

# Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis

Yoon K Loke,<sup>1,2</sup> Chun Shing Kwok,<sup>1</sup> Alagaratnam Niruban,<sup>2</sup> Phyo K Myint<sup>1,2</sup>

► An additional table is published online only. To view this file please visit the journal online (<http://thorax.bmj.com>).

<sup>1</sup>School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, Norfolk, UK

<sup>2</sup>Directorate of Medicine, Norfolk and Norwich University Hospital, Norwich, Norfolk, UK

## Correspondence to

Dr Yoon K Loke, School of Medicine, Health Policy and Practice, Chancellors Drive, University of East Anglia, Norwich, Norfolk NR4 7TJ, UK; [y.loke@uea.ac.uk](mailto:y.loke@uea.ac.uk)

Received 23 December 2009

Accepted 25 June 2010

Published Online First

20 August 2010

## ABSTRACT

**Background** Several scoring systems have been used to predict mortality in patients with community-acquired pneumonia. The properties of commonly used risk stratification scales were systematically reviewed.

**Methods** MEDLINE and EMBASE (January 1999–October 2009) were searched for prospective studies that reported mortality at 4–8 weeks in patients with radiographically-confirmed community-acquired pneumonia. The search focused on the Pneumonia Severity Index (PSI) and the three main iterations of the CURB (confusion, urea nitrogen, respiratory rate, blood pressure) scale (CURB-65, CURB, CRB-65), and test performance was evaluated based on 'higher risk' categories as follows: PSI class IV/V, CURB-65 (score  $\geq 3$ ), CURB (score  $\geq 2$ ) and CRB-65 (score  $\geq 2$ ). Random effects meta-analysis was used to generate summary statistics of test performance and receiver operating characteristic curves were used for predicting mortality.

**Results** 402 articles were screened and 23 studies involving 22 753 participants (average mortality 7.4%) were retrieved. The respective diagnostic odds ratios for mortality were 10.77 (PSI), 6.40 (CURB-65), 5.97 (CRB-65) and 5.75 (CURB). Overall, PSI had the highest sensitivity and lowest specificity for mortality, CRB-65 was the most specific (but least sensitive) test and CURB-65/CURB were between the two. Negative predictive values for mortality were similar among the tests, ranging from 0.94 (CRB-65) to 0.98 (PSI), whereas positive predictive values ranged from 0.14 (PSI) to 0.28 (CRB-65).

**Conclusions** The current risk stratification scales (PSI, CURB-65, CRB-65 and CURB) have different strengths and weaknesses. All four scales had good negative predictive values for mortality in populations with a low prevalence of death but were less useful with regard to positive predictive values.

## BACKGROUND

Community-acquired pneumonia (CAP) is a common cause of hospital admission and a leading cause of death in the UK.<sup>1</sup> The overall mortality of CAP varies, with higher rates of death seen in those judged to have severe pneumonia.<sup>2–5</sup> A reliable method of assessing the severity of pneumonia may potentially improve the triage or initial management of patients by helping clinicians determine whether close monitoring and aggressive treatment is more appropriate than conservative management. However, accurately assessing the severity of pneumonia can be challenging, so a number of scales based on prognostic factors have been designed to identify patients at

high risk of death, as well as those at low risk who may require less of a watchful eye.

The most notable scales are the CURB-65 (confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years) which is recommended by the British Thoracic Society<sup>5</sup> and the Pneumonia Severity Index (PSI) which originates from the USA.<sup>6</sup> Different iterations of the CURB-65 are available, including the CRB-65 which may be particularly suited for community use as it relies on clinical history and examination without requiring blood urea measurements,<sup>7</sup> and the CURB score which excludes age as part of its criteria, thus reducing emphasis on chronological age as a prognostic factor.<sup>8</sup>

In this review we aimed to systematically retrieve and appraise the available data on the PSI and the three main iterations of the CURB scoring system (CURB, CRB-65, CURB-65) in order to determine the ability of each test to correctly predict mortality in patients with pneumonia. Although we are aware of a multitude of individual studies in different settings, the comparative performance of these scales has yet to be summarised in a meta-analysis.

## METHODS

### Search strategy

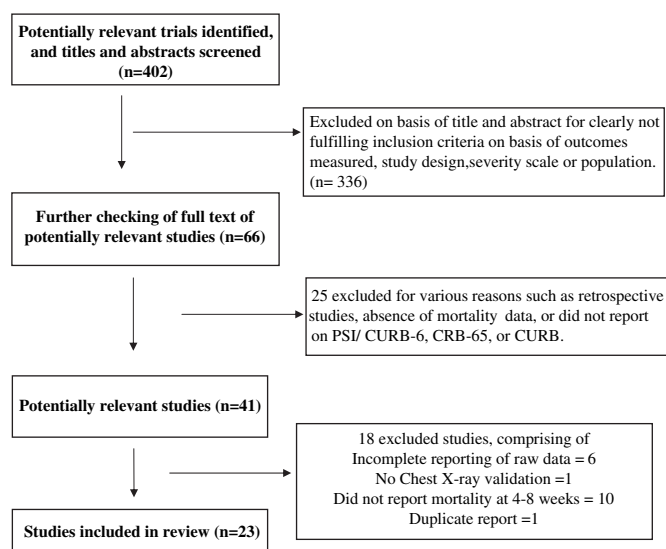
Ovid SP was used to search PubMed and EMBASE from 1999 up to October 2009 using the terms (Pneumonia and (Sever\* or Predict\* or prognos\*) and (scale or score or assessment or index) and (mortality or survival or death) and (community-acquired)).mp. The bibliographies of included studies were checked for any other relevant articles and the authors were contacted for further information where necessary.

