Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis

Christina Wacker, Anna Prkno, Frank M Brunkhorst*, Peter Schlattmann*

Summary

Background Procalcitonin is a promising marker for identification of bacterial infections. We assessed the accuracy and clinical value of procalcitonin for diagnosis of sepsis in critically ill patients.

Methods We searched Medline, Embase, ISI Web of Knowledge, the Cochrane Library, Scopus, BioMed Central, and Science Direct, from inception to Feb 21, 2012, and reference lists of identified primary studies. We included articles written in English, German, or French that investigated procalcitonin for differentiation of septic patients—those with sepsis, severe sepsis, or septic shock—from those with a systemic inflammatory response syndrome of non-infectious origin. Studies of healthy people, patients without probable infection, and children younger than 28 days were excluded. Two independent investigators extracted patient and study characteristics; discrepancies were resolved by consensus. We calculated individual and pooled sensitivities and specificities. We used P to test heterogeneity and investigated the source of heterogeneity by metaregression.

Findings Our search returned 3487 reports, of which 30 fulfilled the inclusion criteria, accounting for 3244 patients. Bivariate analysis yielded a mean sensitivity of 0·77 (95% CI 0·72–0·81) and specificity of 0·79 (95% CI 0·74–0·84). The area under the receiver operating characteristic curve was 0·85 (95% CI 0·81–0·88). The studies had substantial heterogeneity (P=96%, 95% CI 94–99). None of the subgroups investigated—population, admission category, assay used, severity of disease, and description and masking of the reference standard—could account for the heterogeneity.

Interpretation Procalcitonin is a helpful biomarker for early diagnosis of sepsis in critically ill patients. Nevertheless, the results of the test must be interpreted carefully in the context of medical history, physical examination, and microbiological assessment.

Funding Ministry of Education and Research, the Deutsche Forschungsgemeinschaft, Thuringian Ministry for Education, Science and Culture, the Thuringian Foundation for Technology, Innovation and Research, and the German Sepsis Society.

Introduction

Worldwide, sepsis and its sequelae are still a common cause of acute illness and death in patients with community-acquired and nosocomial infections.1,2 The American College of Chest Physicians and the Society of Critical Care Medicine Consensus Conference (Northbrook, IL, USA; August, 1991) defined sepsis as systemic inflammatory response caused by infection.3 However, no gold standard exists for proof of infection. Bacteraemia is identified in only about 30% of patients with sepsis, depending on previous antibiotic treatment.4,5 Furthermore, early clinical signs of sepsis, such as fever, tachycardia, and leucocytosis, are non-specific and overlap with signs of systemic inflammatory response syndromes of non-infectious origin, especially in patients who have undergone surgery. Other signs, such as arterial hypotension, thrombocytopenia, or increased lactate concentrations suggest, too late for life-saving treatment, progression to organ dysfunction. Thus, delay in diagnosis and treatment of sepsis increases mortality, prolongs length of hospital stay, and increases costs,6 highlighting the need for early and reliable diagnostic biomarkers for sepsis.

Several humoral and cellular systems are activated during sepsis, with a subsequent release of various molecules that mediate the host response to infection. Several potential bloodstream biomarkers have been investigated for their ability to diagnose sepsis, estimate its severity, and provide a prognosis. The 116-aminoacid polypeptide procalcitonin had been termed the “champion so far” for identification of bacterial infections because it has several advantages over other potential biomarkers—i.e., wide biological range, short time of induction after bacterial stimulus, and long half-life.7 However, only two meta-analyses have investigated the accuracy of procalcitonin for the diagnosis of sepsis, with conflicting results.8,9 Both were limited by selected populations, did not include a heterogeneous patient population, and, most importantly, were biased by the choice of a gold standard for the definition of sepsis. Additionally, new studies of procalcitonin have been done since the publication of the meta-analyses and our understanding of procalcitonin is still developing.

We did a meta-analysis to investigate the ability of procalcitonin to differentiate between sepsis and systemic inflammatory response syndromes of
non-infectious origin in critically ill patients and address the heterogeneity of patients and the affect of individual covariates.

Methods
Search strategy and selection criteria
We systematically searched Medline (via PubMed), Embase (via OvidSP), ISI Web of Knowledge, the Cochrane Library, Scopus, BioMed Central, and Science Direct for studies that assessed the accuracy of procalcitonin for the diagnosis of sepsis.

Our medical subject heading terms (for Medline), EMTREE terms (for Embase), and text (for others) were “(procalcitonin OR PCT) AND (sepsis OR “bacterial infection” OR “systemic inflammatory response syndrome” OR SIRS)”. To reduce the number of results, for searches in Science Direct, Embase, and Scopus, we also used the search terms “NOT (review OR letter OR editorial OR “animal experiment” OR “meeting abstract” OR “proceeding paper” OR “poster presentation” OR “meta-analysis” OR “case report”). We searched the databases between inception and Feb 21, 2012. We also searched the reference list of each primary study identified and of previous systematic reviews.

Studies were included if they assessed the accuracy of procalcitonin for differentiation between critically ill patients with sepsis from those who have a systemic inflammatory response syndrome without infection.

