CHAPTER 5: Mapping the influenza epidemics of 2004-5 and 2005-6 in Ontario using data from emergency department and Telehealth Ontario utilization

Eric Moore¹

¹ Department of Geography, Queen's University, Kingston, Canada
**Introduction**

Influenza epidemics are an annual occurrence in Ontario, with known cases usually reaching a peak in the early months of the year and declining in the Spring. The specific timing and severity of the epidemic varies each year depending on a variety of factors including the particular influenza strains involved and the effectiveness of vaccination programs[1]. The standard for documenting the rise and fall of influenza in Ontario and in Canada as a whole is the data on laboratory-confirmed cases provided by the Canadian Communicable Disease Report (CCDR). Unfortunately, these data are not immediately available to public health agencies, particularly to Public Health Units, and other sources of data need to be evaluated for their ability to provide near real-time monitoring of the growth and spread of the annual epidemic. Two source offer promise in this regard: (i) the National Ambulatory Care Reporting System (NACRS) data on Emergency Department visits and hospital admissions, which are collected daily and are coded by major symptom permitting patient visits and hospital admissions for respiratory reasons to be identified; (ii) data from use of Telehealth Ontario, where each call is also coded by major symptom. Both Emergency Department (ED) and Telehealth Ontario data are coded by 3-digit postal codes (there were 510 such postal codes, also known as Forward Sortation Areas (FSAs), in Ontario in 2004) and these form the basis for defining the spatial extent of the epidemic over time.

In this document the procedures for generating the data for mapping the influenza data on a weekly basis during the influenza seasons in 2004-5 and 2005-6 are presented. The maps themselves are available in separate documents.

**Overall Comparison of Data Series**

Three series are examined:

a) *Telehealth Ontario calls for respiratory reasons.* All Telehealth Ontario calls are characterized in terms of 486 guidelines; these guidelines were grouped into 32 syndromes by medical experts and data for the respiratory syndrome were used in this study. During the study period from July 1, 2004 to June 30, 2006, 222,787 cases were available for spatial analysis.

b) *ED visits for respiratory reasons.* All NACRS data on Emergency Department visits are given an ICD-10-CA code assigned by an experienced health coder based on a physician’s medical diagnosis. Only ICD codes for a respiratory syndrome were used in this study. The period from July 1, 2004 to March 31, 2006 produced 791,354 respiratory cases for spatial
analysis. The shorter time period reflects the reporting procedures for the Ministry of Health which currently release these data by fiscal year which ends in March.

c) CCDR confirmed cases. The 4313 confirmed influenza cases for the 2 year period were concentrated in 2 distinct influenza seasons: i) from Oct. 31, 2004 to April 30, 2005 and ii) from Dec. 11, 2005 to May 27, 2006.

The decision was made to present each data series aggregated by week to conform to the reporting procedures of CCDR. The total counts by week for the 3 series for Ontario are graphed in Figure 1. From visual comparisons it is clear that the timing of the peaks in these data tend to coincide. One major exception is that the ED data indicate a significant peak over the Christmas – New Year period when family physician services are often not available. It is interesting that this substitution effect is dramatically higher for the ED service than for Telehealth Ontario. A second difference is that the CCDR peak seems to be earlier than the Telehealth/ED peak in 2006. This difference may arise as the relative importance of Type A and Type B influenza varied between the two seasons. Type B occurred earlier in 2006 and generally resulted in less significant health problems than the later Type A[2].

When we examine the relationship between Telehealth Ontario and ED data more analytically, the correlation between the two weekly series is very strong (Table 1). The highest association is with zero lag, indicating that the series are coterminous. The high, but declining, correlations for the other lags reflects the fact that the overall pattern shows only one major peak in a twelve month period with relatively smooth transitions between winter highs and summer lows. Separate analyses shows that these strong temporal relations also exist over this period for individual PHU’S.

