CURB-65 Performance Among Admitted and Discharged Emergency Department Patients With Community-acquired Pneumonia

Adam L. Sharp, MD, MS, Jason P. Jones, PhD, Ivan Wu, MD, Dan Huynh, MD, Keith E. Kocher, MD, MPH, Nirav R. Shah, MD, MPH, and Michael K. Gould, MD, MS

Abstract

Objectives: Pneumonia severity tools were primarily developed in cohorts of hospitalized patients, limiting their applicability to the emergency department (ED). We describe current community ED admission practices and examine the accuracy of the CURB-65 to predict 30-day mortality for patients, either discharged or admitted with community-acquired pneumonia (CAP).

Methods: A retrospective, observational study of adult CAP encounters in 14 community EDs within an integrated healthcare system. We calculated CURB-65 scores for all encounters and described the use of hospitalization, stratified by each score (0–5). We then used each score as a cutoff to calculate sensitivity, specificity, positive predictive value, negative predictive value (NPV), positive likelihood ratios, and negative likelihood ratios for predicting 30-day mortality.

Results: The sample included 21,183 ED encounters for CAP (7,952 discharged and 13,231 admitted). The C-statistic describing the accuracy of CURB-65 for predicting 30-day mortality in the full sample was 0.761 (95% confidence interval [CI], 0.747–0.774). The C-statistic was 0.864 (95% CI, 0.821–0.906) among patients discharged from the ED compared with 0.689 (95% CI, 0.672–0.705) among patients who were admitted. Among all ED encounters a CURB-65 threshold of ≥1 was 92.8% sensitive and 38.0% specific for predicting mortality, with a 99.9% NPV. Among all encounters, 62.5% were admitted, including 36.2% of those at lowest risk (CURB-65 = 0).

Conclusions: CURB-65 had very good accuracy for predicting 30-day mortality among patients discharged from the ED. This severity tool may help ED providers risk stratify patients to assist with disposition decisions and identify unwarranted variation in patient care.

Pneumonia is a leading cause of death in the United States and accounts for over 1.2 million hospitalizations annually, resulting in $10.2 billion in health care costs.1,2 Unnecessary hospitalizations put patients at risk for adverse events and strain the limited resources of an already taxed health care system.3,4 As the key decision makers about hospitalization, emergency department (ED) providers play a key role in assuring appropriate use of the hospital.5 Several studies show that there is unexplained variability among ED physicians’ decisions to admit or discharge patients.6–9

One way for providers to reduce variation in hospitalization would be to use proven decision-making tools.10 To date, pneumonia severity tools have been developed and validated in cohorts of hospitalized patients with very little information in outpatient cohorts.11–13 To
safely use such tools to assist providers with the disposition of ED patients as recommended, an understanding of the outcomes of patients treated outside of the hospital and the performance of such tools for these patients is necessary.

The objective of this paper is to report the accuracy of CURB-65 at predicting 30-day mortality for groups of ED patients that were discharged or hospitalized. We also describe hospitalization practices and patient outcomes for community-acquired pneumonia (CAP) encounters evaluated in community EDs within an integrated health system.

METHODS

Study Setting and Population
We performed a retrospective, observational study of adult ED CAP encounters within Kaiser Permanente Southern California (KPSC), from July 2009 to June 2012. KPSC is a large integrated healthcare system that currently serves over 4 million members who are representative of the diverse population of Southern California. There are 14 EDs ranging in annual volume from 25,000 to 90,000 visits per year, with total of over 900,000 visits annually.

We used existing structured data from electronic health and administrative records collected during routine clinical care and operations. We included data on all visits in the 14 EDs within the KPSC system, but excluded health plan non-members to ensure complete information about postvisit outcomes for those discharged from the ED. Experienced staff performed data collection using information from electronic health records. Human subjects approval was obtained through the KPSC Institutional Review Board.

Measurements
We evaluated the performance of CURB-65 in ED patients, specifically focusing on patients discharged home. We chose to use CURB-65 as the severity assessment tool because it offers a simple six-point scale, using variables readily available for CAP patients evaluated in the ED. CURB-65 has fewer variables than the Pneumonia Severity Index (PSI), which requires 20 separate patient characteristics and does not require obtaining an arterial blood gas, yet has demonstrated similar accuracy in predicting 30-day mortality in studies of patients who have been hospitalized.

