSs and the risk of disease in metropolitan areas [29].

Moreover, previous studies focused on the relationships between UGSs and human health usually used institutional measures of UGSs (e.g., area of green spaces per capita) instead of quantitative spatial characteristics of UGSs (e.g., composition, spatial configuration, and spatiotemporal patterns of UGSs). Green spaces with a complicated shape could bring people closer to nature, improve their physical and mental health [36]. In addition, an elongate park could be accessible by more residents, and a connected green network has higher amenity value than smaller and fragmented ones [37]. The fragmentation and separation of green space contributed to increased mortality of cardiovascular disease through air pollution and temperature [28]. The quantitative indicators, which refer to morphologies and vegetation cover, and spatial distribution, are considered as three main dimensions to assess UGSs effectively [38]. Quantifying the spatial characteristics of UGSs by using a geographic information system (GIS) and landscape metrics provides a more systematic and precise way to analyze the spatial characteristics of UGSs.

The variables of spatial characteristics of UGSs and human health possibly have multiple causes and effects between them. Determining outcomes based on a correlation analysis between characteristics of UGSs and morbidity of specific diseases may be insufficient to indicate the general health benefit of UGSs. Moreover, morbidities of different diseases may have disparate relationships with the characteristics of UGSs. In studies that only consider one criterion measure of human health, a multiple regression analysis can be conducted. However, this study tended to examine the relationships between two variable sets, i.e., the predictor variables (spatial characteristics of UGSs) and criterion variables (morbidities of diseases). Therefore, canonical correlation analysis (CCA) was applied to determine which variables are more or less influential in the model. The National Health Insurance System of Taiwan provides a high-quality dataset of outpatient treatments and hospitalizations with an accurate disease classification scheme. This dataset was the base to derive the morbidity of diseases, i.e., cardiovascular diseases, mental disorders, and respiratory diseases, as observed variables of human health in this study.

The following method section introduces the research methods, the study area, data sets required, and techniques for quantifying spatial characteristics of UGSs. Additionally, this section briefly introduces canonical correlation analysis, the multivariate statistical method used in this study. The section of results reveals the key variables of UGS on human health and the roles of variables of characteristics of UGS based on an exploratory statistical analysis. The discussion section compares the key findings in this study with current knowledge reported in recent studies. The last section concludes with the significant correlations between UGSs and human health and the advantages of using the data derived from the national health insurance system.

2. Methods

2.1. Study Area

In this study, the largest metropolitan area in Taiwan, Taipei Metro Area, was selected for the analysis of the relationships between spatial patterns of green spaces and human health. Figure 1 shows that Taipei Metro Area has three cities, i.e., Taipei City, New Taipei City, and Keelung City, and 48 districts and has a total area of 2457 km^2, with a population of 7.04 million at the end of 2017, representing 30% of the total population of Taiwan [39]. The mean annual temperature of Taipei Metro Area is 23 °C, with the variation between the lowest monthly average temperature of 16 °C in January and the highest monthly average temperature of 29.6 °C in July. The annual rainfall varies by terrain. The rainfall approximates 2400 mm in plain areas and is over 4500 mm in mountain areas (The information of rainfall and temperature data about the Taipei Metro Area was collected from [40–42]) This study analyzed UGSs in 37 districts that had a population larger than 25,000 and population density above 500 pop/km^2. The remote mountain districts and rural districts with lower population or population density were excluded.
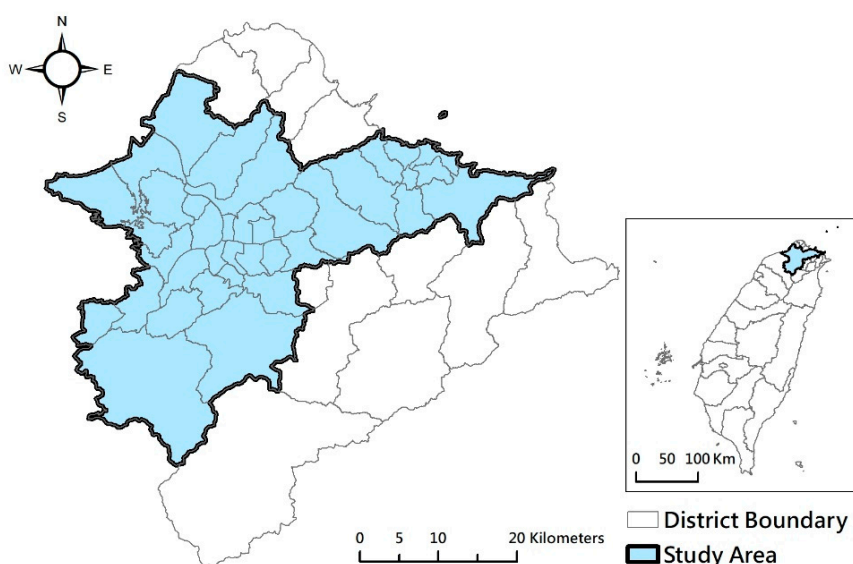

Figure 1. Study area.

2.2. Spatial Characteristics of UGSs

The digital map of UGSs of the study area was derived from the Taiwan land use investigation maps for 2015 produced by the National Land Surveying and Mapping Center, Ministry of the Interior. UGSs were selected by the attributes of land use/land cover and reclassified into five types: farmland, forest, wetland, grassland, and park and open spaces. This study also selected residential land use as the focus area to assess the spatial accessibility to UGSs. The ArcGIS tool was applied to convert the map from vector format to raster format with a pixel size of 30 m × 30 m for metric analysis using the tool for measuring landscape metrics, FRAGSTATS 4.2. Satellite data of SPOT 5 on 8 May 2014 and 5 February 2012 were also used for analyzing the normalized difference vegetation index (NDVI) of UGSs by using ERDAS IMAGINE.

A large set of metrics for analyzing landscape composition and configuration has been developed, modified, and tested in the past three decades [42,43]. In this study, we targeted the landscape metrics of UGSs regarding human health based on the literature review. The proportion of green spaces (PLAND) was calculated to assess the availability of UGSs in the Taipei Metro area. This study also selected the shape of UGS (SHAPE_MN), mean distance of UGS patches to the nearest residential patch (ENN_MN), and continuity of UGS patches (COHESION) as the indicators of the spatial pattern of UGSs. All these landscape metrics were implemented by using the statistical package FRAGSTATS 4.2 [44] (See Table 1).

