Comparison of the prognostic performance of the CURB-65 and a modified version of the pneumonia severity index designed to identify high-risk patients using the International Community-Acquired Pneumonia Collaboration Cohort

Raphae S. Barlas, Allan B. Clark, Yoon K. Loke, Chun S. Kwok, Derek C. Angus, Ane Uranga, Pedro P. España, Dean T. Eurich, David T. Huang, Shin Y. Man, Timothy H. Rainer, Donald M. Yealy, Phyo K. Myint, Maria K. Mor, Michael J. Fine

PII: S0954-6111(22)00149-4

DOI: https://doi.org/10.1016/j.rmed.2022.106884

Reference: YRMED 106884

To appear in: *Respiratory Medicine*

Received Date: 6 January 2022

Revised Date: 13 May 2022

Accepted Date: 14 May 2022

Please cite this article as: Barlas RS, Clark AB, Loke YK, Kwok CS, Angus DC, Uranga A, España PP, Eurich DT, Huang DT, Man SY, Rainer TH, Yealy DM, Myint PK, Mor MK, Fine MJ, Comparison of the prognostic performance of the CURB-65 and a modified version of the pneumonia severity index designed to identify high-risk patients using the International Community-Acquired Pneumonia Collaboration Cohort, *Respiratory Medicine* (2022), doi: https://doi.org/10.1016/j.rmed.2022.106884.

Credit Author Statement

Raphae S Barlas: Conceptualization; Methodology; Formal analysis; Writing - Original Draft; Writing - Review & Editing

Allan B Clark: Data Curation; Writing - Review & Editing

Yoon K Loke: Writing - Review & Editing

Chun Shing Kwok: Data Curation; Writing - Review & Editing

Derek C Angus: Conceptualization; Resources; Writing - Review & Editing

Ane U Echeverria: Resources; Writing - Review & Editing

Pedro P España: Resources; Writing - Review & Editing

Dean T Eurich: Resources; Writing - Review & Editing

David T Huang: Conceptualization; Resources; Writing - Review & Editing

Shin Y Man: Resources, Writing - Review & Editing

Timothy H Rainer: Resources; Writing - Review & Editing

Donald M Yealy: Conceptualization; Resources; Writing - Review & Editing

Phyo K Myint: Conceptualization; Data curation; Writing - Review & Editing; Supervision,

Maria K Mor: Conceptualization; Methodology; Formal Analysis; Writing - Review & Editing

Michael J Fine: Conceptualization; Resources; Writing-original draft; Writing-review and editing; Supervision

The authors wish to remember the work of fellow colleague in the International Community Acquired Pneumonia (CAP) Collaboration Cohort (ICCC), Sumit R. Majumdar, MD MPH, who contributed to the assembly of the ICCC Cohort but passed away prior to the completion of this study. We would also like to thank Dr. Thomas J. Marrie (Dalhousie University) for his expertise in the study design and critical review of the manuscript.

Comparison of the Prognostic Performance of the CURB-65 and a Modified Version of the Pneumonia Severity Index Designed to Identify High-Risk Patients Using the International Community-Acquired Pneumonia Collaboration Cohort

Raphae S Barlas[1], Allan B Clark[2], Yoon K Loke[2], Chun S Kwok[3], Derek C Angus[4], Ane Uranga[5], Pedro P España[5], Dean T Eurich[6], David T Huang[4], Shin Y Man[7], Timothy H Rainer[7], Donald M Yealy[8], Phyo K Myint[1,2], Maria K Mor[10,11], Michael J Fine[9,10]

[1]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

[2]Norwich Medical School, University of East Anglia, Norwich, UK

[3] Keele University, Stoke on Trent, UK

[4]The CRISMA Center, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[5]Servicio de Neumología, Hospital de Galdakao, Galdakao, Bizkaia, Spain

[6]School of Public Health, University of Alberta, Edmonton, Alberta, Canada

[7]Emergency Medicine Unit, Faculty of Medicine, University of Hong Kong, Hong Kong SAR

[8]Department of Emergency Medicine at the University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[9]Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[10]Center for Health Equity Research and Promotion (CHERP), VA Pittsburgh Healthcare System, Pittsburgh, PA, USA

[11]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

Address correspondence to:
Michael J. Fine
CHERP (151C)
Building 30, Suite 2A140
Pittsburgh, PA 15240-1001
Email: Michael.fine@va.gov

Declarations of interest for all authors: None

Word count: 2,894

Abstract

Background: Although the PSI and CURB-65 represent well-validated prediction rules for pneumonia prognosis, PSI was designed to identify patients at low risk and CURB-65 patients at high risk of mortality. We compared the prognostic performance of a modified version of the PSI designed to identify high-risk patients (i.e., PSI-HR) to CURB-65 in predicting short-term mortality.

Methods: Using data from 6 pneumonia cohorts, we designed PSI-HRs a 6-class prediction rule using the original prognostic weights of all PSI variables and modifying the risk score thresholds to define risk classes. We calculated the proportion of low-risk and high-risk patients using CURB-65 and PSI-HR and 30-day mortality in these subgroups. We compared the rules' sensitivity, specificity, positive and negative predictive values for mortality all risk class thresholds and assessed discriminatory power using areas under their receiver operating characteristic curves (AUROCs).

Results: Among 13,874 patients with pneumonia, 1036 (7.5%) died. For PSI-HR versus CURB-65, aggregate mortality was lower in low-risk patients (1.6% vs. 2.2 %, p=0.005) and higher in high-risk patients (36.5% vs. 32.2%, p=0.27). PSI-HR had higher sensitivities than CURB-65 at all thresholds; PSI-HR also had higher specificities at the 3 lowest thresholds and specificities within 0.5 percentage points of CURB-65 at the 2 highest thresholds. The AUROC was larger for PSI-HR than CURB-65 (0.82 vs. 0.77, p<0.0001).

