RISK ANALYSIS OF BREAST CANCER SURVIVORSHIP IN METRO VANCOUVER

CONSIDERING WALKABILITY AND VANDIX

Han Bai, Melanie MacInnes, Aateka Shashank, Terence Wong, Harry Ma

GEOG 455 - Advanced Issues in GIScience, Spring 2014, Simon Fraser University

TABLE OF CONTENTS

WEBSITE

For a more detailed report on our research and findings, please refer to our website at *https://webdav.sfu.ca/web/geog/geog455_1141/Walkability/home.html*

The website contains the following:

- video abstract
- interactive map
- more maps that illustrate our research findings
- more graphic information describing our research findings

From right to left: Aateka Shashank, Han Bai, Melanie MacInnes, Terence Wong, Harry Ma

TEAM

METRO VANCOUVER

The region of Metro Vancouver was chosen as the spatial unit of this research because there is enough health information regarding breast cancer and factors contributing to walkability chosen for this research purpose: land use, slope, street connectivity, population density, proximity to parks and proximity to bus stops. The quantifiable spatial unit of choice was the 6-digit postal code.

WALKABILITY

Walkability, as defined by Janes Walk, an international event-based walking community, is

a quantitative and qualitative measurement of how inviting or un-inviting an area is to pedestrians. Walking matters more and more to towns and cities as the connection between walking and socially vibrant neighbourhoods is becoming clearer. Built environments that promote and facilitate walking - to stores, work, school and amenities – are better places to live, have higher real estate values, promote healthier lifestyles and have higher levels of social cohesion. (Janes Walk, 2014)

Walkability has been measured by several initiatives in Metro Vancouver, some of which are the Walkability Index and Google Walk Scores. The Walkability Index, created by a study conducted by the University of British Columbia measure walkability in neighborhoods based on street connectivity, land use mix, residential density, and commercial density (UBC, 2013). While these studies adequately measure the built environment to measure the ability of a neighborhoods to sustain walking activities, they doesn't say much about the demographics or needs of the populations living in these neighborhoods. Google Walk Scores assumes a similar algorithm to measure walkability, however, Google Walk Scores makes it easy to adapt walk scores to car dependency, real estate and also bikability of neighbourhoods.

The one thing the aforementioned walkability measures have in common is the attempt to define and describe neighborhood designs that can sustain walking as an activity; the focus is entirely on the built environment regardless of the populations living in it. One study finds a severe disconnect between walkability and its ability to measure crime rates (Carr et al, 2010). Another study considers to what extent walkability studies are adequately able to reflect on the walking patterns of populations in certain neighbourhoods (Manaugh & El-Geneidy, 2011).

This research project differs from those walkability studies highlighted above in that it attempts to amalgamate the rift between the built environment and the health of a population based on its access to physical activity. By doing so, this research aims to provide a more comprehensive analysis of how people have access to walkability in neighborhoods, rather than defining how neighborhoods sustain walking through their built environments.

BREAST CANCER SURVIVORSHIP

In a detailed study published in the American Journal of Lifestyle Medicine, physical activity is very clearly linked to an increased survival rate in post-menopausal women with breast cancer (Brunet, Sabiston & Meterissian, 2012). This study calls for “aerobic training at least 3 times/week for 30 minutes” where the study defines ‘aerobic activity’ as “walking, jogging, and running”. Although the study is aimed at providing information regarding realistic goals and recommendations of physical activity for breast cancer patients, it also informs this research project’s aim to connect neighborhood walkability to breast cancer survivorship. The studies also outline some of the benefits of physical activity in breast cancer patients, including “improved physical fitness, overall functioning, quality of life, and reduced fatigue”.

The concern with post-menopausal versus peri-menopausal breast cancer is highlighted in one study that highlights that survivorship is higher in women between the ages of 50 and 59 (Lagerlund et al, 2005) and that survival is higher amongst women with higher socioeconomic status. This would be a valid argument when one thinks about the access to medical services and higher quality of living that this demographic of women have access to. However, another study attributes a higher level of breast cancer occurence in women with a higher socioeconomic status, predicated by their income levels, as compared to women with a lower socioeconomic status (Borugian et al, 2011).

While these studies are paradoxical and oppose each other in nature, this research study attempts to measure, map and overly socioeconomic status on walkability analysis maps in order to comprehend the extent to which breast cancer survivorship can be spatially understood in comparison to the built environment and socioeconomic status concurrently. This study uses The Vancouver Area Deprivation Index (VANDIX) as a measure of socioeconomic status and overlays it on the various factors of walkability.

VARIABLE CHOICE

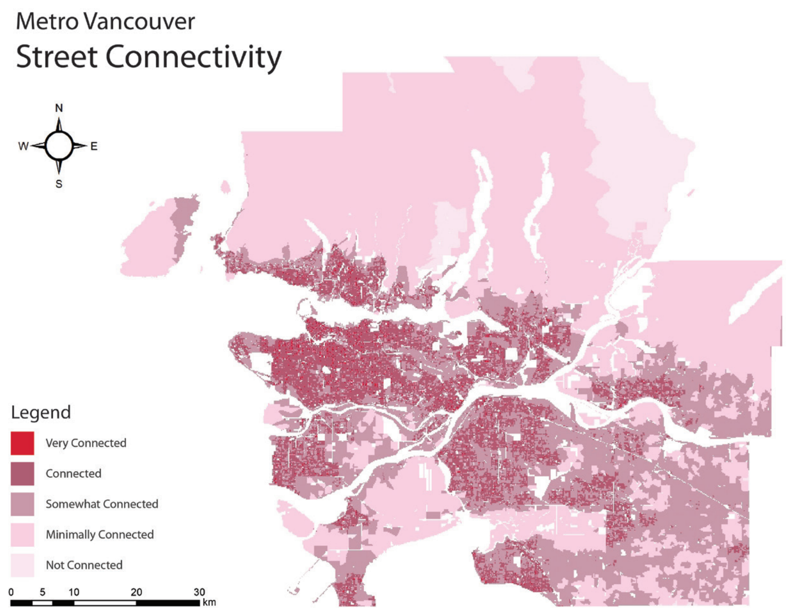
This study uses the following variables (Table 1) to create a walkability algorithm that understands the roles of the built environment, the natural environment, socioeconomic status and perceptions of safety in creating scenarios of access to physical activity, or, walkability.

Table 1. Chart of variables used in walkability algorithm

Variable	Reason	Data used	Data source
Street connectivity	Street Connectivity is associated with increased walkability (Kostova, 2011; Ozbil, 2010; Swinkels & Mu, 2014)	Street Network File Postal Code Data	Statistics Canada (2010) Statistics Canada (2006)
Land use mix (LU)	Increased land use mix and diversity is associated with lower body mass index (BMI), when considering residential diversity, commercial diversity and institutional land, not including rural and industrial land use mix (Smith et al, 2008).	Metro Vancouver Land Use data. The Forward Sortation Area (FSA)	Metro Vancouver (2003) Statistics Canada (2006)
Proximity to transportation	Closer proximity to public transportation is associated with increases in walking as a form of commuting (Bauman et al, 2009; Swinkels & Mu, 2014).	Street Network Dataset (created from Street Network File) Bus Stop Locations Excelsheet Postal Code Data	Translink (2003)
Proximity to parks	Some studies have found that closer proximity to parks is associated with lower body mass index (BMI) (Kostova, 2011; Swinkels & Mu, 2014).	Street Network File Park Location Data Postal Code Data	Clipped from the land use dataset using only the 'recreation and protected natural areas' category.

Variable	Reason	Data used	Data source
Population density	Increased population density is related to the compactness of an area and increased proximity to commercial outlets and destinations as compared to proximities created by urban sprawl (Kaestner & Zhao, 2011; Kostova, 2011; Bauman et al, 2009).	Population Census Data Postal Code Data	Statistics Canada (2006)
Slope	Slope is related to safety and accessibility, for example for wheel chair accessibility, strollers, etc. (Hamer, 2014).	Digital Elevation Model (DEM) of GVRD, derived at 25 meters (TRIM) LMAS 1.64 X 10 meters Postal Code Data	GeoData BC (2013)
Socioeconomic status (SES) and perceptions of safety	SES studies are inconsistent in regards to overall physical activity of residents (Bauman et al, 2007); however, we included SES data by overlaying VANDIX data on one of the analyses completed. Positive perceptions of one's own neighbourhood is associated with increased walkability (Slymen D, 2007).	VANDIX Data, which was originally collected from the 2006 Canada Census data and combined several surveys together. Postal Code Data	Bell, N., PhD., & Hayes, M. V., PhD. (2012)

INDIVIDUAL ANALYSIS



Map 1. Street connectivity

Connectivity

The 'Street Network' dataset was given a 2 meter buffer to ensure roads that bordered the postal codes boundaries were included in the connectivity of both postal codes. The data was then spatially joined and the total lengths of roads were summed and divided by the area of the postal code, resulting in a connectivity in proportion to the area of the postal code. The data was then classified and ranked into 5 ranks of connectivity: 1 representing very connected/walkable and 5 representing not connected at all/not walkable. The steps to analyze connectivity were coded as follows:

