Original Contribution

The Co-Seasonality of Pneumonia and Influenza With *Clostridium difficile* Infection in the United States, 1993–2008

Kevin A. Brown*, Nick Daneman, Paul Arora, Rahim Moineddin, and David N. Fisman

* Correspondence to Kevin A. Brown, Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario, Canada M5T 3M7 (e-mail: kevin.brown@mail.utoronto.ca).

Initially submitted June 5, 2012; accepted for publication November 21, 2012.

Seasonal variations in the incidence of pneumonia and influenza are associated with nosocomial *Clostridium difficile* infection (CDI) incidence, but the reasons why remain unclear. Our objective was to consider the impact of pneumonia and influenza timing and severity on CDI incidence. We conducted a retrospective cohort study using the US National Hospital Discharge Survey sample. Hospitalized patients with a diagnosis of CDI or pneumonia and influenza between 1993 and 2008 were identified from the National Hospital Discharge Survey data set. Poisson regression models of monthly CDI incidence were used to measure 1) the time lag between the annual pneumonia and influenza prevalence peak and the annual CDI incidence peak and 2) the lagged effect of pneumonia and influenza prevalence on CDI incidence. CDI was identified in 18,465 discharges (8.52 per 1,000 discharges). Peak pneumonia prevalence preceded peak CDI incidence by 9.14 weeks (95% confidence interval: 4.61, 13.67). A 1% increase in pneumonia prevalence was associated with a cumulative effect of 11.3% over a 6-month lag period (relative risk = 1.113, 95% confidence interval: 1.073, 1.153). Future research could seek to understand which mediating pathways, including changes in broad-spectrum antibiotic prescribing and hospital crowding, are most responsible for the associated changes in incidence.

*Clostridium difficile*, influenza, human; pneumonia; seasons; United States

Abbreviations: CDI, *Clostridium difficile* infection; CI, confidence interval; RR, relative risk.

*Clostridium difficile*, a toxin-producing bacterium that causes diarrhea, has been identified as the largest single cause of morbidity and mortality among hospital-acquired infections in the Canadian province of Ontario (1). National studies of short-term care hospitals in the United States and Europe show that *C. difficile* infection (CDI) incidence in these jurisdictions has more than doubled since 2000, with increases concentrated in patients over 65 years of age (2–4). Several explanations have been proposed for the increasing rates of CDI and include increasing patient age, acuity, comorbidities, increasing antimicrobial prescribing in both the community and hospitals, and the rapid spread of a previously rare strain (5). Other identified risk factors for hospital-acquired CDI include length of stay, exposure to symptomatic cases of CDI (6), and hospital factors including large hospital size and teaching hospital status (3).

Descriptive studies of health-care admission patterns demonstrate large wintertime surges of pneumonia and influenza hospital admissions, particularly among infants and the elderly (7). Influenza-related admissions are increasingly concentrated in the elderly, and this has been attributed to the increasing frequency of A(H3N2) strain-dominant seasons and the decreasing severity of influenza A(H3N2) seasons for the younger age groups (8). Changes in CDI distribution demonstrate similar patterns (3). Recently, CDI incidence has been shown to be associated with the seasonal variations in the incidence of pneumonia and influenza hospitalizations (9), but this may be attributable to increased wintertime hospitalizations rather than increased risk.

If respiratory disease attributable to influenza is an important driver of wintertime surges in CDI, such a link would have important implications for influenza vaccination programs.
and could change the risk calculus related to antibiotic prescribing for wintertime respiratory illness. We sought to evaluate the link between the seasonality of pneumonia and influenza and CDI using a large, nationally representative, multihospital database from the United States. We sought to do this by 1) estimating the mean annual timing of both the CDI peak and the pneumonia and influenza peak, as well as the time lag between the 2 (model 1); and 2) ascertaining the magnitude of impact of fluctuations in pneumonia and influenza on subsequent CDI incidence using optimal statistical methods (model 2) (10).

