
MICHAEL BUZZELLI
Department of Geography, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2 (e-mail: buzzelli@geog.ubc.ca)

JASON SU
Department of Geography, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2 (e-mail: jasonsu@geog.ubc.ca)

NHU LE
Department of Statistics, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2 (e-mail: nle@bccrc.ca)

TENNY BACHE
Department of Geography, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2 (e-mail: tbache@geog.ubc.ca)

This paper lays the foundation for a research program concerned with the geographical patterning of environmental and population health at the urban neighbourhood scale. Based on the Vancouver metropolitan region, the aim is to better understand the role of neighbourhoods as epidemiological spaces where environmental and social characteristics combine as health processes and outcomes at the community and individual levels. With respect to procedure, this paper builds a cohort of commensurate neighbourhoods across all six census periods from 1976 to 2001, assembles neighbourhood air pollution (total particles) data, and provides an initial analysis to demonstrate how air pollution systematically and consistently maps onto neighbourhood socio-economic markers, specifically education and family status. We conclude with a discussion of how the neighbourhood cohort can be further developed to address emergent priorities in

Cet article jette les bases d'un programme de recherche qui porte sur la structuration géographique de la santé environnementale et humaine à l'échelle du voisinage urbain. À partir du cas de la région métropolitaine de Vancouver, l'objectif est de mieux comprendre le rôle des quartiers comme espaces épidémiologiques dans lesquels se conjuguent les aspects environnementaux et sociaux qui constituent les processus et les résultats en matière de santé tant au niveau collectif qu'individuel. Quant à la méthode utilisée dans le cadre de cette étude, une cohorte de quartiers comparables est constituée à partir de chacune des six périodes de recensement de 1976 à 2001. Elle réunit des données sur la pollution atmosphérique locale (particules totales en suspension) et, au moyen d'une analyse exploratoire, montre comment la pollution atmosphérique concorde de manière systématique et consistante avec les marqueurs
the population and environmental health literatures, namely the need for temporally matched data, a life-course approach and analyses that control for spatial scale effects.

Introduction

As students of health and health care, geographers have infused health research with a range of perspectives, approaches and techniques. An important contribution to the wider health community has been to shed light on the geographical patterning of disease, including the influence of place on health. The purpose of this paper is to lay the foundation for a long-term project concerned with the nature of places as evolving contexts of individual and population health.

A number of as yet disparate trends in social science health research lead us to consider the influence of neighbourhoods over health through time. First, the health inequalities literature features the role played by space and place in patterning and influencing health outcomes. Much of this work is at the neighbourhood scale, yet neighbourhoods have been used mainly to model health outcomes cross-sectionally, at single points in time. How do neighbourhoods themselves change in relative and absolute terms and how does their tracking present new research opportunities? Second, and related, is the emergent marriage of population and environmental health, two hitherto disparate literatures increasingly concerned with health and socio-economic inequalities. How do ambient health hazards and social stratification, separately and in combination, influence health? More generally, how do neighbourhood context and characteristics of residents interact to generate individual and community health outcomes? Focused on Vancouver, Canada, from 1976 to 2001, the objective is to develop a ‘cohort’ of neighbourhoods as the basis for a range of spatial epidemiologic study designs that incorporate explicitly the evolving spatial contexts of health.

Unlike other large cities, Vancouver’s air quality record is very favourable: despite a 60 percent increase in total population and associated economic and transportation activities since 1976, all ambient criterion pollutants have steadily declined (GVRD 2003). It would seem Vancouver continues to earn its reputation as Canada’s urban Eden (Wynn 1992). However, data from the regional air quality monitoring network—among the best the world over in terms of number and distribution of monitoring stations (Figure 1)—and local studies underscore an emergent insight from air pollution epidemiology (Brauer et al. 2002; Henderson et al. 2004): that intra-urban variability in air pollution ranges as widely if not more so than between-city contrasts (Briggs et al. 1997; Brunekreef et al. 1997; Hoek et al. 2001; Brauer et al. 2003). Such variability within the region creates the potential for unequal exposures along neighbourhood socio-economic lines. How do these processes play out among Vancouver’s neighbourhoods of affluence, despair and ethnocultural clustering?

This project is built retrospectively from 1976 to 2001 by developing a geodatabase of neighbourhoods in the Vancouver CMA. These neighbourhoods will be followed prospectively with each census period (2006 onwards) in the future (Figure 1). A commensurate set of 195 census tracts, or neighbourhoods, are constructed from varying numbers of tracts in each of the six censuses over the study period. Our approach is to explore the relationships between ambient health hazards and socio-economic status (SES) as defining features of neighbourhoods. We focus on air pollution and a range of markers of socio-economic status to highlight the neighbourhood-level processes that can generate health outcomes.
among their residents, as individuals and communities, and point the way forward for health studies.