Conclusions: PSI-HR demonstrated superior prognostic accuracy to CURB-65 at the lower end of the severity spectrum and identified high-risk patients with nonsignificant higher short-term mortality at the higher end.

Key Words: Pneumonia, prediction rules, prognosis, severity of illness

Introduction

Accurately assessing illness severity in patients with community-acquired pneumonia (CAP) helps clinicians make important management decisions. The pneumonia severity index (PSI) and CURB-65 are two validated prognostic prediction rules recommended to supplement clinician judgement in such decisions.[1-4] Whereas the PSI was designed to identify low-risk patients who could safely be managed in the outpatient setting,[1] CURB-65 initially sought to identify high-risk patients and was subsequently modified to identify patients in three severity strata who could be managed with increasing acuity levels of medical care.[2, 5]

Prior studies comparing the prognostic performance of PSI and CURB-65 as prediction rules demonstrated that 30-day mortality is lower in the two to three lowest risk classes using PSI and higher in the highest two risk classes using CURB-65.[6, 7, 8] Given the performance trade-offs of prediction rules designed to function optimally at different regions of the severity of illness spectrum, prior individual studies and meta-analyses comparing these rules showed that sensitivity and negative predictive values for mortality are consistently higher for PSI, whereas specificity and positive predictive values are consistently higher for CURB-65.[6-10] Many studies have also demonstrated that the discriminative power for mortality of PSI is larger than CURB-65 across all CAP severity classes.[6,7,9,10]

For CURB-65 and PSI, the observed differences in prognostic performance are largely driven by their variable composition and weighting. Whereas CURB-65 consists of 5 variables with equal one-point prognostic weights summed to define 6 risk classes, the PSI consists of 20 variables with differing prognostic weights summed to define the 4 highest risk classes, with the lowest class defined using a subset of these variables. The simplicity of CURB-65 results in a

finite and fixed maximum number of risk classes, whereas for the PSI, the number of risk classes and/or the risk scores to define them are modifiable depending on the goals of prognostication.

Both the PSI and several CURB-65 score variants have been used as decision aids to guide the initial site of treatment for low-risk patients with CAP. [3, 4, 11] Recent practice guidelines from North America recommend preferential use of the PSI over CURB-65 in guiding the initial site of treatment based on high-quality empirical evidence on the PSI's effectiveness and safety in guiding this decision.[4] In contrast, European guidelines recommend using a simplified version of CURB-65 that omits measurement of urea nitrogen (CRB-65) to identify patients suitable for outpatient treatment, [3, 11] determined in part by the practicality of using a more parsimonious decision aid without the need to obtain laboratory parameters. The prognostic accuracy of CRB-65 can be augmented with the addition of information on medical comorbidities and pulse oximetry, which could improve its performance as a decision aid guiding the initial site of treatment.[12]

The CURB-65 and other prognostic models of severe pneumonia have also been recommended to guide use of higher acuity levels of care for patients hospitalized with CAP.[3, 4, 11] North American guidelines strongly recommend admission to an ICU for patients with hypotension requiring vasopressors or respiratory failure requiring mechanical ventilation (American Thoracic Society [ATS] major severity criteria), and in their absence, conditionally recommend using the presence of three or more ATS minor severity criteria to guide use of higher acuity levels of inpatient care.[4, 13] In contrast, European guidelines recommend using CURB-65 scores of three or more to identify patients who could benefit from management in an ICU.[3, 11] Due to less accurate prediction of mortality at the higher end of the illness severity spectrum for the PSI compared to CURB-65 and the superior performance other models of

severe pneumonia in predicting the need for ICU care, [14, 15] none of these guidelines recommend using the PSI to identify higher risk patients who might benefit from higher acuity levels of inpatient care. [3, 4, 11]

In this study, we redesigned the PSI to better identify high-risk patients. After validating the predictive reliability and discrimination of the new PSI high risk (PSI-HR), we compared its prognostic performance to CURB-65 and the original PSI as three prediction rules for mortality.

Materials and Methods

Study Sites and Patients

In the previously reported International Community-Acquired Pneumonia Collaboration Cohort (ICCC),[16] we aggregated individual-level data for 13,874 patients enrolled in 6 study cohorts from 4 countries. [1,7,17-20] All patients had clinical and radiographic evidence of CAP and were prospectively identified from emergency departments and inpatient and outpatient clinical sites of care. Each study was approved by its local institutional review board, and the parent ICCC was approved by the Faculty of Medicine and Health Sciences Research Ethics Committee at the University of East Anglia, UK.

Assessment of Baseline Patient Characteristics and Mortality

In the ICCC, we assembled baseline data on patient demographic and clinical characteristics, including all prognostic variables comprising the CURB-65 and PSI. Our primary outcome of all-cause mortality 30-days from the initial diagnosis of CAP was available for all patients.

Severity Classification Using PSI, CURB-65, and PSI-HR

We used all prognostic variables comprising PSI and CURB-65 to assign patients to prediction rule specific risk classes **(eFigures 1 and 2)**. For CURB-65, we assigned patients to 6 risk classes based on the presence or absence of the 5 constituent prognostic variables.[2] For PSI, we assigned patients to 5 risk classes based an established two-step algorithm.[1]

To develop PSI-HR, we assigned patients to 6 risk classes based on the sum of the prognostic weights of all PSI variables present **(eFigure 3)**.We used the same number of risk classes as CURB-65 to compare each rule's prognostic performance across an equivalent number of severity thresholds. PSI-HR differs from PSI in the assignment to risk class I, the risk score

cut-points used to define risk classes, and the use of 6 versus 5 severity classes. For PSI, risk class I was defined based on age ≤50 years and absence of comorbid conditions and abnormalities in vital signs and mental status; for PSI-HR, we used a risk score threshold of ≤65. Whereas the highest risk class in the PSI was determined using a total risk score >130, we defined the 2 highest risk classes in PSI-HR (V and VI) using risk scores of 151-175 and >175, respectively. We selected the risk scores to define all severity strata in a 50% random derivation sample of the study population to ensure there were: (1) an adequate number of patients in all risk classes to have stable mortality estimates, and (2) clinically meaningful differences in mortality across all risk classes.

