Longitudinal effects of urban green space on walking and cycling: A fixed effects analysis

Martyna Hogendorf a, b, Joost Oude Groeniger a, c, *, J. Mark Noordzij a, Mariëlle A. Beenackers a, Frank J. van Lenthe a

a Department of Public Health, Erasmus University Medical Centre, PO Box 2040, 3000 CA, Rotterdam, the Netherlands
b Division of Human Nutrition, Wageningen University, 6700 AA, Wageningen, the Netherlands
c Department of Public Administration and Sociology, Erasmus University, PO Box 1738, 3000 DR, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
- Green space
- Active commute
- Walking
- Cycling
- Geographic information systems
- Fixed effects

ABSTRACT

This study examined whether changes in green space within the living environment were associated with changes in walking and cycling frequencies in a cohort of 3,220 Dutch adults between 2004, 2011 and 2014. Data on self-reported weekly time spent walking and cycling for active commute and leisure were linked to geographic information system (GIS) measures of total green areas within 1000 m buffer zones around each participant’s home address, and distance to the nearest green space. First, cross-sectional linear regression models showed no statistically significant associations between green space measures and walking and cycling. Second, fixed effects (FE) models were used to analyze whether changes in green space were associated with changes in walking and cycling, using longitudinal data from respondents who did not relocate over time. As distance to the nearest green area increased by 100 m, individuals spent 22.76 fewer (95% CI: –39.92, –5.60) minutes walking for leisure per week and 3.21 more (95% CI: 0.46, 5.96) minutes walking for active commute. Changes in distance to green space were not significantly related to changes in cycling measures. No clear associations between changes in green areas within 1000 m buffers and changes in walking and cycling were observed. Overall, there was weak evidence of an effect of changes in green space area on changes in walking, and no evidence for cycling.

1. Background

The urban landscape can shape human activity and offer avenues for health promotion. Current trends in overconsumption and sedentary lifestyles contribute to the prevalence of non-communicable diseases (NCDs), accounting for 70% of deaths worldwide and inflicting strain on health, societal, and economic systems. Increased physical activity (PA) is cited as a top priority intervention in curbing the detrimental effects of chronic disease (Beaglehole et al., 2011) by increasing longevity and protective against cardiovascular diseases, site-specific cancers, type 2 diabetes, obesity, osteoporosis, metabolic syndrome, and high blood cholesterol (Wang et al., 2016; Myers et al., 2015). While many public health efforts focus on conscious behavior change to increase PA, the built environment has been shown to have an effective role in encouraging activity (Sugiyama et al., 2018; MacMillan et al., 2018; Brownson et al., 2009). A spatial-analysis of residential vicinities can inform public policies on how best to influence the health of a population.

Walking is recognized among the most common, acceptable, and accessible forms of physical activity across different age groups, gender, and ethnicities (Siegel et al., 1995). Along with cycling, it can be used for commute and leisure purposes to habitually increase daily energy expenditure and improve health (Kerr et al., 2016). The Netherlands offers a unique case study given the high prevalence of commuter walking and cycling, with 25% of all journeys being traveled by bicycle (Gao et al., 2017). Given a cultural predisposition to an active commute, what stimulates or demotivates Dutch adults to walk or cycle? More importantly, how can cities spatially adapt to further increase activity on a population level?

Emerging socio-ecological approaches have focused on the importance of the built environment in shaping health and behavior. Studies often cite street connectivity, land use mix, neighborhood safety, traffic, access to facilities and parks, landscape, and others, as relevant aspects of an active commute (de Vries et al., 2010). However, unlike countries...
like the United States or Australia in which many of these studies have been conducted, the Netherlands offers a pedestrian and cyclist friendly infrastructure featuring extensive cycling rights of way, bicycle lanes and parking, and educational training for cyclists and motorists (Pucher and Buehler, 2008). Exploring other characteristics, such as the availability of green space, may therefore prove more fruitful in decoding the health-place relationship.