Int. J. Environ. Res. Public Health **2020**, *17*, 3227

4 of 14

Table 1. Landscape spatial characteristics of urban green spaces (UGSs).

Four Aspects	Indicators	Description
Availability	PLAND	Percentage of study area comprising all patches of UGSs
Spatial Pattern	SHAPE_MN	Shape complexity of UGSs calculated as patch perimeter (m) of urban green spaces divided by the theoretically minimum perimeter of patch area (m)
	ENN_MN	Mean distance (m) of the patches of UGSs to the nearest neighboring patch based on shortest edge-to-edge distance from cell center to cell center
	COHESION	Connectivity of the patches of UGSs by computing the patch area and perimeter
Quality	NDVI	Indication of exposure and condition of green vegetation in UGSs [27]
Accessibility	Network Distance	Shortest distance between the residential land use and UGSs along a road network [45]

PLAND: Percentage of landscape; SHAPE_MN: Mean patch shape index; ENN_MN: Mean Euclidean nearest-neighbor distance; COHESION: Patch cohesion index; NDVI: Normalized difference vegetation index. Source: [44]. FRAGSTATS v4.2: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the author at the University of Massachusetts, Amherst, MA, USA.

The quality of UGSs was measured by using NDVI to assess the greenness of vegetation in the Taipei Metro Area. The index of NDVI correlates with biophysical properties of the vegetation canopy, such as leaf area index, fractional vegetation cover, vegetation condition, and biomass [46]. NDVI is often used as a metric to assess exposure to surrounding greenness. Previous studies also showed that a higher NDVI value is related to less depression [47] and better mental health [48]. Additionally, the analysis of Network Distance of ArcGIS was used to calculate the shortest network distance from the residential area to UGS as the index of accessibility.

The landscape spatial characteristics of UGSs were analyzed for 37 districts of Taipei Metro Area. The results of the spatial analysis were used to examine the correlation between spatial characteristics of UGSs and morbidities of diseases in the study area.

2.3. National Health Insurance Research Database

The database of Taiwan's National Health Insurance program contains registration files and original claim data for reimbursement. The National Health Insurance Administration, Ministry of Health and Welfare, Taiwan, derives large computerized databases from this program. The databases are maintained by the National Health Research Institutes, Taiwan, and are provided to scientists for research purposes. The data files from the National Health Insurance program are de-identified by scrambling the identification codes of both patients and medical facilities and sent to the National Health Research Institutes to form the original files of the National Health Insurance Research Database (NHIRD). This study applied the Longitudinal Health Insurance Database (LHID), which is a data subset of the NHIRD and contains all original claim data for 1,000,000 beneficiaries randomly sampled from the Registry for Beneficiaries (ID) of the NHIRD. All the registration files and original claim data of these 1,000,000 individuals were collected and constituted the LHID [49].

In Taiwan, according to the Personal Information Protection Act, residential locations of patients are confidential. Dataset of outpatients reveals only the spatial locations of the hospital. To consider the influences of cross-district outpatients, we combined the outpatient ID of the common cold (ICD9 460) and influenza (ICD9 487) with the outpatient ID of cardiovascular disease, mental disorders, and chronic respiratory disease. The locations of medical institutions for common cold and influenza are determined as the living places of outpatients based on the combined data.

This study applied the ambulatory care expenditures by visits (CD) of original claim data as the source of morbidity of diseases. The CD data contains the date of ambulatory care, ICD-9-CM codes, and ID of medical facilities (HOSP_ID). By combining ID of medical facilities (HOSP_ID) and registry

for contracted medical facilities (HOSB) which includes area code of medical facilities (AREA_NO_H), the number of outpatients can be summed up by district level. We included all registered outpatient database between 1 January 2010 and 31 December 2012 and selected the database of outpatients for the three types of disease related to exposure of green space, i.e., cardiovascular diseases (ICD9 390–459), mental disorders (ICD9 290–319), and chronic respiratory diseases (ICD9 470–478, 490–519).

2.4. Canonical Correlation Analysis

In 1936, Hotelling developed canonical correlation analysis (CCA) to examine the correlation between a synthetic criterion variable and a synthetic predictor variable that are weighted based on the relationships between the variables within the sets [50]. Instead of analyzing a single dataset, however, the goal of CCA is to analyze the relationship between a pair of datasets [51]. Each dataset can contain several variables, and CCA calculates a linear combination for each set, called a canonical variable, such that any correlation between two canonical variables is maximized [52]. These correlations are called canonical correlations, and the linear combinations are canonical variates. The fundamental principle behind CCA is to create a number of canonical variates, each consisting of a linear combination of one set of variables (X_i) in the following equation:

$$U_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{iP}X_P$$

Additionally, a linear combination of the other set of variables (Y_i) has the following equation:

$$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \cdots + b_{iq}Y_q$$

The goal of CCA is to determine the coefficients, or canonical weights (a_{ij} and b_{ij}), that maximize the correlation between canonical variates U_i and V_i. The first canonical correlation, Corr. (U_1, V_1), is the strongest possible correlation between a linear combination of variables in the exposure set and a linear combination of variables in the outcome set. Other pairs of maximally correlated linear combinations are chosen in turn and are orthogonal to those previously identified. The process of constructing canonical variates continues until the number of pairs of canonical variates equals the number of variables in the smaller of the two sets.

Because CCA assesses the existence of correlations between two sets of variables, this exploratory statistical method is useful for analyzing relationships between the combinations of urban form and different variables of human health. Although CCA is widely used in social science research [53], the method is rarely used in studies on UGSs [54]. This study applied CCA to examine multivariate relationships between spatial characteristics of UGSs and human health. The proposed canonical correlation model is shown in Figure 2. This study attempted to quantify the independent relationships of spatial characteristics of UGSs and levels of morbidity of three diseases. Table 2 shows the descriptive statistics of the studied variables. It provides information about the number of samples, minimum, maximum, mean, and standard deviation of the studied variables.

