Original Investigation

Identifying Pediatric Community-Acquired Pneumonia Hospitalizations
Accuracy of Administrative Billing Codes

Derek J. Williams, MD, MPH; Samir S. Shah, MD, MSCE; Angela Myers, MD, MPH; Matthew Hall, PhD; Katherine Auger, MD; Mary Ann Queen, MD; Karen E. Jerardi, MD, MEd; Lauren McClain, MD; Catherine Wiggleton, MD, ScM; Joel S. Tieder, MD, MPH

IMPORTANCE  Community-acquired pneumonia (CAP) remains one of the most common indications for pediatric hospitalization in the United States, and it is frequently the focus of research and quality studies. Use of administrative data is increasingly common for these purposes, although proper validation is required to ensure valid study conclusions.

OBJECTIVE  To validate administrative billing data for hospitalizations owing to childhood CAP.

DESIGN AND SETTING  Case-control study of 4 tertiary care, freestanding children’s hospitals in the United States.

PARTICIPANTS  A total of 998 medical records of a 25% random sample of 3646 children discharged in 2010 with at least 1 International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) code representing possible pneumonia were reviewed. Discharges (matched on date of admission) without a pneumonia-related discharge code were also examined to identify potential missed pneumonia cases. Two reference standards, based on provider diagnosis alone (provider confirmed) or in combination with consistent clinical and radiographic evidence of pneumonia (definite), were used to identify CAP.

EXPOSURE  Twelve ICD-9-CM-based coding strategies, each using a combination of primary or secondary codes representing pneumonia or pneumonia-related complications. Six algorithms excluded children with complex chronic conditions.

MAIN OUTCOMES AND MEASURES  Sensitivity, specificity, and negative and positive predictive values (NPV and PPV, respectively) of the 12 identification strategies.

RESULTS  For provider-confirmed CAP (n = 680), sensitivity ranged from 60.7% to 99.7%; specificity, 75.7% to 96.4%; PPV, 67.9% to 89.6%; and NPV, 82.6% to 99.8%. For definite CAP (n = 547), sensitivity ranged from 65.6% to 99.6%; specificity, 68.7% to 93.0%; PPV, 54.6% to 77.9%; and NPV, 87.8% to 99.8%. Unrestricted use of the pneumonia-related codes was inaccurate, although several strategies improved specificity to more than 90% with a variable effect on sensitivity. Excluding children with complex chronic conditions demonstrated the most favorable performance characteristics. Performance of the algorithms was similar across institutions.

CONCLUSIONS AND RELEVANCE  Administrative data are valuable for studying pediatric CAP hospitalizations. The strategies presented here will aid in the accurate identification of relevant and comparable patient populations for research and performance improvement studies.

Published online July 29, 2013.
Community-acquired pneumonia (CAP) is one of the most common causes of childhood mortality worldwide, especially among those younger than 5 years. Although mortality is lower in developed countries, pediatric CAP is still associated with substantial morbidity and remains the most common indication for pediatric hospitalization outside the newborn period in the United States. It is also one of the most expensive, with hospital costs alone approaching $1 billion annually. As a result, pediatric CAP is often the focus of epidemiologic studies and outcomes research and serves as an ideal target for quality benchmarking and improvement efforts.

Administrative data are often used for such assessments because they are convenient and provide ready access to large study populations without the substantially higher costs of prospective studies. However, studies using administrative data often rely on discharge diagnosis codes, such as the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), to identify study populations, and the assignment of these codes may be subject to error or influenced by reimbursement patterns and local coding practices. In the case of pneumonia, many ICD-9-CM discharge codes represent its many microbiologic causes (eg, influenza vs Streptococcus pneumoniae) and potential complications (eg, empyema). Furthermore, a discharge code indicating pneumonia does not readily differentiate CAP from pneumonia that occurs from a hospital-acquired infection, such as ventilator-associated pneumonia. As a result, use of administrative data to identify CAP hospitalizations without appropriate validation is not advisable because of the heterogeneous nature of the disease and the potential for misclassification and erroneous conclusions.

Studies in adults have assessed the validity of ICD-9-CM codes to identify hospitalizations owing to pneumonia, although with varying results depending on the population studied and the codes selected for identification (positive predictive value [PPV], 57%-96%; sensitivity, 48%-98%). The findings from these studies may be limited in their application since they relied on small sample sizes. Two reference standards were used to represent CAP, one highly sensitive (provider diagnosis) and the other highly specific (provider diagnosis plus consistent clinical and radiographic evidence). We assessed each medical record for the presence of the following criteria: (1) provider diagnosis of pneumonia within the first 48 hours of hospitalization (mention of suspected CAP along with consistent management strategy), (2) abnormal temperature (≥38.0°C) or white blood cell count (<5 or >15 × 10^9/L, multiply by 0.001), (3) evidence of a lower respiratory tract illness (eg, cough or increased work of breathing), and (4) chest radiograph indicating pneumonia (eg, infiltrate or consolidation). Any child with a condition precluding CAP (eg, cystic fibrosis or immunodeficiency) was not classified as having CAP regardless of the above-mentioned criteria.