Increased green space has been associated with reduced adult mortality (van den Berg et al., 2015), improved social capital, and lower stress (Mitchell and Popham, 2008; Groenewegen et al., 2012). A recent report by the World Health Organization (WHO) lists pathways linking green space to a multitude of health outcomes (Egorov et al., 2016), and positive associations have been shown between the quantity and quality of urban green areas in relation to small-area life expectancy (Jonker et al., 2014). A review by Hartig et al. details varying and mixed effects of exposure and outcome. Thus, causality cannot be established, putting in doubt the strength and robustness of these observations. Many studies do adjust for confounding factors, but it remains unclear which factors should be included to effectively account for selection (van den Berg et al., 2015). Individuals may choose to live in certain neighborhoods based on lifestyle preferences, environmental considerations, and economic or social factors. The deliberate choice of a physically active person to live in a neighborhood with more green space, for instance, will inflate the association observed between the environment and physical activity. Ultimately, causal evidence may be a cause for action.

This paper aims to decode causal relationships between green space and frequency of walking and cycling by linking comprehensive GIS measures of green space area and proximity to physical activity outcomes from cohort data with 10 years of follow-up. We first describe group-level associations deduced from a cross-sectional analysis. Next, we explore within-subject changes with a fixed effects model. Lastly, we estimate within-subject changes among participants that did not relocate during follow-up.

2. Methods

2.1. Study population

Data was obtained from GLOBE, a prospective cohort study on socioeconomic health inequalities in the Netherlands. The study surveyed adults living in the city of Eindhoven and surrounding areas, a sample representative of the Netherlands as a whole in terms of age, gender, and level of education. Baseline health questionnaires were distributed in 1991 to a random sample of 27,070 individuals aged 15–75 years old, with an overall response rate of 70.1%. The postal questionnaires assessed health, material, behavioral, psychological, and environmental factors indicative of socioeconomic and health disparities. Additional details of the Dutch GLOBE study can be found elsewhere (van Lenthe et al., 2014). The 2004 sample of GLOBE participants representative of the source population of residents aged 25–75 years who resided in Eindhoven and surroundings were selected for the analyses (N = 4,758). Additional questionnaires were administered in 2011 and 2014 (but not in intermediate years). Given that fixed-effects analyses require at least two measurements, respondents who only participated in one year were excluded (30%), resulting in a sample of 3,340 respondents. Analyses were restricted to individuals who resided in Eindhoven and surrounding municipalities at the waves they participated in, and who could be successfully geocoded, resulting in a final sample of 3,220 participants of which 62.8% had measures for all three waves (2004: N = 3,220; 2011: N = 2,884; 2014: N = 2,382).

2.2. Outcome measures of walking and cycling

Self-reported measures of walking and cycling were assessed using the validated SQUASH (Short Questionnaire to Assess Health enhancing physical activity), a tool created by the Dutch National Institute of Public Health and the Environment to measure habitual physical activity levels in an adult population. This simple questionnaire offers a reliable evaluation of physical activity in large populations (Wendel-Vos et al., 2003). Participants reported average number of days per week, and hours and minutes per day, spent walking and cycling as part of an active commute and for leisure purposes. Following SQUASH-guidelines, it was assumed that all participants who filled in hours or minutes per week, but omitted ‘days per week,’ had been active for at least one day. Further, if the number of days was provided without a corresponding time frequency, the median minutes per day of all respondents was substituted, and a final measure of minutes per week was computed. Variables were recoded into separate measures for walking and cycling for active commute, and leisure, as well as total frequencies.

**Abbreviations**

| Abbreviation | Description |
|--------------|-------------|
| FE | Fixed effects |
| GIS | Geographic Information System |
| GLOBE | Gezondheid en LevensOmstandigheden Bevolking Eindhoven en omstreken |
| ISCED | International Standard Classification of Education |
| NCD | Non-communicable disease |
| PA | Physical activity |
| SQUASH | Short Questionnaire to Assess Health enhancing physical activity |
| WHO | World Health Organization |
### 2.3. Exposure measures of green space

The main explanatory variables used included the total area of green space in the living environment, and distance to the nearest green space. GLOBE cohort data from the 2004 and 2011 waves was linked with geographical data from 2003 and 2010 respectively, keeping in line with an appropriate chronology of exposure preceding outcome measures. The 2014 GLOBE cohort data was linked with 2012 geographical data as 2013 geographical data was not available. Respondent addresses were geocoded using geographical software package QGIS and a geocoding plug-in developed by the Dutch National Spatial Data Infrastructure (PDK) (Dutch National; QGIS Development Team, 2017). To maintain respondent privacy, addresses were extracted and geo-coded using a process previously described (Rodgers et al., 2012; Beenackers et al., 2018). In total, 98% of addresses were successfully geo-coded. Movement to a different address between follow-up years was recorded.