A canonical correlation analysis was performed with IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp. released 2013, Armonk, NY, USA). Six variables of spatial characteristics of UGSs were used as predictors of three observed variables of morbidity to evaluate the multivariate shared relationship between these two variable sets. Before the analysis results show the fitness of the model, the Pearson correlation matrix describes the associations between the variables of the spatial characteristics of UGSs. It also detects the possibility of multicollinearity of studied variables. Table 3 shows that all the correlations (r) < 0.765 (in absolute value). The general rule of thumb is that if the correlation (r) > 0.8, then severe multicollinearity may be present. Therefore, it is not necessary to remove the redundant variables.

This study uses a multivariate test of significance of CCA (i.e., Pillai's, Hotelling's, Wilk's, and Roy's multivariate criteria) to show the general fit of the model. Furthermore, the canonical correlation coefficients and the eigenvalues are used to explain the fitness of the canonical roots of the model.

Based on the tests, the significance of each of the roots in the dimension reduction analysis of SPSS, the possible significant roots are found. Since our model contains three disease morbidities and six indicators of spatial characteristics of UGSs, SPSS extracts three canonical roots or dimensions. In the dimension reduction analysis, the first test of significance tests all three canonical roots of significance; the second test excludes the first root and tests roots two to three; the last test tests root three by itself.
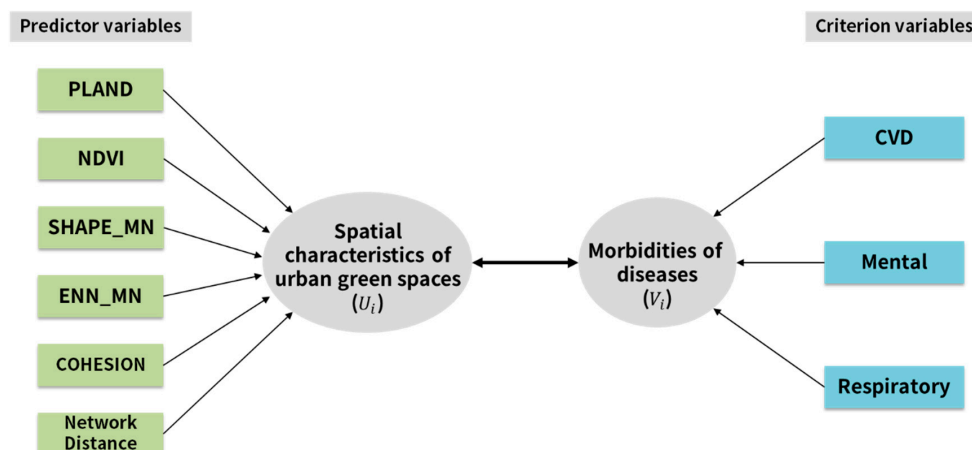

Figure 2. The canonical correlation model that was used in our study to assess the relationship of the predictor variables (spatial characteristics of urban green spaces (UGSs)) with criterion variables (morbidities of CVD (cardiovascular diseases), mental disorders, and respiratory diseases).

Table 2. Descriptive statistics of the studied variables.

Diseases	N	Minimum	Maximum	Mean	Std. Deviation
CVD (per 100,000 person-year)	37	2770	45,566	11,090.81	9870.864
Mental (per 100,000 person-year)	37	821	14,735	3683.97	3214.976
Respiratory (per 100,000 person-year)	37	1185	21,270	5140.27	4694.280
PLAND (%)	37	0.0431	0.8442	0.427227	0.2415632
SHAPE_MN (none)	37	1.26	2.15	1.6136	0.16357
ENN_MN (m)	37	65.65	195.36	102.4700	35.77035
NDVI (none)	37	−0.07	0.37	0.1578	0.12079
Network distance (m)	37	207.79	888.71	443.2970	162.14926
COHESION (none)	37	76.84	99.96	97.2686	4.67171
Valid N (listwise)	37				

Table 3. Correlation analysis of the variables of spatial characteristics of UGSs.

Spatial Characteristics of UGSs.	PLAND	SHAPE_MN	ENN_MN	NDVI	Network Distance	COHESION
PLAND	1	0.234	−0.633	0.674	−0.756	0.708
SHAPE_MN	0.234	1	−0.241	0.162	−0.160	0.446
ENN_MN	−0.633	−0.241	1	−0.408	0.526	−0.427
NDVI	0.674	0.162	−0.408	1	−0.712	0.671
Network Distance	−0.756	−0.160	0.526	−0.712	1	−0.396
COHESION	0.708	0.446	−0.427	0.671	−0.396	1

3. Results

3.1. Spatial Pattern of the Characteristics of UGSs and the Morbidity of Diseases

We analyzed the spatial characteristics of UGSs of the districts in the study area based on the land use investigation map and the FRAGSTATS 4.2 software. Figure 3a shows that the districts with a low percentage of green spaces roughly clustered in the center of the Taipei Metro Area. The level of NDVI of the districts (Figure 3b) shows that UGS quality was lower in the urban central districts than that in