For each reference standard, we assessed the performance of 12 ICD-9-CM coding algorithms to identify CAP hospitalizations among our study population. These strategies incorpo-
rate a number of previously described CAP algorithms\textsuperscript{9,10,13,18,19} as well as the classification of complex chronic conditions (CCCs) described by Feudtner et al.\textsuperscript{20} Strategies that included codes representative of pneumonia symptoms (eg, cough) were also considered\textsuperscript{10}; however, these algorithms did not identify additional CAP cases and were excluded from further analyses.

**Statistical Analysis**

Characteristics of the study population were summarized using proportions for categorical variables and median and interquartile range for continuous variables. Characteristics of children classified as having definite or provider-confirmed CAP were compared using $\chi^2$ tests for categorical variables and Wilcoxon rank sum for continuous variables.

To determine the performance of each identification strategy for CAP, we calculated the sensitivity, specificity, PPV, and negative predictive value (NPV) for each reference standard.\textsuperscript{21} We randomly sampled from all discharges with a pneumonia-related discharge code; thus, the PPV estimates are assumed to reflect the entire population under study, including disease prevalence within the population. To maximize the possibility of identifying false negatives (algorithm negative and reference standard positive), we selected a random sample of hospitalizations without a pneumonia-related discharge code matched on admission date to those hospitalizations with a pneumonia-related discharge code.

To describe interhospital variation in coding practices for CAP, we characterized the performance of each algorithm for provider-confirmed CAP for each hospital. To assess differences in patient characteristics and outcomes across identification strategies for children with provider-confirmed CAP, we summarized several patient-level variables, including concurrent lower respiratory tract diagnoses, care use measures, and severe outcomes (intensive care admission, mechanical ventilation, or empyema).

**Results**

**Study Population**

We reviewed 998 medical records of a possible 3646 discharges with a pneumonia-related \textit{ICD-9-CM} discharge code and 1000 discharges without a pneumonia-related discharge code (Figure 1). Among discharges with a pneumonia-related code, 677 (67.8\%) were classified as provider-confirmed CAP; 545 (80.5\%) were further classified as definite CAP. Children with definite CAP were less likely to have a concurrent discharge diagnosis of asthma but more likely to have a parapneumonic effusion or empyema compared with those without definite CAP (Table 1). Although hospital length of stay was slightly longer, children with definite CAP were less likely to have a complex chronic condition or require admission to in-
Algorithm Performance for Provider-Confirmed CAP

With the reference standard of provider-confirmed CAP (n = 680), sensitivity ranged from 60.7% to 99.7% (NPV, 82.6%-99.8%). Specificity ranged from 75.7% to 96.4% (PPV, 67.9%-89.6%). A pneumonia diagnosis code in any position (algorithm 1) identified nearly all children with provider-confirmed CAP (sensitivity, 99.7%), although this strategy had the lowest specificity (75.7%) (eTable in the Supplement and Figure 2A). An identification strategy that only included children with a primary diagnosis of pneumonia (algorithm 2) improved specificity to 90.9% but reduced sensitivity to 71.0%. Results were similar for the additional inclusion of children with a primary diagnosis of pneumonia and/or a primary diagnosis of a pneumonia-related complication plus a secondary diagnosis of pneumonia (algorithm 3). We also assessed the performance of algorithms 1 through 3 after excluding children with at least 1 CCC discharge code.20 These algorithms (1b, 2b, and 3b) improved specificity (89.7%-95.7%) compared with the analogous algorithms that included CCCs, although at the expense of decreased sensitivity.

Since diagnosis codes indicating pleural effusion or empyema are not exclusive to CAP, we assessed the performance of identification strategies that included effusion or empyema codes only when coupled with more explicit pneumonia codes (eTable in the Supplement and Figure 2A). Algorithms 4 through 6 demonstrated improved specificity with a minimal effect on sensitivity compared with each of the analogous algorithms 1 through 3. The most substantial improvement in specificity was noted for algorithm 4 (84.2%; 95% CI, 82.2-86.1) compared with algorithm 1 (75.7%; 95% CI, 73.3-78.0). Excluding discharges that had 1 or more CCC codes from algorithms 4 through 6 (4b, 5b, and 6b) further improved specificity (91.5%-96.1%).