MATERIALS AND METHODS

Study design and setting

We conducted a retrospective cohort study using the National Hospital Discharge Survey sample that consists of diagnosis and demographic data collected from a probability sample of patient discharge records from nonfederal, short-stay hospitals. Because of unequal selection probabilities resulting from the 3-stage cluster sample design, weights are used to make the sample nationally representative. Between 1993 and 2007, approximately 270,000 discharges were sampled annually from a panel of more than 500 hospitals (11). In 2008, the number of sampled discharges was halved. The National Hospital Discharge Survey has been used previously for the study of CDI incidence (3). The source cohort consisted of all patients ≥40 years of age admitted to an acute care hospital in the United States in the period from January 1, 1993, to December 31, 2008. Younger individuals were excluded because of low risk of CDI. Patients with a length of stay of ≤2 days were excluded because these patients were at negligible risk of being diagnosed with nosocomial CDI.

Outcome measurement

CDI case patients were identified from the full cohort of hospitalized patients using the International Classification of Diseases, Ninth Revision, Clinical Modification, discharge code 008.45 for CDI. Compared with toxin-positive test results, the National Hospital Discharge Survey results have good concordance (κ = 0.64) and neither under- nor overestimate hospital CDI incidence (12). All patients in the National Hospital Discharge Survey with a primary or secondary diagnosis of CDI, who were not excluded because of young age or ≤2-day length of stay, were included. Patients were considered to have acquired CDI in the month of discharge.

Measurement and validation of hospital pneumonia and influenza prevalence

Hospitalized patients with a diagnosis of pneumonia and influenza (International Classification of Diseases, Ninth Revision, Clinical Modification, codes 480–488) were also identified from the National Hospital Discharge Survey data set. The hospital prevalence of pneumonia and influenza was calculated as the proportion of monthly discharges having a diagnosis of pneumonia and influenza listed as a primary or secondary diagnosis among all admitted patients ≥40 years of age. Annual excess pneumonia and influenza hospitalization rates are a strong predictor of excess pneumonia and influenza mortality (8). Also, we validated the correlation between virological influenza H3N2 strain (%) predominance and maximum hospital pneumonia and influenza prevalence (in the November 1–March 31 interval) of a given 1-year period to ensure that H3N2 was positively associated with the hospital pneumonia and influenza burden (13).

Other covariates

We derived clinical and hospital variables including patient age at admission (categorized into 10-year age bins), sex, period of discharge, census region of hospital (Northeast, Midwest, South, and West), hospital bed size, the Charlson comorbidity index, and the following 17 specific comorbidities: myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatological disease, peptic ulcer disease, mild liver disease, diabetes, hemiplegia or paraplegia, renal disease, malignant cancer, liver disease, metastatic tumor, acquired immunodeficiency syndrome, and pneumonia and influenza (14). Discharge month was considered as a continuous variable ranging from 1993 to 2008 by increments of 1 in 12.

Statistical analyses

Several time-series analyses were conducted in order to ascertain whether hospital pneumonia and influenza prevalence preceded and predicted CDI incidence. For these analyses, Poisson regression was used; weighted counts per month were modeled with an offset term corresponding to the sum of the weights of all discharges during the month. For each year of data, survey weights (w) were rescaled to have a mean of 1 and then divided by the design effect (Appendix) to obtain rescaled weights (w*), so that monthly case counts equaled the equivalent number of cases under a simple random sample design, also known as the effective sample size (15). We present customary 2-sided 95% confidence intervals.

Effect of pneumonia and influenza peak timing. Two harmonic regression models were fitted for both the hospital pneumonia and influenza (P&I) prevalence and the CDI counts, to determine peak timing and amplitude of average seasonal incidence. Harmonic regression models are a type of statistical model that explicitly include time as a covariate; they are used to characterize seasonal disease incidence in terms of amplitude (equal to the ratio of peak incidence to trough incidence) and phase shift (16, 17) (Figure 1). The harmonic regression models were as follows:

\[
N_{P&I}(t) \sim \exp[\beta_0 + \beta_1 \sin(2\pi t) + \beta_2 \cos(2\pi t) + \beta_3 t + \beta_4 t^2 + \log(w^*)]
\]
vectors. Specifically, we performed a harmonic analysis combining the current and the preceding 23 months, where \( q \) is a vector of polynomial basis vectors, and \( \cdot \) is the dot product of the 2 vectors. Specifically, the polynomial basis vectors were \( e_1 = \{1, 1, \ldots , 1\} \), \( e_2 = \{1, 2, 3, \ldots , 24\} \), and \( e_3 = \{1, 4, 9, \ldots , 576\} \). For all distributed lag models, we removed secular trends from the pneumonia and influenza time series using a quadratic polynomial and used the residuals as predictor variables.