peripheral districts. The level SHAPE_MN of the districts (Figure 3c) similarly shows that the shape of UGSs in the urban central districts was more regular compared to peripheral districts. In highly urbanized districts, the distances between green spaces (ENN_MN) were longer than those in the lightly urbanized out-ring districts (Figure 3d). Conversely, the continuity of green spaces (COHESION) was higher in peripheral districts compared to central districts (Figure 3e). The Network Distances from residential areas to green spaces were longer in the central districts (Figure 3f). Generally, availability, quality, and accessibility of UGS were higher in southern and northern west parts of the Taipei Metro Area. Inversely, availability, quality, and accessibility of UGSs of the districts in the Taipei basin area and Keelung city were poorer. The connectivity of UGS was lower in highly urbanized districts.
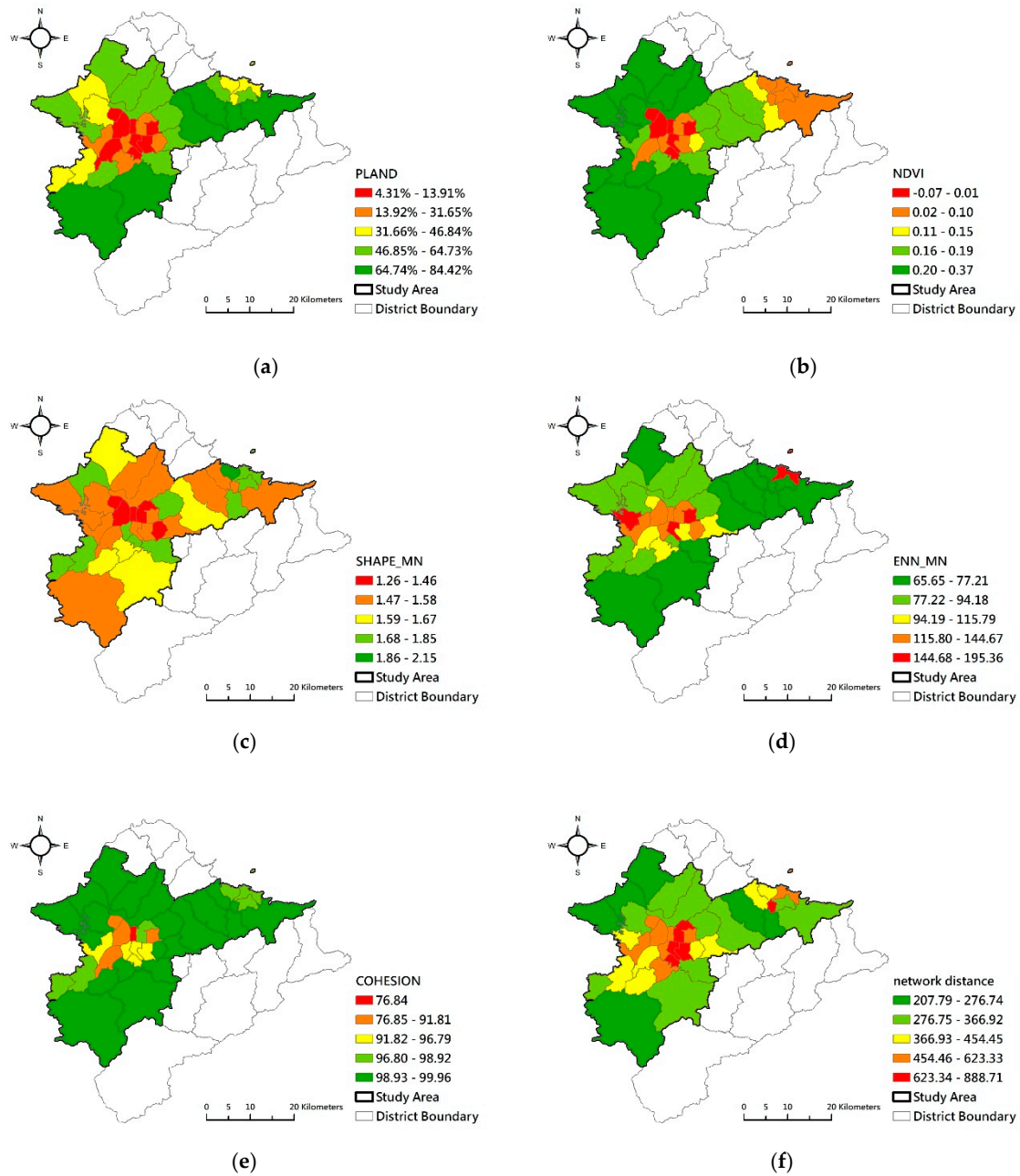

Figure 3. Landscape spatial patterns and distribution of urban green spaces. (**a**) proportion of green spaces (PLAND); (**b**) normalized difference vegetation index (NDVI); (**c**) shape of UGS (SHAPE_MN); (**d**) mean distance of UGS patches to the nearest residential patch (ENN_MN); (**e**) continuity of UGS patches (COHESION); (**f**) Network Distance.

*Int. J. Environ. Res. Public Health* **2020**, *17*, 3227

8 of 14

Morbidities of cardiovascular disease, mental disorder, and chronic respiratory disease were high, and the quality of green spaces was low in the central districts of the Taipei Basin. Conversely, western and southern districts had low morbidity and high quality and connectivity of green spaces. The spatial patterns of morbidity of cardiovascular disease, mental disorder, and chronic respiratory disease (Figure 4) were roughly opposite to the quality, greenness, and connectivity of green spaces in our study area.
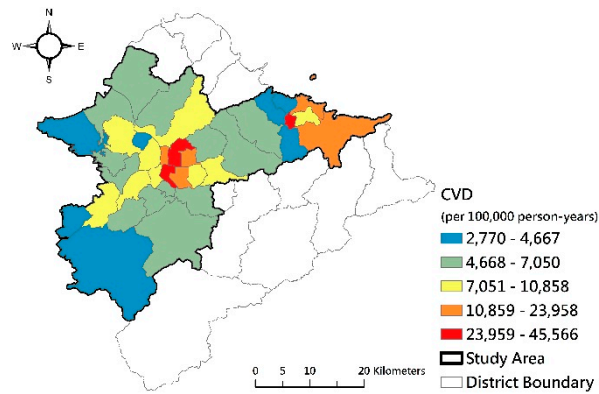
(**a**)
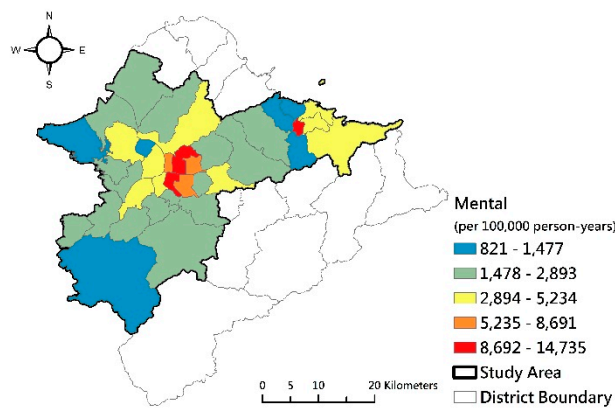
(**b**)
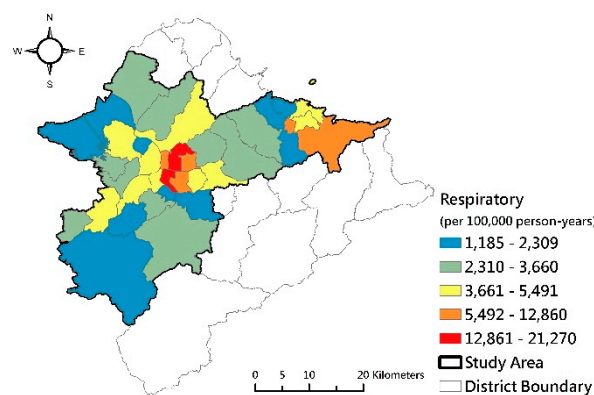
(**c**)

Figure 4. Levels of the morbidity of the three types of diseases. (**a**) Cardiovascular diseases; (**b**) Mental disorders; (**c**) Respiratory diseases.

