A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992–2001)

Eric J. Crighton\textsuperscript{a,}\textsuperscript{*}, Susan J. Elliott\textsuperscript{b}, Rahim Moineddin\textsuperscript{c,d}, Pavlos Kanaroglou\textsuperscript{b}, Ross Upshur\textsuperscript{c,d,e}

\textsuperscript{a}Department of Geography, University of Ottawa, 60 University Avenue, Simard Hall Room 06, Ottawa, Ont., Canada K1N 6N5
\textsuperscript{b}School of Geography and Geology, McMaster University, 1280 Main St. West., Hamilton, ON, Canada L8S 4L7
\textsuperscript{c}Department of Family and Community Medicine, University of Toronto, 256 McCaul Street, 2nd Floor, Toronto, ON, Canada M5T 2W5
\textsuperscript{d}Department of Public Health Sciences, University of Toronto, McMurrich Building, 12 Queen’s Park Crescent W., Toronto, ON Canada, MSS 1A8
\textsuperscript{e}Primary Care Research Unit, Sunnybrook and Women’s College Health Sciences Centre, 2075 Bayview Ave, #E-349, Toronto, ON, Canada M4N 3M5

Available online 23 January 2007

Abstract

Previous research on the determinants of pneumonia and influenza has focused primarily on the role of individual level biological and behavioural risk factors resulting in partial explanations and largely curative approaches to reducing the disease burden. This study examines the geographic patterns of pneumonia and influenza hospitalizations and the role that broad ecologic-level factors may have in determining them. We conducted a county level, retrospective, ecologic study of pneumonia and influenza hospitalizations in the province of Ontario, Canada, between 1992 and 2001 (N = 241,803), controlling for spatial dependence in the data. Non-spatial and spatial regression models were estimated using a range of environmental, social, economic, behavioural, and health care predictors. Results revealed low education to be positively associated with hospitalization rates over all age groups and both genders. The Aboriginal population variable was also positively associated in most models except for the 65+ year age group. Behavioural factors (daily smoking and heavy drinking), environmental factors (passive smoking, poor housing, temperature), and health care factors (influenza vaccination) were all significantly associated in different age and gender-specific models. The use of spatial error regression models allowed for unbiased estimation of regression parameters and their significance levels. These findings demonstrate the importance of broad age and gender-specific population-level factors in determining pneumonia and influenza hospitalizations, and illustrate the need for place and population-specific policies that take these factors into consideration.

Keywords: Canada; Pneumonia; Influenza; Hospitalizations; Determinants; Spatial analysis

Introduction

Pneumonia and influenza are the leading causes of death from infectious disease and among the leading causes of death overall in Canada (Health Canada, 2001). Accounting for over 60,000 hospitalizations or approximately a third of all respiratory disease...
hospitalizations (Health Canada, 2001), pneumonia and influenza represent a significant public health and health care system burden. While it is well established that health and illness are influenced by broad environmental, social, and economic factors (e.g. Evans & Stoddart, 1990; Labonte, 1987), research trying to explain variability in pneumonia and influenza has typically focused on individual-level biological and behavioural risk factors, resulting in partial explanations and largely curative approaches to reducing the disease burden (Loeb, 2003). In an effort to address this shortcoming, we have conducted an ecologic-level study of pneumonia and influenza hospitalizations in Ontario using spatial and non-spatial analytic techniques to examine patterns of these illnesses, and the factors that determine them. Furthering our knowledge in this area is critical for the development of more population- and place-specific preventative strategies that take social, economic, health care, and environmental factors into consideration. This study adds to our previous descriptive temporal (Crighton, Moineddin, Upshur, & Mamdani, 2004) and spatial (Crighton, Elliott, Moineddin, Kanaroglou, & Upshur, 2006) analyses of pneumonia and influenza hospitalizations by showing how hospitalization rates are associated with population-level indicators measuring socioeconomic, behavioural, health care, and environmental factors.

Background

There is a significant body of literature examining various pneumonia and influenza outcomes, including mortality (Fine et al., 1997), hospitalizations (Morris & Munasinghe, 1994) and acquisition of illness (Baik et al., 2000). Risk factors frequently identified with pneumonia and influenza morbidity and mortality include age, gender, chronic diseases, recent viral infections and antibiotic use (e.g. Kaplan et al., 2003). People at the extremes of age, in particular the elderly, have been found to be among those at the greatest risk of acquiring the illnesses and experiencing the worst outcomes (Loeb, 2003). Chronic illnesses associated with pneumonia include asthma (Lange, Vestbo, & Nyboe, 1995), cardiovascular disease (Kaplan et al., 2003) and HIV (Plouffe, Breiman, & Facklam, 1996). Identified lifestyle risk factors include smoking (Almirall et al., 2000), obesity (Baik et al., 2000) and excessive drinking (Nuorti et al., 2000).

While the role of these physiological and behavioural factors in determining pneumonia and influenza is significant, considerable variance remains unexplained (Loeb, 2003). The above studies take a largely biomedical approach to understanding disease which is relevant primarily for clinical decision making, but do not address the influence of other determinants including social, economic, cultural and physical environments (Evans & Stoddart, 1990). There is, however, a well established and growing body of evidence pointing to the importance of these factors for numerous, typically chronic, health conditions including, heart disease, respiratory diseases, diabetes as well as all cause mortality (e.g. Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997; Ross et al., 2000; Wilkinson, 1996). Recent studies suggest that such factors are similarly important in determining infectious diseases such as pneumonia and influenza (Farr, Bartlett, Wadsworth, & Miller, 2000; Wood, Sallar, Schechter, & Hogg, 1999), resulting in calls for further research in this area (Loeb, 2003).