Mapping Respiratory Illness at the Sub-Provincial Level

Although the strong cyclical pattern of both respiratory illness and confirmed influenza incidence is evident in the provincial data, we need more detailed spatial data to explore the nature of the spread of respiratory illness across the province in each of the influenza seasons during our study period. Given that Public Health Units (PHUs) are an important administrative entity with respect to local management of epidemics, it would be desirable to present the spatial element of epidemics at the PHU level as well as producing more general models of the epidemics with continuous surfaces of incidence rates. The ability to produce such maps is dependent on the way in which geographic locations are coded in the Telehealth Ontario and ED records.
Basic Coding

Telehealth Calls:
Each call is coded by the first 3-digits of the Postal Code (PC) where the call originated, also known as the Forward Sortation Area (FSA). FSAs are not the most satisfactory of units as they vary greatly in size (Figure 2A and 2B). There were 510 Forward Sortation Areas (FSAs) in Ontario in 2004. Clearly, the sizes reflect the differences in population density in urban and rural areas with the sparsely populated Northern region containing FSAs which have areas of more than 20,000 sq. km. Unfortunately, there are two additional problems with FSAs: i) the rural areas surrounding urban areas are often one single FSA (see Figure 3 for an example), and ii) the FSAs do not map simply onto Public Health Unit boundaries (Figures 4 & 5).

Emergency Department visits
Each visit to the ED is coded in the NACRS database by both the FSA and the PHU in which the patient resides. PHUs also vary in size in response to variations in population density (Figure 4) but not in as dramatic a way as FSAs (Figure 4). The more serious issue is that there are significant overlaps with single FSAs often crossing the borders of 2 or more PHUs. Figure 5 illustrates the issue in the Ottawa region. The FSA ‘K0K’ is a continuous area around the southern boundary of the city and it overlaps 3 different Public Health Units, so if we have a particularly record for ‘K0K’ we cannot tell where it is located with any degree of accuracy. If we are to produce maps of Telehealth Ontario calls related to respiratory problems at the PHU level we need an algorithm for assigning calls to PHUs based on their FSA location.

The FSA-PHU Allocation Procedure
Given that we are primarily concerned with aggregate measures (weekly counts of Telehealth Ontario cases by PHU), we do not need to allocate each record for a given FSA to a specific PHU. What we need is an estimate of the probability $p_{ij}$ that a record in the $i^{th}$ FSA belongs in the $j^{th}$ PHU. Then, if $N_i$ is the number of calls generated in the $i^{th}$ FSA in a given time period (usually a week), then

$$F_j = \sum_i N_i \cdot p_{ij}$$

(1)

where $F_j$ is the estimated number of calls to the $j^{th}$ PHU in the same period.

One method of generating allocation estimates would be to calculate the proportion of the area of the $i^{th}$ FSA which lies in the $j^{th}$ PHU and use this proportion as a direct estimate of the probability $p_{ij}$. This is a reasonable approach if the population density is relatively uniform across the FSA. However, such uniformity is highly unlikely in outer suburbs and rural areas,
where population tends to occur in well-spaced clusters. We adopted a different approach using all NACRS records, not just those for respiratory reasons, for the 2 year period from July 2004 to March 2006. There were over 9 million eligible records containing valid FSA and PHU codes (about 2 percent of all records had invalid FSA codes or incompatible codes for FSA and PHU). On average there were approximately 18,000 records for each FSA which were empirically assigned to a destination PHU over the 2 year period. It was assumed that these were sufficiently large samples to generate reliable estimates of the underlying probabilities. Thus if \( M_{ij} \) was the total number of visits from the \( i^{th} \) FSA to the \( j^{th} \) PHU during the 2 year period then the estimated probability \( p_{ij} \) is given by

\[
p_{ij} = \frac{M_{ij}}{\sum_{j} M_{ij}}
\]  
(2)

These estimates can then be inserted in (1) to give the estimated number of calls to the \( j^{th} \) PHU in any selected time period, which in this study is the week by week counts from July 2004 to June 2006.

