Clinical Scoring Tools
Which Is Best to Predict Clinical Response and Long-Term Outcomes?

Timothy Wiemken, PhD, MPH, CIC*, Robert Kelley, PhD, Julio Ramirez, MD

THE NEED FOR PREDICTION SCORES

During the initial management of patients with community-acquired pneumonia (CAP), physicians need to assess severity of the disease and predict the likely clinical outcomes of the patient. This information is used to make important clinical decisions, such as site of care, extent of laboratory work-up, and therapeutic interventions. Clinical judgment has been the primary tool to define severity of disease and likely clinical outcomes in hospitalized patients with CAP, but has poor predictive value.1 Because of this, CAP prediction scores were developed to help physicians define severity of disease and likely clinical outcomes of their patients. CAP prediction scores are also important in clinical research. They are used to stratify patients on disease severity or the likelihood of a particular outcome in the design of the study, and to adjust for confounding bias during the analysis phase of the study.

* Corresponding author.
E-mail address: tlwiem01@louisville.edu

http://dx.doi.org/10.1016/j.idc.2012.11.015
id.theclinics.com
0891-5520/13/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
The first prediction score for patients with CAP was the Pneumonia Severity Index (PSI). This prediction score was developed to predict 30-day mortality. Because of the complexity in its calculation, investigators began to develop more simple scores to predict this outcome. Today, there are a multitude of prediction scores for CAP, and most of these scores were developed to predict 30-day mortality. Table 1 outlines a few selected CAP prediction scores described in this article. Data on the use of these scores to predict clinical outcomes other than 30-day mortality are limited. This article reviews the most relevant clinical outcomes in hospitalized patients with CAP and outlines the role of these scores as tools to help physicians predict outcomes.

**CAP CLINICAL OUTCOMES**

The clinical outcomes of hospitalized patients with CAP can be classified as (1) outcomes during hospitalization (usually within the first 7 days after hospital admission); (2) outcomes during 30-day follow-up; and (3) outcomes occurring years after hospital discharge (long-term outcomes). Fig. 1 depicts each outcome in chronologic order.

**Outcomes During Hospitalization**

**Clinical failure**

Clinical failure is associated with increased complications, length of hospital stay, and mortality (see Fig. 1, point 1). These complications also increase the total direct cost of care for hospitalized patients with CAP, adding to strain on the health care system. Definitions of clinical failure vary significantly from study to study. The simplest definition of failure is the lack of response to therapy associated with clinical deterioration. This early outcome can occur in up to one-quarter of patients with CAP and nearly one-third of patients with severe CAP. Although investigators increasingly examine clinical failure as an outcome for hospitalized patients with CAP, few studies have maintained a constant, comprehensive definition. Aliberti and Blasi provide an excellent overview of various failure definitions used in many of these studies. Our group recently reported a comprehensive definition of clinical failure consisting of the following criteria: (1) acute pulmonary deterioration with the need for invasive or noninvasive mechanical ventilation; (2) acute hemodynamic deterioration with the need for aggressive fluid resuscitation (eg, >40 mL/kg colloids or crystalloids), vasopressors, or invasive procedures (eg, pericardial drainage or electrical cardioversion); and (3) in-hospital mortality up to 28 days after hospital admission. It is hoped that this will be used as a basis for a consistently reported definition of clinical failure.

The ability to predict clinical failure in hospitalized patients with CAP is hampered by the lack of consistent definitions. Therefore, little work has been done to predict failure in hospitalized patients with CAP. One study suggested that the PSI and CURB-65 were inferior to SCAP in the prediction of treatment failure, although the areas under the receiver operating characteristic (ROC) curves (AUC) for all of these scores were very low (0.52–0.61). Using data from a subset of 500 patients enrolled in the Community-Acquired Pneumonia Organization (CAPO) international cohort study, we used ROC curves to examine the ability of PSI and CRB-65 to accurately predict clinical failure using the definition described previously. Fig. 2 depicts the results of this analysis. We calculated AUC to be 0.62 and 0.63 ($P = .786$), respectively. These data do not support the use of these scores to accurately predict clinical failure.

