Practice of Epidemiology

Are Neighborhood Health Associations Causal? A 10-Year Prospective Cohort Study With Repeated Measurements

Markus Jokela*

* Correspondence to Dr. Markus Jokela, Institute of Behavioural Sciences, University of Helsinki, Siltavuorenpenger 1A, P.O. Box 9, 00014 Helsinki, Finland (e-mail markus.jokela@helsinki.fi).

Initially submitted October 27, 2013; accepted for publication May 7, 2014.

People who live in disadvantaged neighborhoods tend to have poor physical and mental health, but this might be due to selective residential mobility rather than causal neighborhood effects. As a test of social causation, I examined whether persons were less healthy when they were living in disadvantaged neighborhoods than at other times when they were living in more advantaged neighborhoods. Data were taken from the 10-year Household, Income and Labour Dynamics in Australia (HILDA) prospective cohort study, which had annual follow-up waves between 2001 and 2010 (n = 112,503 person-observations from 20,012 persons). Neighborhood disadvantage was associated with poorer self-rated health, mental health, and physical functioning, higher probability of smoking, and less frequent physical activity. However, these associations were almost completely due to between-person differences; the associations were not replicated in within-person analyses that compared the same persons living in different neighborhoods over time. Results were similar when using neighborhood remoteness as the exposure and when focusing only on long-term residence. In contrast, poor health predicted selective residential mobility to less advantaged neighborhoods, which provided evidence of social selection. These findings provide little support for social causation in neighborhood health associations and suggest that correlations between neighborhoods and health may develop via selective residential mobility.

Abbreviation: SLA, statistical local area.

Editor’s note: An invited commentary on this article appears on page 785, and the author’s response appears on page 788.

Several studies on neighborhood effects have shown that people’s mental and physical health correlate with neighborhood characteristics, such as socioeconomic status, crime rate, or ethnic composition of the residential area (1–6). Most of these studies are based on cross-sectional data (1, 2, 7), and the methodological problem of selection bias in identifying causal neighborhood effects has long been acknowledged (8–11). In terms of causal inference, people may not be exchangeable between neighborhoods (12), and neighborhood characteristics may not be exogenous exposure variables. People’s health may therefore correlate with neighborhood characteristics because people with better health are more able to move to more affluent neighborhoods than are those with poorer health and not because neighborhood qualities have a causal impact on health (13). If this were the case, neighborhood associations could be considered as the “neighborhood consequences” of social inequalities that determine people’s ability to move to desired locations.

Multilevel studies have attempted to identify the separate neighborhood and individual associations with health outcomes (14), but these studies can only demonstrate that a certain neighborhood characteristic is associated with individual health when taking into account specific individual traits. Causal interpretations are hampered by the difficulty of correctly adjusting for relevant person-level confounders while not overadjusting for individual characteristics that mediate the neighborhood associations (8, 11, 15, 16). Social
experiments have provided some evidence of causal neighborhood effects (17). For example, in the Moving to Opportunity randomized community study, the subjective wellbeing of persons who moved away from poor neighborhoods remained modestly improved several years after moving compared with that of controls (18). The generalizability of these results to natural settings of the general population is unknown. Longitudinal data from observational studies can help to determine the temporal order between neighborhood exposure and health outcomes. However, the problem of selection bias is not directly addressed by longitudinal studies of disease incidence (19, 20) or by prospective follow-up studies with only 1 baseline measurement of neighborhood exposure (21, 22) because people may get selected to different neighborhoods on the basis of long-term latent disease incidence risk and health trajectories related to aging.

Evidence of a causal association between neighborhood and health could be strengthened considerably by demonstrating in observational studies that people’s health changes as they move across different neighborhoods. However, very few studies have used longitudinal data with repeated measurements of both neighborhoods and health to examine such within-person changes. A 6-year longitudinal study with repeated measurements showed that urban sprawl was associated with a higher prevalence of obesity, but this association was not observed with first-difference regression of persons who moved to less or more densely populated areas, suggesting that the association was not causal (23). Similarly, data from a Swedish study in which siblings were compared did not support a causal interpretation of the association between neighborhood deprivation and the rate of violent crime (24).

