Neighbourhood effects in health behaviours: a test of social causation with repeat-measurement longitudinal data

Jaakko Airaksinen1,*, Christian Hakulinen1, Laura Pulkki-Råbäck1,2, Terho Lehtimäki3,4, Olli T. Raitakari5,6, Liisa Keltikangas-Järvinen1, Markus Jokela1

1 Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland
2 Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki, Finland
3 Department of Clinical Chemistry, School of Medicine, University of Tampere, Tampere, Finland
4 Fimlab Laboratories, Tampere, Finland
5 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
6 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland

*Correspondence: Jaakko Airaksinen, IBS, University of Helsinki, P.O. Box 9, FIN-00014 University of Helsinki, Finland, Tel: +358 50 3415662, Fax: +358 9 191 29521, e-mail: jaakko.airaksinen@helsinki.fi

Background: Neighbourhood characteristics have been associated with health behaviours of residents. We used longitudinal data to examine whether neighbourhood characteristics (level of urbanization and socioeconomic status) are related to within-individual variations in health behaviours (alcohol consumption, smoking, exercise and self-interest in health) as people live in different neighbourhoods over time. Methods: Participants were from the Young Finns prospective cohort study (N=3145) with four repeated measurement times (1992, 2001, 2007 and 2011/2012). Neighbourhood socioeconomic status and level of urbanization were measured on the level of municipality and zip code area. Within-individual (i.e. fixed-effect) regression was used to examine whether these associations were observed within individuals who lived in different neighbourhood in different measurement times. Results: People living in more urban zip code areas were more likely to smoke ($b=0.06$; CI = 0.03–0.09) and drink alcohol ($b=0.11$; CI = 0.08–0.14), and these associations were replicated in within-individual analysis—supporting social causation. Neighbourhood socioeconomic status and urbanization were associated with higher interest in maintaining personal health ($b=0.05$; CI = 0.03–0.08 and $b=0.05$; CI = 0.02–0.07, respectively), and these associations were also similar in within-individual analysis. Physical exercise was not associated with neighbourhood characteristics. Conclusions: These data lend partial support for the hypothesis that neighbourhood differences influence people’s health behaviours.

Introduction

People’s health and social behaviours are influenced by characteristics of the residential areas where they live. Neighbourhoods may influence how people adopt and maintain important health behaviours, such as smoking, physical activity and alcohol consumption. For example, areas characterized by low socioeconomic status have higher rates of obesity and unhealthy dietary patterns in Canada. In a Finnish study, elderly rural dwellers were reported to walk less and eat more unhealthy foods than their urban counterparts. Neighbourhood associations have also been reported for various health outcomes.

However, the issue of causality in neighbourhood effects has not been thoroughly addressed in previous studies because of important methodological limitations. First, most studies of neighbourhood effects have been observational studies, as experimental neighbourhood studies are rarely feasible. Results from observational studies can be confounded by residual confounding due to unmeasured variables. Second, most of the neighbourhood-effect studies have been cross-sectional, and conclusions from cross-sectional data can be biased by people’s selective residential mobility to specific areas based on their health-behaviour profiles or other traits that correlate with health behaviours, such as socioeconomic status.

Experimental and quasi-experimental neighbourhood studies have been carried out to test causality of neighbourhood effects. The ‘Moving to Opportunity’ study, which included 4608 families from five US cities, is one of the best-known experimental neighbourhood studies, where volunteer participants living in high-poverty neighbourhoods were given the opportunity to move to low-poverty neighbourhoods. Those who moved had better mental health in the follow-up study than those who did not move, supporting social causation of neighbourhood effects. However, it is unknown whether the results from such specific experimental studies can be generalized to the general population. Quasi-experimental studies have also provided evidence for potentially causal neighbourhood effects. For example, a review of studies on the effects of air pollution to children’s health concluded that reduced pollution levels lowered infant death considerably.