Several studies using individual level comparisons have examined relationships between pneumonia (various endpoints) and a range of socioeconomic status (SES) measures, including social class, income and education (Farr et al., 2000; Stelianides, Golmard, Carbon, & Fantin, 1999), although not always coming to the same conclusions. A UK study revealed that unemployed individuals were at a significantly greater risk of being hospitalized with pneumonia (Farr et al., 2000). Similarly, Wood et al. (1999) found a strong negative association between socioeconomic classes and avoidable mortality from pneumonia. On the other hand, Stelianides et al. (1999) found that low SES, defined as long-term unemployment, homelessness or poor living conditions, was not a significant risk factor for pneumonia morbidity; and Vrbova, Mamdani, Moineddin, Jaakimainen, & Upshur (2005) reported no significant association between income and mortality among the elderly.

Cohen (1999) suggests that there are two general mechanisms as to how SES might predispose to pneumonia. The first is through increased exposure to infectious agents from, for example, crowding or institutional living. The second is through increased susceptibility from a weakened immune system related to factors including higher levels of stress, poor or no housing, poor diet, or other life-style factors associated with deprivation. Wilkinson (1996), however, argues that in most Western nations, absolute deprivation is somewhat rare, and that it is in fact relative deprivation that is more
important because of its impact on social capital. Typically measured in terms of social networks, support, and cohesion, social capital has been found to decrease as income inequality increases (Kawachi et al., 1997). Social capital is then linked to health through a range of mediating factors including physiological stress responses, self-esteem, and health behaviors (see Veenstra et al., 2005). While numerous studies have linked social capital to various health outcomes, it has not been studied in the context of infectious diseases such as pneumonia to our knowledge.

Health care system factors have also been identified as being associated with pneumonia and influenza. Although the ability of health care to improve population health is questioned in the public and population health literature (e.g. Marmor, Barer, & Evans, 1994; McKeeown, Record, & Turner, 1975), there is evidence that access to basic health services may reduce the burden of diseases like pneumonia, and the risk of complications leading to hospitalization (Macinko, Starfield, & Shi, 2003; Morris & Munasinghe, 1994). For example, in a study examining the contributions of Primary Care systems in OECD countries, Macinko et al. (2003) found that the strength of the system, measured using a variety of indicators including service access, system financing, and comprehensiveness of care, was significantly negatively associated with premature mortality from, among other diseases, pneumonia and influenza ($p < 0.01$).

A further independent risk factor identified in the literature is race. Being African American, for example, has been found to be independently associated with higher incidence of pneumococcal infection in the USA (Harrison, Dwyer, Billman, Kolczak, & Schuchat, 2000). In Alberta, hospitalization rates for pneumonia among Aboriginal populations were found to be 5 times higher than among non-Aboriginal populations (Marrie, Carriere, Jin, & Johnson, 2004). While genetic endowment may be a factor in these cases, it is expected that underlying socioeconomic factors such as inadequate housing, public sanitation and high unemployment, all of which are chronic problems in many Canadian Aboriginal populations, are at the root of this inequity (Stavenhagen, 2005).

Missing from much of the work done on the determinants of pneumonia and influenza is an understanding of the geographical nature of these illnesses and their determinants. A knowledge of the geographical variation in infectious disease has, in the past, significantly increased our understanding of disease distribution and diffusion, as well as the risk factors for infection or developing the disease (e.g. Cliff, Haggett, & Ord, 1986; Cliff & Smallman-Raynor, 1992; Morris & Munasinghe, 1994). An historical study by Cliff et al. (1986), for example, examined the introduction, spread and disappearance of multiple epidemics of influenza in Iceland, finding associations with various structural features including road connections and population size. Cliff and Smallman-Raynor (1992), using an ecologic study design, examined the spatial distribution of AIDS in Uganda, revealing positive associations with proximity to trucking routes, and army recruitment rates.

There remains a paucity of studies examining the geographical variation in pneumonia and influenza and their determinants. A notable exception is an ecologic-level study conducted by Morris and Munasinghe (1994) looking at the geographic variability and determinants of acute respiratory illness hospitalizations (including pneumonia and influenza) among the elderly in the US. Findings revealed marked regional elevations in rates that were associated with socioeconomic and health care system factors including education, household crowding, income and physicians per capita. However, no control was provided in this analysis for spatial dependence in model residuals, thereby potentially biasing the estimation of regression parameters and their significance levels. More recently, our group conducted a county-level spatial analysis of pneumonia influenza hospitalizations in Ontario (Crighton et al., 2006), revealing significant geographic variability in hospitalization rates. A moderate yet significant level of positive spatial autocorrelation ($I = 0.21; p < 0.05$) was found in the global data, with significant age-specific ‘hot spots’ in several northern counties. Determinants of these patterns were not examined. These studies point to the importance of place in health.