In the present study, I examined social causation and social selection in neighborhood health associations in a large Australian prospective cohort study that had a 10-year follow-up with annual repeated measurements (25). Social causation was assessed using fixed-effect regression, which is based on within-person variation in the exposure and thereby removes confounding caused by stable differences between persons. This provides a strong test for causality in neighborhood associations in observational settings. Social selection was examined by using health status and behaviors to predict subsequent moves to more or less disadvantaged (or remote) neighborhoods.

METHODS

Participants

The Household, Income and Labour Dynamics in Australia (HILDA) Survey is an annual household-based panel study developed to collect information about economic and subjective wellbeing, labor market dynamics, and family dynamics (25). The survey began in 2001 with a large national probability sample of Australian households that occupied private dwellings (n = 7,682 households with 19,914 persons at baseline). All members of the households who provided at least 1 interview in wave 1 formed the basis of the longitudinal panel to be pursued in each subsequent wave. The sample has been gradually extended to include any new household members that resulted from changes in the composition of the original households. Through wave 10, which was carried out in 2010, a total of 28,547 persons had participated in at least 1 study wave.

The present study included all available person-observations from participants for whom data on all study variables in at least 1 study wave were available. The final sample included 112,503 person-observations from 20,012 unique persons across the 10-year follow-up period (an average of 4.4 (standard deviation, 1.7) person-observations per participant). There were no appropriate longitudinal sampling weights for the analysis used in the present study, so all models were fitted without sampling weights.

Measures

Neighborhood characteristics were determined at the level of statistical local areas (SLA), which is the general-purpose spatial unit used to collect and disseminate statistics (n = 1,353 SLAs in 2001). In years in which a census is not conducted, the SLA is the smallest unit defined in the Australian Standard Geographical Classification (see www.abs.gov.au for details of geographic hierarchy). The median population count of SLAs was 5,908 (interquartile range, 2,743–14,517), and the median area size was 74.5 km² (interquartile range, 75–1,944.0). Web Figure 1 (available at http://aje.oxfordjournals.org/) shows the map of SLA boundaries. Household addresses of participants were geocoded at each wave, and the participants’ SLAs were determined from these data.

Two neighborhood indicators were derived from 2001 census data. Neighborhood disadvantage was determined based on the decile index of relative socioeconomic advantage/disadvantage as calculated using the Socio-Economic Indexes for Areas (26) indicators. The index is a continuum of advantage to disadvantage, and it takes into account variables such as the proportion of families with high incomes, people with a tertiary education, and people employed in a skilled occupation. For the present analysis, the scale was coded so that higher scores indicated higher neighborhood disadvantage. Web Figure 1 shows the distribution of neighborhood disadvantage across Australia. At the level of SLAs, the correlation between disadvantage deciles in 2001 and 2011 was 0.89 (calculated from census data available at www.abs.gov.au), which suggests a high rank-order stability of neighborhood disadvantage over the study period. Neighborhood remoteness was measured using Accessibility/Remoteness Index of Australia scores (27). Remoteness is determined on the basis of accessibility to various services, that is, a weighted score of road distances to “service centers” with smaller and larger populations. The scale ranges from 1 for a major city (indicating relatively unrestricted access to a wide range of goods and services and to opportunities for social interaction) to 5 for a very remote/migratory area (indicating very little accessibility of goods and services and few opportunities for social interaction).

Information on health status and health behaviors was collected from participants’ self-reports. Mental health and physical functioning were assessed with the Short Form-36 mental health and physical functioning composite scores (28). Self-rated health was reported on a 5-point scale (1 = poor, 5 = excellent). Smoking was coded dichotomously.
Physical activity level was assessed with a question about how often the person participated in physical activity (without specifying a difference between leisure-time or nonleisure activity), with the following response options: 1 = not at all (11% of all person-observations); 2 = less than once per week (15%); 3 = 1–2 times per week (24%); 4 = 3 times per week (16%); 5 = more than 3 times per week (21%); 6 = every day (13%). Alcohol consumption was determined using the question “Do you drink alcohol?” with the following response options: 1 = I have never drunk alcohol (10%); 2 = I no longer drink (6%); 3 = yes, but rarely (24%); 4 = 2–3 days per month (12%); 5 = 1–2 days per week (19%); 6 = 3–4 days per week (13%); 7 = 5–6 days per week (8%); and 8 = everyday (8%). Age, sex, and country of birth (0 = Australia, 1 = United Kingdom, 2 = other) were included as sociodemographic covariates in all models. Additional covariates included educational level (highest educational degree), income (total household income), and marital status (married or cohabiting vs. not married or cohabiting). These were assessed at each study wave concurrently with the health outcome measures and were modeled as time-varying covariates. The results of the fully adjusted models remained unchanged when the covariates were used as time-lagged (i.e., covariate measured 1 year before the health outcome measure rather than concurrently with the outcome) time-varying variables (data not shown). Table 1 shows additional descriptive statistics of the sample.