### Eligibility criteria

Two reviewers initially checked the title and abstracts against the following inclusion and exclusion criteria.

### Inclusion criteria

- English language journal publication
- At least 100 participants
- Prospective studies
- Reporting patient outcomes after the use of a pneumonia severity scale
- Community-acquired pneumonia, which may include residents of nursing homes
- Randomised controlled trials where any single intervention arm fulfilled inclusion criteria; we did not consider trials comparing different antibiotics.



**Figure 1** Flowchart showing study selection for pneumonia severity scales. CURB, confusion, urea nitrogen, respiratory rate, blood pressure; PSI, Pneumonia Severity Index.

#### Exclusion criteria

- ▶ Less than 100 participants in the study
- ▶ Based in the community (without radiological and laboratory tests), or if recruitment was restricted to patients on the intensive care unit without considering other hospitalised patients. However, we accepted studies that reported fully on all hospitalised patients (in intensive care as well as general wards).
- ▶ Hospital-acquired pneumonia

- ▶ Studies that looked at outcomes of specific types of pneumonia such as legionella or viral rather than community-acquired pneumonia in general
- ▶ Studies that relied entirely on specific biomarkers without using a clinical risk stratification scale
- ▶ Retrospective studies.

From the above we then obtained full-text versions of potentially relevant articles and carried out a more detailed screening process. In conjunction with the above criteria, studies were included only if:

- ▶ Pneumonia was defined by signs, symptoms and chest x-ray
- ▶ Follow-up mortality data were available at 4–8 weeks after presentation
- ▶ Reporting of mortality according to parameters of severity scale (PSI, CURB-65, CRB-65, CURB)
- ▶ Enrolment of patients in the 10-year period spanning 1999–2009 (the CURB-65 and its different iterations only became available after 1999 and we wanted to ensure that comparisons of the PSI and CURB-65 scales were based on patients recruited within the same time period; this was aimed at reducing the possibility of confounding where any differences in the performance of the severity scales might have stemmed from time-related changes in the epidemiology, microbiology or treatment of pneumonia).

#### Validity assessment

Study validity was assessed using a checklist based on published opinion regarding the key components of prognostic studies.<sup>9</sup> The list was based on clear reporting of the following items:

- ▶ Patient selection criteria
- ▶ Diagnostic criteria for pneumonia
- ▶ Loss to follow-up

**Table 1** Study characteristics

Study ID	Country	Setting	Severity scales assessed	Time of outcome ascertainment	No of patients	Age (years)	% Male
Aujesky <sup>14</sup>	USA	32 hospital emergency departments	PSI, CURB, CURB-65	30 days	3181	63 (mean)	48
Barlow <sup>15</sup>	UK	University teaching hospital	CURB-65, CRB-65	30 days	419	74 (median)	47
Bauer <sup>16</sup>	Germany	10 hospitals and outpatient clinics	CRB-65, CURB	14, 30 and 180 days	2184	62.6 (mean)	56.6
Capelastegui <sup>17</sup>	Spain	Teaching hospital	CURB-65, CRB-65	30 days	1776	62 (mean)	63
Charles <sup>18</sup>	Australian	6 university teaching hospitals	PSI, CURB-65	30 days	882	76% >50 years	61
Etzion <sup>19</sup>	Israel	University hospital-based medical centre	PSI	30 days	591	63.5 (mean)	44.5
Ewig <sup>8</sup>	Spain	University hospital	PSI, CURB	30 days	592 for CURB, 489 for PSI	67.8 (mean)	66
Gutierrez <sup>20</sup>	Spain	University teaching hospital	PSI	30 days	493	56.6 (mean)	62.5
Huang <sup>21</sup>	USA	Community, teaching hospital and ED	PSI, CURB-65	30 days	1651	65 (mean)	52
Johnstone <sup>22</sup>	Canada	6 hospitals	PSI	30 days	3284	69 (mean)	53
Kruger <sup>23</sup>	Germany	10 clinical centres	CRB-65	14, 30 & 180 days	1671	61 (mean)	55
Lim <sup>5</sup>	UK, New Zealand, Netherlands	University teaching hospital	PSI, CURB, CURB-65, CRB-65	30 days	1068	64 (mean)	51
Man <sup>24</sup>	Hong Kong	Emergency department	PSI, CURB-65, CRB-65	30 days	1016	72 (mean)	57
Menendez <sup>25</sup>	Spain	University teaching hospital	PSI, CURB-65, CRB-65	30 days	453	67 (mean)	62
Myint <sup>4</sup>	UK	University teaching hospital	CURB-65, CRB-65, CURB	42 days	195	75 (median)	57
Myint <sup>26</sup>	UK	3 UK hospitals	CURB-65	42 days	190	76 (median)	53
Pilotto <sup>27</sup>	Italy	Hospital	PSI	30 days	134	78.7 (mean)	66
Renaud <sup>28</sup>	France	16 emergency departments	PSI	28 days	925	71 (median)	64
Renaud <sup>29</sup>	France, Spain, USA	University teaching hospital	PSI	30 days	853	65% >51 years	56
Reyes Calzada <sup>30</sup>	Spain	4 public hospitals	PSI	30 days	425	69 (mean)	64.5
Schuetz <sup>31</sup>	Switzerland	University teaching hospital	PSI, CURB-65, CRB-65	30 days	373	73 (median)	60
van der Eerden <sup>32</sup>	Netherlands	Teaching hospital	PSI	30 days	260	64 (mean)	53.8
Zuberi <sup>33</sup>	Pakistan	University teaching hospital	CURB-65, CRB-65	30 days	137	60.4 (mean)	74

CRB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years; PSI, Pneumonia Severity Index.