**Age-region stratified analyses.** For the harmonic and distributed lag models above, we considered age-region stratified analyses; specifically, we fit models with the same covariates described above to each of the 20 age-region strata, separately. For these analyses, monthly age-region CDI and pneumonia and influenza (rescaled) case counts and weights were extracted. We assessed the pooled effects using inverse-variance weights and a fixed-effects approach (19) and the heterogeneity of study results by use of the \( I^2 \) statistic (20). In order to consider predictors of age-region–specific CDI timing, we created 3 bivariate linear regression models where age group, region, and pneumonia and influenza phase shift were each entered as a covariate for predicting CDI phase shift; measurement error of the pneumonia and influenza phase shift was accounted for using Deming regression.

Analyses were conducted in R using the dlnm and metafor packages (10, 21).

**RESULTS**

Over the 16-year period from January 1, 1993, to December 31, 2008, there were a total of 4.39 million discharges, of which 1.94 million (44.2%) discharges met the study inclusion criteria (Figure 2).

Over the 16-year period, CDI was identified in 18,465 discharges (8.52 per 1,000 discharges). The incidence of CDI was highest in the northeastern census region (11.50 per 1,000 discharges) and lowest in the southern census region (6.84 per 1,000 discharges; relative risk (RR) = 0.59, 95% confidence interval (CI): 0.56, 0.63) (Table 1). Incidence increased through the 4-year periods: from 4.70 per 1,000 discharges between 1993 and 1996 to 13.24 per 1,000 discharges between 2005 and 2009 (RR = 2.82, 95% CI: 2.53, 3.13).

\[
N_{CDI}(t) = \exp[\beta_0 + \beta_1 \sin(2\pi t) + \beta_2 \cos(2\pi t)] + \beta_3 t (t < 2,000) + \beta_4 t (t \geq 2,000) + \log(w^+)],
\]

where \( t \) is month of discharge and January 2000 corresponds to the approximate timing of NAP1 emergence in the United States (18), \( w^+ \) is the rescaled weighted discharge count, and \( N \) is the rescaled weighted count of cases. For each harmonic model, peak week number (in weeks since January 1st) was obtained by computing \( 52 \tan^{-1} (\beta_1/\beta_2) \), the inverse tangent of the ratio of the sine function coefficient divided by the cosine function coefficient. The amplitude was equal to \( 2(\beta_1^2 + \beta_2^2)^{1/2} \), where \( 2(\beta_1^2 + \beta_2^2)^{1/2} \) is the amplitude in log scale. The difference in phase shift parameter between the pneumonia and influenza time series and the CDI time series was computed by subtraction. Confidence limits for amplitude, phase shift, and phase shift difference were computed on the basis of the \( \delta \) method (17).

**Effect of pneumonia and influenza fluctuations.** In order to consider the impact of fluctuations of pneumonia and influenza hospitalizations on CDI incidence, we developed a second Poisson regression model, in which CDI incidence was dependent on lagged pneumonia and influenza discharges (10), as well as linear spline trend terms. As such,

\[
N_{CDI}(t) = \exp[\beta_0 + \beta_1 q \cdot c_1 + \beta_2 q \cdot c_2 + \beta_3 q \cdot c_3 + \beta_4 t (t < 2,000) + \beta_5 t (t \geq 2,000) + \log(w^+)],
\]

where \( q \) is a vector of pneumonia and influenza prevalence combining the current and the preceding 23 months, \( c_n \) are polynomial basis vectors, and \( \cdot \) is the dot product of the 2 vectors. Specifically, the polynomial basis vectors were

---

**Figure 1.** Measurement of a difference in peak timing for 2 annual harmonic time series. For each time series, the timing of peak incidence or phase shift \((w_1, w_2)\) is calculated by using the coefficients of the sine and cosine terms. The phase shift difference \((w_2 - w_1)\) can then be computed by subtraction. On the x-axis, weeks 0 and 52 correspond to the date of January 1 in consecutive years.