Int. J. Environ. Res. Public Health **2020**, *17*, 3227

9 of 14

3.2. Relationships Between Spatial Characteristics of UGSs and Morbidity

The output of CCA shows the general fit of the model reporting Pillai's, Helling's, Wilk's and Roy's multivariate criteria. The commonly used test is Wilk's, but we found that all of these tests were significant, with $p < 0.001$ (see Table 4).

Table 4. Multivariate tests of significance.

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	1.13385	3.03793	18.00	90.00	0.000
Hotellings	2.49326	3.69372	18.00	80.00	0.000
Wilks	0.19831	3.41680	18.00	79.68	0.000
Roys	0.62941				

Table 5 reports the canonical correlation coefficients and the eigenvalues of the canonical roots. The CCA analysis yielded three canonical functions with squared canonical correlation (Sq Cor.) of 0.62941, 0.40733, and 0.09710 for each root. The first pair of variates, a linear combination of spatial characteristics of UGSs and a linear combination of the variables of morbidities of diseases, had a canonical correlation coefficient of 0.79336 and an eigenvalue of 1.698. The analytical results indicate that the spatial characteristics of UGSs and human health are correlated. The first root was created to maximize the canonical correlation (Pearson r) between the two synthetic variables. The second pair of canonical variates was constructed out of the residuals of the first pair to maximize the correlation between them. The first and second roots had explained variances of 68.12% and 27.57%. The third root had a low explained variance of 4.31% (see Table 5).

Table 5. Eigenvalues and canonical correlations.

Root No.	Eigenvalue	%	Cum. %	Canon Cor.	Sq. Cor.
1	1.69843	68.12088	68.12088	0.79336	0.62941
2	0.68728	27.56565	95.68653	0.63823	0.40733
3	0.10755	4.31347	100.00000	0.31161	0.09710

SPSS extracted three canonical roots or dimensions for our model. The dimension reduction analysis tested the significance of each of the roots (see Table 6). The first test of significance showed the full model across root 1 to 3 was significant (Wilk's λ = 0.19831, F = 3.41680, $p < 0.01$). The second test excluded the first root and tests roots 2 to 3 and showed it was also significant (Wilk's λ = 0.53512, F = 2.21872, $p = 0.036 < 0.05$). The root 3 itself was not significant (Wilk's λ = 0.90290, F = 0.80659, $p = 0.531$) (see Table 6).

Table 6. Dimension reduction analysis.

Roots	Wilks L.	F	Hypoth. DF	Error DF	Sig. of F
1 TO 3	0.19831	3.41680	18.00	79.68	0.000
2 TO 3	0.53512	2.12872	10.00	58.00	0.036
3 TO 3	0.90290	0.80659	4.00	30.00	0.531

Next, we turned to the research question of which variables of the spatial characteristic of UGSs contribute to this relationship between the variable sets across the two functions. Table 7 shows the statistics of the canonical solution of Function 1 and Function 2 for the relationships between spatial characteristics of UGSs and morbidities of diseases. The table presents the structure coefficients (r_s), the squared structure coefficients (r_s^2), and cross-structure coefficients (r_c) between a set of variables and the synthetic variable created by another set of variables.

Int. J. Environ. Res. Public Health **2020**, *17*, 3227

10 of 14

Table 7. Canonical correlation solution.

Predictor Variables	Function 1			Function 2		
	r_s	$r_s{}^2$	r_c	r_s	$r_s{}^2$	r_c
PLAND	−0.53151	0.28250	−0.42168	0.53563	0.28690	0.34186
NDVI	−0.60826	0.36998	−0.48257	0.52014	0.27054	0.33197
SHAPE_MN	−0.48613	0.23632	−0.38568	−0.17472	0.03052	−0.11151
ENN_MN	0.53458	0.28578	0.42411	−0.03399	0.00115	−0.02169
COHESION	−0.34953	0.12217	−0.27730	0.75119	0.56428	0.47943
Network Distance	0.88403	0.78150	0.70135	−0.28024	0.07853	−0.17886

Criterion Variables	Function 1			Function 2		
	r_s	$r_s{}^2$	r_c	r_s	$r_s{}^2$	r_c
Cardiovascular diseases	0.98556	0.97139	0.78190	−0.11529	0.01344	−0.07358
Mental Disorders	0.98352	0.96731	0.78029	−0.17786	0.03163	−0.11351
Respiratory Diseases	0.99449	0.98901	0.78899	−0.06647	0.00442	−0.04242

r_s = structure coefficient; $r_s{}^2$ = squared structure coefficient; r_c = cross-structure coefficient. For emphasis, absolute value of structure coefficients and cross-structure coefficients above 0.45 are underlined.

In Function 1, the first pair of canonical variates groups the variables in the way that the correlation between them is maximized. The structure coefficients (r_s) of PLAND, NDVI, SHAPE_MN, ENN_MN, and Network Distance indicated that they were the primary contributors (with r_s = −0.53151, −0.60826, −0.48613, 0.53458, and 0.88403) to the synthetic variate of characteristics of green spaces. The cross-structure coefficients (r_c) of PLAND, NDVI, and SHAPE_MN showed that they were negatively related to the synthetic variate of morbidity of diseases. Inversely, the cross-structure coefficients (r_c) for ENN_MN and Network Distance show that they were positively related to the synthetic variate of morbidity of diseases. The other side of Function 1, the structure coefficients (r_s) of criterion variables indicate that cardiovascular diseases, mental disorders, and respiratory diseases were positively correlated with the synthetic variate of morbidity (with very high r_s = 0.98556, 0.98352, and 0.99449). The cross-structure coefficients (r_c) of the three criterion variables, all of which were higher than 0.7, revealed that morbidity of cardiovascular diseases, mental disorders, and respiratory diseases all had positive correlations with the synthetic variate of spatial characteristics of UGSs.