The original CURB-65 defined confusion as an Abbreviated Mental Test score of < 8, or new disorientation to person, place, or time.[2]. Because identical component variables were not uniformly available in the ICCC database, we used altered mental status as a proxy for confusion.[6] For PSI and PSI-HR risk score calculations, a partial pressure of oxygen < 60 mm Hg and an oxygen saturation < 90% were considered equivalent. [6, 20] Missing values for any variable comprising CURB-65, PSI, and/or PSI-HR were assumed to be normal. This strategy for handling missing data was used in the original derivation and validation of these prediction rules for CAP and for similar prediction rules for other common medical conditions. [1, 2, 21].

Statistical Analyses

We compared baseline patient characteristics in the randomly selected derivation (50%) and validation (50%) samples using chi-square statistics. For each prediction rule, we used Fisher's exact tests to compare risk-class specific mortality rates in the derivation and validation samples.

We used Fisher's exact tests to compare the proportion of patients in the overall study sample classified as low-risk using standard definitions for each prediction rule (PSI risk classes I-III, CURB-65 scores 0 and 1, and PSI-HR risk classes I and II) and compared the mortality in the low-risk subgroups. We conducted similar comparisons for patients classified as high risk by PSI (risk classes IV and V), CURB-65 (scores 4 and 5), and PSI-HR (risk classes V and VI). For the PSI-HR, we assessed a more continuous association between mortality and illness severity by calculating this outcome for one-hundredths of the study sample ordered by average total risk scores. We validated the prognostic reliability and discrimination of PSI-HR by comparing risk class specific mortality and the areas beneath the receiver operating characteristic curves (AUROCs) in the derivation and validation samples.

After confirming similar performance of PSI-HR in the derivation and validation samples, we evaluated the accuracy of all prediction rules in predicting mortality in the overall study sample by calculating sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) at all risk class thresholds. We assessed the discriminatory power of all rules by calculating their respective AUROCs with 95% confidence intervals. To calculate AUROCs, we used all risk classes for all rules and a continuous version of the PSI-HR based on total risk score. We performed pairwise comparisons of the AUROCs for the PSI, PSI-HR, and CURB-65, using established methods.[22]

We used a two-tailed p-value of <.05 to define statistical significance and STATA version 14.0/MP (Stata Corp., College Station, TX) to perform all analyses.

Results

Patient Characteristics

Among the 13,874 patients in the study sample, 7,472 (53.9%) were men, 1,216 (8.8%) were nursing home residents, and the mean age was 65.5 (±SD 19.2) years **(Table 1)**. Except or small differences in age and serum glucose, there were no other differences in baseline characteristics for patients in the 50% derivation and validation samples.

Prognostic Reliability and Discrimination of PSI-HR in the Derivation and Validation Samples

Risk class specific mortality for the PSI-HR ranged from 0.4% to 53.4% overalls, with similar ranges in the derivation and validation samples **(Table 2)**. There were no significant differences in PSI-HR risk class specific mortality in the derivation and validation samples **(Table** 2), and both samples' AUROCs were identical (0.82; 95% CI, 0.80-0.83).

Comparisons of Risk Class Distribution and Mortality in the Total Sample

Overall, 1,036 (7.5%) patients died within 30 days, ranging from 4.5% to 11.1% across the 6 study cohorts. PSI-HR classified a larger proportion of patients as low-risk (55.2% versus 52.7%, $p<0.001$) and high-risk (5.6% versus 4.7%, $p<0.001$) than CURB-65 **(Table 2**). Aggregate mortality was lower in low-risk patients identified with PSI-HR than CURB-65 (1.6% versus 2.2 %, $p=0.005$) and was non-significantly higher in high-risk patients identified with PSI-HR than CURB-65 (36.5% versus 32.2%, $p=0.27$). Although patients with the highest CURB-65 score of 5 had a higher mortality than those in the highest PSI-HR risk class VI (59.5% versus 53.4%), only 79 (0.6%) patients were in the highest CURB-65 risk class compared to 206 (1.5%) for PSI-HR. Mortality ranged from 0.0% to 58.7% for patients in the lowest to the highest hundredth of the total study sample ordered by PSI-HR risk score **(Figure**

1). For the 79 patients with the highest PSI-HR risk score (equivalent to the number in the highest CURB-65 risk class), 50 (63.3%) died.

Compared to the PSI, PSI-HR classified an identical proportion of all patients as low risk (55.1%) and there was only one additional death in those identified as low-risk by PSI-HR **(Table 2)**. Although PSI-HR identified a smaller proportion of patients as high risk than PSI (5.5% vs. 44.9%, p<0.0001), patients classified as high-risk using PSI-HR had more than a two-fold higher mortality than those classified using PSI (36.5% vs. 14.7%, p<0.0001). Likewise, mortality was more than 25 percentage points higher for patients in the highest PSI-HR versus PSI risk class (53.4% vs. 28.0%, p<0.0001).

Comparisons of Accuracy and Discriminatory Power in the Total Sample

PSI-HR had a higher sensitivity than CURB-65 at all risk class thresholds and a higher specificity than CURB-65 for the 3 lowest thresholds **(Table 3**). Although CURB-65 had a higher specificity than PSI-HR for the 2 highest thresholds, the differences were <0.5 percentage points for each of these comparisons (96.6% versus 96.2% and 99.8% versus 99.3%, respectively). Consistent with these findings, the negative predictive values for mortality were higher for PSI-HR than CURB-65 at all thresholds, and the positive predictive values were higher for PSI-HR than CURB-65 at the three lowest thresholds. At the thresholds used to define low-risk patients, both sensitivity (88.3% vs. 84.3%) and specificity (58.7% vs. 55.7%) were higher for PSI-HR than CURB-65.