**Figure 2.** Composition of the study population, United States, 1993–2008. All numbers (n) correspond to unweighted counts of sampled hospital discharges. CDI, *Clostridium difficile* infection; LOS, length of stay; NHDS, National Hospital Discharge Survey; P&I, pneumonia and influenza.
CDI risk factors

CDI incidence varied from 4.53 per 1,000 discharges in patients aged 40–49 years to 12.10 among patients aged 80 or more years; each 10-year increase in patient age was associated with a 28% increase in incidence (95% CI: 1.25, 1.31) of CDI. Female patients were 1.20 (95% CI: 1.13, 1.27) times more likely to have CDI during a hospitalization. Relative to patients with a length of stay of 3–5 days, patients with a length of stay of 6–10 days were 1.94 times (95% CI: 1.81, 2.10) more likely to be diagnosed with CDI. The Charlson index was significantly associated with CDI ($F = 7.5$, $P < 0.001$); however, the association was not linear. When comorbidities were considered individually, we found that pneumonia (RR = 1.50, 95% CI: 1.38, 1.63), heart failure (RR = 1.17, 95% CI: 1.09, 1.26), chronic renal failure (RR = 1.86, 95% CI: 1.67, 2.08), malignant cancer (RR = 1.33, 95% CI: 1.21, 1.45), and acquired immunodeficiency syndrome (RR = 1.90, 95% CI: 1.42, 2.55) were associated with increased risk of CDI, while cerebrovascular disease (RR = 0.60, 95% CI: 0.53, 0.69), dementia (RR = 0.74, 95% CI: 0.56, 0.98), and hemiplegia (RR = 0.58, 95% CI: 0.43, 0.78) were associated with decreased risk. The most common comorbid conditions in patients with a CDI diagnosis were urinary tract infections (20.3%), congestive heart failure (19.1%), and pneumonia (10.9%).

**Effect of pneumonia and influenza peak timing**

The Poisson model of monthly CDI incidence included a piecewise linear spline with a change-point in January 2000, in addition to the sine and cosine oscillators (Table 2). Peak incidence occurred in week 12 (95% CI: 8.09, 17.02) and was 1.17 times higher than trough incidence (95% CI: 1.07, 1.27). Comparatively, the pneumonia and influenza wave peaked earlier and had a larger amplitude: Peak pneumonia and influenza prevalence was 1.63 times higher than trough prevalence (95% CI: 1.56, 1.71) and peaked in week 3 (phase shift = 3.42 weeks, 95% CI: 2.63, 4.21). Peak pneumonia and influenza prevalence preceded peak CDI incidence by 9.14 weeks (95% CI: 4.61, 13.67). The harmonic time-series model for CDI explained 76.3% of the variation in CDI.

We applied the same harmonic regression model of CDI to the 20 age-region strata. The pooled estimate of the amplitude was larger than the unstratified estimates (amplitude = 1.24, 95% CI: 1.16, 1.34). No heterogeneity in either amplitude ($I^2 = 0\%$, 95% CI: 0, 44.4; $P = 0.46$) or phase shift ($I^2 = 41\%$, 95% CI: 0, 48.0; $P = 0.46$) was detected. When pooled across age-region strata, the CDI peak was estimated to occur in late February (phase shift = 7.70, 95% CI: 5.47, 9.94). Phase shift of the pneumonia prevalence peak was a significant predictor of CDI phase shift (Figure 3; slope = 4.64, 95% CI: 1.31, 7.97), while region ($\chi^2_{df, P} = 0.85$) and age ($\chi^2_{df, P} = 0.83$) were not. Based on the adjusted differences in phase shift, the $C. difficile$ waves for the 20 age-region strata were estimated to peak 4.26 (95% CI: 1.99, 6.53) weeks after the pneumonia waves.