<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>52,707</td>
<td>9,614</td>
<td>46.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>59,796</td>
<td>10,398</td>
<td>53.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
<td>43.8 (18.0)</td>
<td>2.6</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>88,871</td>
<td>15,742</td>
<td>79.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>7,572</td>
<td>1,190</td>
<td>6.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>16,060</td>
<td>3,080</td>
<td>14.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-rated healthb</td>
<td></td>
<td></td>
<td></td>
<td>3.4 (0.97)</td>
<td>0.52</td>
</tr>
<tr>
<td>SF-36 mental health scorec</td>
<td></td>
<td></td>
<td></td>
<td>4.7 (2.47)</td>
<td>1.44</td>
</tr>
<tr>
<td>SF-36 physical functioning scored</td>
<td></td>
<td></td>
<td></td>
<td>3.08 (1.14)</td>
<td>0.6</td>
</tr>
<tr>
<td>Physical activity levele</td>
<td></td>
<td></td>
<td></td>
<td>4.6 (1.55)</td>
<td>1.01</td>
</tr>
<tr>
<td>Alcohol consumptionf</td>
<td></td>
<td></td>
<td></td>
<td>4.34 (2.00)</td>
<td>0.81</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>89,501</td>
<td>16,851</td>
<td>79.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>23,002</td>
<td>5,932</td>
<td>20.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood disadvantageg</td>
<td></td>
<td></td>
<td></td>
<td>5.46 (2.90)</td>
<td>1.01</td>
</tr>
<tr>
<td>Remoteness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major city</td>
<td>69,158</td>
<td>13,262</td>
<td>61.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner regional</td>
<td>28,115</td>
<td>5,726</td>
<td>25.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer regional</td>
<td>13,023</td>
<td>2,868</td>
<td>11.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote</td>
<td>1,794</td>
<td>473</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very remote/migratory</td>
<td>413</td>
<td>114</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood dissatisfactionc</td>
<td></td>
<td></td>
<td></td>
<td>2.08 (1.75)</td>
<td>1.18</td>
</tr>
<tr>
<td>Neighborhood problemsb</td>
<td></td>
<td></td>
<td></td>
<td>2.68 (0.59)</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Abbreviations: SD, standard deviation; SF-36, Short Form-36.

* 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 standard deviations, and within-person standard deviations.
To empirically test the validity of the fixed-effect regression models in the present context, we included additional outcome measures of self-reported neighborhood dissatisfaction (“How satisfied are you with the neighborhood you live in?”: 0 = totally satisfied, 10 = totally dissatisfied) and perceived neighborhood problems (9 items each rated on a 5-point scale, with higher values indicating more perceived problems, such as noise, vandalism, and hostile residents). If the fixed-effect regression models can accurately capture individual variations that accompany neighborhood changes, the within-person analysis should yield support for causal neighborhood associations at least for the measures of neighborhood satisfaction and neighborhood problems, because these are expected to be sensitive to within-person changes across locations.