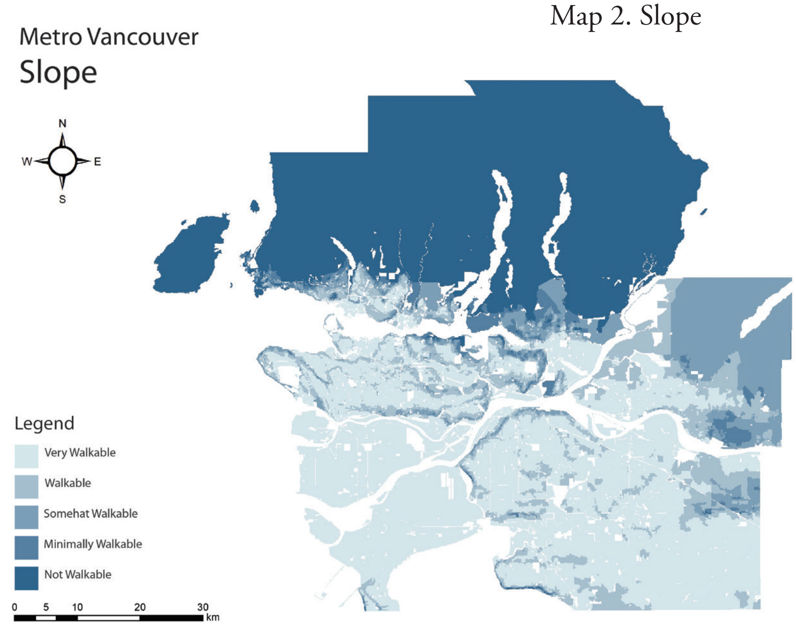
- Start with '6-digit postal code' dataset.
- Add a 2 meter buffer to the 'Road Network' dataset.
- Run Spatial Join to join both datasets and count roads per 6-digit postal code.
- Create a proportion based on number of roads per area.
- Normalise data to avoid skewness.
- Rank from 1 to 5

Slope

In order to accurately convert the DEM slope dataset from raster to vector several steps had to be followed that aren't listed here in detail. However, in order to maintain a high standard of data preservation, a spatial join was run in order to average the slope per postal code boundary. The steps to analyze slope were coded as follows:

- Start with DEM of Metro Vancouver.
- Create a zonal statistics map using the vector 6-digit postal code boundaries.
- Use the INT function to convert the float data attributes into integer data attribute
- Convert the zonal statistics raster map into vector map
- Run Spatial Join to join postal code data and the zonal statistic data
- Rank from 1 to 5

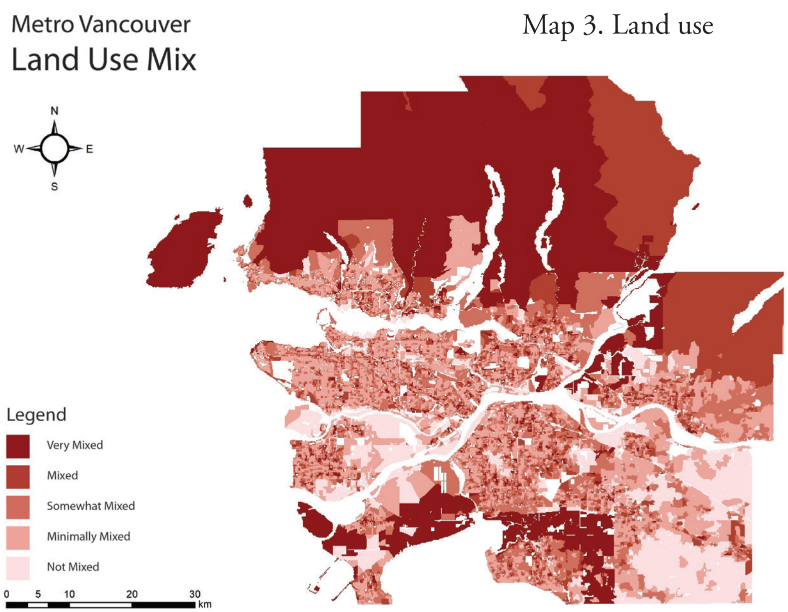
Map 2. Slope



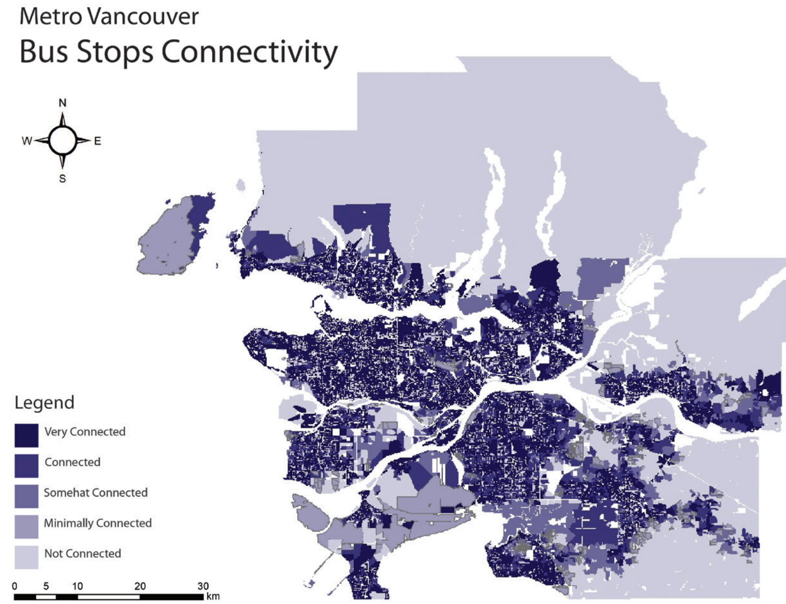
Land use

The land use data collected from Metro Vancouver included land use mix metadata that was useful to walkability research. The 16 land use classes included in the data were narrowed down to include the following: commercial, all types of residential uses, commercial-residential mixed, institutional, recreation and naturally protected areas, lakes and water bodies. The steps to analyze land use mix were coded as follows:

- Start with the Metro Vancouver 'land use mix' dataset.
- Dissolve the land use areas together.
- Run a spatial join to join land use with FSA boundaries
- Create a count of land use categories per FSA
- Rank from 1 to 5.



Map 3. Land use



Map 4. Proximity to transportation

Transit stops

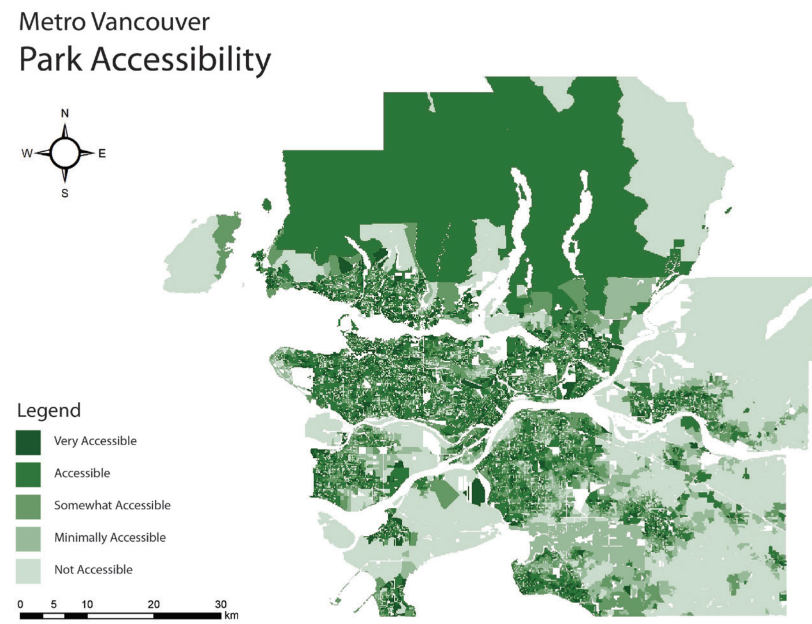
In order to create a route from the centroid of a 6-digit postal code to the nearest transit stop, a network dataset needed to be created from the 'Road Network' dataset. The network analyst tools were able to accurately draw routes with a buffer specification of 2km for this analysis. The resulting distances were then ranked from 1 to 5: 1 represented a short distance/walkable while 5 represented a long distance greater than 2km/not walkable. The steps to analyze connectivity to transit stops were coded as follows:

- Start with 'road network' dataset.
- Create network dataset from 'road network' dataset.
- Run 'closest facility' network analysis from centroid of postal codes to the nearest transit stops and measure distance.
- Multiply distance by walk speed using field calculator
- Rank from 1 to 5

Population density

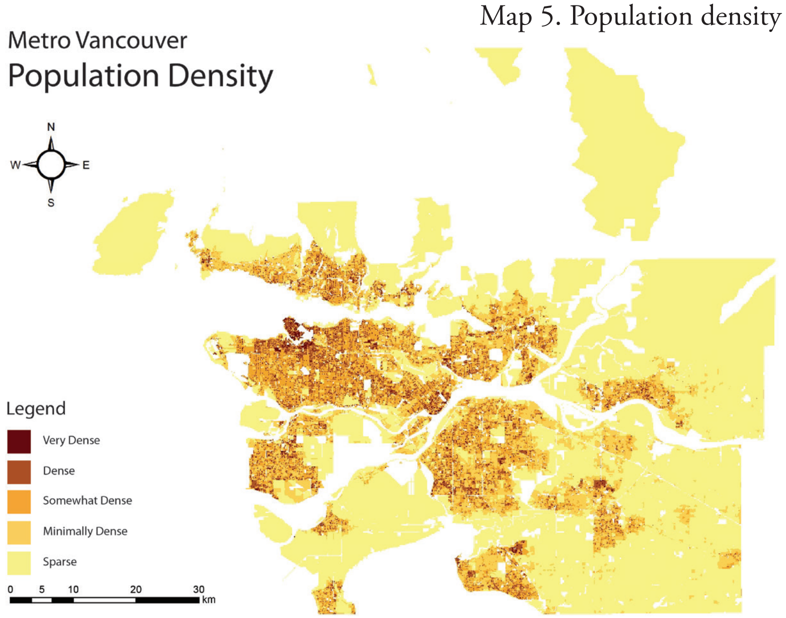
The smaller the area, the more chances that density measures can get skewed because density is a function of area size. Therefore, when measuring population density of 6-digit postal codes, the measures had to be normalised. The steps to analyze population density were coded as follows:

• Start with 'Census population' dataset.
• Clip the dataset to 6-digit postal code boundaries.
• Spatially Join the two data sets together
• Calculate density by dividing the population by areas.
• Normalise data.
• Rank from 1 to 5.