**Effect of pneumonia and influenza fluctuations**

Lag effects in this Poisson model were characterized by a quadratic polynomial cross-basis in order to measure the effect of pneumonia and influenza as a curve across increasing lags. Figure 4 presents the impact of pneumonia and influenza prevalence on current and future CDI incidence. A 1% absolute increase in pneumonia and influenza

---

**Table 1.** Incidence of CDI (per 1,000 Discharges) in Hospitalized Inpatients, United States, 1993–2008*

<table>
<thead>
<tr>
<th>Age, years</th>
<th>Incidence</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>40–49</td>
<td>4.53</td>
<td>1 Referent</td>
</tr>
<tr>
<td>50–59</td>
<td>5.70</td>
<td>1.26</td>
</tr>
<tr>
<td>60–69</td>
<td>7.52</td>
<td>1.66</td>
</tr>
<tr>
<td>70–79</td>
<td>9.46</td>
<td>2.09</td>
</tr>
<tr>
<td>≥ 80</td>
<td>12.10</td>
<td>2.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospital region</th>
<th>Incidence</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>11.50</td>
<td>1 Referent</td>
</tr>
<tr>
<td>Midwest</td>
<td>8.76</td>
<td>0.76</td>
</tr>
<tr>
<td>South</td>
<td>6.84</td>
<td>0.59</td>
</tr>
<tr>
<td>West</td>
<td>7.60</td>
<td>0.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period</th>
<th>Incidence</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993–1996</td>
<td>4.70</td>
<td>1 Referent</td>
</tr>
<tr>
<td>1997–2000</td>
<td>5.71</td>
<td>1.22</td>
</tr>
<tr>
<td>2001–2004</td>
<td>9.81</td>
<td>2.09</td>
</tr>
<tr>
<td>2005–2008</td>
<td>13.24</td>
<td>2.82</td>
</tr>
</tbody>
</table>

**Abbreviations:** CDI, *Clostridium difficile* infection; CI, confidence interval.

* For statistical methods relating to bivariate analyses, refer to the Appendix.
prevalence was associated with a 1.9% relative increase (RR = 1.019, 95% CI: 1.012, 1.026) in CDI incidence at lag-month 0. The effect in the months following decreased asymptotically; pneumonia and influenza prevalence had no significant impact on CDI incidence from lag-month 13 onward. A 1% increase in pneumonia was associated with a cumulative increase in CDI incidence of 11.3% (RR = 1.113, 95% CI: 1.073, 1.153) at 6 months and 19.1% (RR = 1.191, 95% CI: 1.084, 1.338) at 24 months. The model explained 78.7% of variation in CDI. The fitted values of the distributed lag prediction model are plotted alongside the overall time series in Figure 5.

When we applied the same distributed lag model to each of the 20 age-region–specific strata, we found an attenuated effect at lag month 0 (RR = 1.009, 95% CI: 1.004, 1.013) and a similar trend of decreasing effect size. The 6-month cumulative effect of a 1% increase in CDI was 4.8% (95% CI: 1.028, 1.067) and 6.2% (95% CI: 1.023, 1.104) at 24 months. In the meta-analysis of effect sizes across age-region strata for the 6-month cumulative effect, we found no evidence of heterogeneity ($I^2 = 0\%$, 95% CI: 0, 1.5; $P = 0.95$), although we noted nonsignificant differences between regions ($X_{MA}^2 = 4.07$, $P = 0.25$); lagged pneumonia and influenza effects were smallest in the Midwest (RR = 1.022,

### Table 2. Harmonic Regression Model Estimates for Monthly Pneumonia and Influenza Prevalence and CDI Incidence, United States, 1993–2008

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted</th>
<th>Adjusted*</th>
<th>Heterogeneity, $I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>95% CI</td>
<td>Estimate</td>
</tr>
<tr>
<td>Pneumonia and influenza prevalence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude</td>
<td>1.63</td>
<td>1.56, 1.71</td>
<td>1.59</td>
</tr>
<tr>
<td>Peak, week no.</td>
<td>3.42</td>
<td>2.63, 4.21</td>
<td>3.44</td>
</tr>
<tr>
<td>CDI incidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual change, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993–1999</td>
<td>6.8</td>
<td>4.2, 9.5</td>
<td>5.5</td>
</tr>
<tr>
<td>2000–2008</td>
<td>10.8</td>
<td>9.4, 12.1</td>
<td>11.0</td>
</tr>
<tr>
<td>Amplitude</td>
<td>1.17</td>
<td>1.07, 1.27</td>
<td>1.24</td>
</tr>
<tr>
<td>Peak, week no.</td>
<td>12.56</td>
<td>8.09, 17.02</td>
<td>7.70</td>
</tr>
<tr>
<td>Interval between pneumonia and influenza peak and CDI peak, weeks</td>
<td>9.14</td>
<td>4.61, 13.67</td>
<td>4.26</td>
</tr>
</tbody>
</table>

Abbreviations: CDI, *Clostridium difficile* infection; CI, confidence interval.