**Intensive care unit admission**

Early identification of hospitalized patients with CAP in need of intensive care is important to improve outcomes and reduce health care costs (see Fig. 1, point 2). Improving
### Table 1
Selected CAP prediction scores

<table>
<thead>
<tr>
<th>Name</th>
<th>Variables/Points</th>
</tr>
</thead>
</table>
| Pneumonia Severity Index²   | - If male/age (y)  
- If female/age (y) – 10  
- Nursing home resident/10  
- Neoplastic disease/30  
- Liver disease/20  
- Congestive heart failure/10  
- Cerebrovascular disease/10  
- Renal disease/10  
- Altered mental status/20  
- Heart rate $\geq$ 125 beats per min/20  
- Respiratory rate $\geq$ 30 breaths per min/20  
- Systolic blood pressure $< 90$ mm Hg/15  
- Temperature $< 35^\circ$ C or $\geq 40^\circ$ C/10  
- Arterial pH $< 7.35/30$  
- Blood urea nitrogen $\geq 30$ mg/dL/20  
- Sodium $< 130$ mmol/L/20  
- Glucose $\geq 250$ mg/dL/10  
- Hematocrit $< 30%$/10  
- PaO₂ $< 60$ mm Hg/10  
- Pleural effusion/10                                                                                                                                 |
| CURB-65³                    | - Confusion/1  
- Urea $> 7$ mmol/L/1  
- Respiratory rate $\geq 30$ breaths per min/1  
- Systolic blood pressure $< 90$ mm Hg or diastolic blood pressure $\leq 60$ mm Hg/1  
- Age $\geq 65$/1                                                                                                                                  |
| CRB-65³                     | - Confusion/1  
- Respiratory rate/1  
- Systolic blood pressure/1  
- Age $\geq 65$/1                                                                                                                                  |
| SMART-COP⁴                  | - Systolic blood pressure $< 90$ mm Hg/2  
- Multilobar infiltrates/1  
- Albumin $< 2.5$ g/dL/1  
- Respiratory rate $\geq 25$ breaths per min if $\leq 50$ y; $\geq 30$ breaths per min if $> 50$ y/1  
- Heart rate (tachycardia) $\geq 125$ beats/min/1  
- Confusion/1  
- PaO₂ $< 70$ mm Hg, O₂ saturation $< 93\%$ or PaO₂/FiO₂ $< 333$ if $\leq 50$ y; $< 60$ mm Hg, $\leq 90\%$, $< 250$ if $> 50$ y/2  
- Arterial pH $< 7.35/2$                                                                                                                         |
| SCAP (severe CAP)⁵          | - Arterial pH $< 7.30/2.38$  
- Systolic blood pressure $< 90$ mm Hg/2.19  
- Respiratory rate $> 30$ breaths per min/1.83  
- Altered mental status/0.87  
- Blood urea nitrogen $> 30$ mg/dL/0.92  
- PaO₂ $< 54$ mm Hg or PaO₂/FiO₂ $< 250$ mm Hg/1.12  
- Age $\geq 80$ y/0.86  
- Multilobar infiltrates/0.68                                                                                                                      |
outcomes is a necessary goal for all practitioners; however, the reduction in health care costs is only of recent interest.

The “need for intensive care” can be defined directly, in terms of the need for admission or transfer to the intensive care unit (ICU) per physician’s orders, or indirectly

Fig. 1. Chronology of outcomes for hospitalized patients with CAP. ICU, intensive care unit.

Fig. 2. Diagnostic accuracy of the PSI and the CRB-65 for the prediction of clinical failure in hospitalized patients with CAP. AUC, area under the receiver operating characteristic curve.
through assessing the patient’s need for intensive respiratory or vasopressor support.\textsuperscript{4} These indirect variables are proxy measures because they are often only provided in the setting of the ICU.