**Population and Environmental Health: Gaps and Opportunities**

**Health inequalities**

The literature on health inequalities grows out of foundational studies that demonstrated a correlation between health status and social structure. Among the most important, the ‘Whitehall studies’ of the British civil service (Marmot et al. 1978, 1991) and the Black Report (DHSS 1980) popularized the correlation between occupational class and health, showing, for example, that mortality rates drop significantly with each step up the occupational classification (unskilled, semi-skilled, skilled manual and non-manual, managerial, professional) of the British civil service. Since this early work the health inequalities literature has mushroomed.

At an international scale health inequalities follow a health-and-development pattern such that more developed market economies exhibit better health indicators (e.g., longer life
expectancy, lower infant mortality rates). But beyond a certain level of development, little is added to a population's health: life expectancy rises with wealth gains in poor countries but in developed economies wealth gains bring diminishing—almost flat—health benefits (Preston 1975; Rodgers 1979). Rather, it is social inequalities within developed countries that correlate with health. For example, Whitehall-I (Marmot et al. 1978) found that the relative risk of mortality by coronary heart disease among clerical workers was 3.2 times that of administrative workers in the British civil service (for which occupational classes and health data constitute a robust cohort) over several decades up to the 1970s. Using aggregated data, rather than individuals, Kaplan et al. (1996) and Kennedy et al. (1996) were among the first to demonstrate the relationship between all-cause and cause-specific mortality and income inequality at the state level in the United States. Canadian research has also played a fundamental role in formalizing the population health perspective, particularly through the Canadian Institute for Advanced Research (Evans et al. 1994).

Growing out of this corpus of research is the central tenet of the health inequalities literature, namely that developed societies exhibit a SES-health gradient whereby population health status diminishes with rises in socio-economic status (Feinstein 1993). In the Whitehall studies, only 28 percent of the raised incidence of coronary heart disease (CHD) among clerical workers was explained by the traditional risk factors of cholesterol, smoking and blood pressure; the remainder was ‘unexplained’. Further, life expectancy is not the only measure of a population’s health: a range of diseases, including cardiovascular disease and cancer—the two leading causes of death in the developed world—underscore the SES-health gradient (Wilkinson 1996).¹

Seemingly plausible explanations of the SES-health gradient have been offered, but refuted, in a number of powerful statements on population health inequalities. Individual health risk behaviours, health care, genetic predisposition, level of economic development and material deprivation are not the principal reason why we see the SES-health gradient within developed countries (Rose 1992; Evans et al. 1994; Wilkinson 1996; Davey Smith 2003). Instead, some have argued that social structure is itself a ‘fundamental cause’ of a population’s health (Link and Phelan 1995). This line of argumentation has spawned a number of responses and debate (Kaplan and Lynch 1997; Evans and Kantrowitz 2002) but agreement has settled around the influences of psychosocial stress and social cohesion (Evans et al. 1994; Wilkinson 1996). Seen as mediators in the SES-health relationship, one’s social comparison with others, sense of social cohesion and social capital—all feeding into socially stratified psychosocial stress—produce measurable biological markers in human development and endocrine and immune system response to day-to-day life (Brunner 2000). What is important for our purpose is that the SES-health gradient is necessarily social, not individual.

Within this context, the geographical contribution to this literature has been to reveal the health inequalities between places within societies (Gatrell 2002; Curtis 2004 provides good overviews of the geographical patterning of disease and health care). Thus the Kaplan et al. (1996) and Kennedy et al. (1996) studies noted above demonstrated health and income inequality correlations across U.S. states. The Canadian east-to-west gradient in health status across provinces and metropolitan areas is well established (Gilmour 2004). At the local scale Ross et al. (2000) compared the relationship between age-adjusted all-cause mortality and income inequality (percent income of poorest 50 percent of households) between Canadian and U.S. jurisdictions in 1990. In general, mortality and income inequality were both lower in Canada.

¹ It is important to note here, however, that in Canada breast and prostate cancer do not follow the SES-health gradient and much of the SES-cancer gradient is associated with behaviourally linked cancers such as lung and cancers of the digestive system. See Wilkins (2002).

² Some kinds of health care can contribute to improved health for some groups, such as primary care among American urban whites, though notably not African Americans. Specialty care does not bear the same effect. A more important general observation is that access to and use of health care is stratified similarly to health outcome measures (Shi 1999). But in general international health care provision comparisons (health care spending per capita) show little influence over population health.
than in the United States and a significant relationship was found only for U.S. metropolitan areas. Wilkins et al. (2002) examined changes in mortality by income quintiles among Canadian metropolitan neighbourhoods (census tracts) from 1971 to 1996. Based on age-standardized mortality rates of major causes of death as well as life expectancy at birth, potential years of life lost (PYLL), and income-related excess PYLL before the age of 75 years, the disparity in disease between the richest and poorest neighbourhoods diminished for most health measures though nearly one-quarter of the PYLL in 1996 could still be attributed to income inequalities between neighbourhoods. 3

Beyond spatial patterning, geographical research has, perhaps more uniquely, elucidated the potential influence of place and space on population health. Do neighbourhoods influence health independently of the attributes of individuals, and to what degree? Do residents of the same SES (ceteris paribus), but located in different neighbourhoods, display different health statuses? Geographers have sought answers to these sorts of questions via multi-level modelling (MLM) (Macintyre et al. 1993; Duncan et al. 1998). MLMs are a type of regression analysis in which observations are hierarchically nested to delineate the health effects of variables at different levels. The premise of MLM is to overcome both the atomistic fallacy of individual risk factor epidemiology and the ecological fallacy inherent in aggregated data, such as those summarizing neighbourhood populations (Schwartz 1994; Macintyre and Ellaway 2000).