- Methods used to ascertain outcome
- Management protocol used in treating pneumonia (eg, antibiotic regimen)

### Data abstraction

Two reviewers assessed the eligibility and extracted numerical outcomes data from the included studies. The reviewers obtained full consensus on inclusion of the studies and data extraction after resolving any discrepancies through discussion with team members. Authors were contacted if any items required clarification.

### Study characteristics

Geographical location and setting, sample size, age and gender of participants, type of prognostic scale and mortality rate according to classification of severity were recorded.

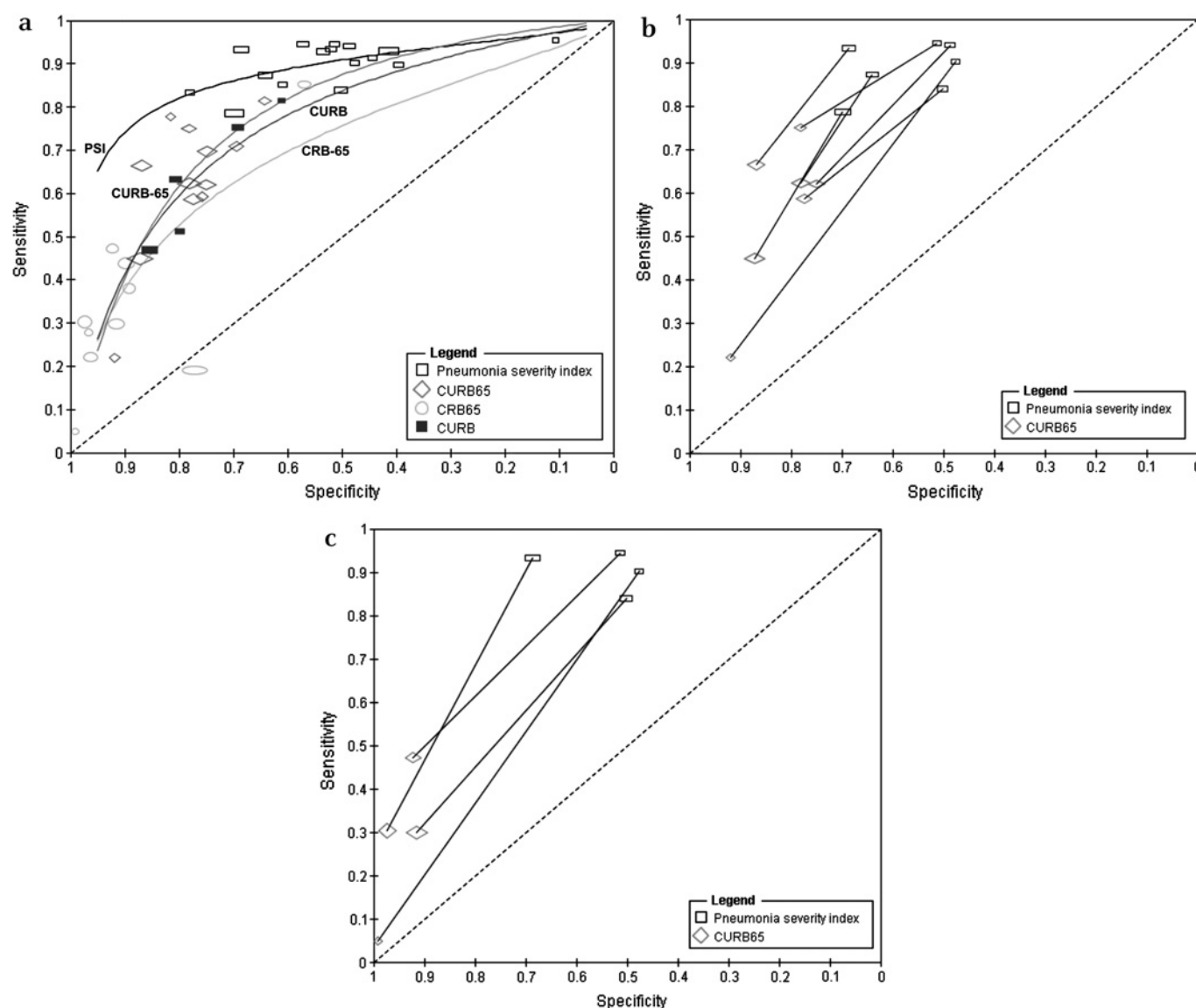
### Quantitative data synthesis and sensitivity analysis

Patients were classified according to risk score and data on mortality were extracted. For each study the numbers of patients

and deaths were recorded, with 'higher risk or severe' dichotomised categories comprising PSI (class IV and V), CURB-65 (score  $\geq 3$ ), CURB (score  $\geq 2$ ) and CRB-65 (score  $\geq 2$ ). Statistical analysis was carried out using R-DiagMeta,<sup>10</sup> MetaAnalyst<sup>11</sup> and RevMan 5.024 (Nordic Cochrane Center, Copenhagen, Denmark). Pooled sensitivities, specificities, diagnostic odds ratios (OR) and positive and negative predictive values were calculated using the random effects model (which takes into account the variability between studies).<sup>12</sup> The summary receiver operating characteristic curves (SROC) for each severity scale were generated with the bivariate random effects approach.<sup>10</sup> Statistical heterogeneity was assessed using the  $I^2$  statistic, with  $I^2$  values  $>50\%$  indicating a substantial level of heterogeneity.<sup>13</sup>

