Time-Series Analysis of the Relation between Influenza Virus and Hospital Admissions of the Elderly in Ontario, Canada, for Pneumonia, Chronic Lung Disease, and Congestive Heart Failure

R. E. G. Upshur, Keith Knight, and Vivek Goel

This study examined the relation between the presence of circulating influenza virus and all hospital admissions of people over age 65 years in Ontario, Canada, for pneumonia, congestive heart failure, and chronic lung disease in 1988–1993. Autoregressive integrated moving average (ARIMA) transfer function models were used to perform a time-series analysis. These models were compared with simple cross correlations by using Pearson's product moment correlation. The results showed statistically significant correlations between the presence of influenza virus and admissions of the elderly for pneumonia (in all 5 years under study) and chronic lung disease (in 4 of the 5 years under study). The relation between circulation of influenza virus and admissions for congestive heart failure was inconsistent. The simple cross correlation tended to overestimate the association between the presence of circulating influenza strains and hospital admissions. Measures of the impact of influenza should include chronic lung disease as an outcome. Further studies, with greater covariate control, are required to delineate more precisely the relation between influenza and hospital morbidity in the elderly. This study demonstrates the power and utility of using time-series methods in the epidemiologic study of communicable diseases. Am J Epidemiol 1999; 149:85–92.

Influenza is a major cause of morbidity and mortality. While a large body of literature exists concerning the relation between influenza epidemics and mortality, the extent of morbidity caused by influenza remains largely unexplored. It has recently been proposed that administrative data, such as hospital admissions databases, may provide a more sensitive measure of the impact of influenza epidemics (1). One method of exploring the impact of influenza is to use time-series techniques to quantify the relation between increased circulation of influenza virus and increased numbers of admissions for influenza-sensitive conditions.

Understanding the epidemiology of influenza and its impact on morbidity and mortality is crucial if control of influenza is to be achieved. In both Canada and the United States, influenza control efforts center on immunization of those at highest risk for complications from influenza infection. The elderly are at the highest risk for mortality and morbidity due to influenza. Considerable resources are devoted to implementation of immunization programs to ensure that a large proportion of those at highest risk are vaccinated. Analyzing the morbidity associated with influenza is one means of assessing the extent to which morbidity has been reduced.

Autocorrelation is an important consideration in time-series regressions. Data collected over time may not satisfy the assumptions used to calculate a simple correlation coefficient. Specifically, the values of variables may be predicted by using previous values. Therefore, the observations may not be independent; thus, the residual terms will not necessarily be independent and identically distributed with a normal distribution. Time-series methodology has a long history of application in econometrics, particularly in the domain of forecasting. It has recently captured the attention of epidemiologists (2–5). Time-series analysis enables the comparison of two sets of data collected over the same time period while controlling for autocorrelation. Influenza epidemics display features of autocorrelation, in that daily and weekly counts of
cases are not independent. The same holds true for admission rates associated with evolving influenza epidemics.

Time-series studies have focused on mortality due to influenza and pneumonia to forecast mortality or to calculate excess mortality (6, 7). Time-series methods have also been used extensively to study the effects of air pollution on mortality while controlling for the effects of influenza (8–11) and to examine the relation between the seasonal patterns of influenza and meningococcemia (12, 13).

To our knowledge, no full-scale time-series analyses have looked at the relation between circulating influenza viruses and hospital admissions of the elderly. Perrotta et al. (14) reported a strong correlation between circulation of virus and increases in hospital admissions of adults. Their correlation coefficient of 0.74 is impressive and has been quoted widely. However, the Perrotta et al. study did not indicate the method by which the cross correlations were calculated. The tabular data in the article indicate a number of strongly positive correlations at different lags, which gives rise to the suspicion that the series were simply cross correlated. As Helfenstein (4) and Bowie and Prothero (5) have noted, simple cross correlations between two seasonally related variables will result in artifactual elevation of the cross-correlation coefficients.

There is some controversy in the literature concerning the association between circulating influenza virus and increased admissions of the elderly for congestive heart failure and chronic lung disease (15–17). The use of transfer function models enables a more precise estimate of the strength of the correlation between the presence of influenza and hospital admissions, and it renders an accurate estimate of the temporal relation between the two series. Our study extended the use of time-series methodology to the analysis of congestive heart failure and chronic lung disease. We are unaware of any published studies that use time-series methodology to assess the relation between these conditions and the presence of circulating influenza virus. Therefore, our study sought to estimate the strength of the correlation between circulation of influenza virus and admissions for pneumonia, chronic lung disease, and congestive heart failure during each of five influenza seasons.