The problem of self-selection via selective residential mobility in observational studies can be mitigated by longitudinal data with repeated measurements. This study design can test whether people are less healthy when they are living in a poorer neighbourhood compared with another time when the same people are living in more affluent neighbourhoods. In this within-individual analysis, all the stable confounding characteristics of individuals are adjusted for, as individuals are compared with themselves when living in different neighbourhoods. A recent Australian study observed no within-individual associations in various health outcomes, providing no evidence for causal neighbourhood effects in the general population. Another 6-year longitudinal study of 5661 participants living in USA examined the association between urban sprawl and obesity, as urban sprawl has been associated with higher obesity rates. Using first-difference regression that examined how residential moves across different
levels of urban sprawl were related to changes in body mass index over time, Eid et al. found no support for the hypothesis that urban sprawl causes obesity. The within-individual analysis cannot account for all the effects of selective residential mobility. For example, individual characteristics that change over time can still confound the within-individual associations, and it is difficult to determine the direction of such bias. However, the within-individual analysis provides less confounded estimates than simple cross-sectional associations.

In this study, we used repeated-measurement longitudinal data from the Young Finns prospective cohort study to examine whether neighbourhood characteristics were associated with health behaviours, following the approach by the Australian study cited earlier. The repeated measurements allowed us to assess whether the same participants had poorer or better health behaviours when they were living in different residential locations. If neighbourhood socioeconomic status or urbanicity have causal effects on health behaviours, one should observe within-individual associations.

Neighbourhood characteristics were measured with indices of rural/urban continuum and neighbourhood socioeconomic status. It is often emphasized that the geographic unit used to measure characteristics may influence results of neighbourhood effects. To test this, the neighbourhood characteristics were assessed at the level of municipalities (i.e. lowest-level administrative units of Finland) and zip code areas. Health behaviours included alcohol consumption, smoking, frequency of exercising and participants’ self-rated interest in maintaining personal health.

**Methods**

**Participants**

The participants were 3145 individuals (1664 women) who participated in the ongoing Young Finns prospective cohort study. The original sample consists of 3596 Finnish healthy children and adolescents derived from six birth cohorts (aged 3, 6, 9, 12, 15 and 18 years at baseline). In order for the sample to be broadly representative in terms of socioeconomic background, Finland was divided into five areas according to locations of university cities with a medical school (Helsinki, Kuopio, Oulu, Tampere and Turku). In each area, urban and rural girls and boys were randomly selected on the basis of their unique social security number. The study began 1980 and participants have been followed in two and 17.4% in one). The number of participants in each study wave were randomly selected on the basis of their unique social security number. The study began 1980 and participants have been followed over time, Eid et al. found no support for the hypothesis that urban sprawl causes obesity. The within-individual analysis cannot account for all the effects of selective residential mobility. For example, individual characteristics that change over time can still confound the within-individual associations, and it is difficult to determine the direction of such bias. However, the within-individual analysis provides less confounded estimates than simple cross-sectional associations.

In this study, we used repeated-measurement longitudinal data from the Young Finns prospective cohort study to examine whether neighbourhood characteristics were associated with health behaviours, following the approach by the Australian study cited earlier. The repeated measurements allowed us to assess whether the same participants had poorer or better health behaviours when they were living in different residential locations. If neighbourhood socioeconomic status or urbanicity have causal effects on health behaviours, one should observe within-individual associations.

Neighbourhood characteristics were measured with indices of rural/urban continuum and neighbourhood socioeconomic status. It is often emphasized that the geographic unit used to measure characteristics may influence results of neighbourhood effects. To test this, the neighbourhood characteristics were assessed at the level of municipalities (i.e. lowest-level administrative units of Finland) and zip code areas. Health behaviours included alcohol consumption, smoking, frequency of exercising and participants’ self-rated interest in maintaining personal health.

