PRACTICE OF EPIDEMIOLOGY

Identifying Pediatric Age Groups for Influenza Vaccination Using a Real-Time Regional Surveillance System

John S. Brownstein1,2,3, Ken P. Kleinman4, and Kenneth D. Mandl1,2,3,5

1 Division of Emergency Medicine, Children's Hospital Boston, Boston, MA.
2 Children's Hospital Informatics Program, Children's Hospital Boston, Boston, MA.
3 Department of Pediatrics, Harvard Medical School, Boston, MA.
4 Department of Ambulatory Care and Prevention, Harvard Medical School, Harvard Pilgrim Health Care, and Harvard Vanguard Medical Associates, Boston, MA.
5 Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA.

Received for publication March 10, 2005; accepted for publication June 10, 2005.

Evidence is accumulating that universal vaccination of schoolchildren would reduce the transmission of influenza. The authors sought to identify target age groups within the pediatric population that develop influenza the earliest and are most strongly linked with mortality in the population. Patient visits for respiratory illness were monitored, using real-time syndromic surveillance systems, in six Massachusetts health-care settings, including ambulatory care sites and emergency departments at tertiary-care and community hospitals. Visits from January 1, 2000, to September 30, 2004, were segmented into age group subpopulations. Timeliness and prediction of each subpopulation were measured against pneumonia and influenza mortality in New England with time-series analyses and regression models. Study results show that patient age significantly influences timeliness (p = 0.026), with pediatric age groups arriving first (p < 0.001); children aged 3–4 years are consistently the earliest (p = 0.0058). Age also influences the degree of prediction of mortality (p = 0.036), with illness among children under age 5 years, compared with all other patients, most strongly associated with mortality (p < 0.001). Study findings add to a growing body of support for a strategy to vaccinate children older than the currently targeted age of 6–23 months and specifically suggest that there may be value in vaccinating preschool-age children.

disease transmission; Fourier analysis; influenza; influenza vaccines; mass immunization; population surveillance; sentinel surveillance; vaccination

Abbreviations: ANOVA, analysis of variance; CI, confidence interval; SD, standard deviation.

Each year in the United States, an influenza epidemic causes hundreds of thousands of hospitalizations (1–4) and tens of thousands of deaths (5, 6), and it has an enormous economic impact (7, 8). Vaccination against influenza is the mainstay of prevention efforts and was initially targeted at older persons and those at high risk of complications (3). In 2003, the Advisory Committee on Immunization Practices recommended universal vaccination of infants and children aged 6–23 months (9, 10). This committee continues to recommend influenza vaccination of only those children aged ≥23 months who have high-risk medical conditions. Given the evidence that vaccination of schoolchildren significantly reduces influenza transmission (11–13), expanding the recommended target population to include healthy children has been suggested (14).

In this paper, we take a novel approach to identifying high-value populations for influenza vaccination. We leverage a real-time population health monitoring system that...
acquires and processes clinical data collected in the routine process of internal medicine, pediatric, and emergency care (15–19). Specifically, we identify, within a regional health-care-based population of patients with respiratory illness, the age cohorts that present the earliest and have patterns of illness most strongly associated with adverse outcomes from influenza.

**MATERIALS AND METHODS**

**Patient populations**

We performed time-series analysis of five health-care populations, identified retrospectively, with respiratory illness syndromes. Four populations consisted of patients presenting to emergency departments that share overlapping catchment areas in eastern Massachusetts but differ in the age distribution of the patients. The first is a pediatric emergency department at a large children’s hospital and has an average patient age of 6.8 (standard deviation (SD), 6.3) years. The second is an adult emergency department, with an average patient age of 52.2 (SD, 22.7) years. The populations from these two departments included encounters from January 1, 2000, to August 1, 2004. The third is a general emergency department that sees both children and adults, with an average patient age of 44.8 (SD, 27.1) years, and includes patients seen from October 1, 2002, to September 30, 2004. The fourth health-care population is a group of community emergency departments that comprises patients seen at three affiliated community-based emergency departments and includes both children and adults, with an average patient age of 37.8 (SD, 21.1) years, seen from July 1, 2001, to June 30, 2004. Emergency department presenting complaints were used to classify patients with respiratory illness, as described previously (20, 21).