Map 6. Proximity to greenspace/parks

Map 5. Population density



Distance to parks

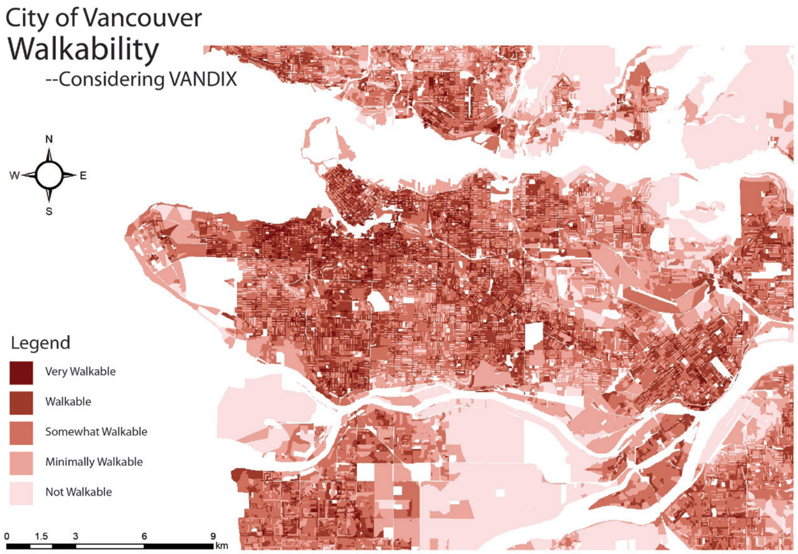
Very similar to the distance to transit stops, the distance to parks was analyzed using the 'Closest Facility' Network Analyst. The steps to analyze distance to parks were coded as follows:

• Start with 'land use' dataset.
• Crop 'recreation and protected natural areas' from 'land use' into a new shapefile called 'parks'.
• Run 'closest facility' layer using the previously created 'road network' from centroid of postal code to nearest park and measure distance.
• Convert distance to cut-off value using field calculator.
• Rank from 1 to 5.

VANDIX

VANDIX data represents socioeconomic data in our research and was joined to the 6-digit postal code data. There are 21 Variables for measuring VANDIX which were all relabeled into 7 categories and joined to the 6-digit postal code data. The data was finally ranked from 1 to 5 in order to assign values representing walkable and non walkable.

Measurement of VANDIX is primarily focused on the correlation between material conditions and social factors that relate to health concerns. However VANDIX ignores the diversity of social factors which may provide alternative views about material conditions (Hall et al, 2012).



Map 7. City of Vancouver, considering VANDIX

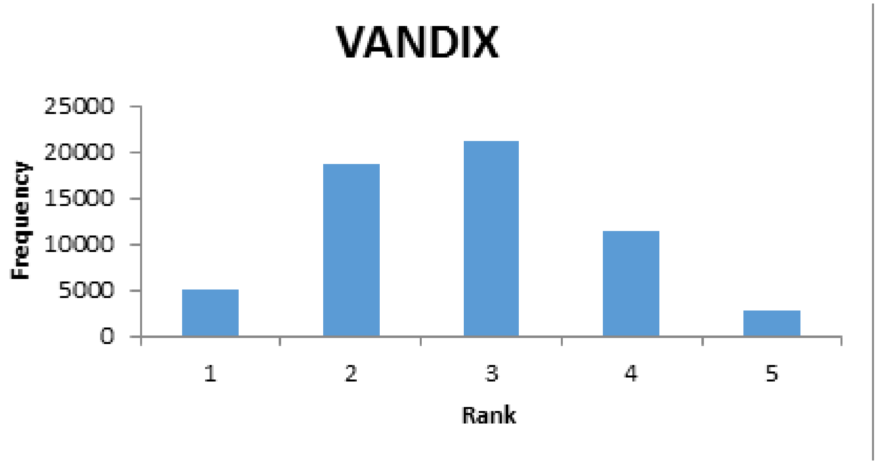


Diagram 1. Histogram showing distribution of VANDIX across all 5 ranks.

RESEARCH STRATEGIES

The purpose of this research project is to determine whether areas of high walkability that facilitate greater physical activity have a positive correlation with breast cancer survivors. The research project measures access to physical activity by calculating walkability in 6-digit postal codes in Metro Vancouver. Socioeconomic data, represented by VANDIX, is overlayed on walkability data and finally compared to Google Map Walk Scores to provide a comprehensive analysis of walkability measures that are considerate of health concerns, particularly access to physical activity for breast cancer patients. The research goals are the following:
- to comprehend the list of factors, qualitative and quantitative, that contribute towards the walkability algorithm
- to map areas that facilitate versus areas that hinder access to physical activity in Metro Vancouver, thereby creating risk scenarios
- to analyze the influence of socioeconomic status on walkability scores
- to compare and contrast our walkability scores with those of Google Maps.

PROJECT STEPS

The direction of the project (conducted over the course of 10 weeks) was separated into the following steps:
1. Vision and goal setting of research
2. Background research
3. Agorithm design
4. Data collection
5. Primary data analysis
6. Field study
7. Secondary analysis

Step 1 and 2: Vision, goals and background research
To eliminate bias in the design and analysis of the project, the team collectively reviewed and analyzed the data and discussed results. A new algorithm needed to calculate a new walkability index needed new variables determined according to the goals of the research.
The two concepts of walkability and breast cancer were separated to dig deeper into the particularities of each that would contribute to this research. Walking was broken down into two purposes: walking for leisure and walking for commuting or walking.

Studies show that the most physically active adults in walkable neighborhoods walk to commute compared to those living in lesser walkable neighborhoods (University of British Columbia, 2010). Furthermore, parks and open spaces encourage this form of walking greatly (University of British Columbia, 2010). Another study shows that measures of walkability must consider street connectivity, residential density and access to public modes of transportation (Carr et al, 2010). In turn, greater residential density and street connectivity support an increased use of public transportation (University of British Columbia, 2010). Studies show that a greater density of street intersections and a smaller average block length encourages walking as a form of commuting as well (Carr et al, 2010). Based on these articles, the following variables were used to calculate our walkability:

1. Land use mix: A neighbourhood with mixed land use encourages walking as a form of commuting.
2. Population density: Greater residential density supports higher use of public transportation and therefore results in more walking.
3. Road connectivity: Smaller block lengths, greater street connectivity result in increased walkability scores.
4. Distance to transit stop and parks: Closer distance to public transit stop and parks induce more walking activity.
5. Slope: People are more willing to walk on a flat street.

Step 2: Algorithm design

Once the variables were chosen, a multi-criteria evaluation (MCE) was chosen as the best GIS analytic tool to calculate a walkability score. An MCE is a fundamental step of the rational decision-making process especially in cases when multiple variables must be evaluated together (Klinkenberg, 2007). In this MCE each variable was standardized, given a weight according to its relational importance, and finally, all variable weights were combined to generate a final result. The weights were generated based on responses from client surveys and group discussion. The final result of our algorithm is a walkability score of Metro Vancouver Area with 1 representing a more walkable area and 5 representing a less walkable area.

Step 3: Data collection and Data analysis

Each group member was assigned a section of the data collection and analysis section of the research. Data was collected and stored on a common folder on a cloud sharing service, Dropbox, in order to facilitate ease of access for all researchers. A detailed commentary on the data analysis will be provided in the 'Data' section of this report.

Step 4: Field Study

After a primary round of data analysis, a field study was conducted to assess the accuracy of the results, and to better inform a secondary analysis of the walkability variables. One of our primary research goals was to assess walkability and its relationship to breast cancer survivorship which directed a research consideration of socio-economic status (SES) in the secondary analysis.

People living in neighborhoods with a higher socioeconomic status may pay more attention to their heath and have a better access to resources in order to mitigate health concerns. Socioeconomic status also influences peoples perceptions of neighborhood safety in relation to walking (Wilson et al, 2004) and thus has a positive correlation to walkability (Carr et al, 2010).