MLM modelling has been applied to a range of health issues including age-stratified all-cause and cardiovascular mortality, morbidity studies of childhood asthma and low birth weight, self-rated health, hypertension and mental health (see Pickett and Pearl 2001 for an overview). The Canadian literature is less developed than elsewhere but studies show expected patterns: after controlling for individual and family-level characteristics, area-level (neighbourhood) features such as income and unemployment are associated—albeit only weakly—with such factors as health care utilization (Shen et al. 2003), child development (Kohen et al. 2002) and deprivation (Beland et al. 2002).

The health inequalities literature has produced a range of SES–health evidence, though it also comes with some caveats—issues we hope to address with a retrospective–prospective neighbourhood cohort. Descriptively, associations from place to place may be sensitive to the spatial scale of analysis. This is not necessarily a problem if we are concerned with processes at the population (ecological) level (Macintyre and Ellaway 2000) though appropriate steps should be taken to address the modifiable areal unit problem that underlies potential ecological fallacy (e.g., King 1997). This is particularly challenging in the health field, as it is dominated by a traditional focus on individual risk factors (cf. Morgenstern 1995; Wakefield 2001; Ross et al. 2004). In addition, within the MLM literature there is a need to ensure that data for individuals and their neighbourhood contexts are temporally matched since neighbourhood change, even within a few years time, can significantly alter social composition (Buzzelli and Su 2005).

Environmental health inequalities

The merger of environment and health with the health inequalities literature seems natural on several levels. Both share an explicit concern with questions of equity and fairness as well as the social processes that reproduce the SES–health gradient—in this case environmental health hazards. How do we bring environmental hazards into the fold? A narrow materialist explanation suggests that because air pollution—our focus here—is unevenly distributed it is an environmental health hazard that generates health inequalities. This all forms part of the wider health inequalities movement that includes emergent calls for a combined population and environmental health agenda (Eyles 1999; Macintyre and Ellaway 2000; Starfield and Paganini 2000; http://www.iseqh.org/en/index.html). Air pollution has figured centrally (Eyles 1997; O’Neill et al. 2003; Bell et al. 2004). Indeed, so influential is the health inequalities movement that mainstream environmental epidemiology has begun to

---

3 It bears noting here that, substantively, these Canadian studies run counter to the more typical pattern in market economies of widening health disparities congruent with widening social disparities, such as in the UK and U.S. (Hofrichter 2003). Nonetheless health and socio-economic status remain correlated in Canada.
consider the question of effect modification: how SES influences the air pollution dose–response relationship. Several recent studies have shown that socio-economic status can insulate advantaged individuals and communities from its impacts when status is high, while disadvantaged communities are relatively more susceptible (Krewski et al. 2000; Evans and Kantrowitz 2002). More specifically, explicit incorporation of SES may reveal several potential influences: (1) confounding, where the impact of a hazard is over-estimated (bias) but there is a lack of control for SES (model misspecification); (2) modifying (joint, relative), where the hazard might have an amplified effect among those of lower SES; (3) an independent influence, where the significance of a hazard is reduced once SES enters the picture (Bell et al. 2002; O’Neil et al. 2003).

Through these examples we see how population and environmental health are converging. But for the merger to be fruitful further methodological issues will have to be addressed. As in geographical research on health inequalities, the spatial scale of analysis is critical for exposure assessment, both of the hazard and the at-risk population (Bowen 2001). The new science of human exposure assessment (Ott 1995) comes with a long list of attempts to refine exposure assessment (see Buzzelli and Jerrett (2003) for a review and Most et al. (2004) for a recent example) and warnings about willy-nilly applications of new ‘turnkey’ computing techniques, including GIS (Maantay 2002). Some recent advances in exposure analysis pave a way forward for research at the neighbourhood scale (Briggs et al. 1997; Jerrett et al. 2003). Based on geocomputation, high-resolution spatial databases and small area socio-economic and health data, recent exposure assessment work has been validated and shown to associate with health effects at the city block scale (Brauer et al. 2003). This reduces exposure misclassification and the potential for confounding effects.