**Statistical analysis**

Associations between neighborhood characteristics and health were assessed with random-intercept multilevel models to take into account the nonindependence of repeated measurements of the same persons over time (linear regression for continuous variables and logistic regression for dichotomously coded variable of smoking). The total regression coefficient is estimated as a weighted average of both between-person and within-person variations in the exposure associated with the outcome (29). With repeated measurements, these 2 components can be estimated separately with the linear regression model \( y_{it} = \alpha + \beta_0 + \beta_w(x_{it} - \bar{x}_i) + \beta_B x_i + \epsilon_{it} \), where \( \alpha \) is the overall intercept, \( \beta_0 \) is the participant-specific intercept, \( x_{it} \) is the exposure variable for the \( it \)th participant at the \( it \)th measurement time of the participant, \( \bar{x}_i \) is the mean value of the exposure variable averaged across all measurement times separately within each participant, and \( \epsilon_{it} \) is the error term. Then the regression coefficient \( \beta_w \) gives the within-person (or fixed-effect) estimate and \( \beta_B \) gives the between-person estimate. The difference between the total and within-person regression coefficients was tested using the Wald test (30). Robust estimation with household clustering was used in all models to account for the nonindependence of household members.

Neighborhood associations may require long exposure periods to develop, in which case short residence times only add unnecessary noise in the data. To test this, the above analyses were repeated by including only person-observations of subjects who had the same level of neighborhood disadvantage (or remoteness) in at least 3 consecutive survey years. The neighborhood associations were thus assessed only when the participant had lived 3 or more years in both the old and the new location. For example, a person who had lived in a neighborhood with a disadvantage score of 5 for 3 years and then moved to a neighborhood with a disadvantage score of 7 for 4 years would contribute 1 person-observation from the first neighborhood (the last year in that neighborhood) and 2 person-observations from the second (the 2 last years in that neighborhood). All the regression models were recalculated with this reduced data set.

Social selection in neighborhood associations arises if persons who move to more affluent neighborhoods have better health than those who move to poorer neighborhoods. To test this, logistic regressions were fitted among persons for whom the level of neighborhood disadvantage changed between 2 consecutive follow-up waves, so that each participant could contribute more than 1 person-observation to the data set. The outcome was the direction of change in neighborhood disadvantage between follow-up waves coded dichotomously as 0 (move to less disadvantaged neighborhood) or 1 (move to more disadvantaged neighborhood). To establish correct temporal ordering for social selection hypothesis, health covariates were assessed at data-cycle baselines, that is, 1 study wave before the move. Each variable was assessed in a separate model that was adjusted for sex, age, and baseline neighborhood disadvantage. The corresponding analyses were fitted for remoteness. Because the present focus was not on the associations between health and overall residential mobility, persons who remained in the same location over time (or for whom the level of neighborhood disadvantage did not change when moving) were not included in this analysis.

The Short Form-36 scales were negatively skewed. The mental health scale was transformed using cubic transformation and then divided by 100,000 to have a range between 0 and 10. The physical functioning scale was too heavily skewed to be transformed to even close to a normal distribution, so the scale was recoded into 4 categories (0–60 = 1; 60–80 = 2; 80–90 = 3; and 90–100 = 4). Using the original scale, the mean was 83.5 and standard deviations were 23.0 across all person-observations, 21.5 between persons, and 12.0 within persons. The corresponding values for the recoded variable were 3.08, 1.14, 1.03, and 0.60, respectively, indicating that the ratio of within-person variation to total variation in the recoded scale (0.60:1.14) remained very similar to the original scale (12.0:23.0). The correlations of the original and recoded scales with other covariates remained largely unchanged (data not shown).

**RESULTS**

Of the 20,012 participants, 4,284 (21%) lived in 2 or more different neighborhoods with different levels of disadvantage, 1,679 (8%) lived in 3, and 774 (4%) lived in 4; 2,038 (11%) lived in 2 neighborhoods with different levels of remoteness and 191 (1%) lived in 3 or more. These were the participants who contributed data to estimate the within-person associations in the fixed-effect models. Residential stability was high, as indicated by intraclass correlations of 0.85 for neighborhood disadvantage and 0.89 for remoteness. Over the 10-year follow-up, there were 11,992 moves across levels of neighborhood disadvantage between consecutive follow-up waves (6,078 moves to more disadvantaged neighborhoods and 5,914 moves to less disadvantaged neighborhoods) and 3,032 moves across levels of remoteness (1,517 moves to more remote neighborhoods and 1,515 to less remote neighborhoods). At the level of person-observations, the correlation between neighborhood disadvantage and remoteness was 0.39.

