eAppendix

Data Sources

The British Columbia Linked Health Database (BCLHD) is a large data resource that contains multiple electronic health insurance linked administrative databases including provincial health insurance registration records, hospital separation records, outpatient medical service records, vital statistics records, and other health data sources. All these data sources are linked at an individual level using anonymous individual-specific study identifiers.[1]

The Medical Services Plan (MSP) of British Columbia governed by the Ministry of Health Services, Government of British Columbia, provides universal health care insurance for the residents in the province. Since enrolment in the health insurance plan is mandatory, almost all residents (over 95%) in the province registered with the plan except for a minority of the residents including federal employees and some First Nations people covered by Canadian federal health insurance.[2] MSP health insurance registration data provided socio-demographic data including date of birth, sex, and 6-digit residential postal code as well as MSP premium registration information such as start date, and days registered in each fiscal year.

Hospitalization data included all admission/discharge records from hospitals in British Columbia. This data source provided individual 6-digit residential postal code, hospital code, admission date, length of stay, separation date, discharge status (i.e. alive or dead), up to 16 diagnoses (ICD-9 before 2001) or 25 diagnoses (ICD-10 since 2001) to describe the reasons of a hospital admission, and types of diagnosis to explain the role of each diagnosis for a hospital admission and the length of stay. Principal diagnosis was the most responsible for a hospital admission and the length of stay.

Outpatient medical service data included medical service records provided by general practitioners, specialists, and some supplementary health care practitioners for outpatients registered with the MSP health insurance plan. This data source provided 6-digit residential postal code, service date, practitioner identifier, specialty code, service code, and one diagnosis code (ICD-9) for each medical service.

Death registration database from the British Columbia Vital Statistics Agency included all death records in the province. This data source provided the date of death, sex, 6-digit residential postal code, and the cause of death (ICD-10 code).

Due to privacy restrictions, residential address information was not available for study subjects; accordingly, 6-digit residential postal codes were used in exposure assessment. Those with the same residential postal code were assigned to the same residential exposure levels as the center of the geographic area (area centroid). In metropolitan Vancouver, a postal code represents one side of a city block in urban areas but may represent a larger area in less densely populated regions.

Residential History Construction

A longitudinal residential history during the exposure period and the follow-up period was constructed for each subject using their residential postal codes and associated dates recorded in provincial health insurance registry, hospitalization separation records, and outpatient medical service records. About 10% of these records contained invalid postal codes or nonresidential postal codes (such as post office boxes) and were excluded from residential history construction. For each study subject, about 114 (median; interquartile range 68-181) records were used to construct his/her residential history. When a subject changed his/her residence as indicated by changes in residential postal codes, transition was set as the midpoint between two dates if there was no overlap, or at the first date of the new postal code if there was overlap.

Self-reported residential postal codes were available for a subgroup of the study subjects (n = 1328) who participated in the Canadian Community Health Survey (CCHS) (cycle 1.1, 2000-2001). A comparison analysis based on the first three digits (Forward Sortation Area) of 6-digit postal codes showed that in 2001 about 91% of these study subjects had residential postal codes which were consistent with those recorded in the CCHS.

Neighborhood Socioeconomic Status

The neighborhood-income quintile variable from the 2001 Statistics Canada Census was used to estimate individual socioeconomic status. For the 2001 Census, the dissemination area was the smallest census geographic unit (400-700 persons) for which all census data were disseminated. Within a census metropolitan area, all dissemination areas were ranked by household size-adjusted average family income and divided into approximate quintiles, with 1 representing the lowest and 5 representing the highest income quintile. The neighborhood-income quintiles were assigned to study subjects through their residential postal codes.

Exposure Assessment

Residential Proximity to Road Traffic

According to the DMTI ArcView street file dataset (version 2006.3) for British Columbia (DMTI Spatial, Markham, Ontario, Canada) and actual traffic investigation,[3] road types in this study included:

(1) Highway (DMTI type 1 and 2 road): expressway (mean 114,000 vehicles/day) or principal highway (mean 21,000 vehicles/day);

(2) Major road (DMTI type 3 and 4 road): secondary highway (mean 18,000 vehicles/day) or major road (mean 15,000 vehicles/day).

Traffic volume for each road type in the brackets was average daily traffic counts, which were derived from actual traffic investigation in this study region.[3]

According to the previous findings that the concentrations of traffic-related air pollutants typically decrease exponentially from major roadways and begin to approach background concentrations within about 150 m, we chose 150 m from a highway or a major road as a surrogate for exposure to traffic-related air pollution.[4 5 6] We also chose 50 m from a highway or a major road to represent different exposure levels to traffic-related air pollution.