CAP prediction scores have been evaluated or developed to predict the need for intensive care on admission to the hospital. The SMART-COP score developed by Charles and colleagues\textsuperscript{4} is the best studied of these scores. In the original study, the investigators found that SMART-COP was very good at predicting the need for ICU admission (AUC, 0.87). This prediction was much better than the next best score evaluated, the PSI, with an AUC of 0.69. Other studies have also shown that SMART-COP is a useful predictor of the need for ICU admission.\textsuperscript{17,18} Investigators evaluating scores to predict ICU admission, such as REA-ICU, CURXO-80, PSI, CURB-65, and the 2007 International Diseases Society of America (IDSA)/American Thoracic Society (ATS) minor criteria for predicting ICU admission, have reported widely varied results.\textsuperscript{5,17,18} Although many of these scores have relatively high AUCs for the prediction of the need for ICU admission, they often result in very low positive predictive values because of the low prevalence of severe CAP.\textsuperscript{19} Although patients in the ICU may obtain the highest level of care, using a score with a low positive predictive value to define if a patient should be admitted to the ICU leads to unnecessary ICU admissions and increased costs of care.

**In-hospital mortality**

Mortality is arguably the worst possible outcome for patients with CAP (see Fig. 1, point 3). The burden on society is also significant with respect to lost productivity and cost. Mortality can be evaluated at any time point, although in-hospital mortality is the first time point of interest for hospitalized patients with CAP. However, in-hospital mortality has not been evaluated as a clinical outcome to a great extent.

In-hospital mortality is simply defined as death during hospitalization. Few investigators separate specific causes of death (eg, related to CAP vs not related to CAP) because of the difficulty in defining the actual cause. Some investigators have examined the predictive accuracy of various scores to predict this outcome. Ewig and colleagues\textsuperscript{20} reported an AUC of 0.68 for the CRB-65 prediction of in-hospital mortality. Furthermore, they suggested that varying the age groups of the CRB score could enhance its ability to predict in-hospital mortality.\textsuperscript{20} Similarly, Richards and colleagues\textsuperscript{21} reported an AUC of 0.65 for CURB-65 and PSI for the prediction of in-hospital mortality. Karmakar and Wilsher\textsuperscript{22} reported a dose-response relationship of CURB-65 scores on increasing in-hospital mortality rates (1% for CURB-65 of 0 or 1, 2% for CURB-65 of 2, and 13% for CURB-65 of >2).

The prediction of various outcomes, including in-hospital mortality, may also vary based on the cause of pneumonia. We reported for hospitalized patients with CAP caused by 2009 H1N1 influenza A virus, the PSI, CRB-65, and the CURB-65 were not adequate for predicting in-hospital mortality.\textsuperscript{23} Although those scores, which weight advanced age heavily, should intuitively predict mortality at a later date, the prediction of in-hospital death may be more closely linked to acute decompensation rather than advanced age. During the time of the initial wave of the pandemic, we created a score using data from the CAPO international cohort study to assist with the prediction of in-hospital death for our patients with 2009 H1N1 influenza A virus pneumonia: the CROMI score (comorbidity, respiratory rate, oxygen saturation, mental status changes, and infiltrates at chest radiograph) (Figs. 3 and 4). This score had an AUC for the prediction of in-hospital mortality of 0.83, whereas the PSI, CURB-65, and CRB-65 had AUCs in the low 0.70s (data not shown).
Using data from our multicenter study for lower respiratory tract infections titled Rapid Empiric Treatment with Oseltamivir Study, we recently evaluated the diagnostic accuracy of the PSI and CRB-65 scores for predicting in-hospital mortality for all hospitalized patients with CAP in seven hospitals in the Louisville metropolitan area. As Fig. 5 depicts, these two scores do not accurately predict in-hospital mortality (AUC, 0.68 vs 0.59, respectively; $P = .197$).

**Clinical stability**

Clinical stability is one of the most essential outcomes for hospitalized patients with CAP because it can assist in the direction of the management of these patients (see Fig. 1, point 4). Use of clinical stability criteria can also assist the physician in the
timely discharge of hospitalized patients with CAP, leading to lower health care resource use and lower costs of care.

Clinical stability is defined as the time at which the patient has improved enough to be switched from intravenous antibiotics to oral antibiotics. This is also known as the number of days to switch therapy, or the time to clinical stability (TCS).\textsuperscript{24,25} Understanding what factors may predict TCS can be a critical factor in site-of-care decisions, improving clinical outcomes, and decreasing the overall cost of care.