Based on risk class and risk score, PSI-HR had greater (p<.0001) discriminatory power than CURB-65 **(Figure 2)**. The AUROCs were 0.82 (95% CI, 0.80-0.83) for PSI-HR (risk class), 0.83 (95% CI, 0.82-0.84) for PSI-HR (continuous), and 0.77 (95% CI, 0.76-0.78) for CURB-65 (risk class).

At every risk class threshold, the specificity was higher for PSI-HR than PSI and the sensitivity was higher for PSI than PSI-HR. All negative predictive values were higher for PSI than PSI-HR and all positive predictive values were higher for PSI than PSI-HR. The AUROC for the PSI-HR based on six risk classes was larger ($p<0.0001$) than the AUROC for PSI based on 5 risk classes [0.82 (95% CI, 0.80-0.83) versus 0.80 (95% CI, 0.79-0.82)].

Discussion

Using individual-level data from over 13,000 patients prospectively enrolled in six pneumonia cohorts,[16] we developed the PSI-HR to better identify high risk patients and validated its prognostic reliability and discrimination as a prediction rule for short-term mortality. At all risk class thresholds, PSI-HR had higher sensitivities and negative predictive values than CURB-65. PSI-HR also had higher specificities than CURB-65 for the 3 lowest thresholds and specificities minimally lower than CURB-65 for the 2 highest thresholds. Finally, PSI-HR had greater overall discriminatory power than CURB-65 in predicting mortality.

Similar to prior comparisons of the PSI and CURB-65 for CAP, [6,7] PSI-HR classified larger proportions of patients as low-risk for mortality and these low-risk patients had a lower cumulative mortality than those classified by CURB-65. Although we demonstrated that PSI-HR also classified larger proportions of patients as high-risk than CURB-65, our comparisons had limited power to detect significant differences in mortality in the more sparsely populated highest risk strata. Nevertheless, patients in the two highest PSI-HR risk classes had non-significantly higher mortalities than those in the two highest CURB-65 risk classes. The superior prognostic performance of the PSI-HR (versus CURB-65) at both ends of the severity spectrum is likely explained by the clinical richness conferred by using a 4-fold larger number of predictor variables with empirically derived prognostic weights.

The observed differences in the prognostic accuracy of the PSI-HR and original PSI likely stem from the explicit goals of developing the rules. Whereas PSI had higher sensitivities and negative predictive values than PSI-HR at all risk class thresholds, PSI-HR had correspondingly higher specificities and positive predictive values than PSI. The PSI-HR and PSI identified an equivalent number of low-risk patients, but for all three rules, mortality was

lowest for PSI class I, and for PSI class II it was more than two percentage points lower than the corresponding risk class for PSI-HR or CURB-65. In contrast, mortality in the two highest PSI-HR risk classes was substantially higher than in the single highest PSI risk class. Thus, the original PSI provides more accurate, finer severity stratification at the low end of the severity of illness spectrum, and PSI-HR outperforms PSI at the high end of the spectrum.

Although our study compared the performance of three prediction rules for mortality following CAP and not their use as decision aids, our findings have implications for clinical decision-making based on risk stratification and future research on their clinical utility for patients with this condition. The superior performance of the PSI compared to PSI-HR in identifying low-risk patients suggests that the original PSI remains the preferred version of this prediction rule to guide the initial site of treatment in CAP. In addition, the enhanced performance of the PSI-HR in identifying patients at higher risk of death compared to PSI and CURB-65 suggests future studies are needed to compare its prognostic performance to existing prognostic prediction rules for severe pneumonia and its use in guiding the intensity of inpatient management (e.g., ICU admission) for CAP. [23, 24, 25]

Our study has limitations. First, there were missing data for some of the vital signs and laboratory variables comprising the prediction rules. We defined these missing values as normal, the same strategy used in the development of the PSI and CURB scores and for prediction rules for other medical conditions. [1, 2, 21] Second, due to the pooled nature of our data, altered mental status was defined differently than in the original development of CURB models. Instead, we used the available data to operationalize an alternative definition with clinical face validity. Third, we did not have all data required to compare the PSI-HR to other models of severe pneumonia, such as the ATS major and minor severity criteria among others. [13, 23-25] Fourth,

the data used in this study originated up to 25 years ago; however, the age of the data are unlikely to bias our comparisons of prognostic prediction rules in patients with CAP.

In conclusion, we demonstrated the prognostic reliability and discriminatory power of the PSI-HR in predicting short-term mortality following CAP. Compared to CURB-65, PSI-HR had greater overall discriminatory power and superior prognostic performance in low-risk and high-risk patients. By design, the original PSI out-performed PSI-HR at the lower end of the severity spectrum, and PSI-HR outperformed PSI at the higher end of the spectrum. Our findings underscore the need to compare the performance of PSI-HR to previously developed prediction rules for prognosis in severe pneumonia and assess its effectiveness and safety as a decision aid to guide the intensity of care among patients hospitalized with CAP.

References

1. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997;336(4):243-250. doi:10.1056/NEJM199701233360402.

2. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58(5):377-382. doi:10.1136/thorax.58.5.377.

3. Lim WS, Baudouin SV, George RC, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax. 2009;64 Suppl 3:iii1-iii55. doi:10.1136/thx.2009.121434.

4. Metlay JP, Waterer GW, Long AC, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med. 2019;200(7):e45-e67. doi:10.1164/rccm.201908-1581ST.

5. Neill AM, Martin IR, Weir R, et al. Community acquired pneumonia: aetiology and usefulness of severity criteria on admission. Thorax. 1996;51(10):1010-1016. doi:10.1136/thx.51.10.1010.