In addition to spatial scale, time is also important. Life course health research has shown that early childhood experiences, in particular, can condition susceptibility to disease in later life, either directly as biological embedding in this early critical stage or indirectly by raising susceptibility to cumulative life course exposures and/or adult exposures (Barker 1990; Power and Hertzman 1997). In an innovative MLM study, Davey Smith et al. (1997) applied a life course perspective to analyse the prevalence and risk of cardiovascular disease in 5,766 men in 27 workplaces in Scotland. Based on prospective data with 21 years of follow-up, they developed cumulative lifetime social status indicators and tested these against a range of risk factors. Notably, area-level (postal code sector in their study) variables were non-significant when cumulative status entered their models, leading the authors to argue that ‘studies with data on socioeconomic circumstances at only one stage of life are inadequate for fully elucidating the contribution of socioeconomic factors to health and mortality risk’ (Davey Smith et al. 1997, 547; see also Davey Smith 2003).

In addition to issues of temporal data match and spatial scale of analysis, cross-sectional health studies have largely missed the opportunity to uncover these life course effects. It is precisely these kinds of insights, based on socioeconomic and environmental exposures together, that we hope to gain with a neighbourhood cohort.

As a place to develop such a cohort, Vancouver presents both general and unique features. In the context of ongoing research concerned with the uniqueness of the Canadian city, urban social structure and emergent urban forms, Vancouver’s growth patterns and social geography are broadly typical (Goldberg and Mercer 1986; Bourne 1989; Davies and Murdie 1994), especially when compared with trends among the largest metropolitan areas (Broadway and Jesty 1998; Bourne and Lorius 1999). How might this associate with environmental health? Too little research exists for a clear answer. While Vancouver consistently ranks among the healthiest urban populations in Canada (Ross et al. 2000; Gilmore 2004; Ross 2004), air pollution may buck the trend. The region’s air pollution health effects adhere to the empirically observed linear, no-threshold pattern (Vedal et al. 2003) such that even low exposures result in adverse health impacts (Schwartz and Zanobetti 2000). Moreover, a recent study in Vancouver found a significant relationship between asthma hospitalizations and gaseous air pollutants (NO2 for males; SO2 for females), but only for those of lower SES (below the region’s 50th income percentile) (Lin et al. 2004). Can we develop a long-term project to address
these sorts of questions with matched neighbourhood life course environmental and SES data?

Data and Methods

Census data

Census tract data were drawn from the six Canadian national censuses conducted every 5 years from 1976 to 2001. In 1986 Statistics Canada initiated full quinquennial census enumeration to match the comprehensive decennial censuses: from 1986 onward each census has collected broadly similar and comprehensive information. The 1976 collection was therefore a 'mini-census' and does not have census tract data on average dwelling value, labour force occupational categories, government transfer payments (state income assistance) and immigration. The 1981 and 1986 census did not collect data on median household income and number of individuals/families below Statistics Canada’s low income cut-off. The census tract analyses and models presented, therefore, use a smaller set of variables from 1976 to 2001, for which family status and education variables are common, and a larger set of variables beginning in 1981 that include most income, education and housing value variables.

In order to compare cross-sectional time series data, a correspondence file of census tracts was created and a common set of 195 tracts for the entire study period was constructed (Table 1). The correspondence file is useful for reference information and provides a key to the spatial linkages among census tracts. For our purposes this file guided decisions on merging tracts through time to produce commensurate units for analysis. To produce a common set of census tracts is to reduce the number of tracts to as large a set as possible while maintaining both comparability across the study period and the integrity of tract boundaries/populations. The principle behind this exercise was to produce neighbourhoods that are internally consistent across the whole study period such that the population from one tract is not assigned to another tract in some other period. Since most new tracts are subdivisions of predecessors in a previous census (i.e., subdivided tracts maintain their perimeter boundaries) this criterion is easily met most of the time by aggregating from 1976 onward. In some cases, as shown in Table 1, notable boundary changes occurred: rather than a tract subdivision, these are cases where the location of a tract boundary was changed in a subsequent census to include part of another tract. In these cases decisions were made whether to: (1) exclude all affected tracts due to uncertainty about population assignment; (2) return to an earlier census and merge all affected cases; (3) continue with the merging procedure if, using period census tract and land use maps, no residential land use (thereby population enumerated in the census) was affected. Most boundary changes were handled by merging tracts to their boundaries of an earlier census (#2) or maintenance of boundaries of no residential land use was altered (#3).

Nearly all of the original 203 census tracts of 1976 were maintained for subsequent analysis as 195 commensurate neighbourhoods. For each of the 195 neighbourhoods, the raw census data were aggregated and any percentage figures were then produced. This builds in a weighting of the percentage observations based on the total population of the original tracts. For pre-existing percentage data such as unemployment rate the resulting average was also produced via weighting of the original tract populations. All dollar variables such as average income and dwelling values were adjusted to 2001 constant dollars using a standard consumer price index (Statistics Canada CANSIM Table 326-00011,2,3,4,5—Consumer price index [CPI] 2001).