Historical geographic data of Eindhoven and surrounding areas was obtained from the Dutch dataset Bestand Bodemgebruik (BBG), created by Statistics Netherlands (CBS). The BBG is a harmonized dataset based on “Top10NL” digital 1:10,000 topographic maps provided by Dutch mapping agency Kadaster, and is available as free, open source GIS files. Each BBG data release is based on the most recently available topographical data from that year. Furthermore, whenever a new wave of the BBG data is released, all previous data waves are updated using the most recent processing techniques. The time-varying exposure variables of green space were calculated at each wave. The harmonization of the BBG data ensures that observed changes in green spaces are representative of actual changes in the built environment and not related to changes in GIS processing. Extensive land classification data was used to locate categories of green spaces relevant to walking and cycling, change in GIS processing. The time-varying characteristics of marital status, employment, income, smoking, and self-rated health. The following model was used for the analyses:

\[
W_{it} = \mu_t + \beta_1 \text{green space}_{it} + \beta_2 x_{it} + \alpha_i + \epsilon_{it}
\]

where Walking/cycling indicates the walking/cycling frequency for individual i at time t, green space\(_{it}\) represents the green space area within separate buffer zones or distance to nearest green space, \(x_{it}\) is a vector of time-varying control regressors, and \(\epsilon_{it}\) is the error term. \(\mu_t\) accounts for time effects that are fixed for all individuals, while \(\alpha_i\) controls for time-invariant personal characteristics.

Robust standard errors were used to account for non-independence clustering at the individual level. Analyses were performed using STATA 13 (StataCorp. Stata Statisti, 2013).

### 3. Results

#### 3.1. Sample characteristics

Participant demographic characteristics are presented in Table 1. The final sample consisted of 3,220 adults of mostly Dutch origin residing in Eindhoven and surrounding areas. The mean follow-up time was 9.2 years. The baseline mean age in 2004 was 53 years; 56% of the participants were women. A little over half of all respondents completed a middle-to-high level of education. On average, respondents walked for 160 min per week and cycled for 150 min per week, spending 66% more time on leisure travel as compared to active travel to work or school. In 2004, respondents resided an average distance of 220 m from the nearest green space. Participants were surrounded by an average green area of 24.6 ha (15%) within a 1000 m buffer around their home address.

#### 3.2. Cross-sectional analyses

Linear regression models applied to cross-sectional data in 2004 showed non-significant and negligible associations between distance to green space and time spent walking and cycling, as shown in Table 2. Similarly, the area of green space was not significantly associated with outcome measures, and results showed wide confidence intervals.
Health and Place xxx (xxxx) xxx

4

3.3. Within-person changes

Within-person changes were observed for all exposures and outcomes, consisting of both increases and decreases in measures over time (Table 3). For the green space measures, about two-thirds of the 6,158 available person observations exhibited changes in distance to nearest green space and changes in green area within a 1000 m buffer. For walking and cycling outcomes, changes were considerably more frequent (81% changes over time for walking and cycling, respectively). For leisure walking and cycling outcomes, changes were particularly small for only data from respondents who did not relocate between years. An additional minutes spent walking for leisure ($\beta$ 3.39, 95% CI 8.37, 1.60, p = 0.183) and total cycling ($\beta$ 3.48, 95% CI 5.23, 0.528) were observed.