MATERIALS AND METHODS

Ontario influenza surveillance data

Influenza surveillance is conducted annually, from October to April, by the Public Health Branch of the Ministry of Health in Ontario, Canada. The data gathered include the results of laboratory surveillance, reports of influenza obtained from the reportable disease information system, reports of institutional outbreaks, and information obtained as a result of absenteeism and illness surveillance at selected schools.

Laboratory confirmation results from isolation of the virus from nasal or pharyngeal secretions, direct detection of antigen from nasal or pharyngeal secretions, or a demonstrated fourfold rise in hemagglutination titers to influenza virus from serologic specimens collected at the onset of illness and 4 weeks later. The majority of laboratory-confirmed influenza infections are detected by using either of these first two methods. For this study, influenza seasons were defined on the basis of reported laboratory-confirmed cases in Ontario. An influenza season commenced with the isolation of influenza virus during consecutive weeks and ended when no further isolations were made. Weekly totals for the influenza seasons were obtained from the Laboratory Centre for Disease Control in Ottawa.

Hospital discharge abstracts

The Canadian Institute for Health Information collects and analyzes health information derived from a standardized discharge separation abstract. When a patient is discharged from the hospital, a medical records clerk abstracts administrative and clinical variables on that patient's chart, such as age and sex, residence, admission date, discharge date, length of stay, and as many as 15 diagnosis codes and procedures. Discharge diagnoses are classified according to the International Classification of Diseases, Ninth Revision (ICD-9).

All 1988–1993 discharges of Ontario residents aged 65 years or older were abstracted for the fiscal year April 1–March 31. The following ICD-9 codes were included: 480.0–487.0, pneumonia and influenza; 490.0–492.0, chronic lung disease; and 428.0–428.9, congestive heart failure. The codes were aggregated to create diagnosis groups, and a hierarchic rule was used to create data sets for each diagnosis group. Each diagnosis group was constructed by extracting records for each condition when it appeared in the first or second diagnosis code, exclusive of the presence of the other diagnosis codes under consideration.

Analysis

Simple correlations were calculated by using Pearson's product moment correlation (18). Autoregressive integrated moving average (ARIMA) models were fit to the influenza series for each influenza season. The Ljung-Box Q statistic was used to assess residual autocorrelation in the input series (19). The regression coefficients and Q statistics for each season are shown in table 1. A first-order autoregressive model was found to fit each input series adequately. First-order autoregressive models were imposed on the aggregate output series for pneumonia, congestive heart failure, and chronic lung disease for each influenza season. The residuals of the input series were cross correlated to the residuals of the output series and the correlation coefficients, and lags were calculated. Statistical significance was defined as a correlation coefficient of more than twice the standard error. Time-series analysis was performed by using EvIEWS 2.0 software (19).

RESULTS

Descriptive statistics

The number of mean weekly admissions for each diagnosis group and the total number of admissions in each series are shown in table 2. Figure 1 shows the overall series of admissions aggregated for the diagnosis group pneumonia and the number of influenza isolates for the period 1988–1993. Figures 2 and 3 show the overall series of admissions for chronic lung disease and congestive heart failure, respectively, during those same years.

In figures 1 and 2, conspicuous spikes are evident in almost every year shown, corresponding to the presence of influenza virus. However, figure 3, representing the number of admissions for congestive heart failure, lacks the striking seasonal peaks that were shown for pneumonia and chronic lung disease in figures 1 and 2, respectively. As shown in figure 3, the overall

<table>
<thead>
<tr>
<th>TABLE 1. Regression coefficients* and Q statistics for the influenza input series, Ontario, Canada, 1988–1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influenza series</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>1988–1989</td>
</tr>
<tr>
<td>1989–1990</td>
</tr>
<tr>
<td>1990–1991</td>
</tr>
<tr>
<td>1991–1992</td>
</tr>
<tr>
<td>1992–1993</td>
</tr>
</tbody>
</table>

* The unit of analysis is a week and is a continuous variable; these values refer to the first-order autoregressive process as applied to the week value.