We also studied daily counts of respiratory illness from ambulatory care encounters at a large group practice among insureds of a health maintenance organization in eastern Massachusetts. Approximately 175,000 members are included. These cases were identified from physician-assigned International Classification of Diseases, Ninth Revision, diagnosis codes assigned by the clinician at the time of consultation by using a modification of a provisional classification scheme produced by the Department of Defense ESSENSE project (22). The ambulatory care population included patients seen between January 1, 2000, and December 31, 2003. All of these data were obtained from two real-time population health monitoring systems, the Automated Epidemiologic Geotemporal Integrated Surveillance system (23) and the National Bioterrorism Syndromic Surveillance Demonstration Project (24).

The effect of patient age was evaluated by considering separately the following age groups: 0–2 (infant and toddler), 3–4 (preschool age), 5–11 (school age), 12–17 (adolescent), 18–39 (youngest adults), 40–64 (older adults), and over 64 (elderly adults) years.

For comparison with an extant surveillance system, we obtained data from the Centers for Disease Control and Prevention US Influenza Sentinel Providers Surveillance Network. For this system, influenza morbidity data are collected from sentinel health-care providers who report the number of patients they have seen with influenza-like illness symptoms. These symptoms include fever (temperature of >100°F (37.7°C)) plus either a cough or a sore throat. Weekly influenza-like illness counts from September 30, 2001, to October 2, 2004, were obtained for Massachusetts.

We evaluated earliness of presentation and association with adverse outcomes for the health-care populations and the sentinel surveillance by comparing them with pneumonia and influenza mortality data. Deaths due to pneumonia and influenza in New England for all ages combined were obtained from the Centers for Disease Control and Prevention 122 Cities Mortality Reporting System published weekly in table III of the Morbidity and Mortality Weekly Report (25).

**Analysis of timeliness**

The temporal associations between the health-care encounter and the mortality data sets were characterized by using time-series analysis. We first removed linear trends in the data and standardized the residuals from this analysis. Given that we already expected that each of the data sets would display strong yearly components, we used a finite Fourier transform to remove random noise from the yearly signals and produce a smoothed picture of seasonal change (26). We then performed cross-spectral analysis to find the estimated lead time (i.e., phase shift) between the underlying yearly components of each of the monitored patient populations and the pneumonia and influenza mortality time series. The lead time is the lag between two time series of interest. All analyses were carried out by using SAS software, version 9 for Windows (SAS Institute, Inc., Cary, North Carolina).

Cross-spectral analysis was initially applied to data streams: pediatric emergency department, adult emergency department, general emergency department, community emergency departments, ambulatory care, influenza-like illness, and pneumonia and influenza. The sine and cosine coefficients were obtained for the yearly frequency of approximately 52 weeks. The lead time was calculated from all monitored patient populations and influenza-like illness to pneumonia and influenza mortality. The yearly signals of respiratory illness from the monitored patient population partitioned into age subgroups were also obtained and compared with those of overall population mortality due to pneumonia and influenza mortality.

Differences in estimated mean phase shift by age group and site of care were evaluated by analysis of variance (ANOVA). We used randomized complete block design ANOVA, where the blocks are the sites of care and the treatments are the age groups, with the estimated phase shift from each site, by age group, as the outcome (27). The adult emergency department and pediatric emergency department data blocks were treated as one block to account for missing age groups at each of these sites. We tested the hypothesis...
that specific pediatric age groups as well as aggregated pediatric age ranges (0–4 and 0–18 years) were timelier than other age groups.

Analysis of predictive value for mortality

We assessed the relative predictive value of the time-lagged health-care population data streams by fitting generalized linear models to pneumonia and influenza mortality counts (28). A Poisson distribution for pneumonia and influenza was assumed because it is usually appropriate for modeling counts. We ran separate models for each population, where the predictor was the respiratory counts from the prior week suggested by the cross-spectral analysis. The same method was used for each age group of each population and influenza-like illness data. Overall model fit for each of the Poisson regression models was calculated by comparing deviance statistics with their asymptotic chi-square values (29). The value of each population’s respiratory counts in predicting mortality was determined by calculating the proportion of the deviance explained, similar to the $R^2$.

Differences in mean predictive value by age group and site of care were evaluated by ANOVA. We once again used randomized complete block design ANOVA, where the blocks are the sites of care and the treatments are the age groups, with the percent deviance explained at each site by age group as the outcome. We tested the hypothesis that specific pediatric age groups as well as aggregated pediatric age ranges (0–4 and 0–18 years) were more predictive than other age groups.