Air pollution data

Air pollution data were drawn from a possible set of 42 high-volume air quality monitoring stations operated by the Greater Vancouver Regional District (GVRD) and other agencies in the Vancouver region from 1976 to 2001. Monitoring data include a range of criterion pollutants, air toxics and related meteorological data, though these vary by station and period. The focus here is...
Table 1
Vancouver census tract correspondence, 1976–2001

<table>
<thead>
<tr>
<th>Year</th>
<th>Original no. of CTs</th>
<th>No. of new CTs</th>
<th>Notable boundary changes</th>
<th>No. of tracts in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>203</td>
<td>n/a</td>
<td>n/a</td>
<td>195</td>
</tr>
<tr>
<td>1981</td>
<td>246</td>
<td>43</td>
<td>4</td>
<td>195</td>
</tr>
<tr>
<td>1986</td>
<td>273</td>
<td>27</td>
<td>2</td>
<td>195</td>
</tr>
<tr>
<td>1991</td>
<td>295</td>
<td>22</td>
<td>0</td>
<td>195</td>
</tr>
<tr>
<td>1996</td>
<td>298</td>
<td>3</td>
<td>0</td>
<td>195</td>
</tr>
<tr>
<td>2001</td>
<td>387</td>
<td>89</td>
<td>2</td>
<td>195</td>
</tr>
</tbody>
</table>

total suspended particles (TSP). TSP serves as a good marker for health-related pollutants, especially smaller particles (Kim and Jerrett 2000; Burra et al. 2002; Jerrett et al. 2005). TSP is used to build the first phase of this project and other pollutants such as PM$_{10}$ and NO$_x$ are being added to the database (note that these will require different interpolation procedures for ambient estimates, as discussed below, and therefore cannot be included in this study).

Geometric mean monthly average TSP data were provided in digital file form by the GVRD for January 1982 through December 2001, inclusive. Monthly data were key-coded from published reports for 1976 to 1981. All data were combined into a database and data gaps at the month scale necessitated aggregation to annual averages (using arithmetic mean). Once aggregated, a subset of 32 stations was deemed usable (due to elimination of large data gaps and removal of stations that changed location over time) to estimate annual average ambient TSP concentrations for the 195 neighbourhoods/census tracts for each year of the study period (every year from 1976 to 2001). Figure 2 shows estimated TSP values for each census year.5

The choice of an interpolator to estimate neighbourhood ambient annual average TSP was guided by the non-stationary characteristic of the pollutant field. In similar work in Hamilton, Ontario, a universal kriging algorithm was suitable for estimation, perhaps because the city’s northeast end steel mills significantly influence the shape of the ambient concentration surface (Jerrett et al. 2001). In Vancouver, by contrast, a robust variogram estimation with several bandwidths, tolerances and maximum distances produces a very weak semi-variance structure. Vancouver region’s sparse sources (particularly mobile sources) and atmospheric variability create a non-stationary concentration process requiring a more nuanced interpolator, one that can account for the spatial variability of correlations between monitoring stations. A spatio-temporal interpolator was used to take advantage of the annual average temporal data available so that spatial and temporal dependence could be incorporated into the estimation procedure (Le et al. 2001). The end result is an annual average TSP surface for each year of the study period, yielding estimates for each neighbourhood that were incorporated into the geodatabase.

Neighbourhood cohort geodatabase

A temporal and spatial geodatabase is used to merge the estimated TSP exposure data with the socio-demographic data for the 195 neighbourhoods across all six censuses (Table 2). In so doing we can visualize and explore relationships between neighbourhood chronic air pollution exposure and standard markers of urban social structure. In the following section we lay the foundation for a SES-environmental health neighbourhood cohort by illustrating how change in air pollution over time associates with the socio-economic make-up of Vancouver neighbourhoods.6

5 Unless otherwise noted all maps and spatial data were processed with ArcGIS 9.0, ESRI Inc., Redlands, California.

6 Unless otherwise noted, all statistical analyses were undertaken with SPSS v. 12 (SPSS Inc., Chicago, Illinois).
Establishing a Neighbourhood-based Cohort

The data summary in Table 2 helps to set out the analytical framework of air pollution and SES in the Vancouver region from 1976 to 2001. One important trend is the marked reduction in TSP, a decrease of more than 50 percent in mean regional values over 25 years owing to higher emission standards (transportation) and a reduction in industrial emissions due to an eroded manufacturing base (mainly in the forestry sector). This reduction was accompanied by a decrease in neighbourhood-to-neighbourhood disparities,
indicating not only an absolute regional decline but a rise in equity between the most- and least-exposed neighbourhoods. These reductions are remarkable when the more than doubling of the region’s population is taken into account. These kinds of trends led Steyn and others (Steyn et al. 1992, 267) to write: ‘To the untrained eye, Vancouver is one of those rare cities which seems to have achieved a delicate balance between urban development and scenic preservation—a balance more characteristic of small settlements than of major metropolitan areas’. What does a closer look reveal?