**Conceptual framework**

This study builds on past research by adopting a population health perspective, which highlights the importance of broad determinants of health and disease. Evans and Stoddart’s (1990) population health framework represents an important starting point for this work. The framework emphasizes the relationship between health and genetic endowment,
the socioeconomic environment and the biophysical environment. These broad categories are thought, in turn, to condition an individual’s behavioural and biological responses to external stimuli. While a number of critiques of this framework have been published (e.g. Poland, Coburn, Robertson, & Eakin, 1998), a further significant criticism stems from the fact that the framework does not explicitly recognize the importance of the temporal and spatial nature of health and illness.

According to Loytonen (1998), time and space form an inseparable combination in all human activity and yet until recently both have been frequently ignored by health researchers. Acknowledging this shortcoming, we propose a conceptual framework that overtly addresses temporal and spatial dimensions of health, while incorporating the health determinants outlined by Evans and Stoddart (1990) as well as the risk factors identified in the literature discussed above. We have grouped potential health determinants into five broad constructs that concomitantly interact among themselves and over time and space: (1) the physical environment including both natural and human-made factors such as air pollution, aeroallergens and climate factors; (2) the social and economic environment which is comprised of a broad range of factors including occupation, education and income, to levels of community involvement and social support; (3) biological influences such as age, gender, and co-morbidities; (4) health behaviour including smoking and drinking; and finally, (5) health and social services which comprises the availability and utilization of health care services as well as other social services, public health policy and health education. Through a variety of either direct or indirect processes these factors may affect both the likelihood of exposure, infection, or transmission of pneumonia and influenza, and the biological or behavioural response to infection. Pneumonia and influenza morbidity also vary over time and space, and conceptually, link back to the aforementioned constructs.

Informed by the conceptual framework, the principal objective of this paper is to look beyond the traditional biomedical focus of pneumonia and influenza determinants, and to identify instead what role social, economic, environmental, behavioural and health service factors may have in determining pneumonia and influenza rates over space, and how this may vary by age and gender. This study builds on our previous work exploring temporal (Crighton et al., 2004) and spatial patterns (Crighton et al., 2006) of pneumonia and influenza hospitalizations in Ontario, and represents a further step towards understanding the combined spatial and temporal dimensions of infectious disease morbidity and health service use presented in the conceptual framework.

Methods and data

We conducted a retrospective, population-based, ecological level study to assess spatial patterns of pneumonia and influenza hospitalizations in Ontario, and the factors that determine these patterns. The geographical unit of analysis used is the census division (N = 49) (Fig. 1). Census divisions correspond to the political regions, counties and districts of the province. For convenience, census divisions will be referred to as ‘counties’. While the use of ecological level data may limit one’s ability to make firm policy recommendations, ecological study designs allow for nearly complete coverage of a population in the study area, and are well suited for generating hypotheses for future work. It is further argued that many important determinants of health (e.g. socioeconomic and environmental factors) are inherently contextual in nature and require study designs that work at the ecological level (Kawachi et al., 1997; MacIntyre & Ellaway, 2000; Wilkinson, 1996).

There were approximately 12 million residents in Ontario as of 2001 (Statistics Canada, 2004). County populations range from approximately 13,000 in Manatoulin, in the North (Fig. 1), to over 2.5 million in Toronto, in the South. Northern Ontario is the most sparsely populated area in the province, while Southern and Eastern Ontario is made up of both sparsely populated rural agricultural areas and the province’s major urban centres (i.e. Toronto, Ottawa, Hamilton and Windsor). For residents of Ontario, access to health care services is universal through the Ontario Health Insurance Program (OHIP), although northern and rural residents must frequently travel significantly longer distances for health care services as compared to their urban counterparts.

Data

Outcome variables

The Canadian Institute for Health Information (CIHI) Discharge Abstract Database was used to
obtain information on hospitalizations for pneumonia and influenza as the principal diagnosis, by county of patients’ usual residence. This database records discharges from all inpatient hospital stays in Ontario acute care hospitals using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). Nine years of aggregated CIHI hospitalization data were examined, covering the period between April 1, 1992 and March 31, 2001. While the reliability of specific pneumonia aetiologic information is somewhat low (approximately 52%; Marrie, Durant, & Sealy, 1987), in aggregate form, pneumonia and influenza have been found to be reliably coded (81%; Upshur, 1997).

All records with a principal discharge diagnosis of influenza or pneumonia (ICD-9 code: 480–487) were selected (N = 241,803). Average annual hospitalization rates were computed by dividing the total number of hospitalizations over the study period by the population at risk. Rates were adjusted for age and sex by the direct method (Breslow & Day, 1987). Comparative hospitalization quotients (CHQ) were then calculated. CHQs are defined here as a ratio between the observed directly standardized hospitalization rate of a given county, to the expected rate if the outcome had occurred at the provincial rate (see Breslow & Day, 1987 for a description of similar comparative measures). Thus, CHQs below 1 indicate that rates are below the provincial mean while CHQs above 1 indicate rates are above the mean.

**Explanatory variables**

The selection of explanatory variables was guided by the conceptual framework. Biological factors and co-morbidities have been examined in numerous studies and are, therefore, not included here; contributions and limitations of this literature are outlined above. While the exclusion of these variables represents a potential limitation given the mediating role they may play in determining pneumonia and influenza hospitalizations, it was decided that the focus here would be on the less well understood social, economic, environmental and health care system factors. Explanatory variables examined in this study are described in the following paragraphs and in Table 1. While some variables are self-explanatory, several warrant more explanation.