### Table 2

Linear regression models regressing walking and cycling outcomes on green space using cross-sectional data from 2004 (n = 3,220).

| Betas | 95% CI | p-value |
|------|------|--------|
| Walking for transport | 0.32 | -1.16 | 1.79 | 0.674 |
| Cycling for transport | 0.23 | -2.33 | 1.87 | 0.828 |
| Walking for leisure | 0.65 | -4.33 | 5.64 | 0.797 |
| Cycling for leisure | -2.92 | -7.09 | 1.25 | 0.169 |
| Walking and cycling for transport | 0.08 | -2.63 | 2.80 | 0.952 |
| Walking and cycling for leisure | -2.27 | -9.17 | 4.60 | 0.519 |
| Total walking | 0.90 | -4.43 | 6.23 | 0.740 |
| Total cycling | -3.39 | -8.37 | 1.60 | 0.183 |
| Total walking and cycling | -2.48 | -10.20 | 5.23 | 0.528 |

Table 3

| Distance to nearest green space (100 m) | Beta | 95% CI | p-value |
|-----------------------------|------|-------|--------|
| Walking for transport | 0.32 | -1.16 | 1.79 | 0.674 |
| Cycling for transport | 0.23 | -2.33 | 1.87 | 0.828 |
| Walking for leisure | 0.65 | -4.33 | 5.64 | 0.797 |
| Cycling for leisure | -2.92 | -7.09 | 1.25 | 0.169 |
| Walking and cycling for transport | 0.08 | -2.63 | 2.80 | 0.952 |
| Walking and cycling for leisure | -2.27 | -9.17 | 4.60 | 0.519 |
| Total walking | 0.90 | -4.43 | 6.23 | 0.740 |
| Total cycling | -3.39 | -8.37 | 1.60 | 0.183 |
| Total walking and cycling | -2.48 | -10.20 | 5.23 | 0.528 |

### Table 3

Within-person changes in green space and walking and cycling between 2004, 2011 and 2014.

| Participants that did not relocate between years (N = 6,158 person observations) | Decrease | No Change | Increase | Average change |
|-----------------------------|----------|-----------|----------|---------------|
| Distance to nearest GS (m) | 23.34 | 0 | 28.76 | N = 1.81 |
| 1000 m buffer (ha) | 5.45 | 0 | 4.78 | N = 0.22 |
| Walking for active commute | 127.74 | 0 | 147.04 | N = 1.60 |
| Cycling for active commute | 132.79 | 0 | 129.85 | N = 4.28 |
| Walking for leisure | 174.04 | 0 | 176.65 | N = 14.00 |
| Cycling for leisure | 155.42 | 0 | 163.14 | N = 5.83 |
| Total active commute | 150.02 | 0 | 158.52 | N = 2.68 |
| Total leisure | 225.44 | 0 | 230.57 | N = 14.88 |
| Total walking | 180.11 | 0 | 185.05 | N = 5.65 |
| Total cycling | 173.27 | 0 | 179.27 | N = 2.29 |
| Total walking and cycling | 239.42 | 0 | 248.88 | N = 16.84 |

3.4. Fixed effects analyses

Table 4 presents results from fixed effects regression analyses using only data from respondents who did not relocate between years. An increase of 100 m in distance to the nearest green space was related to more walking for commute ($\beta$ 3.21, 95% CI 0.46, 5.96), and less walking for leisure ($\beta$ = -22.76, 95% CI = -39.92, -5.60) and total walking ($\beta$ = -21.37, 95% CI = -38.87, -3.88). Greater distance was related to less time spent walking and cycling ($\beta$ = -22.36, 95% CI = -46.19, 1.48), but confidence intervals included the null.

Walking for commute decreased with each additional hectare of green space in the 1000 m buffer ($\beta$ = -33.84, 95% CI = -67.90, 0.23). Meanwhile, increases in green space area seemed to be associated with additional minutes spent walking for leisure ($\beta$ = 58.42, 95% CI = 74.22, 191.06), but confidence intervals included the null. When combined, the measure of total walking minutes was not significantly related to the area of green space in the 100 m buffer ($\beta$ = 39.46, 95% CI = -98.22, 177.14). Minutes spent cycling, and combined measures of all outcomes,
were also not significantly related to green space.

4. Discussion

This study examined whether changes in green space within the living environment were associated with changes in walking and cycling frequencies over a ten-year period. An initial cross-sectional analysis of baseline data did not show significant associations between green space proximity and the amount of green space within the living environment, and weekly walking and cycling. Fixed effects analysis restricted to participants did not relocate during follow-up (n = 6,158 person observations).

were also not significantly related to green space.