* Adjusted for census region and age.
95% CI: 0.994, 1.050) and largest in the South (RR = 1.065, 95% CI: 1.022, 1.110).

**DISCUSSION**

This descriptive study has demonstrated that 1) mean annual peak pneumonia and influenza prevalence precedes peak CDI incidence, 2) hospital pneumonia and influenza prevalence has a dynamic effect on CDI, wherein the effects of increases in pneumonia and influenza on CDI can be felt immediately, and for a period of 12 months, and 3) that these 2 inferences are also valid within age- and region-specific strata. A novel finding of this study is that substantial heterogeneity in the timing of seasonal CDI peak incidence exists for age-region strata and that it was large enough to mask CDI seasonality. A study of CDI seasonality in a smaller, universal, health-care system found larger variation in CDI incidence between peak and trough months (22); such health-care systems may be more susceptible to large-scale CDI epidemics (23) because of greater patient mixing between hospitals (24). Furthermore, our study suggests that these variations in peak CDI timing may be associated with the timing of age-region–specific pneumonia and influenza peaks. Several studies have found age-specific discharge patterns for pneumonia and influenza, with pneumonia and influenza discharge peaks for younger cohorts earlier than those of working-age cohorts (25). In the United States, pneumonia and influenza discharge peaks in the western states occur 3–4 weeks earlier than in northeastern states (26). In this study, similar patterns were found for CDI.

This study found that increases in hospital pneumonia and influenza prevalence are linearly associated with aggregate CDI incidence. This implies that factors impacting hospital pneumonia and influenza prevalence, including seasonal H3N2 predominance (8) and universal influenza vaccination (27), may impact downstream CDI incidence. Further, this study noted homogeneity of relative risks due to lagged pneumonia and influenza across age and region strata; however, since pneumonia and influenza seasonality is stronger in older age groups, this implies that the risk of infection with *C. difficile* attributable to pneumonia and influenza is substantially larger in older age groups.

In the only previous study of CDI seasonality (9), which also utilized US medicodemographic data, the authors equally found that CDI peaked in March with CDI tending to follow the number of influenza cases at a lag of 1 month but not at longer lags. In the present study, we identified a substantially longer effect of lagged pneumonia and influenza; these findings are likely due to methodological differences. This study focused on pneumonia and influenza peak timing relative to CDI peak timing using several different statistical methods, including harmonic regression (16), distributed lag models (28), and fixed effects meta-analysis; the effects we identified were robust in the face of varying methodological approaches. Previous studies considering the co-seasonality of infectious disease processes have used regression models with lagged covariates (29) or Box-Jenkins transfer function models with measurement of cross-correlation at specific lags (22); in these models, a single parameter or measurement must be made for each lag. Distributed lag models have several advantages over these other models, including reduced collinearity, parsimony, and ease of interpretation.

We hypothesize that the marked CDI seasonality that we observe is likely due to a combination of individual-level factors such as increased uptake of certain antibiotics during wintertime (30) and ecological factors such as increased hospital crowding, CDI pressure (31), herd effects of having many individuals taking more antibiotics, and interhospital transfer (24). Other system-level studies have shown that effects of pneumonia and influenza and respiratory syncytial virus on CDI are only partly mediated by fluoroquinolone and macrolide prescribing (22). This study provides additional evidence that effects of pneumonia and influenza on
CDI incidence exist at the ecological level, and that they are consistent across regions and age groups.