### RESULTS

The study selection flowchart is shown in figure 1. Twenty-three studies were included in the analysis,<sup>14–35</sup> with some studies reporting more than one risk stratification scale. Sixteen studies



**Figure 2** Receiver operating characteristics space illustrating performance of scoring systems at identifying patients at risk of death. (A) Paired sensitivity and specificity with SROC curves giving an indirect comparison of all four scoring systems. (B) Paired specificity and sensitivity of PSI versus CURB-65 in predicting mortality in head to head studies. (C) Paired specificity and sensitivity of PSI versus CRB-65 in predicting mortality in head to head studies. CRB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years; PSI, Pneumonia Severity Index.

covered PSI, 12 CURB-65, 10 CRB-65 and 5 CURB. For direct comparisons of the severity scales in the same patient dataset, 7 tested PSI versus CURB-65, 4 tested PSI versus CRB-65 and 2 studied PSI versus CURB. Table 1 and online table show the characteristics of the studies included in the meta-analysis.

The total sample size from the 23 studies was 22 753 participants with 1680 deaths, giving an average mortality rate of 7.4%. Sample sizes ranged from 134 to 3181 participants. Studies were carried out mainly in emergency department and hospital settings in Europe and North America, with a smaller proportion of data available from Australia, Hong Kong and Pakistan. The average age of participants in the studies was typically around 60 or 70 years. For the validity assessment we found the risk of bias to be a particular issue with ascertainment of mortality. Only 14 of the 23 studies gave details on the methods used in confirming whether participants had died during the follow-up period (see online table). Lack of information on patient management was another major potential source of bias as there was little information on the interventions used in treating pneumonia and whether antibiotic use was consistent both within a study and between different studies.

### Indirect and direct comparisons of the ability to identify patients at risk of death

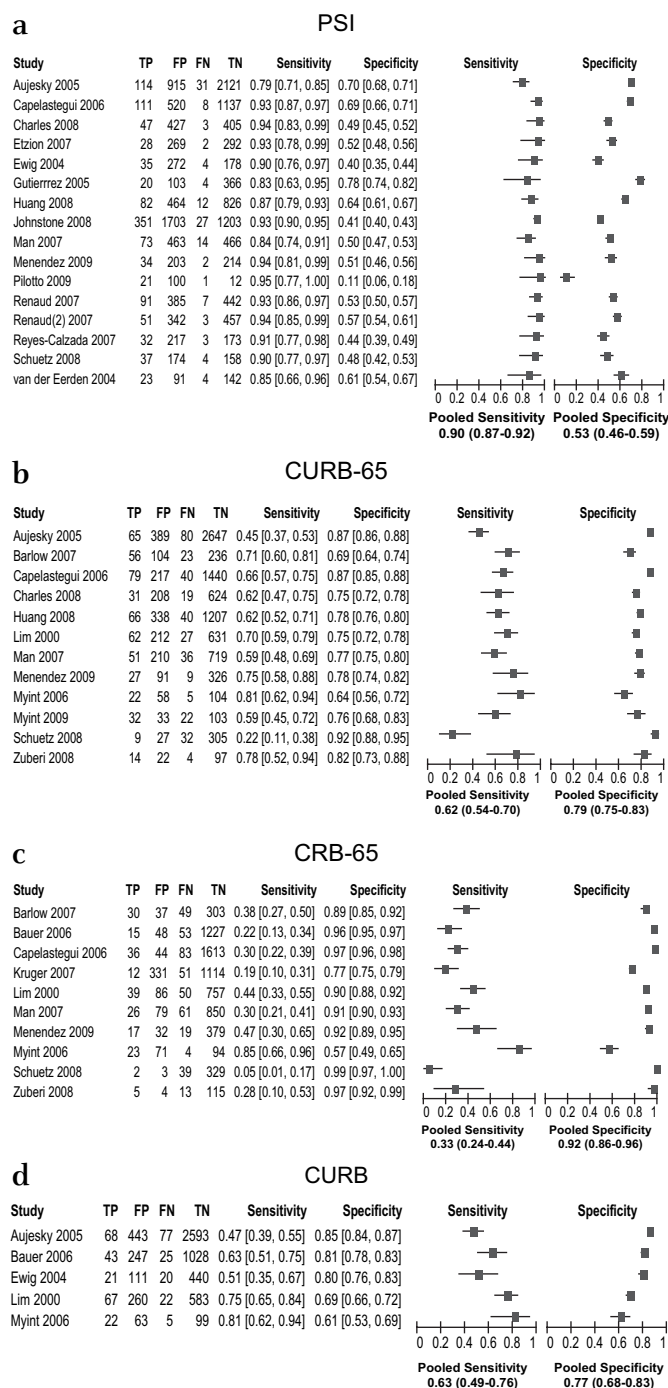
The paired sensitivities (proportion of patients who subsequently die who were correctly classified as 'higher risk') and specificities (proportion of survivors who were correctly classified as not being in the 'higher risk' category) of each study are shown with the SROC curves for the severity scale in figure 2A. Data on the studies with direct comparisons are shown in figure 2B and C, with 7 studies evaluating PSI versus CURB-65 and 4 studies comparing PSI with CRB-65 (the results for other direct comparisons are not shown owing to the small number of studies). Both the direct and indirect comparisons consistently indicate that PSI is more sensitive but less specific than CURB-65, CURB or CRB-65 in identifying those who subsequently die.

Forest plots of the raw data and estimated sensitivities and specificities from each study are shown in figure 3A–D. Pooled estimates of sensitivities, specificities, diagnostic ORs and positive and negative predictive values were measured for each scale and these results are shown in table 2. There was significant heterogeneity for the pooled estimates.