Beginning in 1976 we see a number of expected socio-economic trends. Unemployment follows the ebb and flow of the local economy such that unemployment rates declined in the boom years of the late 1970s/early 1980s and rose in the slower years following. Rising educational attainment and new family forms are also evident in the data. Both of these variables change dramatically over the study period, opening the door to longer-term (in)equity associations with air pollution. Do trends in education and family form map onto air pollution in some systematic way? For the data series starting in 1981, the income/wealth variables remain relatively stable through time (in adjusted real dollars) but the immigration variable ranges widely and may highlight spatially variable relationships with TSP. In this context how does diminished regional air pollution map onto the changing social geography of Vancouver?

We can answer this question in a number of ways. Descriptively, we may compare the percentile groupings of neighbourhoods classed

### Table 2
Descriptive TSP and SES data, Vancouver CMA, 1976–2001

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TSP (annual average $\mu g/m^3$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>53</td>
<td>47</td>
<td>35</td>
<td>31</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Range</td>
<td>26</td>
<td>32</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5,944</td>
<td>6,361</td>
<td>6,922</td>
<td>8,138</td>
<td>9,248</td>
<td>10,031</td>
</tr>
<tr>
<td>Range</td>
<td>21,065</td>
<td>28,050</td>
<td>32,890</td>
<td>32,358</td>
<td>34,201</td>
<td>43,419</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>2,779</td>
<td>3,475</td>
<td>4,240</td>
<td>5,309</td>
<td>6,564</td>
<td>7,739</td>
</tr>
<tr>
<td>Average dwelling value</td>
<td>N/A</td>
<td>326,053</td>
<td>195,365</td>
<td>311,132</td>
<td>366,536</td>
<td>316,120</td>
</tr>
<tr>
<td>Mean</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>54,153</td>
<td>50,196</td>
<td>53,784</td>
</tr>
<tr>
<td>Range</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>875,154</td>
<td>499,995</td>
<td>850,978</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>121,207</td>
<td>78,557</td>
<td>135,068</td>
</tr>
<tr>
<td>Median household income</td>
<td>N/A</td>
<td>121,207</td>
<td>78,557</td>
<td>135,068</td>
<td>163,602</td>
<td>139,888</td>
</tr>
<tr>
<td>Mean</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>121,207</td>
<td>78,557</td>
<td>135,068</td>
</tr>
<tr>
<td>% Less than 9th grade education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Range</td>
<td>61</td>
<td>48</td>
<td>43</td>
<td>36</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Range</td>
<td>36</td>
<td>19</td>
<td>38</td>
<td>33</td>
<td>31</td>
<td>22</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>9</td>
<td>5</td>
<td>12</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>% Employed in manufacturing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Range</td>
<td>23</td>
<td>27</td>
<td>27</td>
<td>30</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>% Lone-parent families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Range</td>
<td>31</td>
<td>30</td>
<td>31</td>
<td>36</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>

1 Indexed to 2001 constant dollars. See text for details.

NOTE: All figures produced from population-weighted merges of census tract data for 195 neighbourhoods/tracts. See text for details.
Figure 3

according to their TSP exposures in each year (Figure 3). Do average TSP values of the lowest and highest quintile groups systematically discriminate neighbourhood SES spatially and temporally? With some nuances, the answer is ‘yes’. The SES series starting in 1976 are analogous to the reduction in absolute and relative neighbourhood TSP exposure. SES disparities diminished through time but the TSP percentile groupings do flesh out the region’s neighbourhood social geography according to educational attainment and lone-parent families. The 1981 series are more nuanced still: in general, the TSP groupings discriminate among neighbourhoods by SES, though immigrant resettlement in the 1980s blurred the boundaries (in fact reversed the expected pattern) while blue-collar employment initially flagged widening neighbourhood disparities and then regressed to the mean.

Model results
Modelling the data can aid in further exploring associations as well as in establishing the magnitude and nature of the TSP-SES relationships discussed thus far, particularly as these may strengthen and decline across space and time. Modelling may also point to critical points in time and space where exposures were significant for particular kinds of groups and thereby provided a rationale for subsequent studies focusing specifically on health outcomes.

As suggested by the relatively small within-year range and variability of TSP (Table 2), particularly as we come forward in time, cross-sectional models for each period produce only weak associations with SES markers: the most consistently significant association in multivariate ordinary least squares (OLS) models was with low education. The best model fit with TSP was for 1981 (adj. $R^2 = 17$ percent) when the education ($b = 0.002, p < 0.01$) and lone-parent families ($b = 0.002, p < 0.01$) were significant. By 2001 the best multivariate model, in which unemployment rate was the sole significant SES neighbourhood marker, explained only 5 percent of the

7 Models built via best subsets and manual forward selection based on highest correlation with TSP. For these and pooled analyses below, TSP was transformed to its natural logarithm (to reduce significant skew in the raw data); standard and spatial model diagnostics were applied to all models.
variation in TSP exposure. Such a low level of explained variance does not mean that we should dismiss unemployment in this period, nor significant covariates in any cross-sectional model: for extreme neighbourhood cases in particular, they may flag short-term social changes—critical health stages as noted earlier—that can be formative in life course health development. But our purpose is to cast a longer-term temporal view.