4. Discussion

This study examined whether changes in green space within the living environment were associated with changes in walking and cycling frequencies over a ten-year period. An initial cross-sectional analysis of baseline data did not show significant associations between green space proximity and the amount of green space within the living environment, and weekly walking and cycling. Fixed effects analysis restricted to participants that did not relocate during follow-up suggested that as distance to the nearest green area increased, individuals decreased their walking frequency, with no relation to changes in cycling measures. No clear associations between changes in green areas within 1000 m buffers and changes in walking and cycling were observed. There was weak evidence overall of an effect of changes in green space area on changes in walking, and no evidence for cycling.

Urban green space has widely been endorsed with health-promoting benefits, with positive associations found between nearby parks and overall health and physical activity (Douglas et al., 2017). Recent policy frameworks, notably the United Nations’ Habitat III New Urban Agenda, also support the greening of urban areas as a means toward physical and mental health promotion (Douglas et al., 2017). However, literature offers mixed results regarding the role of urban green space on physical activity due to variation in methodological approaches, measurement of physical activity (Kaczynski et al., 2008), and the characterization of relevant green space (Markeyvich et al., 2017; Meurs and Haaijer, 2001). The current study is one of few longitudinal analyses which models estimated effects of green space change on the most common, and accessible forms of physical activity: walking and cycling. It fills an important methodological gap by aiming to interpret the relationship between health and place in a way that has more potential for evidence-based action.

Our baseline analysis found weak, non-significant associations between green space and activity levels, which is comparable to findings of Maas et al. (2008). In contrast, the longitudinal fixed effects analysis among participants that did not relocate during follow-up showed that changes in residential proximity to green space significantly impacted walking frequency; an increase of 100 m to the nearest green space resulted in 21 fewer minutes per week spent walking overall, 23 fewer minutes of leisure walking, but 3 additional minutes walking for commute. Previous research has shown that green space within walking distance of the home generally supports human health (Ekkel and de Vries, 2017), while parks located further away are not as likely to be used (Coombes, 1982). While no official cut-off distance is supported by empirical evidence, Annerstedt van den Bosch et al. have proposed a guideline of 1 ha within a 300 m absolute distance to the nearest green space as a green space indicator for public health (Annerstedt van den Bosch et al., 2016). Other studies also cite distance as a key determinant of green space use, with 100–300 m appearing as the threshold beyond which a decline in use is observed (Nielsen and Hansen, 2007). Our findings suggest that introducing green space closer to one’s residence can encourage people to spend more time walking for leisure, but not as part of their commute. Green space closer to the home may deter individuals from walking to work or school, and instead encourage cycling or driving. This observed effect may also relate to the cohort demographic and nature of the activity; members of an ageing cohort gradually enter retirement, thus eliminating the necessity of walking to work, and this in turn can skew the FE model to produce significant results.

Whereas walking seemed to be affected by changes in green space, cycling was not. Moreover, in relation to the changes in green area in 1000 m buffer, no significant associations were observed for total measures of walking and cycling. This lack of significant associations suggests that additional factors may be more important for physical activity than changes in green space. Walking and cycling can depend on personal preferences and constraints. An aging generation will likely be faced with different demands, for example, familial obligations such as caring for grandchildren. The mechanisms linking walking and cycling to green space availability are also likely influenced by other factors in the home environment. For instance, factors such as crime, safety, deprivation, social interaction, road safety, and particularly a pedestrian and cyclist friendly urban environment in the Netherlands, may affect whether or not people walk or cycle in nearby green areas, and may have limited or tempered any effects of changes in green space on changes in activity. This may be particularly pronounced for cycling, considering the wide availability and use of bike lanes in The Netherlands (Everson et al., 2012; Poster et al. 2016; Weimann et al., 2017; Carver et al., 2005). Furthermore, the choice of buffer sizes in measuring total area of green space may play a role in the strength and significance of the
results.

4.1. Strengths and limitations

The original contribution of this study is the multi-methodological approach and use of detailed GIS data, enabling the linkage of changes in environment and behavior. Much of previous research has relied on wider scale, neighborhood or city-level data that does not accurately depict within-subject changes in exposure. The data provided by the BBG allowed for precise calculations of total green space area and identification of actual changes over time that are not affected by changes in GIS processing. Euclidian buffers around respondents’ homes aided in reducing spatial misclassification faced by other indicators, such as neighborhood boundaries (Hirsch et al., 2014), and the choice of a 1000 m buffer was comparable to other studies. Further, data from the GLOBE study offered detailed measures of personal characteristics that were used to control for time-varying confounding. The use of multiple outcomes based on a validated questionnaire offered insight into how specific activities are affected by factors in the environment, discerning between commuting and leisure activities, and modes of walking and cycling.