**Methods**

**Participants**

The participants were 3145 individuals (1664 women) who participated in the ongoing Young Finns prospective cohort study. The original sample consists of 3596 Finnish healthy children and adolescents derived from six birth cohorts (aged 3, 6, 9, 12, 15 and 18 years at baseline). In order for the sample to be broadly representative in terms of socioeconomic background, Finland was divided into five areas according to locations of university cities with a medical school (Helsinki, Kuopio, Oulu, Tampere and Turku). In each area, urban and rural girls and boys were randomly selected on the basis of their unique social security number. The study began 1980 and participants have been followed subsequently in eight study waves in 1983, 1986, 1989, 1992, 1997, 2001, 2007 and 2010–12. All participants gave written informed consent, and the study was approved by local ethics committees. In this study data from the 1992, 2001, 2007 and 2010–12 follow-ups were used. Participants who had all relevant data for at least one study wave were included. This left us with the current sample of 3145 participants.

Municipality-level information was gathered from SOTKAnet, which contains comprehensive statistical information on welfare and health in Finland. Finland is divided into 320 municipalities ranging from 6 to 15 053 km² in land area size with a median size of 540 km². Median number of residents was 5878 (range: 103–595 384) and median population density 10.90 persons per square kilometre (range: ~0–3051 persons/km²). Municipalities are further divided into zip code areas. Zip code data were derived from the database of Statistics of Finland. There are altogether 3069 zip code areas in Finland. Median number of residents for zip code areas was 483 (range: <100–25 820). Data for zip code areas with population <100 were not available due to privacy reasons. Participants included in our analyses lived on 176 and 1246 unique municipalities and zip code areas, respectively.

**Measures**

Information on participants’ health behaviours were collected from self-reported questionnaires in 1992, 2001, 2007 and 2011/2012. Alcohol consumption was rated on a 6-point scale on the question “How often do you drink 6 units of alcohol or more?” (1 = less than twice a year or never, 2 = 2–6 times a year, 3 = Once a month, 4 = 2–3 a month, 5 = once a week, 6 = twice a week or more often). Smoking was reported on 5-point scale (1 = I have never smoked, 2 = I have quit/I am on a break, 3 = less than once a week, 4 = once a week or more, but not daily, 5 = one or more cigarettes a day). Frequency of exercising was rated on a 6-point scale (1 = never, 2 = once a month, 3 = once a week, 4 = 2–3 a week, 5 = 4–6 a week, 6 = daily). Self-rated interest in health was reported on a 5-point scale (1 = I barely pay any attention on my health, 2 = I only pay a little attention on my health, 3 = Neither little nor a lot, 4 = I pay some attention on my health, 5 = I pay a lot of attention on my health). Education was measured as self-reported years of schooling in study waves 2001, 2007 and 2011/2012.

Socioeconomic status and level of urbanization were used as the neighbourhood indicators. On municipality level, socioeconomic status was measured using tax revenue per capita. On zip code area, median gross income per resident in a year was used. Level of urbanization of municipalities was measured based on population density, which was measured as the number of residents per square kilometre. Urbanization level of zip code areas was measured as a combined measure of the proportion of high-rise buildings on the area and available services. For the combined measure, we standardized the variables and used the average of the two variables. Municipality-level tax revenue data were available for all the study waves, expect for 1992, for which the 2001 data was used as a proxy. Municipality tax revenue measures were corrected for inflation. For zip code, data were not available before 2007, so we used data from 2007 for all measurement times. All municipality and zip code area information were standardized (mean = 0, SD = 1) at the level of municipalities or zip codes, and then matched with participant’s self-reported place of residence for each study wave.

**Statistical analysis**

Associations between neighbourhood characteristics and health behaviours were analyzed using multilevel random-intercept regression. The longitudinal data were structured so that the repeated measurements from the participants (level-1 person-observations) were nested within participants (level-2 units). The random-intercept regression produces essentially the same coefficients as ordinary least squares regression but it takes into account the non-independence of observations, which results in correctly calculated error terms. With repeated measurements from the same individuals, it is possible to focus only on within-individual variations over time by subtracting each individual’s average values in the predictor variable from the repeated measurements. This within-individual analysis is often called ‘fixed-effect regression’ and it eliminates all stable differences between different individuals from the model; only the time-variant variables can account for within-individual changes in the dependent variables across repeated measurements. The fixed-effects regression can provide less confounded estimates than ordinary regression as fixed-effects regression adjusts for all stable between-individual differences.