To adopt a longer-term view we first applied yearly intercept indicators for a pooled OLS regression analysis. Given the trends in TSP over the study period the indicators were by far the most significant in all models: using 1976 as the base year (similar results were obtained when only the 1981 series were used), predictors for each successive year capture the large absolute average neighbourhood decline in TSP exposure. Two substantive covariates entered the model with the expected sign to explain a relatively smaller proportion of the variance in TSP: low education ($b = 0.002$, $t = 8.57$) explained 3.5 percent of the change in TSP across the full range, and carried a corresponding relative risk of 1.03; lone-parent families ($b = 0.001$, $t = 3.14$) explained 12 percent, RR = 1.12. Standard model diagnostics including residuals analysis returned no significant violations of OLS assumptions though model fit (adj. $R^2 = 0.911$) is a strong clue that prediction is subject to spatial dependence. As we would expect, based on the trend lines in Figure 3, we find a high degree of spatial dependence in the OLS pooled model residuals: permutation Moran’s $I$ test statistic = 0.72, $p < 0.01$.

To compensate for the model bias created by this spatial dependence, a generalized additive model (GAM) was applied to the data. A loess smoother of the $(x, y)$ coordinates (5 percent smoothing span, or 10 nearest neighbours) of the census tract centroids was entered first, followed by each of the predictors in the OLS pooled model (as linear predictors). Order of entry was based on bivariate GAM models with the highest $F$-ratios. Each iteration was tested for improvement to $F$-ratios and the final model contained the same predictors as the OLS pooled regression (plus the centroid loess smoother). A test for spatial autocorrelation showed that inclusion of a local smoother in a GAM model significantly removed spatial dependence in model residuals (Moran’s $I = 0.07$, $p > 0.05$). While GAM models provide flexible fitting parameters they do not come with the same interpretability as OLS prediction. For our purposes the GAM model confirms the substantive importance of family status and education as markers of neighbourhoods more exposed to TSP.

Conclusions and Discussion

This paper lays the foundation for a retrospective–prospective project concerned with Vancouver neighbourhoods as contexts of environmental and social determinants of health. To lay this foundation the research is built on an SES and air pollution GIS of 195 commensurate neighbourhoods (CTs) from 1976 to 2001. By following these neighbourhoods and their residents through time we may develop a neighbourhood cohort for a range of study designs aimed specifically at the question of neighbourhoods as ‘epidemiological spaces’.

Against a backdrop of environmental and population health, two disciplines moving toward each other in recent years, the relationship between neighbourhood air pollution and SES was described and analysed. Several striking trends emerged. First, TSP air pollution declined significantly over the period, as did criterion pollutants, despite a more than doubling of the region’s population. Second, neighbourhood-to-neighbourhood air pollution disparities also decreased. Third, TSP-based quintiles consistently delimit neighbourhoods by their SES. Fourth, modelling confirmed that neighbourhood family and educational status markers were significantly associated with TSP even if this relationship became more subtle through time.

What do these results say about the relationship between TSP and neighbourhood SES over the study period? The urban literature tells us that lone-parent families and low education are important neighbourhood-level SES markers leading us to expect their association with TSP, itself an indicator of neighbourhood disamenity, through time. At the same time we should be cautious not to over-interpret this temporal

---

8 Based on a first-order, queen’s case contiguity matrix. All Moran’s $I$ tests run with the GeoDa Software package, Spatial Analysis Laboratory, University of Illinois.
correlation given the general rise in education levels and lone-parent families, suggesting the possibility of spuriousness with TSP given its own downward trend over the study period. In this context Vancouver may be a small window on the ‘Canadian paradox’: that rising urban social and health disparities in other countries may not be as pronounced—indeed may be declining—in the Canadian city (Ross 2004) though increasing spatial segregation may lead to greater SES–health disparities in the future (Ross et al. 2004). To address these issues further future research needs to include other air pollutants and, where possible, a wider range of environmental hazards, as described earlier. Analysis at multiple spatial scales would also be useful. A first step would be to test for sensitivity to spatial scale of analysis within the given years by contrasting covariate significance among the 195 neighbourhoods versus the given set of CTs in a census year (requiring a new layer of point interpolations for TSP and any other pollutants).