Explanatory variables were derived from two major secondary sources. The first is the 1996/1997 Ontario Health Survey (OHS), which is comprised
of a sample of approximately 36,000 individuals, and had a household response rate of over 78% (Ontario Ministry of Health, 1997). OHS variables were directly age and gender standardized to the Ontario population. Variables obtained from the OHS include percentages of daily smokers, passive smokers, heavy drinkers, body mass index $>27$ and influenza vaccination in past year. The variables, social support and social involvement are indexes based on the combination of several related questions. Low levels of social support and social involvement have been found to affect a range of health outcomes (e.g., Crighton, Elliott, van der Meer, Small, & Upshur, 2003; Kawachi et al., 1997; Veenstra et al., 2005), but have not, to our knowledge, been studied in relation to pneumonia and influenza.

Data on social and economic characteristics and residential environment were taken from the 1996 census. Detailed descriptions of all census variables used in this study can be found in the 1996 Census Data Dictionary (Statistics Canada, 1999). Instead of using an absolute measure of income, ‘low income’ is used as a general indicator of poverty levels within the population. Statistics Canada defines a low income person as being someone who spends 20% more of total income on food and shelter, than the amount spent by the average person in the population, adjusting for size of settlement to account for cost of living.

A variable for the proportion of family physicians (FPs) and general practitioners (GPs) actively practicing family medicine (i.e. billing for $>50\%$ family medicine codes) per 100,000 population in 1996 was created using the Canadian Physician Database (CPDB) and OHIP billing database. Also, daily temperature data from Environment Canada’s Daily Climate Database was used to calculate a 10-year mean county temperature variable.

Analysis

Analysis was done to assess the degree of spatial autocorrelation in the outcome variables using the Moran’s $I$ statistic. Significant spatial autocorrelation indicates a regular pattern in the data over space such that a value at a given location depends on, and is similar to, a value of defined spatial neighbours. Neighbour relationships are expressed in a row standardized spatial weights matrix $W$ whose elements $W_{ij}$ represent the binary spatial weights assigned to pairs of units $i$ and $j$, where

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Candidate in dependent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>Physical environment</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>Average annual temperature over study period</td>
</tr>
<tr>
<td>Poor housing</td>
<td>% of homes needing major repairs</td>
</tr>
<tr>
<td>Passive smoker</td>
<td>% living in home with regular inside smoker</td>
</tr>
<tr>
<td>Social and economic environment</td>
<td></td>
</tr>
<tr>
<td>Low income</td>
<td>% low income households of total private households—household spends 20% more of total income on food and shelter than average</td>
</tr>
<tr>
<td>Unemployment</td>
<td>% $&gt;15$ years reporting to be unemployed</td>
</tr>
<tr>
<td>Low education</td>
<td>% 15 years or older with less than high school education</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>% Aboriginal population</td>
</tr>
<tr>
<td>Marital status</td>
<td>% 15 years or older not married</td>
</tr>
<tr>
<td>Social support</td>
<td>% reporting low social support</td>
</tr>
<tr>
<td>Social involvement</td>
<td>% reporting low social involvement</td>
</tr>
<tr>
<td>Behavioural</td>
<td></td>
</tr>
<tr>
<td>Daily smoker</td>
<td>% 12 years and over who currently smoke daily</td>
</tr>
<tr>
<td>Heavy drinker</td>
<td>% 12 years and over drinking in excess 14 drinks/week (males) and 9 drinks/week (females)</td>
</tr>
<tr>
<td>Overweight</td>
<td>% overweight (BMI greater than 27)</td>
</tr>
<tr>
<td>Healthcare</td>
<td></td>
</tr>
<tr>
<td>Doctors</td>
<td>Family medicine practitioners/100,000 population</td>
</tr>
<tr>
<td>Flu shot</td>
<td>% reporting influenza vaccination within 1 year</td>
</tr>
</tbody>
</table>

OHS, 1996/1997 Ontario Health Survey; OHIP, Ontario Health Insurance Program; CPDB, Canadian Physician Database.
$W_{ij} = 1$ for neighbours and 0 for non-neighbours. For this analysis, neighbours were defined using the queen’s case adjacency, which considers all counties with common borders as neighbours.

Initial variable screening included calculating correlation coefficients, and checking for linearity in the expected relationship, between explanatory and outcome variables. Outcome variables were log transformed (natural log) in order to yield approximate normality in the data and to help stabilize the variance. All explanatory variables underwent a deviation from the mean transformation for easier interpretation. In cases where there was a non-linear relationship between the outcome and an explanatory variable, a log transformation was applied to the latter.

For each outcome, the selected independent variables were included in an ordinary least-squares (OLS) regression. A backward stepwise procedure was used with a significance level of $p < 0.1$ required in the partial test for explanatory variables to be retained in the final model. In cases where significant correlation between explanatory variables was detected, each was entered into the model selection procedure alternately to assess their relative contributions. The fit of alternative models were compared. Formal diagnostic tests for multicollinearity and heteroskedasticity were conducted. Residuals were also tested for normality as well as spatial dependence using the Moran’s $I$ test. Further analysis was done using a local indicator of spatial autocorrelation or LISA (Anselin, 1995) to assess the degree of localized clustering of model residuals (for a good description of these techniques see Bailey & Gatrell, 1995).