6. Aujesky D, Auble TE, Yealy DM, et al. Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia. Am J Med. 2005;118(4):384-392. doi:10.1016/j.amjmed.2005.01.006.

7. Man SY, Lee N, Ip M, et al. Prospective comparison of three predictive rules for assessing severity of community-acquired pneumonia in Hong Kong. Thorax. 2007;62(4):348-353. doi:10.1136/thx.2006.069740.

8. Ananda-Rajah MR, Charles PG, Melvani S, Burrell LL, Johnson PD, Grayson ML. Comparing the pneumonia severity index with CURB-65 in patients admitted with community acquired pneumonia. Scand J Infect Dis. 2008;40(4):293-300. doi:10.1080/00365540701663381.

9. Chalmers JD, Singanayagam A, Akram AR, et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. Thorax. 2010;65(10):878-883. doi:10.1136/thx.2009.133280.

10. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. Thorax. 2010;65(10):884-890. doi:10.1136/thx.2009.134072.

11. Lim WS, Smith DL, Wise MP, Welham SA. British Thoracic Society community acquired pneumonia guideline and the NICE pneumonia guideline: how they fit together. BMJ Open Respir Res. 2015;2(1):e000091. Published 2015 May 13. doi:10.1136/bmjresp-2015-000091.

12. Kolditz M, Ewig S, Schütte H, et al. Assessment of oxygenation and comorbidities improves outcome prediction in patients with community-acquired pneumonia with a low CRB-65 score. J Intern Med. 2015;278(2):193-202. doi:10.1111/joim.12349.

13. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis. 2007;44 Suppl 2(Suppl 2):S27-S72. doi:10.1086/511159.

14. Marti C, Garin N, Grosgurin O, et al. Prediction of severe community-acquired pneumonia: a systematic review and meta-analysis. Crit Care. 2012;16(4):R141. Published 2012 Jul 27. doi:10.1186/cc11447.

15. Chalmers JD, Mandal P, Singanayagam A, et al. Severity assessment tools to guide ICU admission in community-acquired pneumonia: systematic review and meta-analysis. Intensive Care Med. 2011;37(9):1409-1420. doi:10.1007/s00134-011-2261-x.

16. Myint PK, Kwok CS, Majumdar SR, et al. The International Community-Acquired Pneumonia (CAP) Collaboration Cohort (ICCC) study: rationale, design and description of study cohorts and patients. BMJ Open. 2012;2(3):e001030. Published 2012 May 21. doi:10.1136/bmjopen-2012-001030.

17. Capelastegui A, España PP, Quintana JM, et al. Validation of a predictive rule for the management of community-acquired pneumonia. Eur Respir J. 2006;27(1):151-157. doi:10.1183/09031936.06.00062505.

18. Kellum JA, Kong L, Fink MP, et al. Understanding the inflammatory cytokine response in pneumonia and sepsis: results of the Genetic and Inflammatory Markers of Sepsis (GenIMS) Study. Arch Intern Med. 2007;167(15):1655-1663. doi:10.1001/archinte.167.15.1655.

19. Majumdar SR, McAlister FA, Eurich DT, Padwal RS, Marrie TJ. Statins and outcomes in patients admitted to hospital with community acquired pneumonia: population based prospective cohort study. BMJ. 2006;333(7576):999. doi:10.1136/bmj.38992.565972.7C.

20. Yealy DM, Auble TE, Stone RA, et al. Effect of increasing the intensity of implementing pneumonia guidelines: a randomized, controlled trial. Ann Intern Med. 2005;143(12):881-894. doi:10.7326/0003-4819-143-12-200512200-00006.

21. Aujesky D, Obrosky DS, Stone RA, et al. Derivation and validation of a prognostic model for pulmonary embolism. Am J Respir Crit Care Med. 2005;172(8):1041-1046. doi:10.1164/rccm.200506-862OC.

22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-845.

23. Charles PG, Wolfe R, Whitby M, et al. SMART-COP: a tool for predicting the need for intensive respiratory or vasopressor support in community-acquired pneumonia. Clin Infect Dis. 2008;47(3):375-384. doi:10.1086/589754

24. España PP, Capelastegui A, Gorordo I, et al. Development and validation of a clinical prediction rule for severe community-acquired pneumonia. Am J Respir Crit Care Med. 2006;174(11):1249-1256. doi:10.1164/rccm.200602-177OC.

25. Renaud B, Labarère J, Coma E, et al. Risk stratification of early admission to the intensive care unit of patients with no major criteria of severe community-acquired pneumonia: development of an international prediction rule. Crit Care. 2009;13(2):R54. doi:10.1186/cc7781.

Acknowledgments Author contributions: MJF and RSB are the guarantors of the content of the manuscript, including the data and analysis. RSB and MJF were responsible for writing and revision of the manuscript. All authors contributed to the conceptualization of the study design, had access to the study data and review and synthesis of the final version of this manuscript. The authors wish to especially thank Dr. Thomas J. Marrie (Dalhousie University) for his expertise and critical review of the manuscript.

Conflict of interest statement: None of the authors have any conflicts of interest to declare.

Funding information: The authors received no financial support for the research, authorship, and/or publication of this article.