\[
F_j = \sum_i N_i \cdot \left( \frac{M_{ij}}{\sum_j M_{ij}} \right)
\]  
(3)

**Adjustments to Raw Counts in the Mapping Process**

While mapping weekly counts does give some idea of the actual loads on the two service components (Telehealth Ontario and Emergency Departments) arising from respiratory problems in each PHU, this is of limited value if the focus is on the nature of the spread of an epidemic. In the first instance, weekly counts are directly influenced by the size of the population at risk in each PHU. When such populations are highly variable across the province, maps based on raw counts will tend to reproduce the overall population distribution. The calculation of counts of calls (or visits) per 100,000 represents the first adjustment.

If \( P_i \) is the population of the \( i^{th} \) FSA, \( N_i \) is the number of calls from the \( i^{th} \) FSA in a given week, the call rate for the \( i^{th} \) FSA, \( R_i \) is

\[
R_i = \left( \frac{N_i}{P_i} \right) \cdot 100,000
\]  
(4)

If \( Q_j \) is the population of the \( j^{th} \) PHU, \( F_j \) is the estimated number of calls from the \( j^{th} \) PHU, the call rate for the \( j^{th} \) PHU, \( S_j \) is

\[
S_j = \left( \frac{F_j}{Q_j} \right) \cdot 100,000
\]  
(5)

A second problem in using both of these data series is that the likelihood of using each service is highly variable across the province and differs significantly between the two services. In general, individuals are 4 to 5 times as likely to use Emergency Departments than Telehealth Ontario for respiratory problems, but there are also strong differences from area to area. To provide the most detailed demonstration of these effects we calculated the total number of
respiratory calls per 100,000 persons for the entire 2 year period for each FSA and for the province as a whole. Each FSA rate was then divided by the Provincial rate to give a scaled value of **Intensity of Use**, where the value of 1.0 is equivalent to the Provincial rate, 0.5 is half the Provincial rate and 2.0 is twice the Provincial rate. If \( N_i \) is the total number of calls from the \( i^{th} \) FSA in the 2 year period, \( N_{++} \) is the total number of calls generated in the Province and \( P_+ \) is the population of Ontario, then the Intensity of use for Telehealth Ontario calls for the \( i^{th} \) FSA, \( I_i \) is

\[
I_i = \frac{N_i}{P_i} \times 100,000
\]

and the Provincial Intensity of use, \( I_+ \) is

\[
I_+ = \frac{N_{++}}{P_+} \times 100,000
\]

The Scaled Intensity of Use, \( U_i \) is

\[
U_i = \frac{I_i}{I_+}
\]

Similar calculations produce scaled intensities of Telehealth Ontario calls for the \( j^{th} \) PHU and the entire process can be repeated to produce scaled intensities for Emergency Department visits for both FSAs and PHUs. Figures 6A and 6B present these scaled values for Emergency Department visits for Ontario as a whole and for the Toronto region. Figures 7A and 7B provide comparable data for Telehealth Ontario calls.

The urban-rural contrast in ED use is dramatic, particularly in the Toronto region. The ratio of utilization rates between the Grey-Bruce area to the North-West of Toronto and central Toronto is as much as 8 to 1. While understanding the causes of these differentials deserves detailed analysis, one of the main determinants is undoubtedly the availability of alternative service points. If an individual needs health care (especially non-critical care), a number of options are available: the family doctor, after-hours clinic, walk-in clinic, emergency department and Telehealth Ontario, with the family doctor being the dominant source. The availability of family doctors, and particularly walk-in clinics, is much higher in the larger urban areas. Other socio-economic and demographic factors may also be important – such as age-structure, immigrant concentration and language competence, household income and education – and should be examined.

The patterns for Telehealth Ontario are not so straightforward. The main differentials are the higher than average use in the more remote suburbs of Toronto, areas which are further from hospitals and have relatively high levels of income and education. South Western Ontario, especially along the Lake Huron shoreline has a markedly lower than average use of Telehealth
Ontario, which may reflect the strong network of small hospitals which provide alternative service points.