Algorithm Performance for Definite CAP

With the reference standard of definite CAP (n = 547), sensitivity ranged from 65.6% to 99.6% (NPV, 87.8%-99.8%). Specificity ranged from 68.7% to 93.0% (PPV, 54.6%-77.9%) (eTable in the Supplement and Figure 2B). The performance characteristics of the algorithms relative to one another remained similar.

Hospital and Patient-Level Characteristics

We also explored the performance of identification strategies by hospital (eFigure 1A and B in the Supplement). We noted
few differences when comparing the individual algorithms across hospitals. Small differences in sensitivity for several algorithms occurred at 1 institution, but no other differences were noted.

Finally, we described patient characteristics, including other lower respiratory tract diagnoses, complications, and outcomes for CAP, for each algorithm to assess potential differences (Table 2). Several differences were noted according to the algorithm selected, including the proportion of children with a concurrent diagnosis of asthma or bronchiolitis and those experiencing a severe outcome (intensive care admission, mechanical ventilation, or empyema).

**Discussion**

This study determined the performance of 12 **ICD-9-CM** coding algorithms in identifying pediatric CAP hospitalizations at 4 tertiary care, freestanding children’s hospitals in the United States. Unrestricted application of the pneumonia-related codes does not accurately identify CAP in children, although several coding strategies offer improvements in specificity while also retaining good sensitivity. Administrative data can be used to study pediatric CAP hospitalizations, although the study population must be carefully defined and the specificity of the diagnosis understood to ensure valid study conclusions.

Use of the pneumonia and effusion or empyema codes in any position (algorithm 1) was not accurate, although this strategy identified nearly all children with CAP. This is an important attribute for studies aimed at identifying all episodes of CAP, such as assessments of disease frequency or monitoring for adverse drug effects. To avoid misclassification, application of this algorithm requires a multistep identification strategy, such as **ICD-9-CM**-based screening with medical record review to confirm the diagnosis.

In contrast to algorithms 1 through 3, which allowed the effusion or empyema codes to represent pneumonia without the requirement of an additional pneumonia code, algorithms 4 through 6 required a more explicit pneumonia code, with only algorithm 6 including the effusion or empyema codes (as a pneumonia-related complication). This resulted in an 11% increase in specificity for pneumonia codes in any position (algorithm 4) and identified nearly all CAP cases. Predictive performance was further improved when restricting the population to those with a primary diagnosis of pneumonia (algorithm 5) or a primary diagnosis of a pneumonia-related complication plus a secondary diagnosis of pneumonia (algorithm 6), similar to validation studies in other populations. However, these algorithms identified only approximately 70% of provider-confirmed CAP, and specificity was only marginally improved over the analogous algorithms 2 and 3. Nonetheless, these identification strategies may be a better choice for assessments that seek to maximize specificity over sensitivity and are unable to include confirmation using medical records. For example, these algorithms could be used to improve the validity of outcomes data in a study comparing the effectiveness of antimicrobial treatment choices for CAP.

The algorithms with the best performance characteristics (ie, both highly sensitive and specific) were 1b and 4b. These identification strategies, which do not restrict use of the pneumonia-related codes in any position (algorithm 1), are relatively straightforward and require less information. This strategy identified nearly all children with CAP, which is an important attribute for studies aimed at identifying all episodes of CAP.
monia codes, excluded children with CCCs identified with a previously reported classification scheme.\textsuperscript{20} The sensitivity and specificity of both algorithms were approximately 90% for provider-confirmed CAP. Overall, application of the CCC restriction yielded meaningful improvements in specificity for all algorithms studied without large reductions in sensitivity. Likely, restriction of these codes excludes children with a high probability of complicating factors, which may alter the true risk of CAP (eg, frequent hospitalizations or technology dependence) while minimally affecting children without these risk factors. As a result, these strategies are ideal for identifying and studying pediatric CAP hospitalizations, especially if the presence of complex comorbid conditions is thought to confound hypothesized relationships.