In this section of the research, the Vancouver Area Neighbourhood Deprivation Index (VANDIX) was used to visualize the general socioeconomic status of people in Vancouver. VANDIX uses Canada Census data combined with sample survey to evaluate people's social and economic condition. It uses 21 variables in 7 major aspects to generate an index of the following variables: proportion without high school completion, proportion without university completion, unemployment rate, proportion of lone-parent families, average income, proportion of home owners, and employment ratio.

One postal code was randomly selected per rank (1 to 5) of the walkability score and each was visited to evaluate the correctness of each variables result in that postal code. The field study also elaborates on the safety of the selected environment for the purposes of walking. By comparing the results of the field study and the VANDIX report, this research was able to consider the accuracy of VANDIX in determining the socioeconomic status of certain postal code areas. Finally, the results of the primary analysis was adjusted and a secondary analysis was conducted.

PRIMARY ANALYSES

Multi Criteria Evaluation Steps
· Normalize values
· Assign Rank (1 to 5)
· Assign weights

Normalize Values
After the data for all the factors where normalised using Log to ensure the data was normally distributed and thus not skewed each factor was ranked one to five.

Ranks
Increase land use diversity is linked to increase walkability (Brown, Fan, Kowaleski-Jones, Smith, Yamada and Zick,2008).. Numbers 7 and 8 represent that there are a total 7 or 8 different land classes, not including rural, industrial or streets, suggesting increased land use mix and thus ranked 1 for most walkable and a land use mix of only 1 to 2 is not very well mixed, thus receiving a rank of 5.

Proximity to transportation increases the walkability of a place (Bauman, Coffee, Frank, Hugo, Leslie and Owen, 2009; Swinkels and Mu, 2014 and Ozbil A, 2010). Closer the proximity to transit, the more walkable an area is for walking as a form of commuting; thus proximity with 200 m is rank 1, as most walkable and 1000+ m is ranked 5 for least walkable.

According to walkability studies, proximity to parks is an asset(Kostova, 2011; Ozbil, 2010; Stevens R, 2005); with proximities within 500 to 1000 meters being optimal (Stevens R, 2005). Walking to parks may be more relative for walking for leisure, thus willingness to walk further is factored in. A less then 500 m walk is ranked with a 1 and distances over 2000 m are ranked with a 5.
Increased population density is associated with increased walkability, thus high population density was given a value of 1 and low density areas were given a rank of 5 (Kaestner and Zhao, 2011; Kostova, 2011; Bauman, Coffee, Frank, Hugo, Leslie and Owen, 2009). The population density data was normalised using Log and fit a normal distribution. To fit the data into 5 ranks, natural breaks were used.

Increased connectivity I associated with increased walkability; there fore areas with more connectivity were given a value of 1 and decreased connectivity were given a value of 5 (Kostova, 2011; Ozbil, 2010; Swinkels and Mu, 2014).

Factor	1	2	3	4	5	Method
Land use	7 to 8	6	5	4 to 3	2 to 1	Manual
To bus stops	<200m	200-500m	500-800m	800-1000m	>1000m	Manual
To parks	<500m	500-1000m	1000-1500m	1500-2000m	>2000m	Manual
Population density	1.99	-0.57	-1.02	-1.48	-2.23	Standard deviation
Connectivity	> 0.77	-0.25	-1.13	-2.01	< 2.89	Standard deviation
Slope	0-3%	3-6%	6-9%	9-11%	>11%	Manual
VANDIX	-0.67	-0.30	0.0	0.38	2.10	Natural breaks

Table 2. The ranking system to organize the variables.

CLIENT CONSULTATIONS

Very early on in our project, we sought guidance regarding the context of the research project from Dr. Nadine Schuurman regarding the kinds of GIS analysis and ranking we should consider in this project. Given that walkability and breast cancer survivorship are such large topics of discussion, we acquired a breadth or scope of research that we would be able to conduct given our time limitations.

Dr. Schuurman was also requested to rank the walkability factors so that we could incorporate her views in our analysis (Table 3).

Dr. Schuurman's Weights	LU mix	Transit	Parks	Pop density	Connectivity	VANDIX	Slope	Totals	Weights
LU mix	-	x	x	x	x	x	.5	.5	.024
Transit	1	-	x	x	x	.5	1	2.5	.119
Parks	1	1	-	1	x	1	1	5	.227
Pop density	1	1	x	-	x	1	1	4	.238
Connectivity	1	1	1	1	-	1	1	6	.286
VANDIX	1	.5	x	x	x	-	1	.5	.024
Slope	.5	x	x	x	x	x	-	2.5	.119
							TOTAL	21	1.0

Table 3. Weights assigned to the variables by Dr. Nadine Schuurman

We also sought guidance regarding the context of breast cancer survivorship and its connection to walkability from Dr. Miriam Rosin. This would be the first step toward any future research that could be conducted to create algorithms for walk scores that consider health initiatives and walking as a form of commuting and leisure as well.

Dr. Rosin was also requested to rank the walkability factors so that we could incorporate her views in our analysis (Table 4).

Dr. Rosin's Weights	LU mix	Transit	Parks	Pop density	Connectivity	VANDIX	Slope	Totals	Weights
LU mix	-	1	1	1	1	1	1	6	.3
Transit	x	-	.5	1	1	1	.5	4	.2
Parks	x	.5	-	1	1	1	.5	4	.2
Pop density	x	x	x	-	.5	.5	1	2	.1
Connectivity	x	x	x	.5	-	.5	1	2	.1
VANDIX	x	x	x	.5	.5	-	x	1	.05
Slope	x	.5	.5	x	x	x	-	1	.05
							TOTAL	20	1.0

Table 4. Weights assigned to the variables by Dr. Miriam Rosin

FIELD STUDY

The purpose of this field study is to randomly screen the accuracy of the walkability score we computed through ArcGIS and to rank them from the perspective of our group to compare the differences or similarities between the two.

For the field study, we randomly selected ten postal codes as our samples using stratified sampling in Microsoft Excel. There are two strata: one is walkability including the VANDIX as a factor, and the other is walkability without VANDIX as a factor. Since our walkability scores are classified into 5 ranks, we picked one postal code from each of them. This process was repeated twice in order to ensure the randomness of the samples.

The samples selected are well distributed over the GVRD. Since the destinations are so dispersed, our information collected from the data can be less biased and more precise. Table 5 shows the list of the postal codes visited during our field study.

During the field study, walkability in each postal code was ranked based on observations and perspectives according to 7 factors of walkability. Table 4 shows the scores assigned to each postal code during this field study. These scores were then inserted into the walkability algorithm.

PC	LU	POP DEN	TO PARK	TO BUS	CONNECT	SLOPE	SAFETY
V7H 1N9	5	3	1	3	3	4	2
V7L 3K7	2	2	1	1	2	3	2
V5Y 2G7	2	2	1	1	1	2	1
V6B 2K4	1	1	1	1	1	1	3
V5R 1Y4	3	3	1	1	1	3	3
V3J 4J9	4	3	2	2	3	4	1
V3J 5X7	4	3	1	3	3	4	2
V3S 7M8	5	4	5	5	4	1	2
V3X 3N9	5	3	1	3	4	3	4
V4C 1L6	4	3	1	3	3	1	3

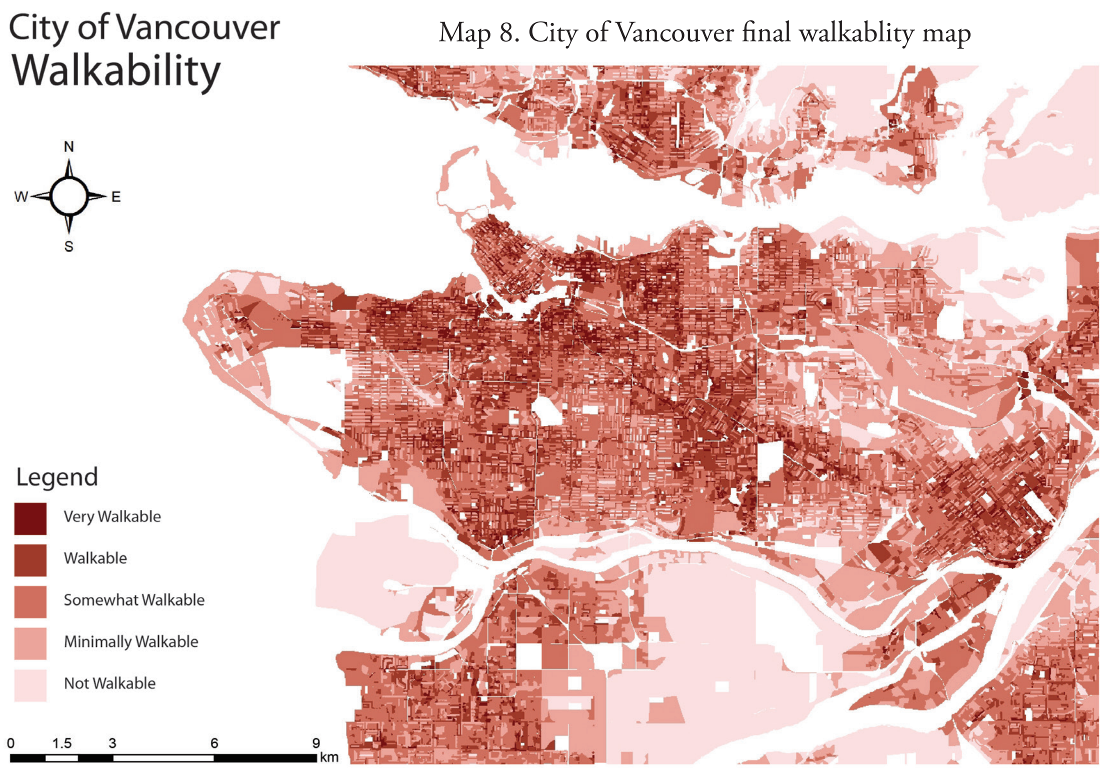
Table 5. Walkability factor rankings based on field observations.