If we are to construct maps of the evolving severity of an epidemic across the province, we need to compensate for the differential propensity to use the two different services. Using the weekly population adjusted rates, the simplest procedure is to adjust each weekly call rate for the $i^{th}$ FSA, $R_i$, by the Scaled Intensity Rate $U_i$ to produce the Intensity adjusted rate, $V_i$

$$V_i = \frac{R_i}{U_i} \quad (9)$$

These calculations can be replicated for PHUs and for ED visits for FSAs and PHUs. The resulting Intensity adjusted rate represents the number of calls/visits per 100,000 persons in a given week in the $i^{th}$ area if the propensity to use the given service (Telehealth Ontario or Emergency Department) is the same as the overall provincial rate in every FSA or PSU. The resulting maps then provide an estimate of the relative severity of the influenza epidemic in different locations across the province for the given week.

**Map Sequences**

The primary focus of the sub-provincial analysis of the spread of influenza is at the scale of the PHU, which is the scale at which many local public health decisions are made. Further, it is desirable to make analyses compatible with the standard reporting by the Canadian Communicable Disease Report (CCDR). Therefore the time scale is the weekly CCDR schedule and the influenza seasons and also defined by CCDR data: i) from Oct.31, 2004 to April 30, 2005 and ii) from Dec. 11, 2005 to May 27, 2006. These criteria led to the production of 4 weekly series:


2. *Weekly Emergency Department Visit Intensity Adjusted Rates by PHU* from the week ending Dec. 11, 2005 to the week ending April 1, 2006 (note that this is the end of the available data from the Ministry of Health not the end of the influenza season).


The full sequences of maps are available in Powerpoint files. Here we provide examples for selected weeks in the Emergency Department sequences for each season (Figures 8 and 9). A number of features are worthy of note:

- The epidemic was stronger and lasted for a longer period in 2004-5 than 2005-6.
- In 2004-5 the epidemic started in Eastern Ontario and then migrated westward and northward. Although it spread rapidly to the northern districts, it died out earlier in the main urban areas and lingered longer in the South-West.
- In 2005-6 the epidemic started later and had a much shorter high intensity period. There was not the same early focus in the East while the South-West (particularly the rural areas) ended with the highest levels.
- The spread in both seasons does not show a classic contagious spread from one area to contiguous areas. The spread is very rapid and moves as quickly from southern cities to the north as to other cities in the same region. The overall pattern of modern travel by air, train and automobile is a far more significant factor than mere proximity.
- With a significantly higher number of cases, the ED based maps are a more reliable representation of the spread of the epidemics than the Telehealth Ontario-based maps, although the latter do reproduce the essential features.

**Modeling the Respiratory Illness Surfaces**

It is also possible to consider the spread of influenza to comprise a continuous surface of infection which moves across the province. This spread is reflected in elevated respiratory illness incidence across the province. The question is then how to use the type of data available to this study to estimate the nature of these changing surfaces. The standard approach is to treat rate data for small areas as being focused on the centroids of those areas so that the value associated with a centroid can be used as a measure of the height of the surface at that point. One needs a relatively large number of points to represent a surface so it is better to use the 510 FSAs than the 36 PHUs. However, problems arise if the density of centroids varies substantially across the province as is the case for Ontario (Figure 10). Using standard mapping software to model a surface for the whole of Ontario produce highly unreliable estimates for many parts of northern Ontario where observation points are very sparse. As a compromise maps were produced for Southern Ontario only although data for the whole of Ontario were used in the estimation of the Southern Ontario rates. Even in this case there are still problems with sparseness of observations in the rural areas to the north-east of Toronto.
ESRI’s ArcMap™9.2 was used in the construction of the maps, using the method of kriging. Kriging is a method of interpolating values in spatial data based on spatial variation in observed data and on minimizing prediction errors[3]. For each map a first-order trend was fitted to assess whether the overall surface indicated broad regional trends in the distribution of calls. Model maps of Southern Ontario were constructed for the same 4 map sequences as for the PHU data. Selected maps for key dates for Telehealth Ontario calls in the 2004-5 season are presented in Figure 11. Similar trends are evident in these maps as for the Emergency Department data at the PHU scale (Figure 8). Respiratory calls originally rose in the eastern part of the province but rapidly spread both westward and northward. As the epidemic died down in April, it lingered longest in the South West.