RESULTS

Timeliness by patient population

The populations are described in table 1. The visit rates for the emergency departments were approximately the same, ranging from 99.6 (SD, 24.0) to 174.7 (SD, 43.0) visits per week. Ambulatory care had approximately 10 times more volume, with 1,585.9 (SD, 556.3) visits per week. The sentinel influenza-like illness surveillance system reported cases at an average rate of 71.1 (SD, 47.9) per week. Each population displayed a highly seasonal cycle in which peaks of illness occurred from the beginning of December to the end of February. Figure 1 shows the results of the cross-spectral analysis. The pneumonia and influenza mortality peak was last, at the end of February.

Timeliness was calculated as the lead time from each of the respiratory illness data sets to pneumonia and influenza mortality. The ambulatory care population, with both pediatric and adult patients, had a mean lead time of about 4 weeks (29 days), peaking in mid- to late January. The pediatric emergency department population displayed the earliest peak of respiratory illness, occurring on average 5 weeks (38 days) prior to the peak in mortality, during the first week of January. Sentinel influenza-like illness data showed that it peaked, on average, 20 days prior to influenza mortality, well after both the pediatric emergency department and ambulatory care populations. The adult emergency department, general emergency department, and community populations were the least timely regarding warning about influenza mortality, with a mean lead time of about 2 weeks (12, 10, and 14 days, respectively), peaking during the first week of February.

Timeliness by age

The lead times varied by age (table 2). Separate cross-spectral analysis of the age groups revealed that, among patients presenting to the different health-care settings, children provided the earliest signal of pneumonia and influenza mortality irrespective of site of care. Those aged 3–4 years were seen first, with a mean lead time of 34 days (95 percent confidence interval (CI): 14.5, 53.5). This age group presented to the pediatric emergency department with the longest lead time (50 days). All adult age groups (18–39, 40–64, and >65 years) in the ambulatory care and emergency department settings sought care later than pediatric patients, with mean lead times of 12.0 (95 percent CI: −4.28), 10.5 (95 percent CI: −19.8, 40.8), and 14.5 (95 percent CI: −6.8, 35.8) days, respectively.

Randomized complete block design ANOVA confirmed a significant effect of both age group ($F = 3.19$, df = 6, $p = 0.021$) and site of care ($F = 4.14$, df = 3, $p = 0.026$) on timeliness. Post hoc mean contrast revealed that mean lead time for children aged 3–4 years was significantly greater than that for older age groups ($p = 0.00142$). Overall, pediatric patients (aged ≤18 years) were timelier than adults ($p < 0.001$), and the youngest children, under age 4 years, arrived before all other groups ($p = 0.0058$).

Mortality prediction by patient population

When the lead times defined by the cross-spectral analysis were used, each health-care population was found to be a statistically significant predictor of mortality ($p < 0.0001$). A comparison of the predictive abilities of these populations.
showed that the ambulatory care, general emergency department, and pediatric emergency department (30–31 percent) explained more of the variation than the adult and community emergency departments (24–25 percent) and the sentinel influenza-like illness data (25 percent).

**Mortality prediction by age**

Prediction of influenza mortality varied by age (table 3). Among age groups presenting to the different health-care settings, children’s pattern of illness was most predictive of pneumonia and influenza mortality across sites of care. Children less than age 3 years provided the best prediction of mortality, explaining on average 40.8 percent (standard error, 4.4) of the deviance. This group was followed by children aged 3–4 years, who explained 36.8 percent (standard error, 6.1) of the deviance. Figure 2 plots mortality prediction versus timeliness for age group of each health-care population and reveals that, in general, pediatric age groups had the best combination of the two indicators.

**FIGURE 1.** Predicted yearly seasonality of respiratory illness for five retrospective patient populations with respiratory illness syndromes presenting to health-care sites and the Centers for Disease Control and Prevention sentinel influenza-like illness surveillance system, eastern Massachusetts, 2000–2004. Yearly cycles were obtained by cross-spectral analysis performed on each data stream after linear detrending and standardization (standardized regression residuals). The phase shift of each data stream with the Centers for Disease Control and Prevention influenza mortality surveillance represents the timeliness of the data stream.