**Fig. 4.** Prediction of in-hospital mortality due to 2009 H1N1 influenza A virus by the CROMI score.

**Fig. 5.** Diagnostic accuracy of the PSI and the CRB-65 for the prediction of in-hospital mortality in hospitalized patients with CAP.
Two major definitions exist for evaluating when patients have improved enough to be considered clinically stable. These two criteria have been offered as recommendations by the ATS and IDSA in their guidelines for the management of patients with CAP. In 2001, the Ramirez criteria\textsuperscript{25} were suggested,\textsuperscript{26} and in 2007 the Halm criteria\textsuperscript{27} were suggested.\textsuperscript{28} These criteria can be found in Table 2.

Scores for predicting TCS have not been well defined. Our group previously evaluated the use of the PSI and CRB-65\textsuperscript{3} for predicting TCS within the first week of hospitalization using the Ramirez criteria.\textsuperscript{29} We found statistically equivalent prediction for the two scores (95% confidence interval for the difference in AUC, -0.03 to 0.01). More importantly, we identified the poor predictive accuracy of both scores with AUC values of 0.64 and 0.65, respectively.\textsuperscript{30} Menendez and colleagues\textsuperscript{31} reported that severe CAP, as defined by a PSI risk class of 3, 4, or 5, was associated with a longer TCS; however, the predictive accuracy of the score was not calculated. These investigators used the Halm (ATS/IDSA 2007) definition of clinical stability. These are the only studies evaluating a CAP score for the prediction of clinical stability. To add to this body of evidence, we used data from the more than 7000 patients enrolled in the CAPO international cohort study to construct ROC curves. These curves evaluated the accuracy of PSI and CRB-65 for predicting TCS in the first week of hospitalization using the Ramirez criteria and the Halm criteria (Figs. 6 and 7, respectively). These data suggest that these two scores are very poor methods for predicting both definitions of clinical stability.

\textbf{Nonresolving pneumonia}

Nonresolving pneumonia is a relevant outcome for the practicing physician (see Fig. 1, point 5). Approximately 15% of hospitalized patients with CAP may have nonresolving pneumonia.\textsuperscript{32} Nonresolving or nonresponding pneumonia can be defined as patients who do not improve within 1 week to 10 days after initiation of antimicrobial therapy\textsuperscript{13,33}; however “resolution” of pneumonia may not be easily defined. This creates difficulties for studying this outcome. Furthermore, a noninfectious cause of nonresolving CAP has been documented to play a role in up to 20% of hospitalized patients with CAP.\textsuperscript{12} To date, no investigations have evaluated the accuracy of any score for the prediction of nonresolving pneumonia.

\textbf{Length of hospitalization}

Decreasing the length of hospitalization of patients with CAP is essential for decreasing the overall costs of care (see Fig. 1, point 6). Prompt discharge of patients who are able to improve at home may also decrease the risk of other poor outcomes, such as falls or health care–associated infections.

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical stability criteria</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Variables</th>
</tr>
</thead>
</table>
| Ramirez et al,\textsuperscript{25} 1995 | 1. Subjective cough and shortness of air improving  
2. Afebrile for at least 8 h  
3. White blood cell count returning to normal |
| Halm et al,\textsuperscript{27} 1998 | 1. Temperature $\leq 38.3^\circ$ C  
2. O\textsubscript{2} saturation $\geq 90\%$  
3. Respiratory rate $<24$ breaths per min  
4. Systolic blood pressure $\geq 90$ mm Hg  
5. Normal mental status |
Fig. 6. Diagnostic accuracy of the PSI and the CRB-65 for the prediction of 2001 clinical stability criteria in hospitalized patients with CAP.

PSI AUC = 64%
CRB-65 AUC = 58%
P = .273

Fig. 7. Diagnostic accuracy of the PSI and the CRB-65 for the prediction of 2007 clinical stability criteria in hospitalized patients with CAP.