One must note that the mixed ecological design (32) used in this study, based on data from multiple age and region subgroups, does not represent individual-level associations. Studies using ecological designs are often criticized because they are not validated at the individual level (33). Indeed, our exposure of interest, health care system pneumonia and influenza prevalence, was ecologic and, as such, individual level inference would have been susceptible to cross-level inference bias (34, 35). Instead, we aimed to explore and infer what the aggregate effects that changes in hospital pneumonia and influenza prevalence had on regional CDI incidence. The immediate and persistent effects that increases in pneumonia and influenza prevalence had upon CDI incidence may indicate that both individual-level effects (i.e., effects on patients with pneumonia and influenza, which should be noted immediately) and herd-level effects (i.e., effects on patients with or without pneumonia and influenza, such as hospital crowding, which may have impacts farther into the future since they act on the hospital system) are at play. A carefully composed multilevel model, incorporating both individual- and group-level measures of pneumonia and influenza exposure, could distinguish the relative contributions of these effects.

Like any observational study, ours has several limitations. First, we did not control for the effects of antibiotic exposure in explaining the association between pneumonia and influenza seasonality and CDI incidence. Indeed, aggregate antibiotic prescribing patterns may partly explain the patterns described in this study. However, because pneumonia and influenza illness is an upstream determinant of antibiotic prescribing, and since pneumonia and influenza presence is associated with seasonal respiratory syncytial virus, pneumonia and influenza may be a better predictor for use in early warning systems; in effect, we see intrinsic value in describing the nature of the co-seasonality of pneumonia and influenza and CDI, since the exact causal pathway may not be explained by any single factor (22) and since nonlinear dynamic effects may make it difficult to make equivalent inference at the individual level (35). Second, our study considered the incidence of CDI in patients hospitalized for more than 2 days; despite the exclusion of patients hospitalized for ≤2 days, our incident cases of CDI were not necessarily hospital acquired. However, CDI incidence drawn from the National Hospital Discharge Survey has been shown to accurately measure hospital CDI burden (12), and the effects of pneumonia and influenza on CDI are likely equally large for community-associated disease as they are for nosocomial disease, due to seasonal patterns of outpatient antimicrobial prescribing (36).

In conclusion, this study of CDI over a 16-year period has found that, in the United States, seasonal upswings in hospital pneumonia and influenza prevalence precede similar changes in CDI incidence, and that the effects on CDI incidence are long lasting. Proper recognition of CDI peaks, the points in time at which these regularly occur, and the impacted groups within the population could contribute to early recognition of cases and help to prevent hospital-system outbreaks.

ACKNOWLEDGMENTS

Author affiliations: Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada (Kevin A. Brown, Paul Arora, Rahim Moeneddin, David N. Fisman); and Division of Infectious Diseases, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada (Nick Daneman).

This work was supported by the Frederick Banting and Charles Best Canada Graduate Scholarship to K.A.B. from the Canadian Institutes of Health Research (CIHR). N.D. is supported by a Clinician Scientist Salary Award from CIHR. D.N.F. received a grant from the Institute of Population and Public Health of the CIHR.

Conflict of interest: none declared.

REFERENCES


APPENDIX

Bivariate analyses

For the bivariate analyses (Table 1), the incidence of CDI was calculated for each risk group for each National Hospital Discharge Survey year, and the standard errors were calculated by using the annual, group-specific b parameters provided in the National Hospital Discharge Survey documentation (11); that is,

\[
\text{RSE}(p) = \sqrt{\frac{b \times q}{(p \times SS)}},
\]

where \( p \) is the incidence, \( q \) is 1 – \( p \), and SS is the weighted total of the number of discharges corresponding to the denominator of \( p \). When group-specific \( b \) parameters were not provided for a given risk factor (such as for hospital size groups), we fell back on the overall parameter for the given year. The weighted average of the incidence (weights corresponding to total discharges in each year) and its standard error were then calculated. The standard error of the relative risk was calculated according to the \( \delta \) method.

Design effect

The design effect is a measure of the increase in variance of a given parameter estimate due to sample design. It is equal to the ratio of the actual variance of an estimate using a given sample design divided by the variance from a simple random sample with the same sample size (15). Since the variance of the incidence (equal to RSE\(^2\) multiplied by \( p^2 \)) equals \( b \times p \times q/SS \), and the variance of a proportion given a simple random sample design is \( p \times q/SS \), where \( SS \) is the sample size (unweighted), the design effect is equal to \( b/(SS/ss) \) where \( SS/ss \) is the mean analysis weight.