**Social causation**

The magnitudes of the total, between-person, and within-person regression coefficients of the multilevel models are illustrated in Figure 1 (see Web Table 1 for numerical details).
Neighborhood disadvantage was associated with poorer health and health behaviors, except for alcohol consumption, which was less frequent in neighborhoods with higher disadvantage. However, these associations were largely due to variation between persons; the within-person associations were substantially weaker and mostly statistically nonsignificant. Only the association between neighborhood disadvantage and lower alcohol consumption was observed in the within-person analysis. All the differences of within-person and between-person coefficients were statistically significant (\( P < 0.05 \)). Including only person-observations from residences at which subjects had lived for 3 years or more strengthened some of the overall associations but did not change the conclusions on the dominating role of between-person associations over the within-person associations (Web Table 1).

Neighborhood remoteness was associated with poorer self-rated health, poorer physical functioning, and higher probability of smoking but also with better mental health and higher level of physical activity (Figure 2; Web Table 2). Again, these associations were largely between-person associations, and only the association between remoteness and higher physical activity level was replicated in the within-person analysis. For mental health, the within-person coefficient was not significantly different from the between-person coefficient (\( P = 0.21 \)), which suggests that the total association should be taken as the most efficient estimate. Adjusting the associations of neighborhood disadvantage and remoteness for time-varying indicators of educational level, marital status, and household income attenuated most of the total and between-person associations to some degree but had negligible influence on the within-person associations (Web Tables 3 and 4).

To confirm that the lack of within-person associations for health outcomes was not due to methodological artifacts that would have precluded the demonstration of true within-person associations, the above models were fitted for neighborhood dissatisfaction and perceived neighborhood problems. Neighborhood disadvantage was associated with higher neighborhood dissatisfaction and neighborhood problems, and these associations were replicated in within-person analyses (Figure 1). The within-person associations were stronger than the total or between-person associations. Remoteness was associated with lower levels of neighborhood dissatisfaction and problems, and these associations were also replicated in the within-person analyses with stronger magnitudes compared with the total and between-person associations (Figure 2). These associations provided support for the validity of the within-person regression in the present context.
Social selection

Compared with those who moved to less disadvantaged neighborhoods between follow-ups, persons who moved to more disadvantaged neighborhoods had poorer self-rated health, mental health, and physical functioning, had lower physical activity levels, were more likely to smoke, and were less likely to use alcohol in the study wave preceding the move (Table 2). Persons who moved to more remote neighborhoods had lower self-rated health than did those who moved to less remote neighborhoods, but no other associations with health or health behaviors were observed for remoteness.

DISCUSSION

Evidence from a 10-year prospective cohort study of more than 20,000 participants with annual repeated measurement data suggests that most of the associations between neighborhood disadvantage and health outcomes represent differences between persons rather than dynamic processes within persons. People living in disadvantaged neighborhoods of Australia had poorer mental and physical health than did those living in advantaged neighborhoods. However, a person was not markedly healthier when living in an advantaged neighborhood than when living in a disadvantaged neighborhood at a different time. The results were more heterogeneous for neighborhood remoteness, but between-person associations were nevertheless more important than within-person associations. These findings provide little support for social causation as the explanation for associations between neighborhood characteristics and health outcomes.

The large number of participants and person-observations from 10 measurement times afforded a large sample size to estimate the within-person regressions with sufficient precision, so their interpretation was not hampered by the wide confidence intervals that are often encountered in fixed-effects models. Neighborhood characteristics were assessed with 2 different variables based on objective measures of the person’s residential location defined with the accuracy of SLAs (roughly comparable to census tracts in the United States), which provided relatively detailed data on participants’ residential locations. Health status and health behaviors were assessed using multiple measures, with converging results. The lack of within-person associations in health outcomes was unlikely to be a methodological artifact, as the validity of fixed-effect regression analysis was supported by within-person associations observed for neighborhood dissatisfaction and perceived neighborhood problems. Despite these strengths, the results need to be interpreted taking into account...
consequences of selective residential mobility (31, 32). This suggests that some of the neighborhood correlations with health may be the result of people moving to different neighborhoods. People who moved to more disadvantaged neighborhoods had poorer mental and physical health, were more likely to smoke, and were less physically active. Thus, at least part of the associations between neighborhood characteristics and health status may in fact be mediated by direct mechanisms related to health (e.g., more educated persons having better health and being more likely to move to advantaged neighborhoods). The drift of persons to neighborhoods that match their health status may influence the development of health differentials between residential areas—people create neighborhoods. Gentrification and “white flight” are examples of person-level selective processes that modify neighborhoods (33).