SECONDARY ANALYSIS

Secondary analysis considered the following factors towards a final output map:
- the walkability factors (land use, slope, distance to parks, distance to bus stops, connectivity and population density)
- soecioeconomic status (measured by VANDIX)
- weighting and ranking of the above two factors by Dr. Nadine Schuurman and Dr. Miriam Rosin

The final map (Map 8), although not considered so according to the research, is but an initial attempt to understand the factors contributing to walkability that aren't measured by Google Walk Scores or the Walkabliity Index. While it only shows the City of Vancouver due to problems regarding the proper visualization of data, it can be said to represent the data of Metro Vancouver to a great extent. The algortihm created faced several challenges, namely those regarding skewness of data in both VANDIX and client ranking. These will be discussed in the following sections of this report.

Map 8. City of Vancouver final walkablity map



FIELD STUDY RESULTS

During the field study, walkability in each postal code was ranked based on observations and perspectives according to 7 factors of walkability. Table 4 shows the scores assigned to each postal code during this field study. These scores were then inserted into the walkability algorithm. The results were compared with the walkability values originally generated in the ten postal codes in ArcGIS (diagram 2).

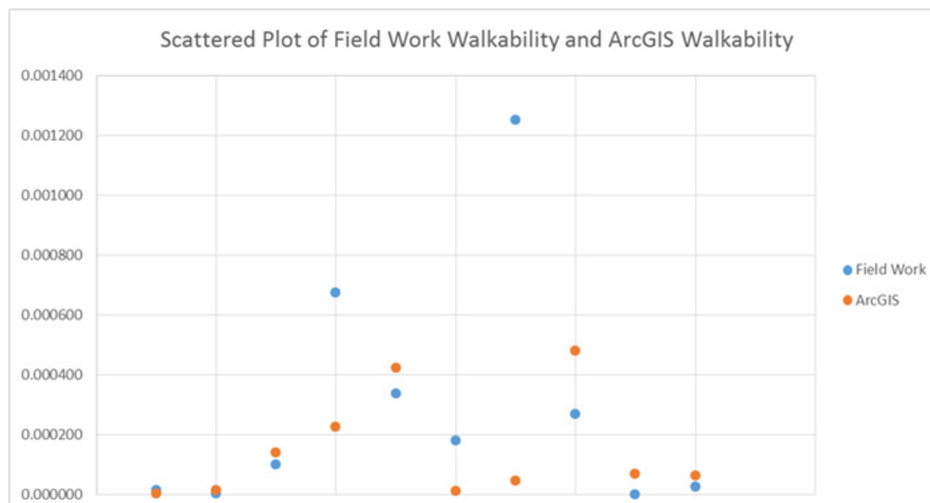
The scatterplot shows the ten postal codes plotted according to their scores both from the fieldwork and the algorithm. The higher the score, the less walkable the area. V3S 7M8 is an outlier we ranked as not very walkable; however, the lab model data shows this area as very walkable. After analyzing a satellite image of the same area in Google Earth we found a commercial complex and a park very near the postal code that was not noticed due to large houses and trees blocking the view (Map 9). Despite this outlier, the results of the field study approximately align with our walkability algorithm's result.

The research team observed that group members who lived in more rural areas deemed some areas more walkable and safer then group members who lived in more walkable urban areas. Thus, varying perceptions on urban/rural differences should be considered; walkability and safety measures may benefit from dividing urban and rural areas and using different factors to rate walkability and safety. The research team concluded that comparatively rural areas felt less safe at night as compared to more urban areas. The most observable reasons for this conclusion were the lack of street lights, side walks and neighbour surveillance.

After investigating the postal codes associated with the skew portion of walkability distribution we could see that the postal code areas were very large in comparison to the other, more walkable postal codes as well as lower population densities.

Map 9. A Google Earth capture of the postal code V3S 7M8 (Surrey).

Diagram 2. Scatterplot comparing field study and ArcGIS walkability algorithm.



VANDIX RESULTS

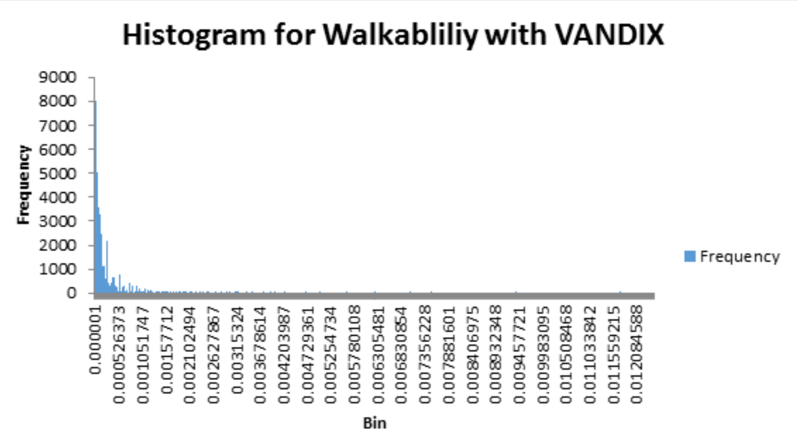


Diagram 3. Positive skew of walkability and VANDIX.

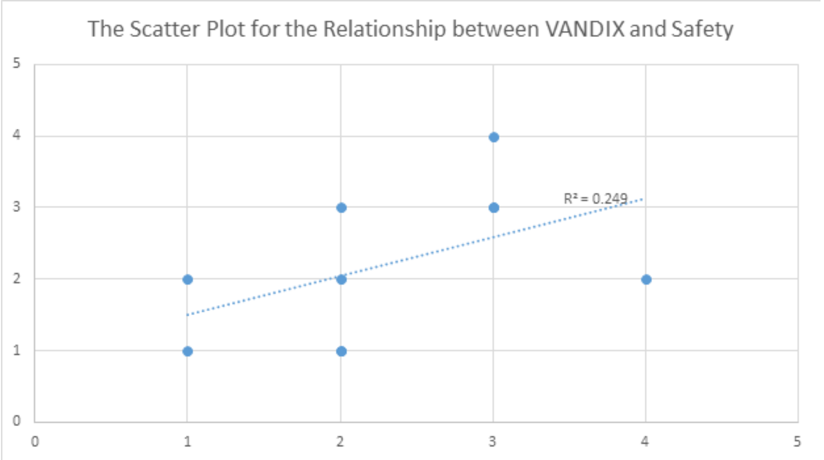


Diagram 4. Correlation between safety and VANDIX (r = 0.499.

The field study also aimed to study the relationship between VANDIX and safety in postal codes. Therefore, we included "safety" in our categorization for the grading scheme. Subsequent research attempted to find whether there was a correlation between our perceptions of safety and VANDIX.

To normalize data that was positively skewed when all factors were multiplied, data was logged. Further, to understand why all the data was positively skewed regardless of the allocated weights, the tail portion of the distribution was investigated (refer to Map 8); postal codes that encompassed the tail portion of the distribution, representing the least walkable areas, were extracted using natural breaks and then imported into Google Earth along with the corresponding database file in order to examine the different factors ranked scores associated with the skewed area. After examining the areas in Google Maps it became apparent that the areas in the skewed portion lacked any housing development or roads and appeared to be heavily forested.