Spatial dependence in the model residuals represents a violation of OLS assumptions. If spatial dependence was identified using the Moran’s $I$ test, spatial lag and spatial error models, two alternative models that incorporate spatial dependence, were considered. These are discussed in detail by Anselin (1992) and Bailey and Gatrell (1995). Based on the results of Lagrange Multiplier tests (see Anselin, 1992), spatial error models were found to be the appropriate alternative. Following Anselin’s (1992) notation, the spatial error model takes the form:

$$Y = X\beta + \varepsilon = \lambda W_\varepsilon + \xi,$$

where $Y$ is a vector of observations on the dependent variable, $W$ is a spatial weights matrix, $\varepsilon$ is a vector of error terms, $W_\varepsilon$ are spatially lagged errors, $\lambda$ is an autoregressive coefficient, and $\xi$ is a vector of independent random errors. Models, using the same explanatory variables as in the OLS regression, were fit using maximum likelihood estimation. Formal diagnostics tests were again performed to assess the suitability of these models. $R^2$ and log likelihood values are used to compare the fit of the models.

To interpret the influence of significant explanatory variables in the models on the log CHQ, we used the regression coefficient as the exponent of the base of the natural log to assess a 1-unit increase in $x$ above its mean: $e^b - 1$. All significant variables examined in this study, other than the Aboriginal variable, are interpreted in the same way. The Aboriginal variable was natural log transformed and thus its coefficient is interpreted differently, using the following formula:

$$\left(\frac{x + 1}{x}\right)^b - 1.$$
and Ottawa. Although there is somewhat more heterogeneity in CHQs among females, the overall spatial pattern of hospitalizations is consistent for both sexes.

The Moran’s $I$ statistic indicates that there is a moderate, statistically significant degree of spatial autocorrelation in the data for females (Moran’s $I = 0.242; p < 0.015$), indicating that counties with similar hospitalization rates are clustered together. Somewhat lower, but still significant autocorrelation is seen for males (Moran’s $I = 0.173; p < 0.044$).

Table 2 summarizes the results for the total population and by gender models. There was no evidence of heteroskedasticity or multicollinearity in these models. For the total population, the results of the initial OLS model (data not shown) were deemed unsuitable as the Moran’s $I$ statistic testing for spatial error dependence was significant (Moran’s $I = 0.162; p = 0.043$). This indicates that the model did not adequately account for spatial dependence in the CHQs and the OLS assumption of residual independence was not met. A spatial

![Comparative morbidity figures (CHQs) of pneumonia and influenza hospitalizations by gender in Ontario counties: 1992–2001.](image)
error model, which incorporates spatial dependence, was therefore used. Three variables display significant, positive relationships with the log CHQ (natural log): Aboriginal population, low education and drinking. The model explained approximately 67% of the variation ($R^2 = 0.668$). For the education variable, the coefficient indicates that a 1-unit (%) increase in low education from its mean is associated with an increase of 0.0544% or 5.4% in the CHQ on average. In the case of the heavy drinking variable, a 1-unit (%) increase from its mean is associated with an increase in the CHQ of 0.0283% or approximately 2.8%. For the same increase in the aboriginal variable, the CHQ increases by 0.0249% or 2.5%. The model log likelihood, a measure of model fit, which can be used to compare OLS and spatial error models, was 18.28 for the OLS models, and 19.67 for the spatial error model. While this indicates that the fit is better in the spatial versus the OLS model, the likelihood ratio test (LR test) indicates that this difference is not statistically significant.

For females, significant autocorrelation in the OLS model residuals was again identified (Moran’s $I = 0.176; p = 0.027$) and a spatial error model was, therefore, employed. The model explained a substantial portion of the variation ($R^2 = 0.67$). CHQs were found to increase with higher percentages of low education, poor housing, and inside smokers. The regression coefficient indicates that for every 1-unit (%) increase in the low education variable from its mean, the CHQ increases by 0.0376% or 3.8% on average. For the same increase in either the poor housing or passive smoking variables, the CHQ increases by 3.8% or 1.2%, respectively. The model log likelihood value was higher for the spatial error model as compared to the OLS model (20.19 versus 18.85) indicating a better model fit. Again, however, this difference was not significant (LR test $p = 0.153$).

For males, the OLS model was determined to be appropriate as no significant error dependence was found (Moran’s $I = 0.106; p = 0.139$). The model explained a substantial portion of the variation in the CHQs for pneumonia and influenza hospitalizations ($R^2 = 0.64$). CHQs were found to increase with higher percentages of Aboriginal population, low education and heavy drinkers. The regression coefficient indicates that every 1% increase in the Aboriginal population variable from the mean is associated with an increase of the CHQ by 0.0226% or approximately 2.3% on average. For every 1% increase in the low education or heavy drinkers variables, the CHQ increases by 4.5% and 2.0%, respectively.