Am J Epidemiol Vol. 149, No. 1, 1999
Statistics for hospital admissions for pneumonia, congestive heart failure, and chronic lung disease among people older than age 65 years, Ontario, Canada, 1988–1993

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>No. of mean weekly admissions (SD)*</th>
<th>Median</th>
<th>Range</th>
<th>Total no. of admissions in series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>248.93 (81.60)</td>
<td>235</td>
<td>118–689</td>
<td>77,418</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>347.18 (51.13)</td>
<td>345</td>
<td>225–489</td>
<td>107,972</td>
</tr>
<tr>
<td>Chronic lung disease</td>
<td>178.77 (37.14)</td>
<td>173</td>
<td>116–359</td>
<td>55,598</td>
</tr>
</tbody>
</table>

* SD, standard deviation.

The correlation coefficients and lag times for each pneumonia season are shown in table 3. Statistically significant positive correlations were found for each season. However, the lags varied from season to season. During 1989–1990, 1990–1991, and 1991–1992, the series were related instantaneously, and there was no lag time between increased influenza circulation and increased numbers of admissions of elderly Ontarians. The Pearson $R$ ranged from 0.42 to 0.64. For the 1992–1993 series, pneumonia admissions were significantly correlated at lag -1 ($R = 0.48$) and lag -3 ($R = 0.49$) weeks. The 1988–1989 season showed a significant correlation at lag 2 weeks ($R = 0.55$), indicating that the number of admissions increased before the number of influenza isolations increased.

Congestive heart failure

The correlation coefficients and lag times for each congestive heart failure season are shown in table 4. This series was found to have no significant relation during the 1988–1989 and 1990–1991 influenza seasons. In 1989–1990 and 1990–1991, admissions were significantly negatively correlated with influenza circulation at lag 2 ($R = -0.46$) and lag 10 ($R = -0.44$) weeks. The 1992–1993 season showed a significant positive correlation at lag -3 weeks ($R = 0.54$).

Chronic lung disease

The correlation coefficients and lag times for each chronic lung disease season are shown in table 5. No significant relation was found during the 1992–1993 season. The other four influenza seasons showed significant positive correlations, with $R$ values ranging from 0.39 to 0.67. The lag times varied from -3 to 1.

DISCUSSION

Pneumonia

Our time-series analysis indicated a moderately strong correlation between circulating influenza virus and hospitalizations for pneumonia among the elderly. The explained variance between the two series changed from year to year. The range of $R^2$ values from 23 to 41 indicates that the presence of circulating influenza virus explains some but not all of the variance related to hospitalizations. In 3 of the 5 years, the series were related instantaneously, and no lag between influenza circulation and hospital admissions for pneumonia occurred. An instantaneous lag indicates that both series peak at the same time, which can be interpreted as suggesting that a very strong, possibly causal relation exists between the circulation of influenza and increased numbers of admissions of the
elderly for pneumonia. Given the extensive background knowledge of the association of influenza with pneumonia, this relation is highly plausible biologically, and the temporal relation is consistent. A positive lag indicates that the admission rate follows the increase in influenza isolates by the value of the lag in weeks. A negative lag shows that the admission series increases before the numbers of influenza isolates increase. The 1992–1993 influenza season was the only one associated with a lag of hospital admissions and circulating influenza strains. During the 1988–1989 season, influenza isolates peaked 2 weeks after the number of admissions began to rise. The inconsistent relation during this season may have been due to variability in when isolates commenced to be reported. There is some evidence of year-to-year variations in when physicians begin to report influenza (20).

Our correlation coefficients were lower and the lag times were different from those reported by Perrotta et al. (14). Their study showed a correlation coefficient of 0.74 at lag -1 week between admissions of adults with acute respiratory disease and circulation of influenza strains. There are several potential reasons for the differences in findings.

The Perrotta et al. (14) study used data from a series of isolates collected randomly by a stratified sample of community physicians. This surveillance method is more likely to capture the onset and peak of the circulation of virus than is the Ontario system, which relies largely on passive reporting from a selected set of participating laboratories. The majority of Ontario isolates were derived from pediatric hospitals and from outbreaks in nursing homes (21). In years with substantial nursing home outbreaks, it is likely that the peak circulation of influenza will correspond more closely to hospitalizations of the elderly, as it is reasonable to believe that the two events would be closely related. Therefore, because surveillance in Ontario relies substantially more on isolates obtained from the elderly than does a stratified community sample, it is possible that the reported lags may vary from the true relation.

Only a study in which surveillance physicians collect data from random samples could the relation be verified between lag times for admission and circulating influenza in the community.