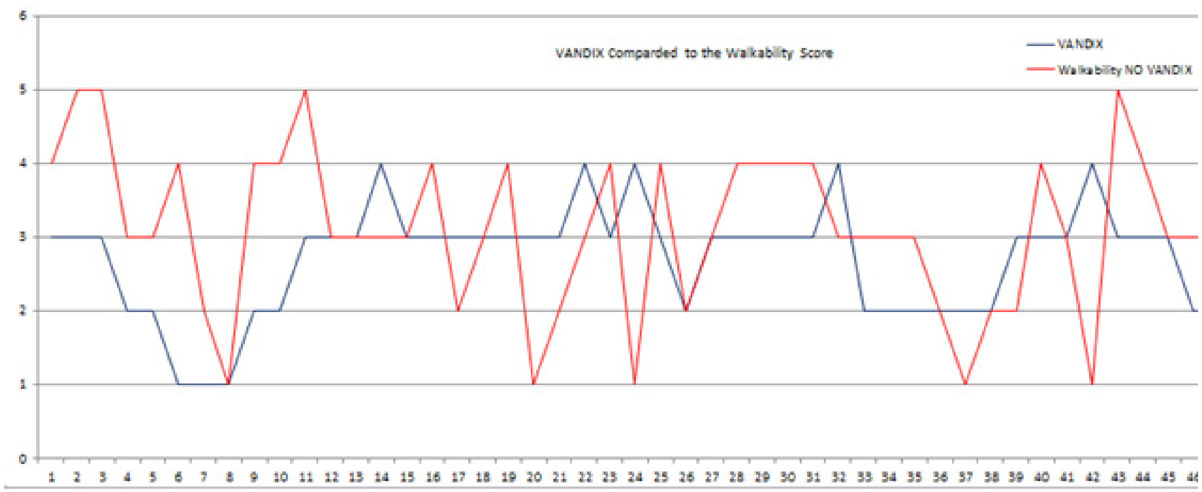
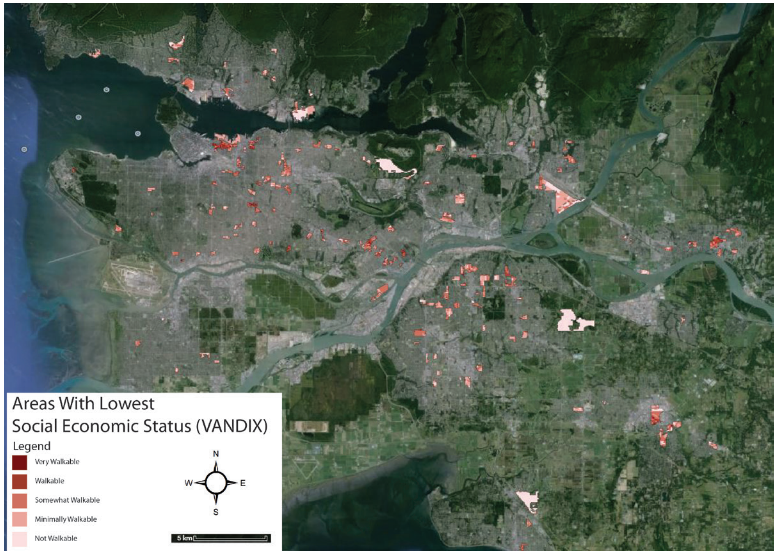
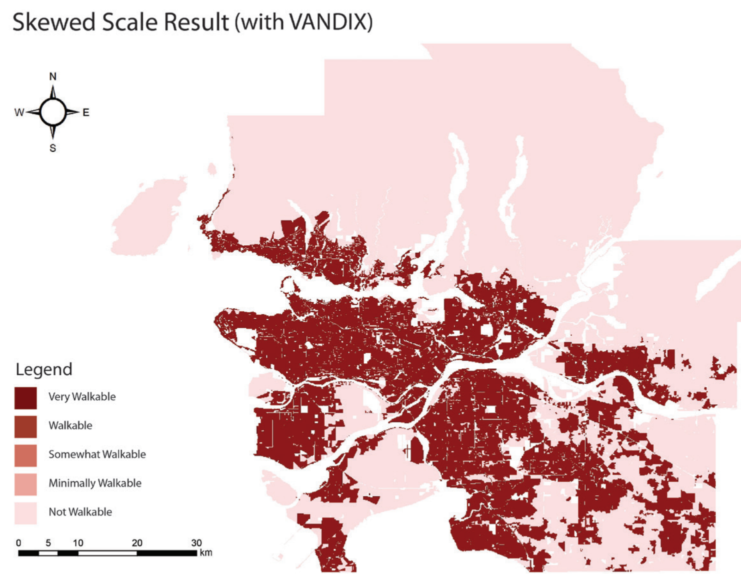


Diagram 5. Correlation between walkability and SES.

VANDIX RESULTS



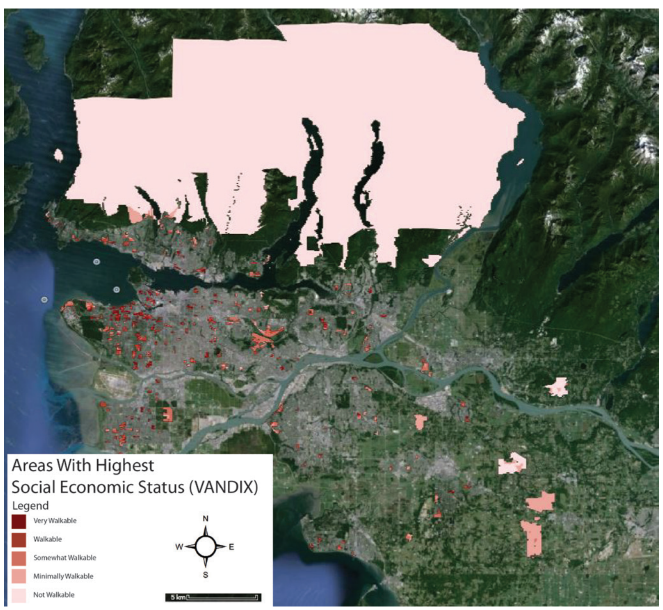
Map 11. Low SES areas might be closer to walkable areas.



Map 10. Skewed result of walkability using VANDIX as a factor.

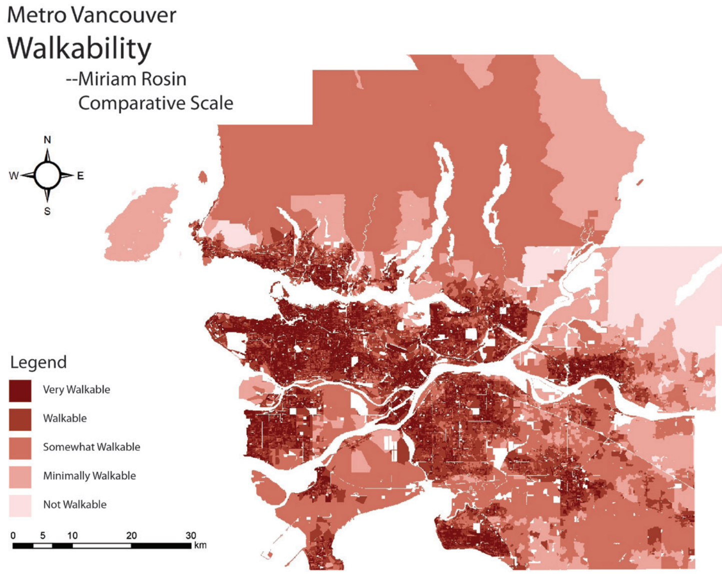
Another interesting comparison of results, using 50 randomly selected postal codes from our data, is the comparison of the VANDIX ranked score and the Walkability Index, which did not include VANDIX in the algorithm. Diagram 6 shows little correlation between increased walkability and high SES (derived using VANDIX).

To further compare VANDIX to walkability, the distribution of high socioeconomic status areas was considered and ranked 1 for the VANDIX factor in the algorithm, against the walkability map. It is interesting that some of the most affluent areas were in the least walkable areas. Some of these areas have a very low population density and may cause some misrepresentation in the VANDIX scores. Furthermore, when analysing areas with low SES in comparison to the standard walkability score, it was clear that lower SES areas are sometimes closer to more walkable areas.

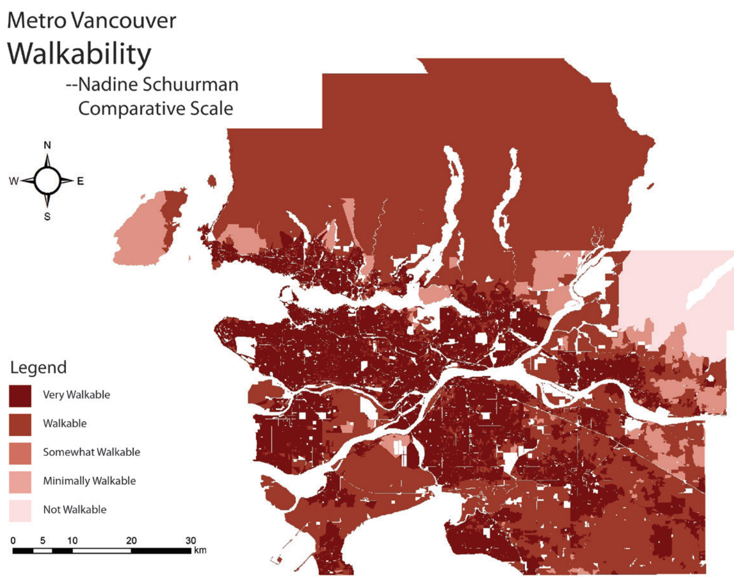


Map 12. High SES areas might be large estates or parks.

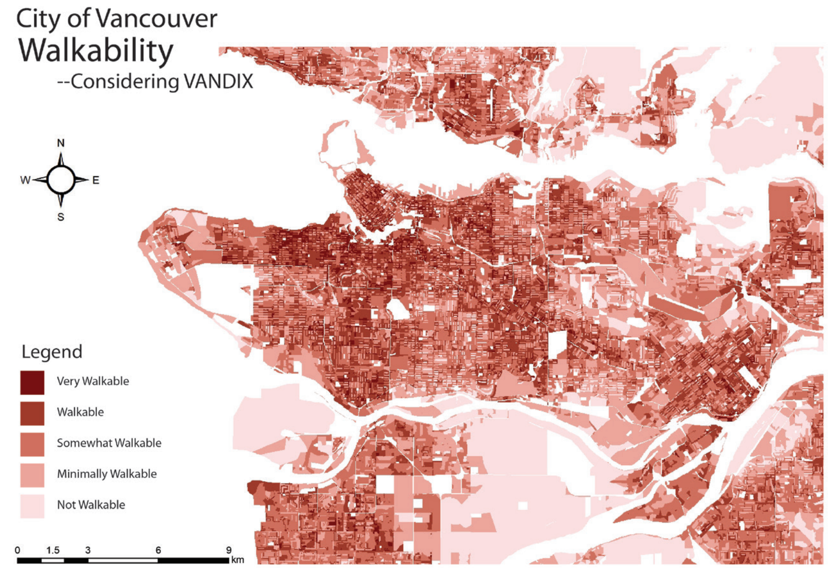
CLIENT RESULTS



Map 13. Miriam's weights spatially displayed in Metro



Map 14. Nadine's weights, also called the 'reference map'.