As there were only a relatively small proportion of deaths, the negative predictive values of all four scales were similarly impressive, ranging from 0.94 (CRB-65) to 0.98 (PSI). The low prevalence of death also accounts for the somewhat low positive predictive values ranging from 0.14 (PSI) to 0.28 (CRB-65) (table 2).

### Estimates of clinical impact of each severity scale

Measures of the clinical impact of risk stratification (such as positive and negative predictive values) depend on the underlying rate of pneumonia-related deaths. Figure 4A shows the absolute number of false negatives (number of patients wrongly classified as non-severe) according to severity scale per 1000 patients with pneumonia. The PSI had the lowest false negative rate, meaning that the test is able to correctly identify patients who have non-severe pneumonia and are at low risk of death. Figure 4B shows the absolute number of false positives (number of patients misclassified as 'high risk') per 1000 patients with pneumonia, according to test. This illustrates that the PSI errs on the side of caution by judging relatively more survivors as being at 'high risk', whereas the CURB-65 and its iterations are more specific in correctly classifying patients who have a greater likelihood of death. These figures also confirm that the absolute



**Figure 3** Forest plots of sensitivity and specificity according to study and prognostic scale. (A) Pneumonia Severity Index (PSI). (B) CURB-65 (confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years). (C) CRB-65 (confusion, respiratory rate, blood pressure, age  $\geq 65$  years). (D) CURB (confusion, urea nitrogen, respiratory rate, blood pressure).

impact of test performance varies with the mortality rate and that the differences between tests become more apparent in populations with high proportions of pneumonia-related deaths.

### DISCUSSION

To our knowledge, this is the first systematic review and meta-analysis covering the comparative test performance of pneumonia severity scales that are in common clinical use. Current risk stratification scales (PSI, CURB-65, CRB-65 and



**Table 2** Summary statistics\* of test performance for the pneumonia risk stratification scales

Severity scale	PSI (n = 16 519)		CURB-65 (n = 11 199)		CRB-65 (n = 8143)		CURB (n = 6237)	
	Summary statistic (95% CI)	I <sup>2</sup>	Summary statistic (95% CI)	I <sup>2</sup>	Summary statistic (95% CI)	I <sup>2</sup>	Summary statistic (95% CI)	I <sup>2</sup>
Pooled sensitivity	0.90 (0.87 to 0.92)	59%	0.62 (0.54 to 0.70)	81%	0.33 (0.24 to 0.44)	84%	0.63 (0.49 to 0.76)	88%
Pooled specificity	0.53 (0.46 to 0.59)	98%	0.79 (0.75 to 0.83)	96%	0.92 (0.86 to 0.96)	98%	0.77 (0.68 to 0.83)	98%
Positive predictive value	0.14 (0.13 to 0.16)	75%	0.24 (0.19 to 0.30)	88%	0.28 (0.18 to 0.41)	92%	0.17 (0.14 to 0.22)	78%
Negative predictive value	0.98 (0.98 to 0.99)	46%	0.95 (0.93 to 0.97)	91%	0.94 (0.92 to 0.95)	89%	0.97 (0.96 to 0.97)	52%
Diagnostic OR	10.77 (8.29 to 13.97)	39%	6.40 (5.05 to 8.10)	56%	5.97 (3.41 to 10.44)	86%	5.75 (4.59 to 7.21)	0%

\*Summary statistics consist of the pooled sensitivity, pooled specificity, positive and negative predictive values and diagnostic OR for the performance of each test in differentiating between survivors and patients who subsequently die.

CRB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years; PSI, Pneumonia Severity Index.

CURB) have different trade-offs and no single scale is clearly superior on all counts. The PSI is the most sensitive test with a low false negative rate (figure 4A), thus giving clinicians greater confidence in identifying patients who may not need hospital admission.<sup>34</sup> Conversely, the CURB-65, CRB-65 and CURB scales are more specific and have higher positive predictive values than the PSI, which means that a greater proportion of patients in the 'higher risk' categories are correctly classified.