Patients who were kept in the hospital under observation status were included among patients admitted to the hospital. CURB-65 scores were calculated based on the established severity scoring system for each adult ED pneumonia encounter, using elements from the electronic health record. The CURB-65 score is composed of five elements, with each element receiving a dichotomous one or zero score, added together to capture an overall score from 0 to 5. The five elements include: confusion, uremia: blood urea nitrogen (BUN) > 20 mg/dL, respiratory rate ≥ 30 breaths/min, blood pressure systolic < 90 mm Hg or diastolic ≤ 60 mm Hg, and age ≥ 65 years. For an operational definition of confusion, we used nursing documentation of the Schmid fall risk assessment.

The Schmid fall risk assessment is a part of regular nursing processes and is the most reliable assessment of mentation for ED patients in our health system. This nursing assessment provides four options for mentation: 1) alert, oriented to person, place, and time; 2) periodic confusion; 3) confusion at all times; or 4) comatose/unresponsive. Patients were given one point for confusion if 2, 3, or 4 was documented during the ED encounter. When encounters resulted in more than one set of vital signs or laboratory values, the initial values were used to calculate the CURB-65 score, as these performed best in predicting mortality in the reported literature. We also performed a sensitivity analysis comparing the performance of CURB-65 using initial vital signs versus the “worst” vital signs recorded during the encounter (i.e., lowest blood pressure and highest respiratory rate).

Cohort Assembly
We included all adult ED encounters with a primary diagnosis of pneumonia (International Classification of Diseases, Ninth Revision [ICD-9] 480–483 or 485–486 or 487.0) or with pneumonia as a secondary diagnosis if the primary diagnosis was respiratory failure (ICD-9 518.8) or sepsis (ICD-9 995.9). We excluded encounters for patients <18 years of age and those with healthcare-associated pneumonia (HCAP). We defined HCAP patients as those who arrived from somewhere other than home (i.e., skilled nursing facility, hospital transfer), had an inpatient encounter in the 30 days prior, or had a past history of HIV infection or organ transplantation. The following additional patient characteristics were gathered to describe the population: age, sex, encounter vital signs, and the comorbidities included in the Charlson Comorbidity Index (CCI).

Outcomes
The primary outcome was 30-day all-cause mortality. This was identified from California state death files, Social Security Administration information for out-of-state deaths, internal KPSC records, and non-KPSC claims. The secondary outcome was hospital admission within 7 days of ED discharge identified by internal KPSC records and claims for care received outside of KPSC.

Analysis
Similar to previously reported derivations and validations of CURB-65, associations between the CURB-65 score and 30-day mortality were used to calculate sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, and negative likelihood ratio using every possible score as a cutoff. This was done for the entire sample and then stratified by disposition to assess performance of CURB-65 in each subgroup. In our analysis, missing values (21% of mental status values and 15% of BUN laboratory values) were considered to be “normal,” similar to other previous reports. Additionally we performed a sensitivity analysis assuming all missing values “abnormal” for comparison. Unpublished evaluation of similar data in our system has shown that approximately 95% of missing values are likely to be normal.
We constructed receiver operating characteristic curves and computed the C-statistic for the full sample and both subgroups. All analyses were performed at the level of the ED encounter, as some patients had more than one encounter.

RESULTS

The sample included 21,183 ED encounters for adults with CAP. Those admitted were older (mean ± SD age = 69.3 ± 16.8 years vs. 53.9 ± 19.0 years) with more comorbidities (median CCI = 2 vs. 5). Table 1 shows the characteristics for all patients in the study stratified by disposition.

Among all encounters, 7,952 (37.5%) resulted in ED discharge and 13,231 (62.5%) resulted in admission. Patient characteristics differed between discharged and admitted groups for all reported variables, with the exception of sex. Every chronic condition was observed at a higher frequency among admitted patients. The distribution of CURB-65 scores was more severe among encounters resulting in admission (Table 2).

For all ED CAP encounters (admitted and discharged), the C-statistic, describing the accuracy of CURB-65 to predict 30-day mortality, was 0.761 (95% confidence interval [CI] = 0.747–0.774; Figure 1). A CURB-65 threshold of ≥1 (n = 13,920), a low-risk score that has previously been suggested to support outpatient management14, was 92.8% sensitive and 38.0% specific for identifying patients who died within 30 days (Table 3).

CURB-65 was more accurate among discharged patients (C-statistic = 0.864, 95% CI = 0.821–0.906) than admitted patients (C-statistic = 0.689, 95% CI = 0.672–0.705). Sensitivity analysis in our overall cohort showed that our model had a higher C-statistic (0.761) than “worst” vital signs (0.759) and “abnormal” missing data assumptions (0.739). Similarly the CURB-65 threshold of ≥1 demonstrated higher sensitivity (94.8% vs. 92.7%) and specificity (62.4% vs. 22.3%) among those discharged (n = 6,982) than for those admitted (n = 6,938; Table 3).