Age, sex and education were included as covariates in all analyses. Analyses were done using Stata 12.

**Results**

Descriptive statistics are shown in table 1. Of the 3145 participants, 36.4% took part in all the four study waves (26.7% in three, 19.5% in two and 17.4% in one). The number of participants in each study...
Analyses were effectively the same as those described earlier (data analyses on dichotomized outcome variables. Results from those observed (data not shown). We also ran multilevel logistic regression analyses on the combined measure of proportion of high-rise buildings and available services.

This study used repeated-measurement data from a prospective cohort study with 3145 participants to examine the possible causal association between neighbourhood characteristics and health behaviours. People living in urban zip code areas were more likely to smoke and drink alcohol than those living in rural areas, and these associations were partly replicated in within-individual analysis—supporting social causation. Neighbourhood socioeconomic status and urbanization were associated with higher interest in maintaining personal health, and these associations were similar in-between-individuals and within-individual analysis. Physical exercise was not associated with neighbourhood characteristics. The within-individual associations were generally ~50% weaker compared with the overall associations, suggesting that stable differences between individuals may introduce upward bias in estimating potentially causal neighbourhood effects.

Table 1 Descriptive statistics of the 9174 person-observations from 3145 unique individuals from the Young Finns Study (1992–2012), and municipality and zipcode area level measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total no.</th>
<th>No. of Persons</th>
<th>%</th>
<th>Mean (SD)</th>
<th>Within-Person SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>4111</td>
<td>1481</td>
<td>47.1</td>
<td>33.4 (9.0)</td>
<td>7.5</td>
</tr>
<tr>
<td>Women</td>
<td>5063</td>
<td>1664</td>
<td>52.9</td>
<td>3.5 (1.0)</td>
<td>0.6</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest in healtha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol consumptionb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smokingc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of exercisеб</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipality</td>
<td>Urbanizationа</td>
<td>471 (875)</td>
<td>405</td>
<td>2.2 (1.6)</td>
<td>0.6</td>
</tr>
<tr>
<td>SES</td>
<td>3026 (730)</td>
<td>368</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zipcode area</td>
<td>Urbanizationd</td>
<td>1.1 (1.2)</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>15 221 (3083)</td>
<td>1554</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SD, standard deviation; SES, socioeconomic status. For categorical variables, the values are the number of total person-observations, number of unique persons and percentages calculated from person-observations. For continuous variables, the values are means, overall SDs and within-person SDs.

Discussion

This study used repeated-measurement data from a prospective cohort study with 3145 participants to examine the possible causal association between neighbourhood characteristics and people’s health behaviours. People living in urban zip code areas were more likely to smoke and drink alcohol than those living in rural areas, and these associations were partly replicated in within-individual analysis—supporting social causation. Neighbourhood socioeconomic status and urbanization were associated with higher interest in maintaining personal health, and these associations were similar in-between-individuals and within-individual analysis. Physical exercise was not associated with neighbourhood characteristics. The within-individual associations were generally ~50% weaker compared with the overall associations, suggesting that stable differences between individuals may introduce upward bias in estimating potentially causal neighbourhood effects.

Previous findings on the association between neighbourhood socioeconomic status and consumption of alcohol have been mixed. A study involving 93 747 Canadians reported that the association is u-shaped, so that people in affluent and in deprived areas drink more than people in mid-range neighbourhoods. We found no such quadratic association between neighbourhood socioeconomic status and consumption of alcohol in our sample (data not shown). Another study on people (n = 8197) living in California found that those living in the least deprived areas were most likely to be heavy drinkers. There are also studies showing that people living in highly disordered neighbourhood tend to drink more than people in more peaceful neighbourhoods. However, all these results were from cross-sectional studies. Our results suggest that people living in more affluent neighbourhoods of Finland drink alcohol more frequently than those living in poorer neighbourhoods but this association is due to differences between different individuals living in different areas—moving across affluent and poor neighbourhoods does not influence people’s drinking habits. The degree of neighbourhood urbanization may be more relevant for social causation, as individuals moving to more urban neighbourhoods drank more frequently compared with another time when they were living in a more rural neighbourhood.