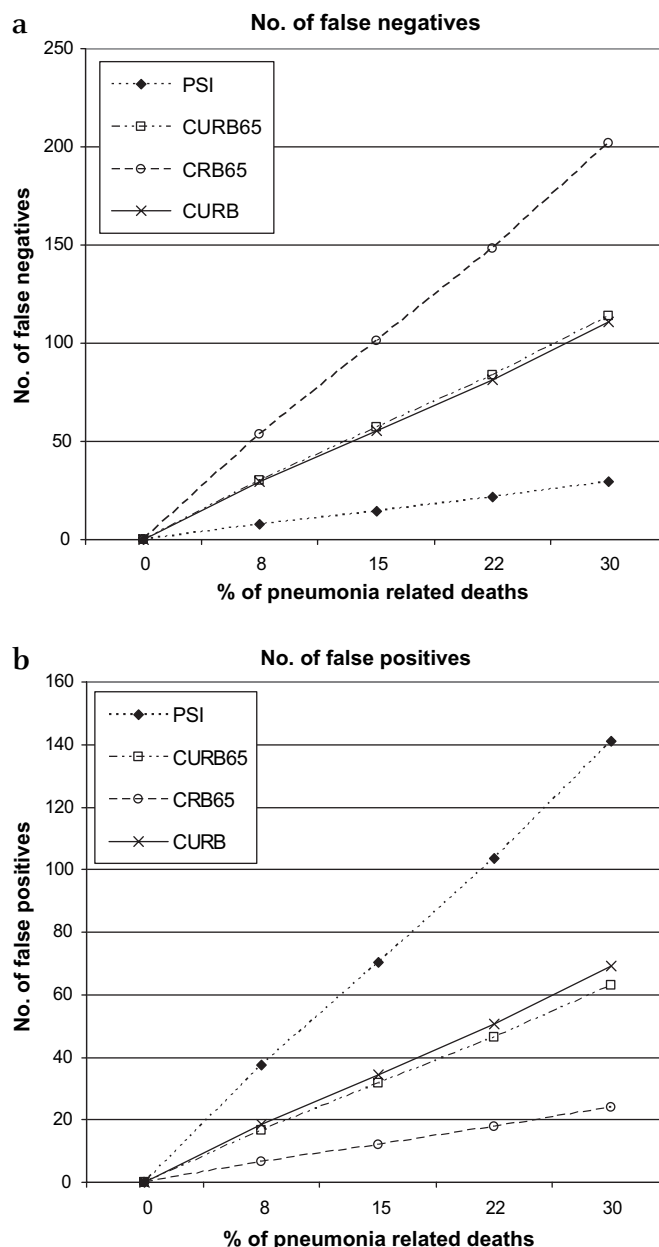
In theory, the poorer sensitivity of the CURB-65-based scales means that some patients may be incorrectly diagnosed and managed as non-severe even when they are actually at higher risk of death. However, as there was only a relatively small percentage of deaths (7.4%) among study participants, the negative predictive values of these CURB-65-based tests are very similar to those of PSI, indicating that the clinical differences between tests are likely to be small. In situations with a low prevalence of adverse outcomes, we can be more confident that a 'non-severe' test result reliably predicts that the patient will have a good outcome, but we are also less certain that a 'severe' rating genuinely predicts death.<sup>35</sup> This is illustrated by the range of positive predictive values from 0.14 (PSI) to 0.28 (CRB-65).

The practical aspects and resource implications of the chosen scoring system should also be considered. The PSI involves a detailed history, physical examination, venous blood sampling, arterial blood gas measurements and chest x-ray, thus requiring the physician to gather a total of 12 parameters from the history and examination as well as 7 parameters derived from further investigations.<sup>6</sup> Although most of the parameters for the PSI are available in hospital settings and can be worked out with a web-based PSI calculator,<sup>36</sup> busy clinicians in emergency departments may end up not bothering to estimate the PSI or may rush through the task inaccurately. However, the CURB-65-based scales do not incorporate potentially important parameters such as hypoxaemia and bilateral pneumonia in their scores, while the CRB-65 (a potentially useful scale for community use) omits the urea measurement, thus possibly reducing sensitivity.

The choice of test depends on the attitude of the health providers regarding healthcare and resources use as well as the rate of pneumonia-related mortality. The most sensitive and labour-intensive test (such as the PSI) may be preferred in resource-rich healthcare settings, particularly where pneumonia mortality is relatively high. In certain community settings with limited resources and where pneumonia mortality is relatively low, the lower sensitivity of the CRB-65 is not a major disadvantage and its ease of use and higher specificity may help clinicians to focus on those requiring more clinical attention.

### Study limitations

There are a number of limitations in our review. There is clearly considerable heterogeneity in the performance of the severity



**Figure 4** (A) Number of patients wrongly classified as non-severe as a function of incidence of mortality per 1000 patients with pneumonia. (B) Number of patients wrongly classified as severe as a function of incidence of mortality per 1000 patients with pneumonia. CRB-65, confusion, respiratory rate, blood pressure, age  $\geq 65$  years; CURB, confusion, urea nitrogen, respiratory rate, blood pressure; CURB-65, confusion, urea nitrogen, respiratory rate, blood pressure, age  $\geq 65$  years; PSI, Pneumonia Severity Index.

scales, which may be related to the diverse populations evaluated and differences in microbiological spectrum and antibiotic sensitivity. This affects the validity of the pooled estimates which we presented as a secondary analysis, even though we used a random effects model which incorporates study level variability into the meta-analysis. Other researchers have suggested that large degrees of heterogeneity are commonly seen in meta-analyses of diagnostic studies.<sup>37</sup> Closer examination of our Forest plots did not demonstrate any consistent source to account for the substantial heterogeneity, and we believe that the data points for each severity scale are actually fairly consistently clustered together in the SROC plane. We also specifically chose to include studies with direct head-to-head evaluation of PSI and the CURB-65-related scoring systems within the same patient population so that heterogeneity would be minimised when comparing different severity scales.

Although we used a comprehensive search strategy, some retrieved papers had to be excluded because raw mortality data were not given or there was insufficient information for us to work out the number of deaths from the reported sensitivity or specificity. Two notable studies had to be excluded for not fulfilling our eligibility criteria: the very first PSI-based article studied patients from 1989 to 1994 which is a very different time period from that of the CURB-65-related studies that began around 1999,<sup>6</sup> while a recent German study of 388 406 patients reported only on inpatient deaths rather than mortality at 4–8 weeks of follow-up.<sup>7</sup> We did not have the resources to evaluate non-English publications and we chose to include only peer-reviewed published literature.

We were unable to properly assess the quality of the included studies owing to the lack of detail in reporting of key areas such as the methods used in confirming the outcomes and the antibiotic regimens. Only a few studies provided information on the treatment pathways, and it is possible that different treatment regimens may have contributed to the substantial heterogeneity seen here. For instance, variations in antibiotic use and antibiotic resistance, availability of chest specialists and intensive care beds could have affected mortality outcomes and performance of the severity scale.