Additional information: e-Figures and e-Tables can be found in the Supplemental Materials section of the online article

Table 1: Baseline Patient Characteristics in the Derivation, Validation, and Total Study Samples

Characteristics	Total Sample (N=13,874)		Derivation Sample (N=6,937)		Validation Sample (N=6,937)		P-value†
Demographics	**n**	**%***	**n**	**%**	**n**	**%**	
Age in years							0.03
< 50	3,444	24.8	1,697	24.5	1,747	25.2	
51-64	2,217	16.0	1,073	15.5	1,144	16.5	
65-80	5,137	37.0	2,652	38.2	2,485	35.8	
> 80	3,076	22.2	1,515	21.8	1,561	22.5	
Sex (male)	7,472	53.9	3,709	53.5	3,763	54.3	0.36
Nursing home resident	1,216	8.8	618	8.9	598	8.6	0.55
Comorbid conditions							
Heart failure	1,985	14.3	1,023	14.8	962	13.9	0.14
Malignancy	912	6.6	448	6.5	464	6.7	0.59
Cerebrovascular disease	1,276	9.3	667	9.7	609	8.8	0.09
Renal disease	1,106	8.0	558	8.0	548	7.9	0.75
Liver disease	305	2.2	149	2.2	156	2.3	0.69
Physical examination findings							
Temperature <35º or >40º C	252	1.9	131	2.0	121	1.8	0.52
Heart rate ≥ 125/minute	1,353	11.6	694	11.8	659	11.3	0.31
Systolic blood pressure < 90 mm Hg	377	2.8	178	2.7	199	3.0	0.27
Diastolic blood pressure < 60 mm Hg	3,069	22.9	1,510	22.6	1,559	23.3	0.31
Respiratory rate ≥30/minute	1,743	15.3	886	15.5	857	15.1	0.45
Confusion	1,478	10.7	741	10.7	737	10.6	0.92
Laboratory and radiographic findings							
Blood urea nitrogen > 30 mg/dl	2,621	22.2	1,323	224	1,298	22.1	0.58
Glucose ≥ 250 mg/dl	686	6.8	378	7.5	308	6.1	0.01
Haematocrit < 30%	681	7.4	339	7.4	342	7.5	0.90
Sodium < 130 mmol/l	644	6.2	327	6.2	317	6.1	0.68

SO_2 <90% or PaO2 < 60 mm Hg	3,780	32.6	1,887	32.5	1893	32.6	0.97
Arterial pH < 7.35	855	9.9	436	10.1	419	9.8	0.58
Pleural effusion	1,942	14.3	971	14.3	971	14.3	1.00

Abbreviations: PaO2 denotes arterial partial pressure of oxygen, and SO2 denotes oxygen saturation.

*In the overall study population, there were no missing data for the 3 demographic characteristics and 4 of the comorbid conditions (heart failure, malignancy, renal and liver disease); cerebrovascular disease was missing in less than 1%. Data were missing for more than 1% overall for temperature (3.4%), heart rate (15.7%), systolic blood pressure (3.2%), diastolic blood pressure (3.4%), respiratory rate (17.8%), blood urea nitrogen (15.0%), glucose (27.3%), haematocrit (33.8%), sodium (24.9%), SO_2 or PaO_2 (16.6%), arterial pH (38.0%), and pleural effusion (1.8%). To calculate the frequency of each characteristic, we removed missing data from the denominator.

†P-values compare the proportions with baseline characteristics in the validation and derivation cohorts and were calculated using chi-square statistics.

Table 2: Risk Class Distribution and Mortality by Prediction Rule in the Derivation, Validation, and Total Study Samples

	Total Sample (N=13,874)				Derivation Sample (N=6,937)				Validation Sample (N=6,937)				
Risk Classes by Prediction Rule	**Distribution**		**Mortality**		**Distribution**		**Mortality**		**Distribution**		**Mortality**		**P-value***
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**%**
PSI													
I	1,696	12.2	3	0.2	853	12.3	1	0.1	843	12.2	2	0.2	0.62
II	3,093	22.3	23	0.7	1,504	21.7	14	0.9	1,589	22.9	9	0.6	0.30
III	2,862	20.6	95	3.3	1,444	20.8	49	3.4	1,418	20.4	46	3.2	0.84
IV	4,424	31.9	412	9.3	2,224	32.1	225	10.1	2,200	31.7	187	8.5	0.09
V	1,799	13.0	503	28.0	912	13.1	261	28.6	887	12.8	242	27.3	0.65
PSI-HR													
I	4,134	29.8	16	0.4	2,030	29.3	9	0.4	2,104	30.3	7	0.3	0.62
II	3,517	25.4	105	3.0	1,771	25.5	55	3.1	1,746	25.3	50	2.9	0.69
III	3,632	26.2	292	8.0	1,816	26.3	155	8.5	1,816	26.2	137	7.5	0.33
IV	1,821	13.1	342	18.8	947	13.7	197	20.8	874	12.6	145	16.6	0.06
V	564	4.1	171	30.3	280	4.0	80	28.6	284	4.1	91	32.0	0.54
VI	206	1.5	110	53.4	93	1.3	54	58.1	113	1.6	56	49.6	0.55
CURB-65													
0	3,419	24.6	28	0.8	1,658	23.9	16	1.0	1,761	25.4	12	0.7	0.44
1	3,897	28.1	135	3.5	1,960	28.3	73	3.7	1,937	28.0	62	3.2	0.43
2	3,790	27.3	314	8.3	1,910	27.5	164	8.6	1,880	27.1	150	8.0	0.55
3	2,124	15.3	351	16.5	1,080	15.6	190	17.6	1,044	15.0	161	15.4	0.27
4	565	4.1	161	28.5	292	4.2	86	29.5	273	3.9	75	27.5	0.72
5	79	0.6	47	59.5	37	0.5	21	56.8	42	0.6	26	61.9	0.85

Abbreviations: PSI denotes Pneumonia Severity Index; PSI-HR denotes Pneumonia Severity Index High-Risk; and CURB-65 denotes Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age 65.

*P-values compare the proportion of patients who died stratified by prediction rule risk classes in the derivation and validation study samples and were calculated using Fisher's exact tests.