Overall, patients with CAP discharged from the ED had a 5.7% chance of returning for hospital admission

Table 1

Characteristics of Patients With Community-acquired Pneumonia Stratified by Disposition of the Patient After Initial ED Evaluation

<table>
<thead>
<tr>
<th>CURB-65</th>
<th>Discharged (n = 7,952)</th>
<th>Admitted (n = 13,231)</th>
<th>Total (n = 21,183)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4,928 (61.9)</td>
<td>2,795 (21.1)</td>
<td>7,723 (36.5)</td>
</tr>
<tr>
<td>1</td>
<td>2,054 (25.8)</td>
<td>4,143 (31.3)</td>
<td>6,197 (29.3)</td>
</tr>
<tr>
<td>2</td>
<td>744 (9.4)</td>
<td>3,869 (29.2)</td>
<td>4,613 (21.8)</td>
</tr>
<tr>
<td>3</td>
<td>200 (2.5)</td>
<td>1,987 (15.1)</td>
<td>2,187 (10.4)</td>
</tr>
<tr>
<td>4</td>
<td>25 (0.3)</td>
<td>407 (3.1)</td>
<td>432 (2.0)</td>
</tr>
<tr>
<td>5</td>
<td>1 (0.0)</td>
<td>20 (0.2)</td>
<td>21 (0.1)</td>
</tr>
</tbody>
</table>

Table 2

A Description of Patient Outcomes for ED Patients With Community-acquired Pneumonia Stratified by Those Discharged From the ED or Admitted to the Hospital

<table>
<thead>
<tr>
<th>CURB-65</th>
<th>Discharged Within 7 Days</th>
<th>30-day Mortality</th>
<th>30-day Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Admitted</td>
<td>n</td>
<td>Percent</td>
</tr>
<tr>
<td>0</td>
<td>206 (of 4,928)</td>
<td>4.2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>159 (of 2,054)</td>
<td>7.7</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>56 (of 744)</td>
<td>7.5</td>
<td>13</td>
</tr>
<tr>
<td>≥3</td>
<td>29 (of 226)</td>
<td>12.8</td>
<td>23</td>
</tr>
</tbody>
</table>

Patients admitted to the hospital within 7 days of ED discharge also reported. Results are categorized by CURB-65 score, a five-point pneumonia severity score increasing from 0 to 5.
The Accuracy of CURB-65 at Each of Five Possible Thresholds in Predicting 30-day Mortality for Patients Discharged or Admitted

DISCUSSION

This study reports the accuracy of the CURB-65 index at predicting 30-day mortality among groups of ED patients that were discharged or hospitalized. We found that CURB-65 performed especially well among patients who were discharged, for whom the accuracy was better than previously reported for inpatient cohorts. Among CAP ED patients with a CURB-65 score of 0, few were subsequently hospitalized (5.7%) and very few (0.1%) died within 30 days. As hypothesized, we found the NPV of a CURB-65 score of 0 is very high among discharged patients (99.9%) as well as for the entire group (99.0%). Despite this, we found that 36% of patients at the lowest risk (CURB-65 = 0) were still admitted to the hospital with CAP.

We do, however, acknowledge the limitations of such instruments and the complexity inherent to the decision to admit a patient. This is shown by the differences seen in the age and comorbidities of patients with similar CURB-65 scores when comparing discharged and admitted groups. Providers may use CURB-65 to assist them in assessing patient risk, but clinical judgment is paramount when making decisions about disposition. Comorbidities, social circumstances, and additional clinical factors, such as hypoxia, may necessitate observation or admission for some patients with CURB-65 scores of 0 or 1. These factors, which are not accounted for in CURB-65, are likely to explain differences in 30-day mortality between patients discharged and admitted when comparing patients with similar CURB-65 scores.