Table 3 summarizes the results of the female models by age group. There was no evidence of heteroskedasticity or multicollinearity in this or any of the age-specific female models. For the female 0–14-year age group OLS model (Table 3), no significant error dependence was identified (Moran’s $I = 0.081; p = 0.218$), and the model was deemed appropriate. The model explains approximately 57% of the CHQ variation, with two variables being retained in the model: Aboriginal population and low education. With a 1% increase in either the poor housing or passive smoking variables, the CHQ increases by 3.8% or 3.8%. For the same increase in the

| Table 3 | Regression coefficients and standard errors for the female pneumonia and influenza hospitalization models by age group (1992–2001) |
|-----------------|--------------------------|--------------------------|--------------------------|
| Explanatory variables | 0–14 years$^a$ | 15–64 years$^b$ | 65+ years$^a$ |
| Intercept | 0.201 | 0.201 | 0.154 |
| SE | 0.044 | 0.054 | 0.027 |
| $p$-value | <0.001 | <0.001 | <0.001 |
| Aboriginal | 0.136 | 0.137 | 0.033 |
| SE | 0.043 | 0.038 | 0.009 |
| $p$-value | 0.002 | <0.001 | <0.001 |
| Low education | 0.078 | 0.085 | 0.033 |
| SE | 0.014 | 0.010 | 0.009 |
| $p$-value | <0.001 | <0.001 | <0.001 |
| Daily smokers | 0.014 | 0.014 | 0.013 |
| SE | 0.007 | 0.007 | 0.017 |
| $p$-value | 0.029 | 0.029 | 0.079 |
| Flu shot | −0.005 | −0.005 | −0.031 |
| SE | 0.003 | 0.003 | 0.017 |
| $p$-value | 0.078 | 0.078 | 0.079 |
| Temperature | 0.466 | 0.144 | 0.466 |
| $\lambda$ | 0.001 | 0.001 | 0.001 |
| Model $R^2$ | 0.567 | 0.733 | 0.431 |
| Model LL | −9.95 | 7.032 | 16.03 |
| LR test | ⎯ | $p = 0.009$ | ⎯ |

$^a$Ordinary least-squares (OLS) model.
$^b$Spatial error model.
low education variable, the CHQ increases by 0.0811% or approximately 8%.

In the female 15–64-year age group OLS model, significant error dependence was identified (Moran’s $I = 0.259; p = 0.002$) making the spatial error model more appropriate (Table 3). The spatial error model accounted for approximately 73% of the variation in CHQs ($R^2 = 0.73$). The LR test indicates that the fit of the error model improved significantly from the OLS ($p = 0.009$). Again, only the Aboriginal population and low education variables were retained in the model. Here a 1% increase in the Aboriginal population variable is associated with an increase in the CHQ by 3.3%. The same increase in the low education variable accounts for an 8.1% increase in the CHQ.

For the 65+-year age group OLS model (Table 3), no significant spatial error dependence was identified (Moran’s $I$ for model residuals $= 0.121; p = 0.081$). Approximately 43% of the variation in the CHQs is explained, with four variables being retained in the model: low education (positive), daily smokers (positive), flu shot (negative), and temperature (negative). Increases of 1% in proportions with low education or in proportions of daily smokers are, respectively, associated with a 3.4% and a 1.4% increase in hospitalizations. A 1% increase in the flu shot is associated with a decrease in the CHQ by 0.5% on average, and a 1° increase in the mean county temperature is associated with a decrease in the CHQ by 3.1%.

Table 4 summarizes the results of the male models by age group. There was no evidence of heteroskedasticity or multicollinearity in any of the male OLS models. Some differences between male age groups can be seen in the fit of the models and the variables that were retained in the models. For the 0–14-year age group model for males (Table 4), OLS was again determined to be appropriate (Moran’s $I = 0.261; p = 0.009$) and a spatial error model was used instead. The model has high explanatory power, accounting for approximately 76% of the variation in CHQs ($R^2 = 0.50$). Again Aboriginal population and low education variables were retained in the model. For every 1% increase in the Aboriginal population variable from the mean, the CHQ increases by 0.0332% or 3.3%, and for the same increase in the low education variable, the CHQ increases by 0.0790% or almost 7.9%.

For the 15–64-year age group OLS model (Table 4), significant error dependence was identified (Moran’s $I = 0.261; p = 0.009$) and a spatial error model was used instead. The model has high explanatory power, accounting for approximately 76% of the variation in CHQs. The LR test indicates that the fit of the error model improved significantly from its OLS equivalent ($p = 0.039$). The model has three significant variables: Aboriginal population, low education and heavy drinkers. The coefficients for these variables suggest that for every 1% increase in either the Aboriginal population, low education, or heavy drinkers variables, a CHQ increases on average by 2.6%, 6.2% and 2.1%, respectively.