Table 3: Prognostic Accuracy of the Prediction Rules for 30-Day Mortality in the Total Study Sample

Risk Classes by Prediction Rule	Sensitivity		Specificity		Predictive Value (+)		Predictive Value (-)	
PSI	**%**	**95% CI**	**%**	**95% CI**	**%**	**95% CI**	**%**	**95% CI**
I	—		—		—		—	
II	99.7	99.2 – 99.9	13.2	12.6 – 13.8	8.5	8.0 – 9.0	99.8	99.5 – 100.0
III	97.5	96.3 – 98.4	37.1	36.3 – 37.9	11.1	10.5 – 11.8	99.5	99.2 – 99.6
IV	88.3	86.2 – 90.2	58.7	57.8 – 59.5	14.7	13.8 – 15.6	98.4	98.1 – 98.7
V	48.6	45.5 – 51.6	89.9	89.4 – 90.4	28.0	25.9 – 30.1	95.6	95.2 – 95.9
PSI-HR								
	—		—		—		—	
II	98.5	97.5 – 99.1	32.1	31.3 – 32.9	10.5	9.9 – 11.1	99.6	99.4 – 99.8
III	88.3	86.2 – 90.2	58.7	57.8 – 59.5	14.7	13.8 – 15.6	98.4	98.1 – 98.7
IV	60.1	57.1 – 63.1	84.7	84.0 – 85.3	24.0	22.4 – 25.0	96.3	96.0 – 96.7
V	27.1	24.4 – 29.9	96.2	95.8 – 96.5	36.5	33.1 – 40.0	94.2	93.8 – 94.6
VI	10.6	8.8 – 12.7	99.3	99.1 – 99.4	53.4	46.3 – 60.4	93.2	92.8 – 93.6
CURB-65								
0	—				—			
1	97.3	96.1 – 98.2	26.4	25.7 – 27.2	9.6	9.1 – 10.2	99.2	98.8 – 99.5
2	84.3	81.9 – 86.4	55.7	54.9 – 56.6	13.3	12.5 – 14.2	97.8	97.4 – 98.1
3	54.0	50.9 – 57.0	82.8	82.1 – 83.4	20.2	18.7 – 21.7	95.7	95.3 – 96.1
4	20.1	17.7 – 22.6	96.6	96.3 – 96.9	32.3	28.7 – 36.1	93.7	93.3 – 94.1
5	4.5	3.4 – 6.0	99.8	99.6 – 99.8	59.5	47.9 – 70.4	92.8	92.4 – 93.3

Abbreviations: PSI denotes Pneumonia Severity Index, PSI-HR denotes Pneumonia Severity Index High Risk; CURB-65 denotes confusion, urea nitrogen, respiratory rate, blood pressure, and age 65; PPV denotes positive predictive value; and NPV denotes negative predictive value.

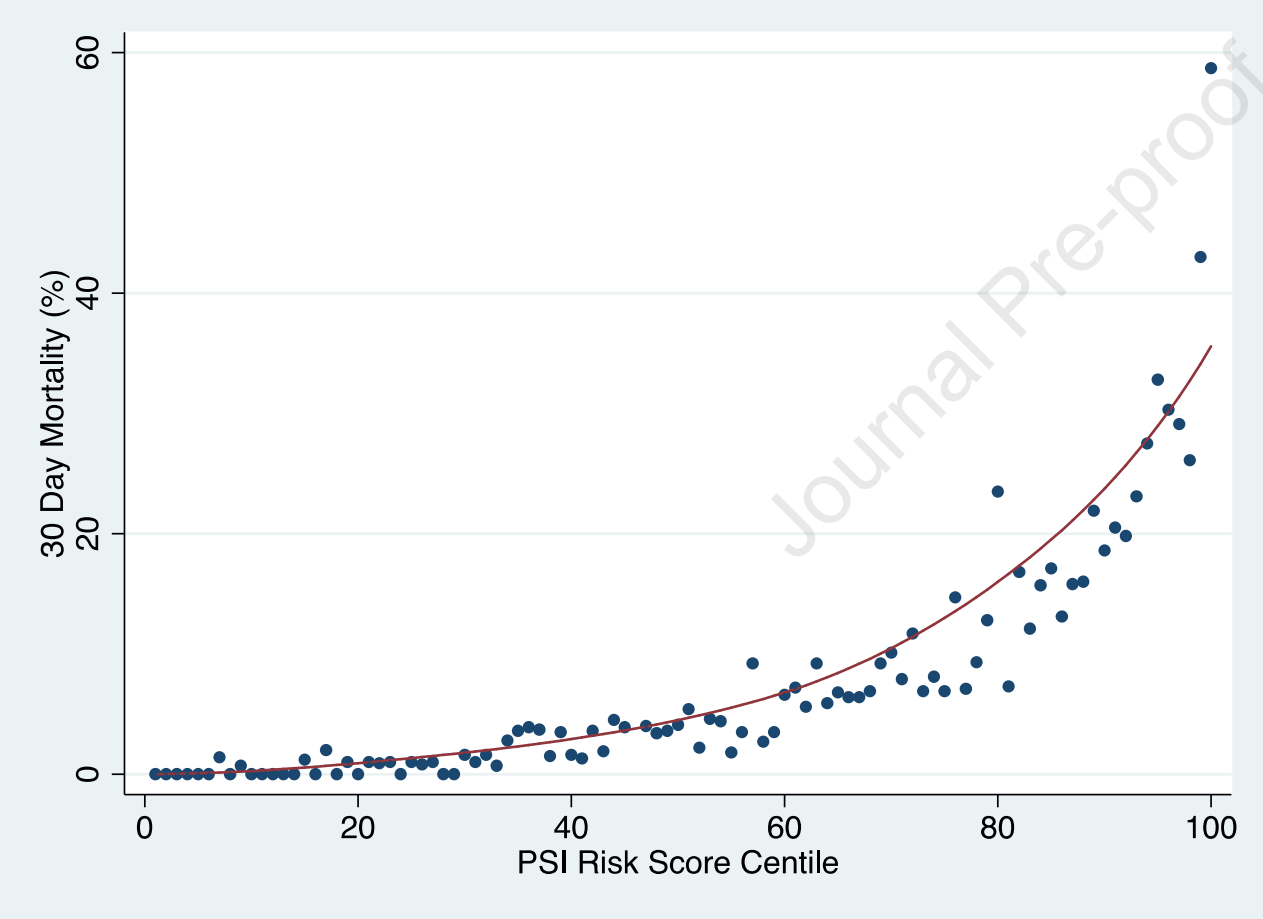

Figure 1: Scatterplot of 30-day Mortality (%) by PSI-HR Risk Score Centile in the Total Study Sample. Mortality ranged from 0% for the six lowest PSI risk score centiles to 58.7% in the single highest centile.