## CONCLUSION

Our findings suggest that the PSI and CURB-65 scoring systems perform well at identifying patients with pneumonia who have a low risk of death. However, it is also clear that all four prognostic scales have limitations and should only be used in conjunction with careful clinical judgement when making a management decision. Further research should focus on identifying the individual characteristics that account for differences among the four scoring systems so that further refinements can be made to enable more accurate risk classification and treatment of patients with pneumonia.

**Competing interests** None.

**Contributors** YKL and PKM conceptualised the review and developed the protocol; YKL, PKM, CSK and AN selected studies and abstracted data; YKL and CSK carried out the data analysis and wrote the manuscript with critical input from PKM and AN. YKL acts as guarantor for the paper.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Guest JF, Morris A. Community-acquired pneumonia: the annual cost to the National Health Service in the UK. *Eur Respir J* 1997;**10**:1530–4.
- Lim WS, Lewis S, Macfarlane JT. Severity prediction rules in community acquired pneumonia: a validation study. *Thorax* 2000;**55**:219–23.
- Kamath A, Pasteur MC, Slade MG, et al. Recognising severe pneumonia with simple clinical and biochemical measurements. *Clin Med* 2003;**3**:54–6.
- Myint PK, Kamath AV, Vowler SL, et al. Severity assessment criteria recommended by the British Thoracic Society (BTS) for community-acquired pneumonia (CAP) and older patients. Should SOAR (systolic blood pressure, oxygenation, age and respiratory rate) criteria be used in older people? A compilation study of two prospective cohorts. *Age Ageing* 2006;**35**:286–91.
- 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**:377–82.
- 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**:243–50.
- Ewig S, Birkner N, Strauss R, et al. New perspectives on community-acquired pneumonia in 388,406 patients. *Thorax* 2009;**64**:1062–9.
- Ewig S, de Roux A, Bauer T, et al. Validation of predictive rules and indices of severity for community acquired pneumonia. *Thorax* 2004;**59**:421–7.
- Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;**323**:224–8.
- Chappell FM, Raab GM, Wardlaw JM. When are summary ROC curves appropriate for diagnostic meta-analyses? *Stat Med* 2009;**28**:2653–68.
- Wallace BC, Schmid CH, Lau J, et al. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC Med Res Methodol* 2009;**9**:80.
- Sousa ML, Ribeiro AR. Systematic review and meta-analysis of diagnostic and prognostic studies: a tutorial. *Arq Bras Cardiol* 2009;**92**:229–38.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–60.
- 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**:384–92.
- Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia. *Thorax* 2007;**62**:253–9.
- Bauer TT, Ewig S, Marre R, et al. CRB-65 predicts death from community-acquired pneumonia. *J Intern Med* 2006;**260**:93–101.
- Capelastegui A, Espana PP, Quintana JM, et al. Validation of a predictive rule for the management of community-acquired pneumonia. *Eur Respir J* 2006;**27**:151–7.
- Charles PG, Whitby M, Fuller AJ, et al. The etiology of community-acquired pneumonia in Australia: why penicillin plus doxycycline or a macrolide is the most appropriate therapy. *Clin Infect Dis* 2008;**46**:1513–21.
- Etzion O, Novack V, Avnon L, et al. Characteristics of low-risk patients hospitalized with community-acquired pneumonia. *Eur J Intern Med* 2007;**18**:209–14.
- Gutierrez F, Masia M, Rodriguez JC, et al. Epidemiology of community-acquired pneumonia in adult patients at the dawn of the 21st century: a prospective study on the Mediterranean coast of Spain. *Clin Microbiol Infect* 2005;**11**:778–800.
- Huang DT, Weissfeld LA, Kellum JA, et al. Risk prediction with procalcitonin and clinical rules in community-acquired pneumonia. *Ann Emerg Med* 2008;**52**:48–58.e2.
- Johnstone J, Eurich DT, Majumdar SR, et al. Long-term morbidity and mortality after hospitalization with community-acquired pneumonia: a population-based cohort study. *Medicine (Baltimore)* 2008;**87**:329–34.
- Kruger S, Ewig S, Marre R, et al. Procalcitonin predicts patients at low risk of death from community-acquired pneumonia across all CRB-65 classes. *Eur Respir J* 2008;**31**:349–55.
- Man SY, Lee N, Ip N, et al. Prospective comparison of three predictive rules for assessing severity of community-acquired pneumonia in Hong Kong. *Thorax* 2007;**64**:348–53.
- Menendez R, Martinez R, Reyes S, et al. Biomarkers improve mortality prediction by prognostic scales in community-acquired pneumonia. *Thorax* 2009;**64**:587–91.
- Myint PK, Sankaran P, Musonda P, et al. Performance of CURB-65 and CURB-age in community-acquired pneumonia. *Int J Clin Pract* 2009;**63**:1345–50.
- Pilotto A, Addante F, Ferrucci L, et al. The multidimensional prognostic index predicts short- and long-term mortality in hospitalized geriatric patients with pneumonia. *J Geriatr A Biol Sci Med Sci* 2009;**64A**:880–7.
- Renaud B, Coma E, Labarere J, et al. Routine use of the Pneumonia Severity Index for the guiding the site-of-treatment decision of patients with pneumonia in the emergency departments: a multicenter, prospective, observational, controlled cohort study. *Clin Infect Dis* 2007;**44**:41–9.
- Renaud B, Coma E, Hayon J, et al. Investigation of the ability of the Pneumonia Severity Index to accurately predict clinical relevant outcomes: a European study. *Clin Microbiol Infect* 2007;**13**:923–31.
- Reyes Calzada S, Martinez Tomas R, Cremades Romero MJ, et al. Empiric treatment in hospitalized community-acquired pneumonia. Impact on mortality, length of stay and re-admission. *Respir Med* 2007;**101**:1909–15.
- Schuetz P, Koller M, Chirst-Crain M, et al. Predicting mortality with pneumonia severity scores: importance of model recalibration to local settings. *Epidemiol Infect* 2008;**136**:1628–37.
- van der Eerden MM, de Graaff CS, Bronsveld W, et al. Prospective evaluation of pneumonia severity index in hospitalized patients with community-acquired pneumonia. *Respir Med* 2004;**98**:872–8.
- Zuberi FF, Khan JA. Prospective comparison of prediction rules of risk for CAP in a developing country. *Int J Tuberc Lung Dis* 2008;**12**:447–52.