In the Function 2, the values of structure coefficients (r_s) indicated that the variables of PLAND, NDVI, and COHESION were the primary contributors to the synthetic variate of predictors. The cross-structure coefficients (r_c) of PLAND, NDVI, and COHESION showed that they were positively related to the synthetic variate of disease morbidities with moderate r_c = 0.34186, 0.33197, and 0.47943. On the other side of Function 2, the structure coefficients (r_c) of criterion variables indicated that cardiovascular diseases, mental disorders, and respiratory diseases had very low correlation coefficients with the synthetic variate of morbidity of diseases (with very low r_s < 0.2). The cross-structure coefficients (r_c) of all the three criterion variables were lower than 0.12, revealing that morbidity of cardiovascular diseases, mental disorders, and respiratory diseases all had very low correlations with the synthetic variate of spatial characteristics of UGSs.

4. Discussion

In this study, we correlated the district-level spatial characteristics of UGSs with the morbidities of specific diseases. The results of CCA analysis showed a significant correlation between spatial characteristics of UGSs and human health in the Taipei Metro Area. The cross-structure coefficients of predictor variables revealed that the morbidity of diseases was decreased by residence in the districts with higher area percentage, denser vegetable cover, more aggregative urban greens, and irregular-shape green spaces. In comparison with the variables of spatial characteristics of UGSs, accessibility (measured by Network Distance) was the major contributor to the relationship between spatial characteristics and human health. The secondary contributors were the variables of quality of UGSs (measured by NDVI)

and spatial pattern of UGSs (measured by SHAPE_MN, ENN_MN, and COHESION). The availability of UGSs (measured by PLAND) was less influential in the relationship between spatial characteristics of UGSs and human health.

The results of correlation analyses validated the evidences of health benefits of exposure to green spaces that have been found by previous studies (e.g., [21,35,55–57]). This study also revealed the key variables of spatial characteristics of UGSs in the relationship between UGSs and human health. The analytical results showed that the variables of quality (measured by NDVI) and accessibility played key roles in influencing the morbidities of diseases in the Taipei Metro Area. These findings confirm the evidence found by [9,12,34] in their study areas. In this study, the strategy of using a database of national health insurance and conducting district-level analyses is similar to the newly published study of [29]. However, the previous study usually concerned only with a single disease, i.e., cardiovascular disease. Our study reveals the correlations between the spatial characteristics of UGSs and the morbidity of three diseases.

A major limitation of our study is that the data of the National Health Insurance Research did not reveal the real addresses of the patients but only the districts where the patients were insured and outpatient clinics located. However, the locations of medical institutions they visited sometimes may be different from that they lived. To reduce the effects of cross-district outpatients, we cross-matched the address of the medical institution which the patients visited for the common cold and flu. This match was based on the assumption that urban residents tend to visit medical institutions near to their dwelling places to get timely treatment for a light ailment, such as cold and flu.

The database of residential health used in this study is different from self-reported data used by the previous studies that investigate the association between exposures to UGSs and health status of interviewees (e.g., [25,27,32,33,58]) We do not know their real contact time to green spaces and individual behavior of urban residents based on the database of the National Health Insurance. However, this approach is better than that using self-reported data in the case of dealing with the relationship between UGSs and health status for whole residents in a big metropolitan area.

5. Conclusions

Many previous studies focusing on the health benefits of urban green spaces relied on self-reported health and socio-economic data from a relatively small number of respondents. Using the database of the National Health Insurance Research of Taiwan and applying the statistical method of canonical correlation analysis distinguish this study from previous ones. This study shows the detailed relationships between spatial characteristics of UGSs and three diseases that humans suffer from based on the data of morbidities of three diseases. The analytical results revealed a significant positive correlation between the spatial characteristics of UGSs and human health. The results also showed that the districts with lower availability, quality, accessibility to green spaces had a high level of morbidity of cardiovascular diseases, mental disorders, and respiratory diseases. Conversely, districts with high aggregation, availability, quality, and accessibility to green spaces apparently had a low level of morbidity.

Since previous studies of UGSs and human health have focused on the risks of a specific disease, this study investigated the relationships between the spatial characteristics of UGSs and levels of general human health. Instead of relying on the morbidity of a specific disease, this study simultaneously investigated the relationships between the morbidities of three diseases and the spatial characteristics of UGSs. The analytical results reveal that human health generally has a strong correlation with the spatial characteristics of urban green spaces. The results also reveal that better biophysical property of vegetation canopy in urban areas is very important to improve human health. We emphasize not only the importance of the availability of UGSs for human health but the contributions of quality, accessibility, and spatial patterns of UGSs to human health. This study revealed that Network Distance and NDVI of UGSs are the main contributors to human health. Respiratory disease is shown by our results as the most influenced disease by spatial characteristics of urban green spaces. In addition to increasing the number of UGSs, improving the quality, accessibility, and spatial patterns of UGSs can

reduce the morbidities of diseases and promote better urban human health. The findings of this study can be useful for maintaining precious green spaces in cities and for urban planning to improve green infrastructure in urban areas.