A main limitation of this study is the low within-person variability in walking and cycling for active commute, which restricts the statistical efficiency of a fixed-effects analysis. Although FE analyses are better able to infer causality, they are dependent on observable changes in exposure and outcome measures. The current FE models may have not been able to depict significant relationships due to limited changes observed in the sample population. Further, baseline characteristics reflect a generally active, healthy, and affluent sample of individuals, which may influence how they react to changes in the built environment. For instance, aspects such as car ownership, or the propensity for an active lifestyle, can minimize the impact of de-greening a neighborhood. In addition, around one quarter of missing baseline data on household income was imputed, with implications for biased effect estimates if data was not missing at random. Our statistical model assumes no correlation of attrition and missingness to unmeasured, time-varying characteristics in the study sample, but, if violated, this correlation may have biased the results. In addition, the assumption that the residuals of the linear regression model are normally distributed was violated in the cross-sectional analysis for the active commute measures. However, using negative binomial regression models did not change the findings. Moreover, the fixed effects models did not suffer from this limitation (changes in walking and cycling were mostly normally distributed). We therefore reported results from the linear regression models only.

Self-reported measures of walking and cycling, though based on a validated questionnaire, are subject to recall bias if older participants struggle to provide accurate measures of physical activity. In addition, while our GIS data offered an accurate measure of existing green space, there is no evidence for the actual use or even awareness of these green areas by participants. Similarly, the nearest green space to an individual’s home address may not be perceived as such, given its size or functionality, and Euclidian distances may not reflect the travel routes taken by participants.

4.2. Future research

To better understand environmental influences on walking and cycling, prospective studies should incorporate both individual and social factors that may affect outcomes, such as self-efficacy, attitude, or social support (Owen et al., 2004). Neighborhood level factors of safety and deprivation may confound the effect of green space on physical activity, and should be considered in future research. While our study focused on adults of mostly Dutch origin, the inclusion of youth and non-Dutch residents would offer a more representative group of green space users. Objective measures of walking and cycling, through the use of accelerometers or GPS trackers, might strengthen the validity of

outcome values. Additional green space indicators, such as network distance, can be used to better evaluate the use of green space, reflecting likely routes of access. Similarly, the number of green spaces present within a residential area, and a specification of the types of changes occurring in green space, may provide a more robust analysis. Testing for interaction would assess the cumulative effects that multiple factors may have on physical activity. Finally, conducting similar studies in diverse geographic settings on large study samples would help build a solid foundation of evidence generalizable to a wider population.

5. Conclusions

The methods used to study relationships between place and health greatly shape the foundation of knowledge that exists in this field. Our approach separately compared group-level associations, and individual within-person effects, of green space on walking and cycling, leveraging longitudinal data to strengthen the basis for causal inference. Our results indicate that walking, and particularly leisure walking, decreases as green spaces are moved further from one’s residence. However, local green space alone may not significantly affect physical activity. Replicating our approach on larger, diverse study samples with more variability across time would strengthen the reliability of these findings, or introduce different patterns of effect. Future research should aim to identify aspects of the quality and quantity of changes required in the built environment to improve physical activity, which can steer urban planning and policy efforts and ultimately guide the prevention of chronic disease in an increasingly urbanized world.

Ethics approval and consent to participate

The use of personal data in the GLOBE study is in compliance with the Dutch Personal Data Protection Act and the Municipal Database Act, and has been registered with the Dutch Data Protection Authority (number 1248943).

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated during and/or analysed during the current study are not publicly available due to privacy regulations but are available from the corresponding author on reasonable request.

Funding

The study was supported by a grant from the European Commission HORIZON 2020 research and innovation action (grant number 667661) and the Netherlands Organisation for Health Research and Development (grant number 200500005). The funders had no role in the study design or the analysis and interpretation of the data. All authors and their institutions reserve intellectual freedom from the funders.

Declaration of competing interest

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.