Conclusion
The overall impact of the mapping exercise is positive. Not only do the respiratory data from NACRS and Telehealth Ontario closely follow the temporal sequence of confirmed influenza cases reported by CCDR for the province as a whole, but the temporal sequence of each data series are very similar at the PHU scale. When the temporal sequences are mapped at the PHU scale, a number of important aspects of the 2004-5 and 2005-6 influenza epidemics emerge:

- The two influenza seasons are not only different in intensity but also in the spatial patterns of their evolution, although some central features of the spread are common to both years.
- Although the epidemic can start in a variety of places in the province it moves to the large urban centres rapidly. It then spread to all parts of the province very quickly. The rapid linkages between the main population centres provided by the transport network are more important than mere proximity. The tail end of the epidemic occupies different spatial locales in each year.
- The implication is that once the epidemic emerges in a specific locale in the province, all areas of the province should be alert for its potential arrival in their area within 1-2 weeks (it could take a little longer but the specifics of the transmission cannot be predicted with any greater accuracy.

There are data implications that arise from this exercise:

- Both the NACRS and the Telehealth Ontario databases are valuable tools in the monitoring of the spread of influenza in the Province. Although both databases identify
respiratory illness rather than influenza *per se*, the data track the temporal evolution of the CCDR confirmed cases so closely that there is much confidence in the ability to identify the upsurge in influenza cases in the community.

- The restriction of the geographic coding of released data to 3-digit postal codes (FSAs) forces the introduction of estimation techniques in generating PHU counts for Telehealth Ontario calls. While reasonably robust, any such estimates require assumptions about local distribution of calls which are not easily testable. It would be useful to explore ways of releasing more detailed geocodes (such as 6-digit postal codes or coding to Statistics Canada Dissemination Areas) while maintaining appropriate protection of individual privacy. This would be an important step if links to other resources such as local socio-economic and demographic profiles available from the Census were desired for further analysis of factors affecting the spread of epidemics.
References


TABLE 5-1: Lagged Correlations of Weekly Data for Emergency Department and Telehealth Ontario Records

<table>
<thead>
<tr>
<th>Province</th>
<th>Tel (-2)</th>
<th>Tel (-1)</th>
<th>Tel (0)</th>
<th>Tel (+1)</th>
<th>Tel (+2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.792</td>
<td>0.883</td>
<td>0.946</td>
<td>0.942</td>
<td>0.895</td>
</tr>
</tbody>
</table>
Figure 5-1: Comparative data series by week – 2004-6
Figure 5-2A: 3-digit postal codes (FSAs) in Ontario
Figure 5-2B: 3-Digit postal codes (FSAs) in the Toronto region
Figure 5-3: Rural FSAs surrounding Ottawa and Belleville
Figure 5-4: Ontario Public Health Units in 2005
Figure 5-5: Overlaps of FSA and PHU boundaries in Ottawa
Figure 5-6A: Scaled intensity of emergency room use – Ontario 2004-6
Figure 5-6B: Scaled intensity of emergency room use – Toronto region 2004-6
Figure 5-7A: Scaled intensity of Telehealth Ontario use – Ontario 2004-6
Figure 5-7B: Scaled intensity of Telehealth Ontario use – Toronto region 2004-6
Figure 5-8: Progress of the 2004-5 influenza season at the PHU scale
Figure 5-9: Progress of the 2005-6 influenza season at the PHU scale
Figure 5-10: Distribution of FSA centroids in Ontario
Figure 5-11: Model surfaces for the 2004-5 influenza epidemic for Southern Ontario