**TABLE 2.** Timeliness of patient populations with respiratory illness regarding signaling of pneumonia and influenza mortality, by site of care and patient age, eastern Massachusetts, 2000–2004

<table>
<thead>
<tr>
<th>Data source</th>
<th>&lt;3 years</th>
<th>3–4 years</th>
<th>5–10 years</th>
<th>11–17 years</th>
<th>18–39 years</th>
<th>40–64 years</th>
<th>&gt;64 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambulatory care</td>
<td>35</td>
<td>37</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>Pediatric emergency department</td>
<td>35</td>
<td>50</td>
<td>40</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult emergency department</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Community emergency department</td>
<td>21</td>
<td>26</td>
<td>12</td>
<td>25</td>
<td>3</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>General emergency department</td>
<td>19</td>
<td>23</td>
<td>15</td>
<td>32</td>
<td>12</td>
<td>–13</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>27.5</td>
<td>34.0</td>
<td>23.0</td>
<td>27.5</td>
<td>12.0</td>
<td>10.5</td>
<td>14.5</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>13.6, 41.3</td>
<td>14.5, 53.5</td>
<td>2.9, 43</td>
<td>22.2, 32.8</td>
<td>–4.0, 28.0</td>
<td>–19.8, 40.8</td>
<td>–6.8, 35.8</td>
</tr>
</tbody>
</table>

* Lead time to pneumonia and influenza mortality calculated by cross-spectral analysis.
Randomized complete block design ANOVA showed a significant effect of both age group ($F = 2.92, \text{df} = 6, p = 0.036$) and site of care ($F = 74.79, \text{df} = 3, p < 0.0001$) on predictive ability. For children less than age 3 years, predictive ability was significantly greater than for all other age groups ($p = 0.0019$). When grouped together, data for pediatric patients, aged 0–18 years, did not explain significantly more of the deviance than those for adults ($p = 0.0906$). However, the youngest children, those under age 5 years, clearly provided the best prediction of all age groups ($p = 0.0012$). When plotted against timeliness, data for these two youngest age groups provided the maximum prediction while also supplying the earliest signal (figure 3).

**DISCUSSION**

Patient age is a key determinant of the timing of visits for respiratory illness; pediatric patients, specifically preschool-age children aged 3–4 years, seek ambulatory and emergency care earliest. Furthermore, respiratory illness in children less than 5 years of age is significantly associated

<table>
<thead>
<tr>
<th>Data source</th>
<th>Mortality prediction (% deviance explained)*</th>
<th>&lt;3 years</th>
<th>3–4 years</th>
<th>5–10 years</th>
<th>11–17 years</th>
<th>18–39 years</th>
<th>40–64 years</th>
<th>&gt;64 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambulatory care</td>
<td></td>
<td>28.3</td>
<td>31.1</td>
<td>28.5</td>
<td>27.8</td>
<td>26.1</td>
<td>32.0</td>
<td>27.6</td>
</tr>
<tr>
<td>Pediatric emergency department</td>
<td></td>
<td>34.4</td>
<td>23.1</td>
<td>17.7</td>
<td>14.7</td>
<td>19.8</td>
<td>17.3</td>
<td>17.2</td>
</tr>
<tr>
<td>Adult emergency department</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community emergency department</td>
<td></td>
<td>42.2</td>
<td>41.9</td>
<td>40.0</td>
<td>41.6</td>
<td>40.3</td>
<td>41.5</td>
<td>39.9</td>
</tr>
<tr>
<td>General emergency department</td>
<td></td>
<td>58.3</td>
<td>50.9</td>
<td>44.2</td>
<td>45.4</td>
<td>49.5</td>
<td>41.2</td>
<td>44.6</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>40.8</td>
<td>36.8</td>
<td>32.6</td>
<td>32.4</td>
<td>33.9</td>
<td>33.0</td>
<td>32.3</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td></td>
<td>20.1</td>
<td>61.5</td>
<td>17.4, 56.1</td>
<td>13.6, 51.6</td>
<td>10.1, 54.7</td>
<td>12.5, 55.4</td>
<td>15.9, 51.1</td>
</tr>
</tbody>
</table>

* Calculated by Poisson regression analyses.
with mortality from pneumonia and influenza, with a lead time of 4–5 weeks. Pediatric populations are sentinels of infection, and they signal the consequent burden of illness. Although this finding does not necessarily prove that preschool-age children are driving the yearly influenza epidemics, they intriguingly suggest that preschool-age children are the initial group infected and may be important in the subsequent spread.