Map 15. Standardized walkability map with VANDIX overlay.

After the data was logged to correct for the distribution for the positively skewed data, we then classified the logged values into 5 classes based on the standard deviation of the reference map of walkability. Maps 13, 14, and 15 show ranked walkability results.

To understand why the ranked data appears to be skewed for Nadine and Miriam's weights, the ranking distribution was visualized for some of the factors that were weighted heavier than in our weights using histograms (refer to website). Land use and slope are positively skewed; proximity to parks and VANDIX are somewhat positively skewed; proximity to bus stops is somewhat negatively skewed. These variances may contribute to the variation in the different client analyses. For example, Miriam weighted land use fairly high but also weighted proximity to bus stops at the same weight of .2 (refer to Table _, page __). The extra weight on bus stop proximity may explain why her data is less positively skewed then Nadine's. Nadine weighted population density, parks and connectivity fairly high, in which populaiton density and connectivity have very little representation in the 5th rank of "not very walkable."

GOOGLE WALKSCORE

Secondary analysis also compared the final analysis of the walkability score with the Google Map Walk Score, as well as against the analyses that use Miriam and Nadine's weights. Google Map Walk Score uses an algorithm that is primarily based on proximity to any commercial building and does not consider the type of commercial outlet within proximity. For example, a postal code may be located close to commercial stores but would be allocated a fairly high walkability score, when in fact they are only located close to gas stations and auto malls.

Table 6 compares the three scores after they are classified into 5 ranks; 1 represents very walkable while 5 represents less walkable. It is visible in this graph that the different analyses of walkability are very similar. A larger field study of this nature would be benefit the validity of our small sampled field study.

Miriam	Nadine	no VANDIX	with VANDIX	Google Walk Scores
1	1	1	1	1
2	2	2	2	1
4	3	4	4	4
3	3	3	4	3
4	3	5	4	2
1	1	2	2	3
2	2	3	3	4
4	3	4	4	3
3	3	3	3	5
3	3	3	3	4

Table 6. Comparison of the 2 ranked scores from our clients, VANDIX scores, and Google Walk Scores.

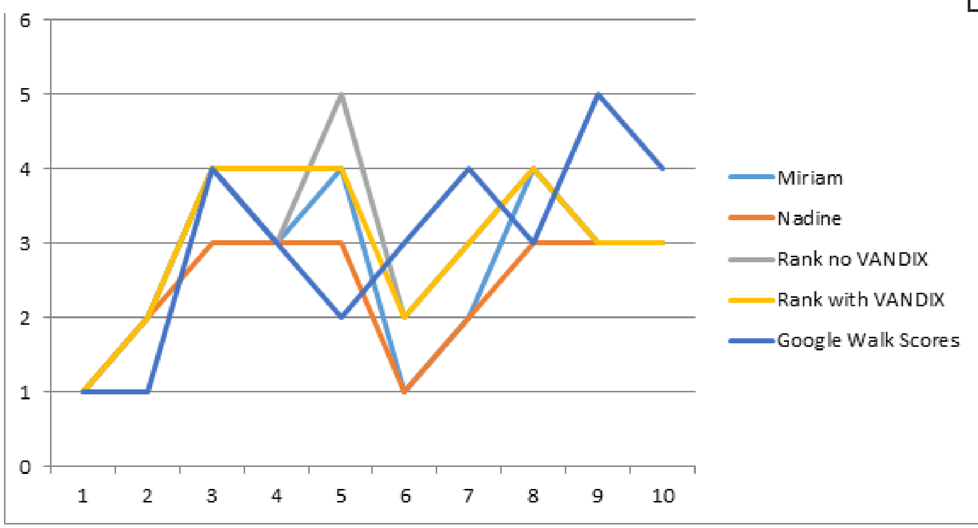


Diagram 6. Chart representing scores shown in Table 6.

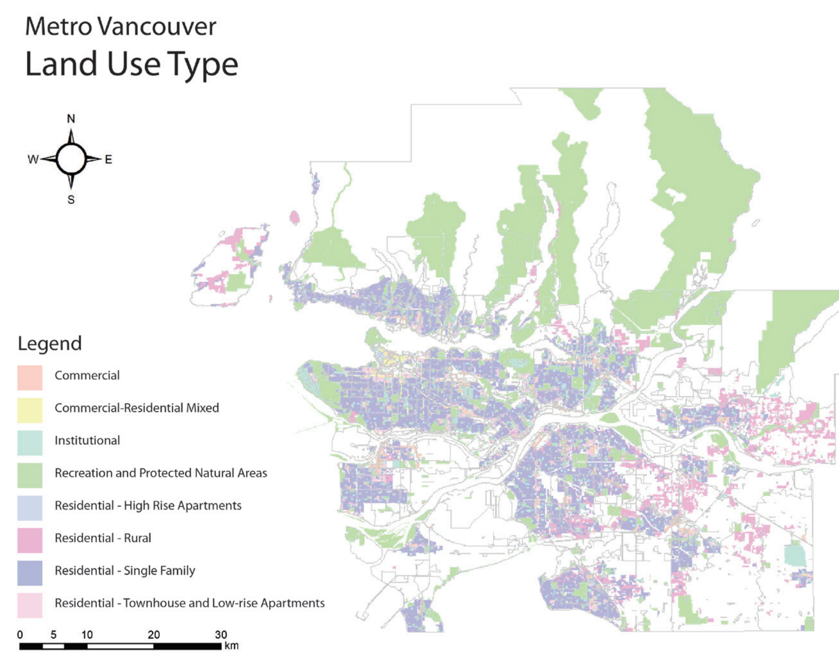
DATA LIMITATIONS

The 6-digit postal code was chosen as the foundation layer of walkability analysis for the purposes of this project because it accurately represents the socioeconomic information to the finest detail without crossing any privacy concerns of individuals. Although there were several occasions throughout the analyses where 3-digit postal code analysis would have been more efficient, such as when considering land use mix or population density, the research consistently used 6-digit postal codes. Using 6-digit postal codes also allows future analysis the ability to overlay breast cancer records, when acquired, on the current analysis and socioeconomic data.

Postal code data was collected from the Simon Fraser University GIS Data Warehouse, but originates from Statistics Canada (StatsCan, 2006). The 6-digit postal code data contained several duplicate postal codes because the postal code data was originally created so that Canadian Census data could correspond with the areas. The 'Postal2Census' tool from the Census of Canada was used to associate historic census geography boundaries (UEP) with Canadian postal code points (MEP) (StatsCan, 2006). In this way, all postal code areas with the same postal code name were merged together to eliminate any duplicate entries.

There were several problems associated with the use of 6-digit postal codes while conducting analysis. Firstly, postal code boundaries are based on population density and the volume of mail the residents receive (Canada Post, 2014). Therefore, some high population density areas have very small postal code areas while some low population density areas have very large postal code areas.

In this way, the varying size of postal code boundaries becomes problematic and skews the individual results for some factors. Thus, one of the problems with using 6-digit postal code boundaries is accurate representation of the spatial patterns of certain factors. For example, the factor 'mixed land use' is not well represented by the 6 digit postal code because even if a 6-digit postal code is surrounded by many different land use types (Map 16), the postal code might only contain one or two land use types within it. Therefore, a small spatial unit of representation can be limiting when considering a practical problematic such as walkability which inherently considers connection to destinations that might be outside of these restrictive boundaries.



Map 16. Land use map representing the various land use types.

DATA LIMITATIONS

Furthermore, large areas, such as postal codes VON 2E0 and VON 3Z0, receive a high score of land use mix. Although it is extremely clear that these areas have a very homogenous land use of protected natural areas, small urban areas within these, such as Lions Bay, can skew the data.

The Forward Sortation Area (FSA) is represented by the first three digits of the postal code and encompasses all areas represented by the last three digits of a postal code as well. For the specific purposes of this study, our research showed that using FSAs as a foundation layer would be more appropriate as it allows those areas with higher population density to consider all land use mixes within close proximity to individual 6-digit postal code boundaries rather than those land uses that only occur within the 6-digit postal code boundary.

Through further visual analyses it became apparent that there were areas missing representation in the postal code dataset and thus appeared blank. In order to explore this issue, the postal code shape file was converted into a Keyhole Markup Language (KML) file and then loaded into Google Earth. Through the visualization of the vector postal codes onto a remotely sensed image of the area, the reasons behind these misrepresentations were identified. Table 1 may help explain why some of the postal codes are misrepresented; large areas with a land use type of 'Recreation and Protected Natural Areas' are usually represented by an FSA only.

Table 7. Missing postal code areas and corresponding land use.