PSI AUC = 59%
CRB-65 AUC = 49%
P = .015
Length of hospitalization for patients with CAP is defined as the number of days between hospital admission and discharge. This outcome can be used as a proxy measure for the cost of care, because a reasonable fixed price can be assigned to a patient per day of hospitalization. However, this measure can be biased significantly because of several factors. For example, if a score is to predict how many days a patient will be hospitalized, a significant confounding effect is the number of days the patient is ill before hospitalization. Regardless, limited data have been reported examining the prediction of length of stay using pneumonia severity scores. In 2010, our group presented data suggesting the poor predictive accuracy of PSI (AUC, 0.65) and CRB-65 (AUC, 0.60) for this outcome.

Outcomes During 30-Day Follow-Up

Clinical cure at 30 days
Clinical cure can be defined as complete resolution of signs and symptoms of respiratory infection (see Fig. 1, point 7). Resolution of signs and symptoms include the lack of fever; shortness of air; leukocytosis; pulmonary infiltrate; and cough (or return to the level of cough present before the episode of pneumonia in patients with chronic obstructive pulmonary disease). This results in the patient’s respiratory function returning to the level present before the development of pneumonia.

In some patients, the primary signs and symptoms of pneumonia may resolve, but organ functions do not return to baseline levels. These patients are clinically cured, but their general well-being is not back to baseline. In this regard, evaluation of “quality of life” after an episode of pneumonia may be a more relevant outcome that “clinical cure” after pneumonia.

The ultimate goal of pneumonia therapy is not only to cure the pneumonia but also to restore quality of life to the level present before the pulmonary infection. Regardless, no data are available for the accuracy of CAP scores to predict clinical cure at 30 days postdischarge.

Quality of life at 30 days
Quality of life, as previously described, is arguably the most important outcome for hospitalized patients with CAP because the goal of quality medical care should be to cure the patient and improve the patient’s overall quality of life (see Fig. 1, point 8). Quality of life for patients with CAP is defined using standardized instruments, such as the Short Form-36. These measures can include absence of pain, ability to perform daily tasks, and various measures of physical abilities. However, little work has focused on the ability to predict the patient’s quality of life at 30 days based on their disease severity on admission. One study reported that patients with pneumonia had significantly worse quality of life (as measured by the Short Form-36) compared with an age- and gender-matched nonpneumonia cohort shortly after hospitalization; however, differences based on initial disease severity were not thoroughly analyzed. Our group is currently collecting data on 30-day quality of life measures for hospitalized patients with CAP as part of a multicenter, randomized, clinical trial entitled Rapid Empiric Treatment with Oseltamivir Study. These data will be used to examine the accuracy of various severity scores to predict quality of life as they become available.

Rehospitalization at 30 days
Rehospitalization has recently become a significant clinical outcome, particularly because of the risk of limited reimbursement from insurance agencies (see Fig. 1, point 9). In 2005, it was shown that the cost of preventable rehospitalizations was approximately $12 billion for Medicare beneficiaries. Because of the extreme costs
of preventable rehospitalizations, this outcome will be a focus of health care for years to come.

Approximately 20% of patients with CAP are rehospitalized within 30 days, with 30% of these patients reporting pneumonia to be the reason for readmission. These data emphasize the need to be able to predict which patients may be at risk for preventable rehospitalization so they can be evaluated before discharge. However, no studies have evaluated the accuracy of CAP scoring tools to predict readmission at 30-days. To help fill this gap, we analyzed data from the CAPO international cohort study. As depicted in Fig. 8, it is clear that neither the PSI (AUC, 0.56) nor the CRB-65 (AUC, 0.49) is able to help assess the need for rehospitalization.

**Thirty-day mortality**

As outlined for in-hospital mortality, 30-day mortality is a critical outcome for hospitalized patients with CAP (see Fig. 1, point 10). This outcome is also used as a quality measure for the Centers for Medicare and Medicaid Services in the United States.