For the male 65+ age group model (Table 4), OLS was determined to be appropriate, demonstrating no spatial error dependence (Moran’s $I = 0.081; p = 0.119$). The model has a relatively weak explanatory power ($R^2 = 0.40$) with two significant variables: low education and temperature. Here a 1% increase in the low education

---

**Table 4**

Regression coefficients and standard errors for the male pneumonia and influenza hospitalization models by age group (1992–2001)

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>0–14 yearsa</th>
<th>15–64 yearsb</th>
<th>65 + yearsa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>SE</td>
<td>p-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.200</td>
<td>0.047</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>0.118</td>
<td>0.046</td>
<td>0.013</td>
</tr>
<tr>
<td>Low education</td>
<td>0.076</td>
<td>0.015</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Heavy drinker</td>
<td>0.024</td>
<td>0.009</td>
<td>0.008</td>
</tr>
<tr>
<td>Temperature °C</td>
<td>0.372</td>
<td>0.158</td>
<td>0.018</td>
</tr>
<tr>
<td>Model $R^2$</td>
<td>0.502</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model LL</td>
<td>-13.37</td>
<td>18.490</td>
<td>22.74</td>
</tr>
<tr>
<td>LR test</td>
<td>—</td>
<td>p = 0.039</td>
<td>—</td>
</tr>
</tbody>
</table>

aOrdinary least-squares (OLS) model.

bSpatial error model.
variable is associated with a 2.6% increase in the CHQ, and a 1° increase in mean county temperature is associated with a 3.8% decrease in the CHQ on average.

Results from the LISA analyses (data not shown) indicate that the four groups with significant \textit{global} error dependence in their OLS residuals (i.e. total population; females all ages, females15–64 years; and males 15–64 years) had similar patterns of \textit{local} error dependence. Clusters of large residuals centred on Grey County in the southwest of the province were identified, indicating that an underlying process is not being accounted for by the explanatory variables in the models. Two significant clusters of small residuals for these models are seen for in Haldimand-Norfolk and Elgin counties in the south indicating the strong predictive power of the models in these areas. The spatial dependence in the OLS model residuals was accounted for in the spatial error models.

**Discussion**

The objectives of this research centred on developing a better understanding of the spatial characteristics of pneumonia and influenza hospitalizations by age and gender, and the broad contextual factors that determine them. Several issues arise from the results presented.

We found that pneumonia and influenza hospitalization ratios varied significantly across Ontario counties, with similar ratios clustering together (Fig. 2) suggesting that common spatial processes may be at play in determining hospitalizations in these areas. The highest rates are seen in the sparsely populated northern and rural areas where they are between 2 and 3 times the provincial average. The lowest rates are seen in southern and urban areas where rates are between 10% and 30% below the provincial average. Efforts to reduce the burden of pneumonia and influenza must be directed accordingly.

Our results indicate that low education is strongly associated with pneumonia and influenza hospitalizations over all age groups and both genders. Education is a good proxy for SES, particularly among the elderly, because it remains mostly unaffected over time as compared to employment or income \cite{Loeb, 2003}. The association identified here is supported by the results of an ecologic level analysis of pneumonia hospitalizations in the USA \cite{Morris & Munasinghe, 1994}. Similar associations have been reported for other health conditions \cite{Chen, Dales, & Krewski, 2001; Joines, Hertz-Picciotto, Carey, Gesler, & Suchindran, 2003; Sundquist & Johansson, 1997}. The process by which education affects hospitalizations is somewhat unclear in terms of whether or not there is a direct relationship with incidence or severity of symptoms, or whether the relationship is mediated entirely through lifestyle, socioeconomic, environmental or health care factors. In the case of health care, for example, low education may affect help seeking behaviour at the primary care level, as well as adherence with medical regimes, thereby leading to higher rates of hospitalizations. Structural equation modelling represents a potential next step towards better understanding such processes.

The Aboriginal population variable was positively associated with pneumonia and influenza hospitalizations in most models except those for the 65+-year age groups. This association has not been previously reported at this level of analysis, however, it is consistent with studies showing that Aboriginal populations experience higher rates of pneumonia morbidity and mortality than non-Aboriginal populations \cite{Marrie et al., 2004}. Identified risk factors such as poor housing and high rates of smoking and drinking, are found at significantly higher rates among Aboriginal populations \cite{Stavenhagen, 2005}. Programs to improve living conditions and access to education, and public health measures to reduce smoking and drinking rates could be expected to reduce pneumonia and influenza hospitalization rates in counties with large Aboriginal populations.

Smoking and drinking both achieved significance in different models. Findings suggest that gender may modulate the effect of these behavioural factors on hospitalization rates. The association between heavy drinking and hospitalizations among men, but not women, could be anticipated, given that heavy drinking rates are higher among men. The significant association between daily smoking among women, but not men, is somewhat surprising, however, as daily smoking rates among women are lower.

We further found that influenza vaccination was weakly, negatively associated with hospitalizations in the female 65+ model. While this finding supports the continuation of influenza vaccination campaigns targeted at the province's elderly, the development of better vaccination exposure variables is required in order to accurately assess their
efficacy in reducing pneumonia and influenza morbidity more widely.

The environmental factors examined in this study were all found to be associated with pneumonia and influenza hospitalization rates in one or more models. Passive smoking was significant in the female all age groups model. This finding contradicts those of Farr et al. (2000) who found no significant association between passive smoking and pneumonia hospitalizations. Temperature was negatively associated in both 65+-year age group models. This result may be explained by the adverse effects that cold can have on the immune system’s resistance to respiratory infection (The Eurowinter Group, 1997) particularly among the elderly who may already have challenged immune systems. This relationship is likely mediated through other factors including poor housing conditions. Poor housing was found to be independently associated with pneumonia and influenza hospitalizations in the female all age groups model. While this variable is categorized in this research as a proxy for indoor environmental conditions (e.g. indoor air quality, inadequate heating, and so on), it also functions as an indicator of low SES.