There is ample prior evidence that children play a primary role in influenza transmission. Given their increased tendency to acquire and shed influenza, children have been identified as predominant vectors in the household spread of influenza (30–32). Our findings support the notion that specifically targeting preschool children may reduce transmission. Children under 5 years of age have higher infection rates than do older children (33–35). In addition, vaccination of this age group has been shown to significantly reduce morbidity among their household contacts (36). For this reason, concentrating immunization efforts on preschool children may eliminate the primary pathway of infection.

Other studies have shown that older children (aged 5–18 years) are the most important targets and that their routine vaccination would reduce disease burden at the community level (11–13, 37–40). Our results suggest that younger children may initiate spread to these older children and therefore may be of value as targets of vaccination out of proportion to their lesser numbers.

While our study suggests that young children are infected first, there are other possible explanations for their early presentation to the health-care system. It may be not just the inherent vulnerability of children but also health-care-seeking behaviors that make them timely sentinels of influenza (41). Family members may have a lower threshold for bringing in febrile young children because of morbidity concerns specific to the pediatric population, and these children will thus have been seen by physicians at the earlier stages of viral illnesses (42, 43). However, we found that the pediatric emergency department populations arrive prior to the pediatric ambulatory populations. Because the emergency department populations are, naturally, more acutely ill (24), the reason for the early presentation of children is likely at least partly rooted in genuine morbidity, not just parental behavior. In addition, if the early arrival of children could be explained primarily by the behavior of worried parents and pediatricians, we would instead expect to see the youngest, most fragile children—infants—arriving before the preschoolers; in preschoolers, simple febrile illnesses simply do not pose the same risks or require as much testing (44).

A limitation of our study is that we were measuring respiratory illness but not virologically confirmed influenza infection. Our findings are confounded by co-circulation with other viruses, including respiratory syncytial virus and parainfluenza virus, for which there are no vaccinations

**FIGURE 3.** For patients with respiratory illness syndromes, prediction and timeliness regarding signaling of pneumonia and influenza mortality, summarized by age group in years (values given with black circles), eastern Massachusetts, 2000–2004. Timeliness is the average lead time to influenza mortality for each age group across all sites of care obtained by cross-spectral analysis. Mortality prediction is the average proportion of variance explained by each age group across health-care sites calculated by Poisson regression. As shown, the two youngest age groups had the best combination of timeliness and prediction.
currently available. Another limitation is that our data are from the Greater Boston Area and may not be entirely generalizable to other regions. However, the patients are seen at seven diverse institutions and are likely to be highly representative of the region; also, a priori, it is not clear why there would be regional differences.

This study has other implications as well. Since the data are available in a real-time population health monitoring system, understanding the temporal dynamics of respiratory illness through different age groups can be used to inform medical practice and enable improved prevention and control efforts by individual clinicians. Monitoring respiratory illness in the ambulatory care and pediatric emergency department populations by using syndromic surveillance systems was shown to provide even earlier detection and better prediction of influenza activity than the current Centers for Disease Control and Prevention sentinel surveillance system. Supplying physicians with a mechanism to identify the earliest and most sensitive warning of respiratory mortality can help them implement prevention strategies that will protect their general patient population.

In this paper, we clearly demonstrated that, across a region, preschool-age children are the first to seek health care for respiratory infections and, furthermore, that there is a strong association between their temporal patterns of illness and subsequent mortality in the general population from influenza. While our findings do not definitively indict preschool-age children as those initially infected and primarily responsible for spread to other age groups, this age group does appear to have an important role in influenza transmission. These results bolster arguments for a recommendation currently under consideration by the Advisory Committee on Immunization Practices to begin to universally vaccinate preschool-age children.

ACKNOWLEDGMENTS

This work was supported by grant R01 LM007677-01 from the National Library of Medicine (National Institutes of Health), contract 290-00-0020 from the Agency for Healthcare Research and Quality, and by contract 52253337HAR from the Massachusetts Department of Public Health.

The authors gratefully acknowledge the input of Drs. Tracy Lieu, Ben Reis, Cecil Wolfe, and Karen Olson. They thank Andrew Ellingson and Drs. John Halamka and Tom Stair for their participation in the Automated Epidemiologic Geotemporal Integrated Surveillance system.

Conflict of interest: none declared.

REFERENCES

22. Lazarus R, Kleinman KP, Dashevsky I, et al. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of