34. **Niederman MS.** Making sense of scoring systems in community acquired pneumonia. *Respirology* 2009;**14**:327–35.
35. **Agency for Healthcare Research and Quality.** *Pneumonia severity index calculator*. Rockville, Maryland, 2003. <http://pda.ahrq.gov/psicalc.asp> (accessed 16 Nov 2009).
36. **Altman DG,** Bland M. Diagnostic tests 2: Predictive values. *BMJ* 1994;**309**:102.
37. **Dinnes J,** Deeks J, Kirby J, *et al.* A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess* 2005;**9**:1–113.

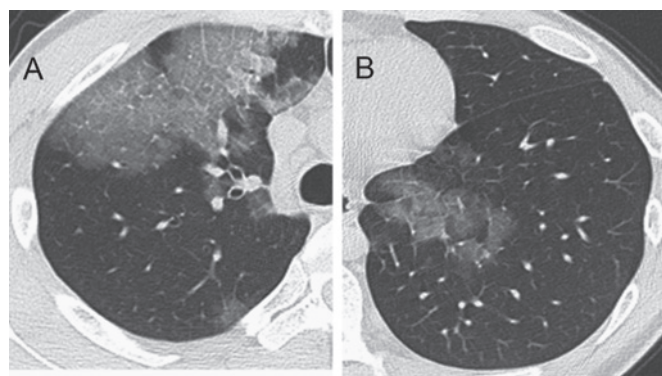
## Pulmonary puzzle

### Dyspnoea, fever, patchy ground-glass opacities and intermittent severe epigastralgia

#### CLINICAL PRESENTATION

A 33-year-old Japanese man experienced intermittent severe epigastralgia 5 days before admission followed by dyspnoea and fever, and he was hospitalised in acute respiratory failure. He had taken no medications before admission. Bilateral ground-glass opacities were seen on the chest radiograph on admission. The patient received BCG in childhood and had no history of atopy, asthma or any allergies, and did not smoke. On examination, his body temperature was 38.3°C, and oxygen saturation on air was 88%. Mild fine inspiratory crackles were detected by chest auscultation. There were no signs of urticarial rash or angioedema. He had epigastric tenderness without signs of peritoneal irritation.

Laboratory tests revealed the following: white blood cell count,  $1.26 \times 10^9/l$ ; proportion of peripheral eosinophils, 2.3% (normal <7.0%); C-reactive protein, 13.6 mg/dl (normal <0.3 mg/dl); amylase, 120 IU/l; immunoglobulin E, 456 IU/ml (normal <170 IU/ml); and *Aspergillus* precipitins and antinuclear cytoplasmic antibody (ANCA) were negative. Arterial blood gas analysis on air showed a pH of 7.40,  $P_{aCO_2}$  of 4.68 kPa (35 mm Hg) and  $P_{aO_2}$  of 7.51 kPa (56 mm Hg). High-resolution CT revealed bilateral non-segmental ground-glass opacities that were distributed dominantly in peripheral lesions. Interlobular septal thickening was also noted (figure 1A,B).



**Figure 1** Chest high-resolution CT on admission revealed bilateral non-segmental ground-glass opacities and interlobular septal thickenings. (A) Right upper lobe view; (B) left lower lobe view.



**Figure 2** Image from oesophagogastroduodenoscopy at the gastric corpus.

We performed immediate bronchoalveolar lavage and trans-bronchial lung biopsy. Severe epigastralgia occurred immediately following the bronchoscopy, so abdominal CT and oesophagogastroduodenoscopy were performed (figure 2). No abnormalities were observed on abdominal CT.

#### QUESTION

What is the cause of the pulmonary infiltrates and respiratory failure? What are the underlying cause and significance of the intermittent severe epigastralgia? What diagnostic test can confirm the diagnosis?

*See page 920 for the answer*

**Naoya Hijikata,<sup>1,2</sup> Noboru Takayanagi,<sup>2</sup> Yutaka Sugita,<sup>2</sup> Yoshinori Kawabata<sup>3</sup>**

<sup>1</sup>Department of Internal Medicine, Saitama Kyoudou Hospital, Saitama, Japan;

<sup>2</sup>Department of Respiratory Medicine, Saitama Cardiovascular and Respiratory Center, Saitama, Japan; <sup>3</sup>Division of Pathology, Saitama Cardiovascular and Respiratory Center, Saitama, Japan

**Correspondence to** Dr Naoya Hijikata, Department of Internal Medicine, Saitama Kyoudou Hospital, 1317 Kizoro, Kawaguchi, Saitama 333-0831, Japan; [n.hijikata-rocaltrain.1016@r7.dion.ne.jp](mailto:n.hijikata-rocaltrain.1016@r7.dion.ne.jp)

**Competing interests** None.

**Patient consent** Obtained.

**Provenance and peer review** Not commissioned; externally peer reviewed.

*Thorax* 2010;**65**:890. doi:10.1136/thx.2010.136655