Also interesting are the variables that were not found to be significant in the models. For example, low income was not significantly associated with pneumonia and influenza hospitalization rates, a finding supported by a number of previous studies (Stelianides et al., 1999; Vrbova et al., 2005), and contradicted by others (e.g. Harrison et al., 2000; Morris & Munasinghe, 1994; Wood et al., 1999).

The relationship must, therefore, not be discounted, and further analysis examining alternative measures should be considered that include absolute and relative measures of income both at individual and ecological levels, and at different spatial aggregations. The FP/GP availability variable was also not significantly associated with pneumonia and influenza hospitalizations. This suggests the following possible explanations: the variable is not an adequate measure of system access; there is a mismatch between the scale of analysis and the scale of the underlying process; or more simply, that greater access to primary care does not equate reduced risk of pneumonia and influenza hospitalization. This last explanation seems unlikely given that Macinko et al. (2003) found that, in a study covering 18 OECD nations, strong primary care systems were negatively associated with premature mortality from pneumonia.

Much of the spatial error dependence identified in the hospitalization data was explained by the covariates in our OLS models. Possible explanations for the remaining spatial error dependence include the omission of a relevant variable from the models, or perhaps, again, a mismatch between the scale of analysis and the scale of the underlying process. In models where spatial dependence remained, spatial error models provided control, and confirmed the independent contribution of identified social, economic and behavioural covariates.

While the spatial effect on the models was not large in any given case, it is important to recognize that spatial effects are frequently found to be significant (e.g., Green, Hoppa, Young, & Blanchard, 2003; Joines et al., 2003), and as such wider use of spatial modelling techniques should be encouraged.

This study has several limitations. Firstly, it is an ecologic study. While this approach lends itself to the examination of contextual determinants of health, complete population coverage, and the generation of hypotheses, it limits our ability to draw conclusions concerning factors that are responsible for variations in hospitalization rates at the individual level. The modifiable areal unit problem (MAUP) (Openshaw, 1984) represents a second potential limitation in that the patterns identified here may depend on the areal aggregations used (i.e. counties). Unfortunately, testing for this is difficult due to population size constraints at larger scales. Thirdly, the use of hospitalizations is not necessarily reflective of morbidity in the population, and does not account for differential access to services (Eyles, Birch, Chambers, Hurley, & Hutchison, 1991). However, given that health insurance is universal in Ontario, hospitalizations are believed to represent a good estimate of severe morbidity. Findings from a British study identifying a strong correlation ($r = 0.69; p < 0.01$) between hospitalizations and morbidity for respiratory illness (Payne, Coy, Patterson, & Milner, 1994) reinforce this point, although this has not been confirmed in a Canadian context. Fourthly, access (as defined by proximity) to hospitals was not controlled for in this analysis despite it being a potential determinant of use. Our results, however, indicate that rural areas typically had the highest hospitalization rates, despite having the least access. Thus, the absence of this variable is not believed to have confounded our findings. Finally, despite aetiological differences, influenza and the pneumonias represented by the ICD-9 diagnostic codes.
of these illnesses and related health services use in influenza increases, developing a better understanding of the underlying causes of spatial variability of pneumonia and influenza outcomes. Its utility is borne out by the range of explanatory variables, which emerged as significant in the models, and the large proportion of geographic variation in pneumonia and influenza hospitalizations that it has helped explain. The framework further illustrates the complexity of the relationships that exist and the need for more sophisticated analytic techniques that can integrate both spatial and temporal dimensions. Methodologically, this research demonstrates how spatial analytic techniques can be applied in studying infectious disease. The use of spatial error regression models allowed for unbiased estimation of regression parameters and their significance levels. Wider use of spatial modelling techniques should be encouraged. Substantively, a better understanding of the geographic variability in pneumonia and influenza hospitalizations in Ontario, and of the importance of the broad range of age and gender-specific ecologic-level factors associated with this outcome, will help inform the development of more region and population-specific social, economic, public health and health care programs. The significance of these findings further extends to various (re)emerging infectious diseases including HIV/AIDS, tuberculosis, and the more recent concern, Avian Influenza, where similar factors play a role in determining disease exposure, susceptibility, and health care system use (Cohen, 1999; Lederberg, Shope, & Oaks, 1992). Efforts made to reduce pneumonia and influenza rates through, for example, improved social welfare programs, could, therefore, be expected to reduce rates of other infectious diseases. As the populations of Canada and most Western nations age, and the burden of pneumonia and influenza increases, developing a better understanding of these illnesses and related health services use is crucial. Further work is needed that includes the examination of different geographical aggregations and alternative pneumonia and influenza outcomes. Multilevel modelling to understand the relative roles of individual and contextual level factors in determining pneumonia and influenza outcomes is also necessary. This paper compliments work previously done examining temporal patterns (Crighton et al., 2004) and spatial patterns (Crighton et al., 2006) of pneumonia and influenza hospitalizations, and sets the stage for analyses that incorporates spatial and temporal dimensions together.

References
and gender. *Epidemiology and Infection* July 7; 1–9 (Epub ahead of print).


Ottawa, Canada: Statistics Canada.