In terms of outcome prediction, 30-day mortality is the outcome for which most of the CAP scores (severity scores) were developed. However, it is clear that up to 50% of CAP deaths are unrelated to the initial severity of disease. These data question the use of predicting 30-day mortality in hospitalized patients with CAP. Regardless, numerous authors have published data on the use of nearly every pneumonia severity score to predict this outcome. Recently, Chalmers and colleagues reported data on a meta-analysis of CAP scores to predict 30-day mortality. Table 3 depicts the wide range of AUCs calculated for many different prediction scores. These data vary significantly and suggest that a maximum AUC of 0.89 for all severity scores falls below the established cutoff of 90% for an excellent discriminatory test. Some authors have recommended recalibrating these scores to account for differences in various

![Fig. 8. Diagnostic accuracy of the PSI and the CRB-65 for the prediction of 30-day rehospitalization in hospitalized patients with CAP.](image-url)
populations under study. Recalibration may change the AUCs for many of the published studies, although the methods and use of recalibration of severity scores in the clinical setting (eg, outside of research) are not well defined.

Clinical cure at 1 year or 5 years
No data are available defining the accuracy of CAP scores to predict this outcome (see Fig. 1, point 11). Nonetheless, the quality of the patient’s life becomes increasingly important. For example, underlying comorbidities may be accelerated because of the pneumonia even if the patient is cured. This may lead to a decrease in quality of life for years after the initial episode of pneumonia, regardless of the timing of clinical cure.

Long-Term Outcomes

Quality of life at 1 or 5 years
Similar to the quality of life at 30 days, no studies have evaluated the long-term quality of life prediction of various CAP scores (see Fig. 1, point 12). However, some data are available to suggest that severe disease has long-term consequences for quality of life. As part of the RETO study, our group is actively collecting data on the long-term quality of life of hospitalized patients with CAP and plan to report the prediction of various severity scores on this outcome.

Rehospitalization at 1 or 5 years
Data have not been reported regarding the accuracy of CAP scores to predict the need for long-term rehospitalization (see Fig. 1, point 13). We are collecting rehospitalization data for 6 months and 1-year postdischarge in the RETO study and will report these data as they become available.

Mortality at 1 or 5 years
Pneumonia may now be considered a chronic disease because of the long-term impact on patient survival (see Fig. 1, point 14). However, there are still very few studies closely examining this long-term impact. Understanding host/environment/pathogen factors that contribute to long-term impact is critical to ensure good patient outcomes. Moreover, predicting which patients with CAP will be more impacted in the long-term is a crucial factor for novel treatment modalities. We were not able to find data evaluating the accuracy of any available CAP score for predicting long-term mortality. As part of the RETO study, we evaluated the accuracy of the PSI and CRB-65 scores for predicting 1-year mortality (Fig. 9). These data suggest that the PSI and CRB-65 are not adequate for predicting 1-year mortality in hospitalized patients with CAP (AUC, 0.73 vs 0.66, respectively; $P = .083$).

<table>
<thead>
<tr>
<th>Severity Score</th>
<th>AUC Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSI</td>
<td>0.70–0.89</td>
</tr>
<tr>
<td>CURB-65</td>
<td>0.73–0.87</td>
</tr>
<tr>
<td>CRB-65</td>
<td>0.69–0.74</td>
</tr>
<tr>
<td>ATS 2001, sensitivity/specificity</td>
<td>41%–94%</td>
</tr>
<tr>
<td>IDSA/ATS 2007, sensitivity/specificity</td>
<td>45%–81%</td>
</tr>
</tbody>
</table>

SUMMARY

This article identifies the most critical clinical outcomes for hospitalized patients with CAP, and assesses the merits of various CAP scores for the prediction of each outcome. In regards to the best scoring tool to predict clinical response and long-term outcomes, PSI is probably more accurate score to predict a wide range of clinical outcomes. Considering the complexity of the PSI calculation and the sensitivity and specificity for any particular outcome, a great need exists for simple and more accurate scoring tools.

Currently available scores have focused on past medical history, signs and symptoms of pneumonia, and basic laboratory work-up. It is unlikely that further combinations of these variables will significantly increase the accuracy of any score. Improving these scores requires at least two areas of research. First, investigators need to incorporate biomarkers, such as proadrenomedullin, C-reactive protein, brain natriuretic peptide, and procalcitonin, which have shown promise for predicting various outcomes and enhancing currently available scoring tools. A second area of research should focus on the identification of new biomarkers associated with early and late outcomes.

For the management of hospitalized patients with CAP, early outcomes are often considered the most important outcomes. There are currently very few accurate, validated scores to assist physicians in the prediction of any clinical outcomes.

REFERENCES