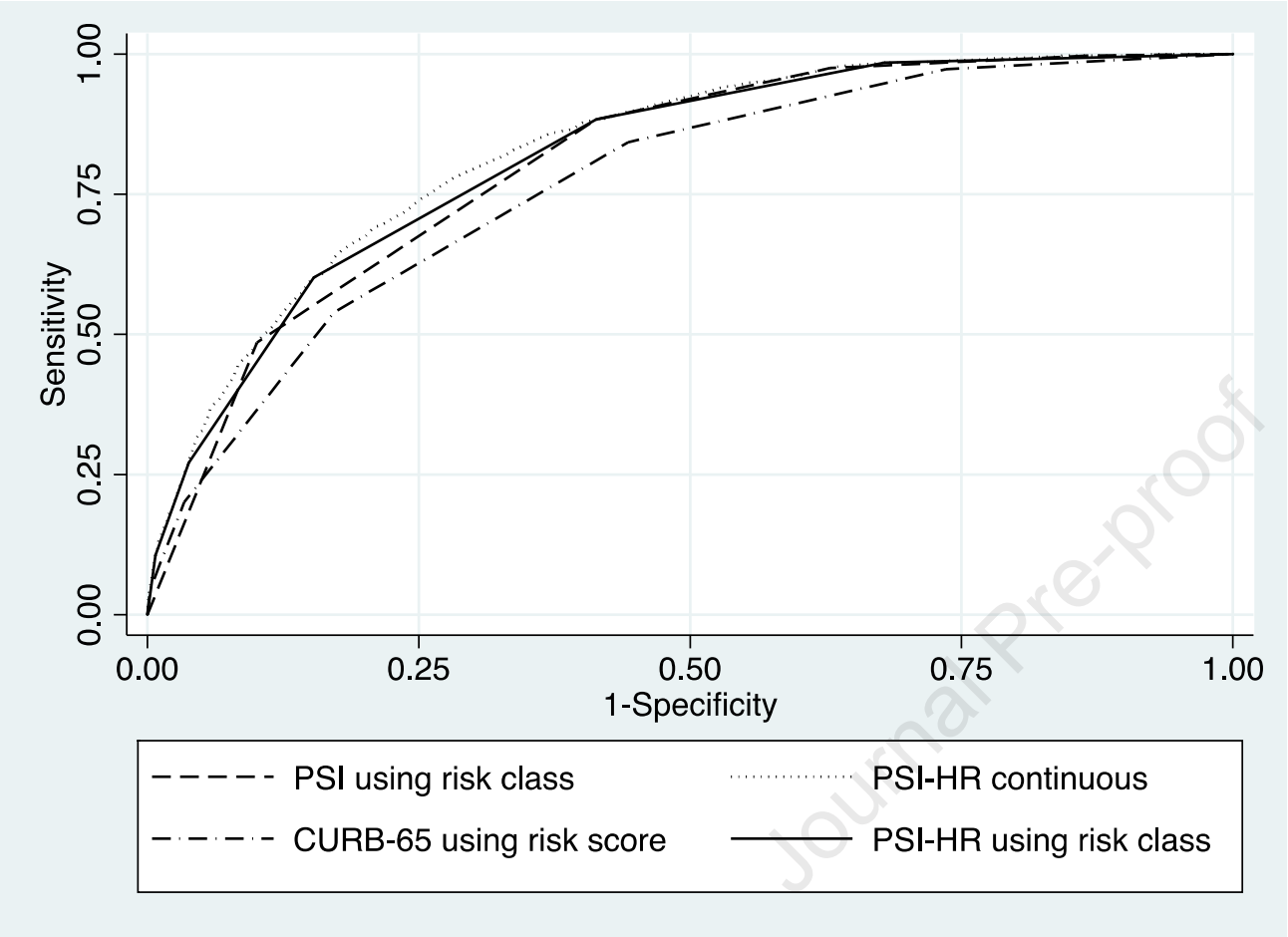

Figure 2: Receiver Operating Characteristic Curves for 30-Day Mortality for PSI, PSI-HR, and CURB-65 in the Total Study Sample. The areas under the receiver operating curves (AUROCs) were 0.80 (95% confidence interval [CI]: 0.79-0.82) for the PSI using risk class, 0.82 (95% CI: 0.80-0.83) for PSI-HR using risk class, 0.83 (95% CI: 0.82-0.84) for PSI-HR using risk score, and 0.77 (95% CI: 0.76-0.78) for CURB-65 using risk class. All pairwise comparisons of the AUROCs for the three PSI-based rules and CURB-65 were statistically significant ($p < 0.0001$).

Clinical Significance

- PSI-HR identified a larger number of low-risk patients with a lower mortality than CURB-6, and PSI-HR identified more high-risk patients with a non-significantly higher mortality than CURB-65.
- Prognostic performance was higher for PSI-HR than CURB-65 in the 3 lowest risk classes and was similar for both prediction rules in the 2 highest risk classes.
- PSI-HR had higher overall discriminatory power in predicting mortality than CURB-65.
- The prognostic performance of PSI-HR was superior to the PSI for high-risk patients and was marginally inferior to the PSI for low-risk patients.