References

1. Forman, R.T.T. *Urban Regions: Ecology and Planning Beyond the City*; Cambridge University Press: New York, NY, USA, 2008.
2. Nowak, D.J.; Crane, D.E.; Stevens, J.C. Air pollution removal by urban trees and shrubs in the United States. *Urban For. Urban Green.* **2006**, *4*, 115–123. [CrossRef]
3. Vailshery, L.S.; Jaganmohan, M.; Nagendra, H. Effect of street trees on microclimate and air pollution in a tropical city. *Urban For. Urban Green.* **2013**, *12*, 408–415. [CrossRef]
4. WHO Regional Office for Europe. *Urban Green Spaces and Health: A Review of Evidence*; WHO Regional Office for Europe: Copenhagen, Denmark, 2016.
5. Hartig, T.; Mitchell, R.; De Vries, S.; Frumkin, H. Nature and health. *Annu. Rev. Hum. Health* **2014**, *35*, 207–228. [CrossRef] [PubMed]
6. Chiesura, A. The role of urban parks for the sustainable city. *Landsc. Urban Plan.* **2004**, *68*, 129–138. [CrossRef]
7. Balram, S.; Dragićević, S. Attitudes toward urban green spaces: Integrating questionnaire survey and collaborative GIS techniques to improve attitude measurements. *Landsc. Urban Plan.* **2005**, *71*, 147–162. [CrossRef]
8. Witten, K.; Hiscock, R.; Pearce, J.; Blakely, T. Neighbourhood access to open spaces and the physical activity of residents: A national study. *Prev. Med.* **2008**, *47*, 299–303. [CrossRef]
9. James, P.; Banay, R.F.; Hart, J.E.; Laden, F. A Review of the Health Benefits of Greenness. *Curr. Epidemiol. Rep.* **2015**, *2*, 131–142. [CrossRef]
10. Barbosa, O.; Tratalos, J.A.; Armsworth, P.R.; Davies, R.G.; Fuller, R.A.; Johnson, P.; Gaston, K.J. Who benefits from access to green space? A case study from Sheffield, UK. *Landsc. Urban Plan.* **2007**, *83*, 187–195. [CrossRef]
11. Wolch, J.R.; Byrne, J.; Newell, J.P. Urban green space, human health, and environmental justice: The challenge of making cities "just green enough". *Landsc. Urban Plan.* **2014**, *125*, 234–244. [CrossRef]
12. Janssen, I.; Le Blanc, A.G. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 40. [CrossRef]
13. WHO. *Global Recommendations on Physical Activity for Health*; World Health Organization: Geneva, Switzerland, 2010.
14. Vienneau, D.; de Hoogh, K.; Faeh, D.; Kaufmann, M.; Wunderli, J.M.; Röösli, M. More than clean air and tranquillity: Residential green is independently associated with decreasing mortality. *Environ. Int.* **2017**, *108*, 176–184. [CrossRef]
15. Kim, J.; Kaplan, R. Physical and psychological factors in sense of community: New urbanist Kentlands and nearby orchard village. *Environ. Behav.* **2004**, *36*, 313–340. [CrossRef]
16. Seeland, K.; Dübendorfer, S.; Hansmann, R. Making friends in Zurich's urban forests and parks: The role of public green space for social inclusion of youths from different cultures. *For. Policy Econom.* **2009**, *11*, 10–17. [CrossRef]
17. Dadvand, P.; Bartoll, X.; Basagaña, X.; Dalmau-Bueno, A.; Martinez, D.; Ambros, A.; Nieuwenhuijsen, M.J. Green spaces and General Health: Roles of mental health status, social support, and physical activity. *Environ. Int.* **2016**, *91*, 161–167. [CrossRef] [PubMed]

Int. J. Environ. Res. Public Health **2020**, *17*, 3227

13 of 14

18. Bowler, D.E.; Buyung-Ali, L.M.; Knight, T.M.; Pullin, A.S.A. systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Hum. Health* **2010**, *10*, 456. [CrossRef] [PubMed]

19. Maas, J.; Verheij, R.A.; De Vries, S.; Spreeuwenberg, P.; Schellevis, F.G.; Groenewegen, P.P. Morbidity is related to a green living environment. *J. Epidemiol. Community Health* **2009**, *63*, 967–973. [CrossRef]

20. Pereira, G.; Christian, H.; Foster, S.; Boruff, B.J.; Bull, F.; Knuiman, M.; Giles-Corti, B. The association between neighborhood greenness and weight status: An observational study in Perth Western Australia. *Environ. Health* **2013**, *12*, 1. [CrossRef]

21. Tamosiunas, A.; Grazuleviciene, R.; Luksiene, D.; Dedele, A.; Reklaitiene, R.; Baceviciene, M.; Nieuwenhuijsen, M.J. Accessibility and use of urban green spaces, and cardiovascular health: Findings from a Kaunas cohort study. *Environ. Health* **2014**, *13*, 20. [CrossRef]

22. Dadvand, P.; Villanueva, C.M.; Font-Ribera, L.; Martinez, D.; Basagaña, X.; Belmonte, J.; Nieuwenhuijsen, M.J. Risks and benefits of green spaces for children: A cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy. *Environ. Health Perspect.* **2015**, *122*, 1329–1335. [CrossRef]

23. Takano, T.; Nakamura, K.; Watanabe, M. Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces. *J. Epidemiol. Community Health* **2002**, *56*, 913–918. [CrossRef]

24. de Vries, S.; Verheij, R.A.; Groenewegen, P.P.; Spreeuwenberg, P. Natural environments—Healthy environments? An exploratory analysis of the relationship between greenspace and health. *Environ. Plan. A* **2003**, *35*, 1717–1731. [CrossRef]

25. Mukherjee, D.; Safraj, S.; Tayyab, M.; Shivashankar, R.; Patel, S.A.; Narayanan, G.; Prabhakaran, D. Park availability and major depression in individuals with chronic conditions: Is there an association in urban India? *Health Place* **2017**, *47*, 54–62. [CrossRef] [PubMed]

26. Lee, H.; Kang, H.M.; Ko, Y.J.; Kim, H.S.; Kim, Y.J.; Bae, W.K.; Cho, B. Influence of urban neighbourhood environment on physical activity and obesity-related diseases. *Public Health* **2015**, *129*, 1204–1210. [CrossRef] [PubMed]

27. Pereira, G.; Foster, S.; Martin, K.; Christian, H.; Boruff, B.J.; Knuiman, M.; Giles-Corti, B. The association between neighborhood greenness and cardiovascular disease: An observational study. *BMC Hum. Health* **2012**, *12*, 1. [CrossRef] [PubMed]

28. Shen, Y.S.; Lung, S.C.C. Can green structure reduce the mortality of cardiovascular diseases? *Sci. Total Environ.* **2016**, *566*, 1159–1167. [CrossRef]

29. Seo, S.; Choi, S.; Kim, K.; Kim, S.M.; Park, S.M. Association between urban green space and the risk of cardiovascular disease: A longitudinal study in seven Korean metropolitan areas. *Environ. Int.* **2019**, *125*, 51–57. [CrossRef] [PubMed]

30. Astell-Burt, T.; Feng, X.; Kolt, G.S. Is neighborhood green space associated with a lower risk of type 2 diabetes evidence from 267,072 Australians. *Diabetes Care* **2014**, *37*, 197–201. [CrossRef]