Distance from each subject's residence indicated by the residential postal code (area centroid) to a highway or a major road was measured using Geographical Information System (GIS) software (ArcGIS 9.2, ESRI, Redlands, CA, USA). The buffers of 50 m and 150 m from a highway or a major road were created, respectively. Subjects living within a specific buffer distance were assigned to the exposure group, those who did not were assigned to the non-exposure group.

Traffic-Related Air Pollution Assessment

The detailed land use regression method and data collection process has been described elsewhere.[7 8 9] Briefly, to better characterize spatial contrasts in the concentrations of traffic-related air pollutants in the study region, we developed land use regression (LUR) models in 2003. During the model building, NO and NO_2 concentrations were measured using passive samplers at 116 sites during two 14-day sampling periods. $PM_{2.5}$ concentrations were measured at a subset of 25 locations during a two-month sampling period. Short-term concentrations of particle light absorbance (a surrogate for black carbon) were measured in a mobile monitoring campaign on one occasion at a subset of 39 sites.[9] We have previously demonstrated strong correlations (R^2 = 0.7-0.8)[10,11] between particle absorbance and traditional measurements of elemental carbon.[12]

Based on these measurements and adjustment for temporal variation, annual average concentrations of these pollutants were calculated for each site. Meanwhile, a total of 55 variables were generated in a Geographic Information System (ArcGIS; ESRI, Redlands, CA, USA) to describe the land use characteristics of each site. Measured air pollutant concentrations and the most predictive land use characteristics were modeled using multiple linear regression techniques. The coefficient of determination (R^2) in the linear regression model was 0.62 for NO, 0.56 for NO_2, 0.52 for $PM_{2.5}$, and 0.56 for black carbon.

Based on these LUR models, smooth spatial surfaces of predicted annual average concentrations for these pollutants were generated for this study region with a resolution of 10 m. We then applied month-year adjustment factors to each surface to calculate monthly concentrations. These monthly air pollution data were assigned to study subjects through their residential postal codes (area centroids). After integrating changes in residences during the 5-year exposure period, average concentrations for NO, NO_2, $PM_{2.5}$ and black carbon were calculated for each subject.

REFERENCES

1. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. *Can J Public Health.* 1998;89:270-273.

2. British Columbia Ministry of Health Services. Medical services plan for BC residents. http://www.health.gov.bc.ca/msp/infoben/index.html Accessed May 10, 2010.

3. Setton EM, Hystad PW, Keller CP. Road Classification Schemes – Good Indicators of Traffic Volume? http://www.cher.ubc.ca/UBCBAQS/SSL05-014-TRAFFIC.pdf Accessed May 10, 2010.

4. Zhu Y, Hinds WC, Kim S, Sioutas C. Concentration and size distribution of ultrafine particles near a major highway. *J Air Waste Manag Assoc.* 2002;52:1032-1042.

5. Zhou Y, Levy JI. Factors influencing the spatial extent of mobile source air pollution impacts: a meta-analysis. *BMC Public Health.* 2007;7:89.

6. Roorda-Knape MC, Janssen NAH, de Hartog JJ, van VLiet PHN, Harssema H, Brunekreef B. Air pollution from traffic in city districts near major motorways. *Atmos Environ.* 1998;32:1921-1930.

7. Brauer M, Lencar C, Tamburic L, Koehoorn M, Demers P, Karr C. A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ Health Perspect*. 2008;116:680-686.

8. Henderson SB, Beckerman B, Jerrett M, Brauer M. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ Sci Technol.* 2007;41:2422-2428.

9. Larson TV, Henderson SB, Brauer M. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land-use regression. *Environ Sci Technol.* 2009;43:4672-4678.

10. Noullett M, Jackson PL, Brauer M. Winter measurements of children's personal exposure and ambient fine particle mass, sulphate and light absorbing components in a northern community. *Atmos Environ.* 2006;40:1971-1990.

11. Rich K. Air pollution and patients with implanted cardiac defibrillators: an epidemiological analysis and assessment of exposure (MSc thesis). University of British Columbia: Vancouver, 2003. http://circle.ubc.ca/handle/2429/14256 Accessed May 10, 2010.

12. Cyrys J, Heinrich J, Hoek G, et al. Comparison between different traffic-related particle indicators: elemental carbon, PM2.5 mass, and absorbance. *J Expo Anal Environ Epidemiol.* 2003;13:134-143.