Although urbanization and neighbourhood socioeconomic status were not associated with better health behaviours, people living more urban and affluent areas paid more attention to their own health than people living more rural and deprived areas. There was no difference between the within-individual and between-individual components of the associations, so a causal association is plausible but not strongly supported by the present data. This
discrepancy between health behaviours and self-reported interest in health in urban and rural areas could be related to how people interpret their level of interest in health. When reporting how interested they are in monitoring and maintaining their health, urban people might think of other health behaviours besides alcohol, smoking and exercise, such as eating healthy food. The association may also reflect reporting bias where urban people are more inclined to report higher interest in health even though they may not act accordingly.

Strengths and limitations
The strengths of our study lie in the population-based sample and longitudinal data with repeated measurements. On the side of limitations, it must be noted that all our health behaviour outcomes measurements were based on single-item for each behaviour. Multiple items could have increased the reliability of those measurements. All outcome measurements were also based on self-reports, which may be subject to social desirability bias.36 Also, our neighbourhood level measurements for zip code areas were not available for each of the study waves. Thus, we could not examine how neighbourhood change at zip code area level might be related to health behaviours among individuals who did not move across neighbourhoods. Furthermore, the zip code areas might not be equivalent to subjective impression of neighbourhood boundaries, which may bias the results. The study was limited only to Finland, which may be a rather homogenous country in terms of demographics compared with many other countries, and so the results of this study cannot readily be generalized to countries with higher sociodemographic diversity. The same issue was noted in a study comparing rural-urban disparities in health between Canada and Australia.37 Finally, it must be emphasized that the within-individual analysis cannot exclude all potential confounding effects (e.g. individual traits that vary over time) and some of the truly causal neighbourhood effects may not be observed with within-individual analysis in adulthood (e.g. exposure to adverse neighbourhood effects in childhood).38 Additional life-course methods are needed to address the limitations of within-individual analysis.

Conclusions
The majority of evidence for neighbourhood effects in health is based on cross-sectional observational studies, which limits causal interpretations of the findings.6,8 Our results from repeated-measurement longitudinal data provide partial support for social causation in neighbourhood effects involving alcohol consumption and smoking. However, the evidence was stronger and more consistent for between-individual differences than for within-individual changes over time—the latter being more important when considering social causation. Therefore, additional longitudinal studies are needed to test whether similar or different dynamics apply to different health outcomes in different countries.

Supplementary data
Supplementary data are available at EURPUB online.
Funding
This work was financially supported by the Academy of Finland (L.K.J., Grant nos. 258711 and 265869; M.J., Grant no. 268388), the Kone Foundation (M.I.) and the Juho Vainio Foundation (L.P.-R.). The Young Finns Study has been financially supported by the Academy of Finland (grants 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), 41071 (Skidii)), the Social Insurance Institution of Finland, Kuopio, Tampere and Turku University Hospital Medical Funds (grant 9N035 for Dr. Lehtimäki), Juho Vainio Foundation, Paavo Nurmi Foundation, Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation, Tampere Tuberculosis Foundation and Emil Aaltonen Foundation (for Dr. Lehtimäki).

Conflicts of interest: None declared.

Key points
- The causal role of neighbourhood health associations is unclear.
- We examined within-individual changes in neighbourhoods and health.
- People moving to urban areas were more likely to start smoking and drinking alcohol.
- People moving to more affluent areas were more likely to be the ones who already drank more alcohol.
- Results lend only partial support for causal neighbourhood effects.

References