31. Bodicoat, D.H.; O'Donovan, G.; Dalton, A.M.; Gray, L.J.; Yates, T.; Edwardson, C.; Hill, S.; Webb, D.R.; Khunti, K.; Davies, M.J.; et al. The association between neighbourhood greenspace and type 2 diabetes in a large cross-sectional study. *BMJ Open* **2014**, *4*, 1–8. [CrossRef]

32. Carrus, G.; Scopelliti, M.; Lafortezza, R.; Colangelo, G.; Ferrini, F.; Salbitano, F.; Sanesi, G. Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landsc. Urban Plan.* **2015**, *134*, 221–228. [CrossRef]

33. Bertram, C.; Rehdanz, K. The role of urban green space for human well-being. *Ecol. Econ.* **2015**, *120*, 139–152. [CrossRef]

34. Akpinar, A. How is quality of urban green spaces associated with physical activity and health? *Urban For. Urban Green.* **2016**, *16*, 76–83. [CrossRef]

35. Akpinar, A.; Barbosa-Leiker, C.; Brooks, K.R. Does green space matter? Exploring relationships between green space type and health indicators. *Urban For. Urban Green.* **2016**, *20*, 407–418. [CrossRef]

36. Alcock, I.; White, M.P.; Wheeler, B.W.; Fleming, L.E.; Depledge, M.H. Longitudinal effects on mental health of moving to greener and less green urban areas. *Environ. Sci. Technol.* **2014**, *48*, 1247–1255. [CrossRef]

37. Tian, Y.; Jim, C.Y.; Wang, H. Assessing the landscape and ecological quality of urban green spaces in a compact city. *Landsc. Urban Plan.* **2014**, *121*, 97–108. [CrossRef]

38. De La Barrera, F.; Reyes-Paecke, S.; Banzhaf, E. Indicators for green spaces in contrasting urban settings. *Ecol. Indic.* **2016**, *62*, 212–219. [CrossRef]

39. Ministry of the Interior, Republic of China (Taiwan). Statistical Yearbook of Interior. Available online: https://www.moi.gov.tw/files/site_stuff/321/2/year/year_en.html (accessed on 17 April 2020).

40. Department of Budget, Accounting and Statistics, Taipei City Government. *2018 Taipei City Statistical Yearbook*; Taipei City Government: Taipei, Taiwan, 2019.

41. Department of Budget, Accounting and Statistics, New Taipei City Government. *2018 New Taipei City Statistical Yearbook*; New Taipei City Government: Taipei, Taiwan, 2019.

42. Department of Budget, Accounting and Statistics, Keelung City Government. *2018 Keelung City Statistical Yearbook*; Keelung City Government: Keelung, Taiwan, 2019.

43. Forman, R.T.T.; Godron, M. *Landscape Ecology*; Wiley: New York, NY, USA, 1986.

44. McGarigal, K.; Ene, E.; Holmes, C. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps, Computer Software Program Produced by the Authors at the University of Massachusetts. 2002. Available online: https://www.umass.edu/landeco/research/fragstats/fragstats.html (accessed on 17 April 2020).

45. La Rosa, D. Accessibility to greenspaces: GIS based indicators for sustainable planning in a dense urban context. *Ecol. Indic.* **2014**, *42*, 122–134. [CrossRef]

46. Jiang, Z.; Huete, A.R.; Chen, J.; Chen, Y.; Li, J.; Yan, G.; Zhang, X. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sens. Environ.* **2006**, *101*, 366–378. [CrossRef]

47. Cohen-Cline, H.; Turkheimer, E.; Duncan, G.E. Access to green space, physical activity and mental health: A twin study. *J. Epidemiol. Community Health* **2015**, *69*, 523–529. [CrossRef]

48. Beyer, K.M.; Kaltenbach, A.; Szabo, A.; Bogar, S.; Nieto, F.J.; Malecki, K.M. Exposure to neighborhood green space and mental health: Evidence from the survey of the health of Wisconsin. *Int. J. Environ. Res. Public Health* **2014**, *11*, 3453–3472. [CrossRef]

49. National Health Research Institutes. National Health Insurance Research Database, Taiwan. 2018. Available online: https://nhird.nhri.org.tw/ (accessed on 1 April 2020).

50. Sherry, A.; Henson, R.K. Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *J. Personal. Assess.* **2010**, *84*, 37–48. [CrossRef]

51. Avron, H.; Boutsidis, C.; Toledo, S.; Zouzias, A. Efficient Dimensionality Reduction for Canonical Correlation Analysis. *SIAM J. Sci. Comput.* **2014**, *36*, S111–S131. [CrossRef]

52. Rencher, A. Interpretation of canonical discriminant functions, canonical variates, and principal components. *Am. Stat.* **1992**, *46*, 217–225.

53. Pugh, R.C.; Hu, Y. Use and Interpretation of Canonical Correlation Analyses in Journal of Educational Research Articles: 1978–1989. *J. Educ. Res.* **1991**, *84*, 147–152. [CrossRef]

54. Kardan, O.; Gozdyra, P.; Misic, B.; Moola, F.; Palmer, L.J.; Paus, T.; Berman, M.G. Neighborhood greenspace and health in a large urban center. *Sci. Rep.* **2015**, *5*, 1–14. [CrossRef]

55. Nutsford, D.; Pearson, A.L.; Kingham, S. An ecological study investigating the association between access to urban green space and mental health. *Public Health* **2013**, *127*, 1005–1011. [CrossRef]

56. Richardson, E.A.; Mitchell, R. Gender differences in relationships between urban green space and health in the United Kingdom. *Soc. Sci. Med.* **2010**, *71*, 568–575. [CrossRef]

57. Lovasi, G.S.; O'Neil-Dunne, J.P.M.; Lu, J.W.T.; Sheehan, D.; Perzanowski, M.S.; Macfaden, S.W.; King, K.L.; Matte, T.; Miller, R.L.; Hoepner, L.A.; et al. Urban tree canopy and asthma, wheeze, rhinitis, and allergic sensitization to tree pollen in a New York city birth cohort. *Environ. Health Perspect.* **2013**, *121*, 494–500. [CrossRef]

58. De Vries, S.; Van Dillen, S.M.; Groenewegen, P.P.; Spreeuwenberg, P. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Soc. Sci. Med.* **2013**, *94*, 26–33. [CrossRef]