6-digit postal codes	Points of lcoation representing postal code	Type of land use
V7X 1M3	Bentall Centre, Vancouver	Commercial
V5A	Burnaby Lake Park, BurnabY	Lakes and Water Bodies
V5A 4M8	Burnaby Mountain Golf Course, Burnaby	Recreation and Protected Natural Areas
V3C 6M3	Colony Farm Regional Park, Port Coquitlam	Recreation and Protected Natural Areas
V5H 4R4	Deer Lake Park, Burnaby	Lakes and Water Bodies
V5J 1A3	Central Park, Burnaby	Recreation and Protected Natural Areas
V5Z 4A2	Charleson Park, Vancouver	Recreation and Protected Natural Areas
V5G 1C7	Intersection of Willington Ave and Highway 1, Burnaby	Transportation, Communication and Utilities
V3J	Mundy Park, Coquitlam	Recreation and Protected Natural Areas
V5A 1S6	SFU & Burnaby Mountain, Burnaby	Recreation and Protected Natural Areas

FUTURE STUDIES & POLICY IMPLICATIONS

When considering the data in this report, it is important to remember that these walkability scores are an attempt to address policies regarding physical activity in neighborhoods. This study hopes to proove some of the assumptions one can make regarding socioeconomic status and access to physical activity. There are several stakeholders for whom this report can be of use, including, public health policy makers, GIS analysts and urban or neighbourhood planners.

Policy makers will be able to find that many of the factors contributing to walkability are of political importance and can be tuned with policy creation and implementation. While individual cities and townships within Metro Vancouver have their own health strategies, there is yet a comprehensive Metro Vancouver health strategy to exist. The strategy of most relevance to this research is the Metro Vancouver Sustainable Region Initiative.

GIS analysts and urban planners will find a large amount of information on how various components of the built environment and people interact with each other. By attuning many factors of the built environment to positively influence population health, urban planners have the power to affect change and create healthier communities. The best example of this is the City of Vancouver's Healthy City Strategy.

Some considerations for future research are outlined in the following paragraphs. Firstly, this research connected VANDIX to walkability and whether socioeconomic status influences access to physical activity or not. Something for future consideration would be to connect not only these two factors to each other, but to overlay a connection between VANDIX and breast cancer survivorship and whether socioeconomic status has a positive or negative spatial connection to breast cancer survivorship.

Another consideration would be a more thorough analysis of safety. As our research suggests, safety might prove a very important factor to the analysis of access to physical activity after all. Safety is currently the missing piece of the puzzle to analyzing subjective concerns to walkability scores in neighbourhoods and perhaps one of the biggest factors, the analysis of which, could greatly change the way walkability scores are measured.

Although our research initially studies the difference between walking for leisure and walking for commuting, our analysis was unable to separate the two due to unavailability of accurate census data or physical activity data for Metro Vancouver. A future study would acquire and analyze the influence of these two modes of walking on varying degrees of access to physical activity as well as the duration and time of the day when such physical activity is common in breast cancer patients.

REFERENCES

Bauman A., Coffee N., Frank L., Hugo G., Leslie E & Owen N., (2007). Neighborhood Walkability and the Walking Behavior of Australian Adults. American Journal of Preventive Medicine 33 (5), 387–395

Bauman, A., Coffee, N., Frank, L., Hugo, G., Leslie, E., & Owen, N. (2009). Mixed Land Use and Walkability: Variations in Land Use Measures and Relationships with BMI, Overweight and Obesity. Health & Place 15 (4), 1130–1141.

Bauman, A., Coffee, N., Frank, L., Hugo, G., Leslie, E., & Owen, N. (2005). Walkability of local communities: Using geographic information systems to objectively assess relevant environmental attributes. Health and Place 13 (1), 111-122.

Bell, N., PhD., & Hayes, M. V., PhD. (2012) VANDIX [Data file]. Retrieved from Blake Walker, MA Student, Department of Geography, Simon Fraser University.

Bell, N., PhD., & Hayes, M. V., PhD. (2012). The Vancouver area neighbourhood deprivation index (VANDIX): A census-based tool for assessing small-area variations in health status. Canadian Journal of Public Health, 103, S28-S32A. Retrieved from http://search.proquest.com.proxy.lib.sfu.ca/docview/1269153595?accountid=13800

Brunet, J., Sabiston, C., & Meterissian, S. (2011). Physical Activity and Breast Cancer Survivorship: Evidence-Based Recommendations. American Journal of Lifestyle Medicine 6 (3): 224-240

Borugian, M., Abanto, Z., Spinelli, J., Xu, C., & Wilkins, R. (2011). Breast cancer incidence and neighbourhood income. Statistics Canada Health Reports 22 (2), 1-8.

Canada Post. (2014, January 13). Addressing Guidelines. Retrieved from http://www.canadapost.ca/tools/pg/manual/PGaddress-e.asp#top

Carr, L. J., Dunsiger, S.I., & Marcus, B.H. (2010). Walk ScoreTM As a Global Estimate of Neighbourhood Walkability. American Journal of Preventative Medicine 39 (5), 460-463. Density, Design, and New Diversity Measures. American Journal of Preventive Medicine 35 (3), 237-244.

City of Vancouver. (2014). A healthy city for all: Vancouver's Healthy City Strategy 2014-2025. Retrieved from https://vancouver.ca/people-programs/healthy-city-strategy.aspx

Devlin A., Frank l., Johnstone S., & Loon J. (2010). Neighbourhood Design, Travel, and Health in Metro Vancouver: Using a Walkability Index. University of British Columbia. Retrieved from http://act-trans.ubc.ca/files/2011/06/WalkReport_ExecSum_Oct2010_HighRes.pdf

Hall A., Hayes M., Oliver L., and Schuurman N. (2011). Assessing thru influence of the built environment on physical activity for utility and recreation in suburban metro Vancouver. BMC Public Health 11 (959), 2011.

Health and Community Design Lab, University of British Columbia. (2013). The Walkability Index. Retrieved from http://health-design.spph.ubc.ca/tools/walkability-index/

GeoData BC (2013). GVRD DEM. Retrieved from SFU SIS Data Warehouse.

Janes Walk. (2014). What is Walkability? Retrieved from http://www.janeswalk.org/information/resources/walkability

REFERENCES

Kaestner R., Zhao Z. (2010). Effects of Urban Sprawl on Obesity. Journal of Health Economics 29 (6), 779–787.

Ken R. Smith, PhD, Barbara B. Brown, PhD, Ikuho Yamada, PhD, Lori Kowaleski-Jones, PhD, Cathleen D. Zick, PhD, Jessie X. Fan, “Walkability and Body Mass Index

Klinkenberg, B. (2007). Multi-criteria evaluation [Powerpoint slides]. Retrieved from http://www.geog.ubc.ca/courses/geog570/notes/mce/mce08.pdf

Kostova, D. (2011). Can the Built Environment Reduce Obesity? The Impact of Residential Sprawl and Neighborhood Parks on Obesity and Physical Activity. Eastern Economic Journal, 37 (3), 390–402.

Metro Vancouver. (2002). Sustainable Region Initiative 2002-2011: A Roadmap To Sustainability. Retrieve from http://www.metrovancouver.org/planning/Docs/CompPlan_Overview.pdf

Metro Vancouver (2003). GVRD Land use. Retrieved from SFU SIS Data Warehouse with permission from Joaquin Karakas (Feb 2003).

Ozbil, Ayse. (2010). Walking to the Station: The Effects of Street Connectivity on Walkability and Access to Transit. (Doctoral Dissertation). Retrieved from ProQuest UMI Dissertations Publishing. (Accession Order No. AAT 3500469)

Saelens, B. E., Sallis, J. F., Slymen, D. J., Frank, L. D., Cain, K. L., Conway, T. L., … Kerr, J. (2007). Neighborhood Walkability and Income. Medicine and Science in Sports & Exercise 39 (5), S31.

Statistics Canada. (2006). Population and dwelling counts A portrait of the Canadian population [Data file]. Retrieved from http://www12.statcan.gc.ca/census-recensement/2006/rt-td/pd-pl-eng.cfm

Statistics Canada (2006). Postal Code Conversion File. Retrieved from SFU SIS Data Warehouse http://www.statcan.gc.ca/mgeo/postalcode-codepostaux-eng.htm

Statistics Canada (2010). Road Network File. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/geo/RNF-FRR/index-eng.cfm

Stevens R. (2005). Walkability Around Neighborhood Parks: An Assessment of Four Parks in Springfield, Oregon. *University of Oregon*. Retrieved from https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/1288/stevens_thesis.pdf?sequence=1

Swinkels, H ., & Mu, L. (2014, March 13). We must plan for a healthy future. Peace Arch News. Retrieved from http://www.peacearchnews.com/opinion/letters/250194481.html

Translink (2013). Google Transit Feed Specification Data. Retrieved from http://www.translink.ca/en/Schedules-and-Maps/Developer-Resources/GTFS-Data.aspx

University of British Columbia. (2010). Neighbourhood Design, Travel, and Health in Metro Vancouver: Using a Walkability Index. Vancouver, Canada: Frank, L., Devlin, A., Johnstone, S., & van Loon, J.

Wilson, D. K., Kirtland, K. A., Ainsworth, B. E., & Addy, C. L. (2004). Socioeconomic Status and Perceptions of Access and Safety for Physical Activity. Annals of Behavioral Medicine 28 (1), 20-28.