It is also likely that the variability in time lags and the explained variance from year to year depend on other plausible causes, such as temperature, ambient air quality, and comorbidities in the persons acquiring influenza, as well as circulation of other respiratory viruses. The moderate correlation coefficients for each influenza season indicate that such factors play a role.

Recent studies have demonstrated that respiratory syncytial virus (RSV) is likely a substantial contributor to morbidity and mortality in the elderly. Studies by Fleming and Cross (22), Nicholson (23), and Falsey et al. (24) showed that it is an important contributing factor. The Nicholson and Falsey et al. studies indicated that the incidence of morbidity and mortality may be greater for RSV than for influenza and probably is not restricted to the institutionalized elderly, although further research to clarify this relation is required. A study by Dowell et al. (25) indicated that RSV is a significant respiratory pathogen in the community. If the association of RSV with morbidity and mortality in the elderly is as significant as these studies suggest, then the exact role of influenza as the major contributor to respiratory morbidity and mortality in the elderly will have to be reconsidered. Certainly, the magnitude of preventable hospitalization and mortality due to influenza may be lower than previously believed.

In Ontario, RSV cocirculates with influenza on an annual basis (26). Given that RSV and influenza are clinically indistinguishable in the elderly and may in fact be present simultaneously in the same person (27, 28), further studies are required to separate the effects of RSV and influenza and clarify their impact on hospital admissions.

The Nicholson study (23) also indicated that temperature variables are significantly associated with winter mortality. Temperature and humidity, and air quality variables such as sulphur dioxide and ozone, are plausibly linked to respiratory morbidity and mortality (8, 10, 29–33). Kim et al. (33) recently demonstrated a positive association between the presence of invasive pneumococcal disease, air pollution, and respiratory virus circulation. It is likely that these variables influence the number of hospitalizations for pneumonia among the elderly. As temperature and air quality data are also collected systematically, a future study including these variables is recommended.

### Congestive heart failure

To our knowledge, no published studies have used time-series analysis to link circulating influenza

### TABLE 5. Cross correlations between influenza and hospital admissions for chronic lung disease among people older than age 65 years, Ontario, Canada, 1988–1993

<table>
<thead>
<tr>
<th>Influenza series</th>
<th>Lag (week)</th>
<th>Pearson $R$</th>
<th>Pearson $R$ at lag 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988–1989</td>
<td>1</td>
<td>0.67</td>
<td>0.53</td>
</tr>
<tr>
<td>1989–1990</td>
<td>0</td>
<td>0.61</td>
<td>0.62</td>
</tr>
<tr>
<td>1990–1991</td>
<td>–3</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>1991–1992</td>
<td>–2</td>
<td>0.56</td>
<td>0.72</td>
</tr>
<tr>
<td>1992–1993</td>
<td>–</td>
<td>–†</td>
<td>0.38</td>
</tr>
</tbody>
</table>

* No significant correlation.
† Not applicable.
strains to hospital admissions for congestive heart failure. The results of our correlation study indicate that there is no consistently strong relation between the two. Only during the 1992–1993 influenza season did we find a plausible significant relation. During the 1989–1990 and 1990–1991 seasons, there were negative correlations; in 1988–1989 and 1991–1992, there were no significant correlations. These results contrast with reports of increases in admission rates for congestive heart failure and corresponding significant decreases in admissions for congestive heart failure among persons who have been vaccinated against influenza (16, 34).

**Chronic lung disease**

As with pneumonia, the time-series study indicated a significant relation between circulating influenza strains and admissions for chronic lung disease. During each influenza season except 1992–1993, a significant positive correlation existed. The Pearson $R$ values for the correlation tended to be slightly higher for the chronic lung disease series than for the pneumonia series. As with the pneumonia series, the presence of influenza explains some but not all of the variance in admissions. The same arguments advanced for pneumonia hold here: Further studies with a more complete and exhaustive set of explanatory variables are required.

**Comparison of raw correlations with ARIMA-modeled cross correlations**

For four of the five pneumonia series, using the raw correlation coefficient overestimated the relation between the presence of circulating influenza virus and the number of hospital admissions. The $R$ values were inflated 8–31 percent. For the three seasons during which the relation between the influenza series and the pneumonia series was instantaneous, the $R$ values were comparable only in 1991–1992. For congestive heart failure, the pattern was inconsistent. During two seasons, negative correlations became positive; two seasons with no significant correlation by ARIMA modeling became significantly negatively correlated. For the chronic lung disease series, the $R$ values were comparable, with the ARIMA-modeled values being slightly higher. They were 26 percent higher for the 1988–1989 season and 23 percent lower for the 1991–1992 season.