Similarly, neighborhood health differentials may represent the downstream consequences of the more fundamental causes of health inequalities (34), that is, socioeconomic resources that determine people’s ability to select residential locations. These processes may even extend to intergenerational continuities, as recent studies have suggested that there is moderate stability in residential characteristics between parents and their children (35, 36). The present results do not yet tell us much about the extent to which social selection may help explain the emergence of health variations across neighborhoods (37). The overall impact of selective mobility depends not only on the magnitude of associations between health and residential mobility but also on specific migration patterns, such as total numbers of people migrating between areas. More detailed spatial modeling is needed to evaluate the plausible long-term associations of selective residential mobility on neighborhood health differences (38).

The present findings call into question the causal interpretation of neighborhood effects in health outcomes. However, the results do not necessarily imply that all neighborhood associations identified in previous studies are not causal. First, the present sample included mostly adults, and neighborhood associations may have a different impact on children and adolescents (6). Some of the noncausal associations might reflect long-term intergenerational continuities in neighborhood disadvantage and poor health that originate in childhood, in which case time-varying associations in adulthood would not be expected (36). Second, neighborhood associations on other outcomes besides health (e.g., criminal behavior or school performance) may be causal even if associations with health and health behaviors are not (39). Third, the assessment of causality in neighborhood associations may depend on various methodological choices, such as the specific measures of neighborhood qualities, measures of health outcomes, geographical level of analysis, and country-specific factors (3, 40–43) that need to be examined in more detail.

Reviews of studies of neighborhoods and health (4, 10, 15) have repeatedly emphasized the problem of deriving causal inferences from cross-sectional studies that tend to dominate the neighborhood research literature (1, 2). Despite the crucial importance of this methodological problem, surprisingly few studies have used longitudinal data and appropriate panel-study methods to assess whether people’s health varies as they move across different neighborhoods. The present findings from fixed-effects regressions suggest that neighborhood associations do not operate within persons but rather reflect stable differences between persons who live in different neighborhoods. This is in contrast to what one would expect if neighborhood disadvantage were causing poor individual health. Future studies of neighborhood associations need to consider more carefully the role of selective residential mobility as a potential mechanism causing geographic health inequalities.

**Table 2.** Associations Between Baseline Covariates and Subsequent Moves to Neighborhoods With Higher Disadvantage (Model 1) or Higher Remoteness (Model 2) in the Household, Income, and Labour Dynamics in Australia Survey, 2001–2010

<table>
<thead>
<tr>
<th>Predictor Variablea</th>
<th>Neighborhood Disadvantageb OR 95% Cl</th>
<th>Neighborhood Remotenessc OR 95% Cl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-rated health</td>
<td>0.85 (0.82, 0.89)</td>
<td>0.88 (0.86, 0.91)</td>
</tr>
<tr>
<td>Mental health</td>
<td>0.96 (0.94, 0.97)</td>
<td>0.98 (0.96, 1.01)</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>0.84 (0.81, 0.88)</td>
<td>0.99 (0.98, 1.12)</td>
</tr>
<tr>
<td>Physical activity level</td>
<td>0.97 (0.94, 0.99)</td>
<td>1.02 (1.00, 1.04)</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>0.95 (0.93, 0.97)</td>
<td>0.99 (0.93, 1.05)</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.67 (1.52, 1.82)</td>
<td>1.17 (0.92, 1.48)</td>
</tr>
<tr>
<td>Neighborhood dissatisfaction</td>
<td>0.96 (0.94, 0.98)</td>
<td>1.05 (1.00, 1.10)</td>
</tr>
<tr>
<td>Neighborhood problems</td>
<td>0.90 (0.83, 0.98)</td>
<td>1.28 (1.03, 1.61)</td>
</tr>
</tbody>
</table>

Abbreviations: Cl, confidence interval; OR, odds ratio.

a All associations are from logistic regression models fitted separately for each predictor variable and outcome and adjusted for sex, age, and baseline neighborhood disadvantage (or remoteness).

b Outcome was coded as 0 (move to neighborhood with lower disadvantage) or 1 (move to neighborhood with higher disadvantage); n = 11,992 person-observations from 6,737 participants who moved to different neighborhoods.

c Outcome was coded as 0 (move to less remote neighborhood) or 1 (move to more remote neighborhood); n = 3,032 person-observations from 2,229 participants who moved to different neighborhoods.