It is not clear what constitutes an acceptable mortality rate when making recommendations regarding disposition, and not all deaths from pneumonia can be prevented. Our expected 30-day mortality frequency for patients at lowest risk (CURB-65 = 0) was 0.1% for those discharged from the ED, but 2.6% for patients admitted. It is possible that the latter frequency could have been higher if large numbers of apparently low-risk patients were not admitted, and this should be

### Table 3
The Accuracy of CURB-65 at Each of Five Possible Thresholds in Predicting 30-day Mortality for Patients Discharged or Admitted From the ED

<table>
<thead>
<tr>
<th>CURB-65 Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>PLR</th>
<th>NLR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discharged</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>94.8% (85.6–98.9)</td>
<td>62.4% (61.3–63.5)</td>
<td>1.8% (1.7–2.5)</td>
<td>99.9% (99.8–99.9)</td>
<td>2.5 (2.4–2.7)</td>
<td>0.1 (0.0–0.2)</td>
</tr>
<tr>
<td>≥2</td>
<td>62.1% (48.4–74.5)</td>
<td>88.2% (87.4–88.9)</td>
<td>3.7% (3.5–3.6)</td>
<td>99.7% (99.5–99.7)</td>
<td>5.2 (4.3–6.5)</td>
<td>0.4 (0.3–0.6)</td>
</tr>
<tr>
<td>≥3</td>
<td>39.7% (27.0–53.4)</td>
<td>97.4% (97.1–97.8)</td>
<td>10.2% (9.0–16.5)</td>
<td>99.5% (99.2–99.6)</td>
<td>15.4 (10.9–21.8)</td>
<td>0.6 (0.5–0.8)</td>
</tr>
<tr>
<td>≥4</td>
<td>13.8% (6.1–25.4)</td>
<td>99.8% (99.6–99.9)</td>
<td>30.8% (29.1–48.6)</td>
<td>99.4% (99.5–99.6)</td>
<td>60.5 (27.4–133.5)</td>
<td>0.9 (0.8–1.0)</td>
</tr>
<tr>
<td>5</td>
<td>1.7% (0.0–9.2)</td>
<td>100.0% (N/A)</td>
<td>100.0% (21.3–100.0)</td>
<td>99.3% (77.5–N/A)</td>
<td>N/A</td>
<td>1.0 (0.9–1.0)</td>
</tr>
<tr>
<td><strong>Admitted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>92.7% (90.9–94.2)</td>
<td>22.3% (21.6–23.0)</td>
<td>9.0% (8.7–11.3)</td>
<td>97.4% (96.7–97.5)</td>
<td>1.2 (1.2–1.2)</td>
<td>0.3 (0.3–0.4)</td>
</tr>
<tr>
<td>≥2</td>
<td>74.0% (71.1–76.6)</td>
<td>54.6% (53.8–55.5)</td>
<td>11.9% (11.5–13.5)</td>
<td>96.2% (95.6–96.3)</td>
<td>1.6 (1.6–1.7)</td>
<td>0.5 (0.4–0.5)</td>
</tr>
<tr>
<td>≥3</td>
<td>41.4% (38.3–44.5)</td>
<td>83.6% (82.9–84.3)</td>
<td>17.3% (16.7–19.2)</td>
<td>94.5% (93.8–94.7)</td>
<td>2.5 (2.3–2.7)</td>
<td>0.7 (0.7–0.7)</td>
</tr>
<tr>
<td>≥4</td>
<td>10.6% (8.8–12.7)</td>
<td>97.4% (97.1–97.7)</td>
<td>25.3% (23.2–29.3)</td>
<td>92.9% (91.4–93.6)</td>
<td>41 (3.3–5.0)</td>
<td>0.9 (0.9–0.9)</td>
</tr>
<tr>
<td>5</td>
<td>0.7% (0.3–1.4)</td>
<td>99.9% (98.8–99.9)</td>
<td>35.0% (23.9–52.7)</td>
<td>92.4% (82.9–95.8)</td>
<td>6.5 (2.6–16.2)</td>
<td>1.0 (1.0–1.0)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>92.8% (91.1–94.3)</td>
<td>38.0% (37.4–38.7)</td>
<td>7.4% (7.2–9.3)</td>
<td>99.0% (98.7–99.0)</td>
<td>1.5 (1.5–1.5)</td>
<td>0.2 (0.2–0.2)</td>
</tr>
<tr>
<td>≥2</td>
<td>73.3% (70.5–75.9)</td>
<td>67.9% (67.1–68.4)</td>
<td>10.8% (10.5–12.2)</td>
<td>97.9% (97.6–98.0)</td>
<td>2.3 (2.2–2.4)</td>
<td>0.4 (0.3–0.6)</td>
</tr>
<tr>
<td>≥3</td>
<td>41.3% (38.3–44.3)</td>
<td>89.0% (88.6–89.5)</td>
<td>16.7% (16.1–18.5)</td>
<td>96.6% (96.2–96.7)</td>
<td>3.8 (3.5–4.1)</td>
<td>0.7 (0.6–0.7)</td>
</tr>
<tr>
<td>≥4</td>
<td>10.8% (9.0–12.8)</td>
<td>98.3% (98.1–98.5)</td>
<td>25.6% (23.6–29.5)</td>
<td>95.4% (94.4–95.8)</td>
<td>6.5 (5.3–7.9)</td>
<td>0.9 (0.9–0.9)</td>
</tr>
<tr>
<td>5</td>
<td>0.7% (0.3–1.5)</td>
<td>99.9% (99.5–100.0)</td>
<td>38.1% (26.5–54.9)</td>
<td>95.0% (89.0–97.3)</td>
<td>11.5 (4.8–27.8)</td>
<td>1.0 (1.0–1.0)</td>
</tr>
</tbody>
</table>