Note that the raw correlations at lag 0 are not the highest correlation coefficients. When cross correlation is performed at various lags, there are higher values at different lags. For example, the 1991–1992 influenza season had correlation coefficients of as high as 0.88 at lag 3. The Perrotta et al. study (14) reported the highest correlation coefficient at lag -1. When their tabular data are examined, it is evident that they chose the highest in a series of lagged correlation coefficients as the most significant. This indicates that performing simple cross correlations will miss relations at different lags and that simply cross correlating lagged series will result in a multitude of difficult-to-interpret results.

This study is subject to data and methodologic limitations. The Ontario influenza surveillance data rely on passive reporting. Community physicians rarely order diagnostic tests for influenza in either a practice or a hospital setting. As a consequence, the number of isolations detected on an annual basis must be regarded as an underestimation of the true extent of influenza circulation. The results of undercounting would make the association less precise. A more accurate estimate of the number of isolations would add more information to the input series to enable cross correlations to be calculated. The second limitation of passive reporting is that it may miss the true onset of circulating influenza virus. A study by Quenel et al. (20), using data from a large sentinel surveillance system, examined the relation between health-service-based indicators and influenza viruses. They found that health service indicators increased earlier than virologic isolations. In other words, only after an index of suspicion was raised were isolations for virus obtained.

Influenza vaccination may also influence admission rates. In our study, the vaccination status of those admitted was unknown. Nichol et al. (17) and Fedson et al. (35) both reported lower admission rates for those who were vaccinated for the conditions targeted in this study. A further limitation of the analysis is that there was no verification of the presence of influenza in those admitted to the hospital.

Time-series analysis is also subject to limitations. In this study, transfer function models were fit to influenza seasons only. The duration of these seasons ranged from 16 to 22 weeks. Many time-series analysts caution that 30–50 data points are required to adequately calculate the autocorrelation and partial autocorrelation functions necessary to specify an appropriate model (36). The concern with small samples is that the $Q$ statistic may lack the power to detect remaining autocorrelation in the residual series. That is, it may proclaim the residuals to be white noise when in fact they are not. However, these constraints should not be regarded as absolute. In our analysis, to get 30–50 data points would have entailed adding zero values to the input series, in effect reducing the mean value of the series and biasing the model.
Certainly, evolving influenza seasons have the hallmark features of a first-order autoregressive process, in that weekly or daily values are not independent but are closely related to the immediately preceding value. In this study, all data were collected using the same unit of measure (week). It makes biologic sense to model hospital admissions as an output only in the presence of influenza circulation. When each season is looked at independently, the presence of other cyclic influences is removed, as is the need for seasonal differencing. Time-series techniques are particularly useful for studying phenomena that vary concomitantly. By assuming that many of the potentially confounding covariates are relatively stable on a week-to-week basis (such as gender, smoking status), their potential influence as explanatory variables can be largely controlled.

This study is unique in two respects: 1) It applied time-series transfer function methodology to five independent influenza seasons for three different diagnosis groups, and 2) it focused exclusively on the impact of influenza on the elderly by using administrative data from a large geographic area. Time-series methodology has been applied to study the association of influenza with pneumonia and acute upper respiratory admissions, but no known published studies exist that examine the association with congestive heart failure and chronic lung disease. The results of our study indicate that hospital admissions for chronic lung disease and pneumonia in the elderly population of Ontario were strongly correlated with the presence of circulating influenza virus. However, no convincing relation existed for admissions for congestive heart failure. The presence of influenza did not completely explain the increase in hospital admissions; other explanatory variables, such as RSV, temperature, and air quality, should be included in future studies.

Time-series methods are a useful addition to the study of communicable disease. The methodology helps to illustrate the temporal aspects of epidemics and is now readily available to epidemiologists. The systematic collection of influenza isolates via sentinel surveillance by physicians would increase the precision, accuracy, and utility of time-series modeling of influenza epidemics. Such a sentinel surveillance system has been established in Canada.

ACKNOWLEDGMENTS

Dr. Goel is supported in part by the National Health Research & Development Program.

REFERENCES


Am J Epidemiol Vol. 149, No. 1, 1999