Account some methodological limitations. The analysis was restricted only to one country, so it is uncertain how the findings from Australia generalize to other countries in which neighborhood influences and patterns of selective residential mobility may be different. All health information was based on self-reported data, so potential reporting biases related to neighborhood characteristics might have confounded the results.

In contrast to the limited evidence for within-person neighborhood associations, the alternative hypothesis of social selection received some support in the case of neighborhood disadvantage but not in the case of neighborhood remoteness. Compared with those who moved to more advantaged neighborhoods, people who moved to more disadvantaged neighborhoods had poorer mental and physical health, were more likely to smoke, and were less physically active. Thus, at least some of the neighborhood correlations with health may be the consequences of selective residential mobility (31, 32). This may be mediated by direct mechanisms related to health (e.g., poor health making it more difficult to move) and by indirect mechanisms of sociodemographic factors correlated with health (e.g., more educated persons having better health and being more likely to move to advantaged neighborhoods).

The drift of persons to neighborhoods that match their health status may influence the development of health differentials between residential areas—people create neighborhoods. Gentrification and “white flight” are examples of person-level selective processes that modify neighborhoods (33). Similarly, neighborhood health differentials may represent the downstream consequences of the more fundamental causes of health inequalities (34), that is, socioeconomic resources that determine people’s ability to select residential locations. These processes may even extend to intergenerational continuities, as recent studies have suggested that there is moderate stability in residential characteristics between parents and their children (35, 36). The present results do not yet tell us much about the extent to which social selection may help explain the emergence of health variations across neighborhoods (37). The overall impact of selective mobility depends not only on the magnitude of associations between health and residential mobility but also on specific migration patterns, such as total numbers of people migrating between areas. More detailed spatial modeling is needed to evaluate the plausible long-term associations of selective residential mobility on neighborhood health differences (38).

The present findings call into question the causal interpretation of neighborhood effects in health outcomes. However, the results do not necessarily imply that all neighborhood associations identified in previous studies are not causal. First, the present sample included mostly adults, and neighborhood associations may have a different impact on children and adolescents (6). Some of the noncausal associations might reflect long-term intergenerational continuities in neighborhood disadvantage and poor health that originate in childhood, in which case time-varying associations in adulthood would not be expected (36). Second, neighborhood associations on other outcomes besides health (e.g., criminal behavior or school performance) may be causal even if associations with health and health behaviors are not (39). Third, the assessment of causality in neighborhood associations may depend on various methodological choices, such as the specific measures of neighborhood qualities, measures of health outcomes, geographical level of analysis, and country-specific factors (3, 40–43) that need to be examined in more detail.

Reviews of studies of neighborhoods and health (4, 10, 15) have repeatedly emphasized the problem of deriving causal inferences from cross-sectional studies that tend to dominate the neighborhood research literature (1, 2). Despite the crucial importance of this methodological problem, surprisingly few studies have used longitudinal data and appropriate panel-study methods to assess whether people’s health varies as they move across different neighborhoods. The present findings from fixed-effects regressions suggest that neighborhood associations do not operate within persons but rather reflect stable differences between persons who live in different neighborhoods. This is in contrast to what one would expect if neighborhood disadvantage were causing poor individual health. Future studies of neighborhood associations need to consider more carefully the role of selective residential mobility as a potential mechanism causing geographic health inequalities.

**ACKNOWLEDGMENTS**

Author affiliations: Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland (Markus Jokela); and...
Department of Psychology, University of Cambridge, Cambridge, United Kingdom (Markus Jokela).
This work was supported by the Kone Foundation (grant 31-225) and the Academy of Finland (grant 268388).
Conflict of interest: none declared.

REFERENCES