The sensitivity, specificity, PPV, NPV, PLR, and NLR for the different thresholds for CURB-65 scoring. (Calculations that are not possible or ill-defined are listed as not applicable [N/A].

NLR = negative likelihood ratio; NPV = negative predictive value; PLR = positive likelihood ratio; PPV = positive predictive value.
addressed in future studies. Alternatively, the higher mortality among those admitted may be related to common risks associated with inpatient care, such as hospital-acquired infections, medication errors, and procedure-related complications. Previous studies of the CURB-65 and the PSI reported observed mortality rates of 2.0% (CURB-65 = 0–1) and 1.6% (PSI I-III) for low-risk inpatient groups, similar to what we found.11

The problem of setting a threshold for admission is not unique to pneumonia, and it is a key question that requires further discussion to optimize clinical outcomes and use of limited healthcare resources. Accurate predictive models will enable healthcare providers to engage their patients in such discussions and to incorporate this information into shared decision-making. At the very least, this tool provides a metric for comparison between providers and hospitals as a marker for evaluation of efficient use of hospital services.

Taking advantage of the electronic health record, a real-time process to automatically calculate CURB-65 for pneumonia patients is one strategy to identify and reduce unnecessary variability and ensure that patients get care in the most appropriate setting. It is hoped that the thoughtful implementation of CURB-65 and future prospective research will clarify the impact this tool may have if implemented into clinical emergency care. We believe that CURB-65 is a potentially useful stratification tool to compare the use of inpatient care for “low-risk” patients with pneumonia between providers and EDs.

LIMITATIONS

This is a retrospective observational study and, consequently, comes with the limitations inherent to this design. Specifically, there is a risk of ascertainment bias, as our cohort reflects the ICD-9 diagnosis of pneumonia, and no expert or independent review of patient charts and chest radiography was performed for confirmation. Furthermore, in many cases, factors other than those captured by the CURB-65 score likely influenced the providers on their disposition decisions, potentially contributing to the differences in outpatient versus inpatient mortality. This includes varying importance of each of the five variables in CURB-65. For example, a CURB-65 score of 1 due to confusion is much more concerning than a patient with score of one due to age of 66. These factors were not well defined or analyzed in this study, but may be the subject of future investigations.

We also recognize the imperfect sensitivity of how we defined HCAP. Although this definition covers the large majority of patients seen in our system who would be classified as HCAP, it is possible that patients within our groups categorized as CAP would be classified as HCAP under more complex definitions currently reported by the American Thoracic Society/Infectious Disease Society of America.27

As with any large data set, we were limited by missing data for patient encounters without complete information to calculate a CURB-65 score; however, we took steps to mitigate this with a sensitivity analysis and followed previously published strategies.12 Our primary outcome of 30-day all-cause mortality may include random unrelated deaths such as by a motor vehicle collision. We chose to use all-cause mortality, instead of only mortality from pneumonia because we understand ED providers must take into consideration other comorbidities when deciding to hospitalize or discharge patients from the ED. Finally, our primary outcome depends on the accuracy of vital status information and therefore we attempted to avoid any false classification by combining various data sources, including California state death files, Social Security Administration for out-of-state deaths, internal KPSC records, and outside claims, this strategy has shown to be complete and highly accurate.31

CONCLUSION

We conclude that CURB-65 had very good accuracy for predicting 30-day mortality for patients discharged from the ED with community-acquired pneumonia. This severity tool may help ED providers to risk-stratify patients, assist with disposition decisions, and identify unwarranted variation in patient care. Prospective research is needed to better understand the impact of CURB-65 on provider decision making and its effect on clinical care and patient outcomes.

References

10. Wilk S, Michalowski W, O’Sulliven D, et al. A task-based support architecture for developing point-of-