Requiring the CAP reference standard definition to include a provider diagnosis of pneumonia coupled with objective clinical evidence of a lower respiratory tract infection with radiographic confirmation (ie, definite CAP) resulted in expected reductions in specificity for all algorithms studied. Several differences were noted between patients classified as having definite CAP and those having provider-confirmed CAP only, underscoring the clinical challenge of diagnosing CAP in children. For instance, nearly half of those with provider-confirmed CAP who were not also classified as having definite CAP had a concurrent discharge diagnosis of asthma. Although asthma is a risk factor for CAP, in the clinical setting, it is often difficult to distinguish a superimposed pneumonia from atelectasis in an acutely wheezing asthmatic child when a chest radiograph reveals an opacity.\textsuperscript{24,25} In contrast, the presence of a parapneumonic effusion in a child with presumed pneumonia offers additional certainty of the CAP diagnosis. Indeed, the presence of effusion or empyema was more common among those diagnosed with definite CAP. Thus, the clinical uncertainty inherent in diagnosing CAP is reflected in the differences in predictive performance for the 2 reference standards presented here. This provides a more objective assessment of the degree of uncertainty likely to exist in a population of children with CAP identified using administrative data. To this end, it is important to note that 80.4% of the children with provider-confirmed CAP were also classified as having definite CAP.

Coding variation across institutions may limit the validity of administrative data.\textsuperscript{26} In our study, specificity across all 4 institutions and sensitivity at 3 of the 4 institutions studied were similar. Sensitivity was lower for several algorithms at 1 institution, although differences were small and likely reflect local variation in coding practices. Nonetheless, our findings demonstrate that coding practices for pediatric CAP are generally consistent across hospitals, suggesting that these identification strategies can be used to accurately benchmark performance and compare patients across institutions.\textsuperscript{27,28}

We also observed that identification strategies that considered either a primary or a secondary diagnosis of pneumonia (algorithms 1, 1b, 4, and 4b) identified a higher proportion of children with a concurrent diagnosis of asthma or bronchiolitis. In contrast, algorithms restricted to a primary diagnosis of pneumonia (algorithms 5 and 5b) included fewer children with these diagnoses but also identified the lowest proportion of children requiring intensive care unit admission or mechanical ventilation. This is an important consideration when designing studies to examine severe pneumonia outcomes.\textsuperscript{10,29} Moreover, strategies that considered a secondary diagnosis of pneumonia only if it was coupled with a primary diagnosis of a pneumonia-related complication (algorithms 6 and 6b) identified a high proportion of children with severe outcomes while minimizing those with concurrent diagnoses of asthma or bronchiolitis. This suggests that these latter algorithms may be particularly relevant for identifying severe CAP while also enriching the study population for bacterial pneumonia.

Strengths of our study include the multicenter design and large sample of children, the use of standardized data collection methods and CAP reference standards, a wide range of CAP identification algorithms, and the exclusive focus on pediat-

### Table 2. Outcomes and Concurrent Diagnoses for Provider-Confirmed Community-Acquired Pneumonia According to Identification Algorithm

<table>
<thead>
<tr>
<th>Algorithm*</th>
<th>Asthma</th>
<th>Bronchiolitis</th>
<th>Intensive Care Admission</th>
<th>Mechanical Ventilation</th>
<th>Empyema</th>
<th>Hospital Length of Stay, Median (IQR), d</th>
<th>Hospital Charges, Median (IQR), $</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>212 (31.1)</td>
<td>53 (8.0)</td>
<td>71 (10.7)</td>
<td>26 (4.2)</td>
<td>102 (15.3)</td>
<td>2 (1-4)</td>
<td>11 662 (8000-19 858)</td>
</tr>
<tr>
<td>1b</td>
<td>200 (33.1)</td>
<td>50 (8.3)</td>
<td>57 (9.4)</td>
<td>21 (3.5)</td>
<td>93 (15.4)</td>
<td>2 (1-3)</td>
<td>11 373 (7900-18 990)</td>
</tr>
<tr>
<td>2</td>
<td>120 (25.1)</td>
<td>12 (2.5)</td>
<td>26 (4.5)</td>
<td>9 (1.9)</td>
<td>88 (18.4)</td>
<td>2 (1-3)</td>
<td>10 608 (7648-17 169)</td>
</tr>
<tr>
<td>2b</td>
<td>111 (25.8)</td>
<td>11 (2.6)</td>
<td>19 (4.4)</td>
<td>6 (1.4)</td>
<td>81 (18.8)</td>
<td>2 (1-3)</td>
<td>10 246 (7583-15 885)</td>
</tr>
<tr>
<td>3</td>
<td>122 (24.9)</td>
<td>13 (2.7)</td>
<td>35 (7.2)</td>
<td>18 (3.7)</td>
<td>91 (18.6)</td>
<td>2 (1-3)</td>
<td>10 784 (7679-17 782)</td>
</tr>
<tr>
<td>3b</td>
<td>113 (25.9)</td>
<td>12 (2.7)</td>
<td>26 (5.9)</td>
<td>13 (3.0)</td>
<td>83 (19.0)</td>
<td>2 (1-3)</td>
<td>10 324 (7612-16 346)</td>
</tr>
<tr>
<td>4</td>
<td>212 (32.3)</td>
<td>53 (8.1)</td>
<td>70 (10.7)</td>
<td>28 (4.3)</td>
<td>93 (14.2)</td>
<td>2 (1-4)</td>
<td>11 523 (7931-19 551)</td>
</tr>
<tr>
<td>4b</td>
<td>200 (33.6)</td>
<td>50 (8.4)</td>
<td>56 (9.4)</td>
<td>21 (3.5)</td>
<td>85 (14.3)</td>
<td>2 (1-3)</td>
<td>11 217 (7870-18 413)</td>
</tr>
<tr>
<td>5</td>
<td>118 (25.9)</td>
<td>12 (2.6)</td>
<td>23 (5.0)</td>
<td>8 (1.8)</td>
<td>65 (14.3)</td>
<td>2 (1-3)</td>
<td>10 247 (7538-15 713)</td>
</tr>
<tr>
<td>5b</td>
<td>109 (26.7)</td>
<td>11 (2.7)</td>
<td>16 (3.9)</td>
<td>5 (1.2)</td>
<td>60 (14.7)</td>
<td>2 (1-3)</td>
<td>10 073 (7522-15 081)</td>
</tr>
<tr>
<td>6</td>
<td>122 (25.3)</td>
<td>13 (2.7)</td>
<td>35 (7.3)</td>
<td>18 (3.7)</td>
<td>84 (17.4)</td>
<td>2 (1-3)</td>
<td>10 611 (7653-17 338)</td>
</tr>
<tr>
<td>6b</td>
<td>113 (26.2)</td>
<td>12 (2.8)</td>
<td>26 (6.0)</td>
<td>13 (3.0)</td>
<td>77 (17.9)</td>
<td>2 (1-3)</td>
<td>10 276 (7583-16 100)</td>
</tr>
</tbody>
</table>