But to target our central aim—to identify and elucidate the influence of places on individual and community health—another key feature needed for neighbourhood cohort development is individual health and SES data.9 Individual health studies can be based on sampling designs informed by ecological studies like the present paper (Ayers 2002). For example, recent landmark air pollution health studies have shown that health impacts are greatest for individuals with lower education (e.g., Krewski et al. 2000). Might our first cut at a Vancouver neighbourhood cohort point to amplified health effects in neighbourhoods marked by residents of low education? Extending the statistical models provides some clues. By introducing interaction terms of our substantive variables and yearly indicators, we can test whether each variable (lone-parent families and low education) significantly captures neighbourhood TSP variation within a given year as well as over the whole period. Both the OLS and GAM models returned low education as significant across the whole period as well as within the 1991 and 1996 cross-sections (the same can be said for the interaction term of lone parents in 2001 though it took the negative sign which should not be surprising given the lone-parent trend lines in Figure 3). What this tells us is that less educated neighbourhoods not only experienced more TSP exposure across the whole period but that raised exposure was especially marked in these neighbourhoods in particular years. In this way it is possible to envision a spatio-temporal sampling frame for a retrospective-prospective study based on education as an individual filter/selection criterion.

With both the present findings and targets for development in mind, let us conclude with a more general statement of the potential for a neighbourhood cohort to throw new light on the place-based influences over individual and population/neighbourhood health. We can envision a number of study designs focused on individuals and neighbourhoods: in effect a $2 \times 2$ table (Figure 4) with individuals and neighbourhoods on both axes, intersecting to produce health processes and outcomes.

The research presented here speaks to neighbourhood-level exposure processes. With aggregated individual health data we may devise community health outcome profiles (e.g., Luginaah et al. 2000) as both research visualization and surveillance/policy tools. Individual data are important for studies of neighbourhood-level processes and outcomes because: (1) they help control for confounders such as smokers in

<table>
<thead>
<tr>
<th></th>
<th>Individuals</th>
<th>Neighbourhoods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals</td>
<td>Ind. risk factor epidemiology</td>
<td>Compositional effects</td>
</tr>
<tr>
<td>Neighbourhoods</td>
<td>Multi-level models</td>
<td>Community health profiles</td>
</tr>
</tbody>
</table>

Figure 4
A neighbourhood cohort approach

---

9 Fortunately in British Columbia the renowned BC Linked Health Database provides comprehensive individual, spatially resolved (mapable individual health and SES data at the 3 or 6 digit postal code level) administrative health records from 1986 forward. We have begun to use some of these data in parallel projects (and they have been used by many other groups in a range of projects in the province) but a full discussion of the data is beyond the scope of this paper (please refer to http://www.chspr.ubc.ca/Bclhd/aboutbclhd.htm).
exposure populations; (2) they help control for compositional effects since we are really interested in the influence of contextual effects—the role of uniquely neighbourhood-level features such as physical environmental quality or investment into public good. For instance, local research has shown that individuals’ housing circumstances significantly influence self-rated health and health status, including neighbourhood-to-neighbourhood differences (Dunn and Hayes 2000; Dunn 2002). For disease mapping of air pollution health effects, for example, controlling for the physical characteristics of housing provides a more realistic picture of the impacts of air pollution exposure.

At the individual level we can envision a series of multi-level studies that address the literature’s shortcomings discussed earlier, especially the need for temporally matched data and life course approaches. By enabling exposure assignment and SES in earlier periods, we may not only decompose individual and neighbourhood effects at the latest follow-up stage but also antecedent exposures and SES contributions. Assigning exposures throughout the life course is a basic necessity if we have to understand the influence of place over health, but we may take it one step further: Given spatially resolved individual health data (see note 9) the residential migration of individuals can be traced permitting exposure classification in intervening periods as well. In this way one could build in contrasts between, say, low-educated residential movers versus those who stay put. What is the subsequent health outcome at a later follow-up period? For a specified sub-population? Between movers and non-movers? Are those who moved to relatively less polluted areas better off at a later follow-up stage, controlling for age and changes in individual and neighbourhood status? Or is the current exposure the driving force behind observed health outcomes? Having a common set of neighbourhoods defined by exposure and SES markers is the first step to answering these questions.

Acknowledgements

The first author wishes to acknowledge funding support provided by the BC Medical Services Foundation and the Social Sciences and Humanities Research Council of Canada.

References

Barker, D. J. P. 1990 ‘Fetal and infant origins of adult disease’ British Medical Journal 301, 1111.
Davey Smith, G., Hart, C., Watt, G., Hole, D., and Hawthorne, V. 1997. Individual social class, area-based deprivation, cardiovascular...


ROSS, N. 2004 What Have We Learned Studying Income Inequality and Population Health? (Ottawa: Canadian Institute for Health Information)


ROSS, N., TREMBLAY, S., and GRAHAM, K. 2004 ‘Neighbourhood influences on health in Montreal, Canada’ Social Science and Medicine 59, 1485–1494


SCHWARTZ, J., and ZANOBETTI, A. 2000 ‘Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death’ Epidemiology 11, 666–672


SCHWARTZ, J., and ZANOBETTI, A. 2000 ‘Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death’ Epidemiology 11, 666–672


VEDAL, S., BRAUER, M., WHITE, R., and PETKAU, J. 2003 ‘Air pollution and daily mortality in a city with low levels of pollution’ Environmental Health Perspectives 111(1), 45–51

WAKEFIELD, J. 2001 ‘A critique of ecological studies’ Biostatistics 1, 1–18