Abbreviation: IQR, interquartile range.

* For algorithm definitions, see the caption to Figure 2.
ric patients hospitalized with CAP, a population that, to our knowledge, has not been validated using administrative data. There are also several limitations. Discharges with and without a pneumonia-related discharge code were reviewed separately, although reviewers were masked to individual billing codes. A comparison of rater agreement also was not performed. There is no universally accepted reference standard for defining CAP, and thus, as with all studies of pneumonia, there is at least some level of uncertainty in classifying the disease. We attempted to minimize misclassification by using 2 reference standards that rely on provider assessment and objectively measured criteria, investigator training and piloting of data collection procedures, and use of a common web-based data collection system. With more than 40,000 discharges in 2010 among the 4 study hospitals, it was not feasible to review all hospitalizations during the study period. However, we selected a random sample of discharges using a wide range of possible pneumonia-related diagnosis codes and reviewed a similar number of discharges without a pneumonia-related diagnosis code. Our matching strategy also maximizes the possibility of identifying missed pneumonia cases, resulting in an accurate estimation of PPV and a most conservative estimate of NPV. However, it is likely that this strategy overestimates sensitivity and underestimates specificity. As an example, applying the PPV and NPV estimates from algorithm 1 for provider-confirmed CAP to the entire study population (using actual numbers of algorithm-positive and total hospital discharges in 2010) results in an estimated sensitivity of 96.1% (reported 99.7%) and specificity of 97.7% (reported 75.7%). Although this exercise is an oversimplification, the true sensitivity and specificity can be expected to fall within this range. Finally, since this study was conducted at tertiary care, freestanding children’s hospitals, our results may not be generalizable to community hospitals. Nonetheless, Children’s Hospital Association-affiliated hospitals account for nearly 20% of pediatric hospitalizations each year, making the Pediatric Health Information System one of the largest and most frequently used administrative data sources for quality and research studies of pediatric hospitalizations in the United States. Moreover, given the high degree of agreement between hospitals, it is likely that our results would be applicable to other pediatric administrative data sets.

In conclusion, to our knowledge, this is the first study to validate ICD-9-CM discharge diagnosis codes for children hospitalized with CAP. We have demonstrated that administrative data are a valuable tool for studying pediatric CAP and provide several strategies that reliably identify this population. Application of these validated algorithms allows for a more accurate identification of relevant and comparable patient populations for research and performance improvement purposes. Understanding the strengths and limitations of these data, as well as the uncertainty occasionally associated with diagnosing CAP, will help ensure valid study conclusions.

REFERENCES


3. Hall, Auger, Queen, and Tieder

4. Lanzky, Auger, Queen, and Tieder

5. Williams, Shah, Auger, Queen, and Tieder