To be eligible, studies had to have a well defined reference standard for sepsis, which included the use of definitions established by the American College of Chest Physicians and Society of Critical Care Medicine Consensus Conference1 or the German Sepsis Society.12 In accordance with these definitions, the presence of infection had to be microbiologically confirmed or at least clinically suspected because of one or more characteristics: white blood cells in a normally sterile body fluid, perforated viscus, radiographic evidence of pneumonia in association with production of purulent sputum, and syndrome associated with a high risk of infection (eg, ascending cholangitis).

Furthermore, the studies had to provide sufficient information to construct the 2×2 contingency table—ie, false and true positives and negatives were provided.

We only included publications written in English, German, or French. Animal experiments, reviews, correspondences, case reports, expert opinions, and editorials were excluded. We also excluded all studies that involved healthy people, patients without probable infection, and children younger than 28 days.

Procedures
Two investigators (CW, AP) independently extracted data, including the quality assessment from the retrieved studies. Discrepancies were resolved in a consensus meeting or, if agreement could not be reached, they were resolved by referral to a third investigator (FMB).

The extracted data were general and detailed methodology characteristics, characteristics of the study population (adults or children), setting (emergency department, general ward, or intensive care unit), admission category (surgical or medical), severity of illness (sepsis, severe sepsis, or septic shock), and details of the procalcitonin assays and cutoffs used.

Each investigator also recorded the number of true and false positives and negatives. We contacted the corresponding authors if further information was needed. If no response was received after sending a reminder, the study was excluded.

We assessed the methodological quality of the studies with the Quality Assessment of Diagnostic Accuracy Studies checklist.13 We tailored the guidelines for scoring each item of the checklist to our review.9 Because overall quality scoring is difficult and should not be included in meta-analyses,15 we included only item 9 (description of the reference standard) and item 11 (diagnostic review bias) of the 14 individual quality-related items as covariates in a bivariate random-effects model to test them as possible sources of variation and bias.

Statistical analysis
We tabulated true positives, false negatives, false positives, and true negatives in patients with sepsis and systemic inflammatory response syndrome, stratified by study. We used the numbers to calculate sensitivity and specificity and a corresponding CI.

To synthesise data, we used an exact binomial rendition16 of the bivariate mixed-effects regression model.

Figure 1: Study selection
Some studies were excluded for more than one reason. *Did not investigate the diagnostic accuracy of procalcitonin as a marker for sepsis.
model developed by van Houwelingen et al. for meta-analysis of treatment trials, modified for synthesis of diagnostic test data.23,24 This model does not transform pairs of sensitivity and specificity of individual studies into a single indicator of diagnostic accuracy, but preserves the two-dimensional nature of the data taking into account any correlation between the two.

Based on this model, we estimated mean logit sensitivity and specificity with their standard error and 95% CIs, the between-study variability in logit sensitivity and specificity, and the covariance between them. We back-transformed these quantities to the original receiver operating curve scale to obtain summary sensitivity, specificity, and diagnostic odds ratios. We then used the derived logit estimates of sensitivity, specificity, and respective variances to construct a hierarchical summary receiver operating curve for procalcitonin with summary operating points for sensitivity and specificity on the curves and a 95% confidence contour ellipsoid (two-dimensional CI).

We calculated I² to assess heterogeneity. If heterogeneity among studies was recorded, the potential source of heterogeneity was investigated by metaregression. Study-level covariates can be used in metaregression to combine results from multiple studies with attention to between-study variation. We used study-specific covariates such as population or admission category. To investigate publication bias, we constructed effective sample size funnel plots versus the log diagnostic odds ratio and did a regression test of asymmetry.25

We calculated κ statistics to assess the agreement between the two investigators for assessment of methodological quality.

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
<th>Admission category</th>
<th>Setting</th>
<th>Procalcitonin assay</th>
<th>Cutoff (ng/mL)</th>
<th>n</th>
<th>Prevalence (%)</th>
<th>Severity</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ED</td>
<td>PCT-Q</td>
<td>0.5</td>
<td>120</td>
<td>59%</td>
<td>No information</td>
<td>63</td>
<td>11</td>
<td>38</td>
<td>8</td>
<td>0.89 (0.79–0.95)</td>
<td>0.78 (0.63–0.88)</td>
</tr>
<tr>
<td>2009</td>
<td>Adult</td>
<td>Medical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>0.5</td>
<td>337</td>
<td>36%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>73</td>
<td>45</td>
<td>170</td>
<td>49</td>
<td>0.60 (0.52–0.69)</td>
<td>0.79 (0.73–0.84)</td>
</tr>
<tr>
<td>2006</td>
<td>Paediatric</td>
<td>Medical and surgical</td>
<td>PICU</td>
<td>PCT-LIA</td>
<td>2</td>
<td>28</td>
<td>50%</td>
<td>No information</td>
<td>12</td>
<td>0</td>
<td>14</td>
<td>2</td>
<td>0.86 (0.57–0.98)</td>
<td>1.00 (0.77–1.00)</td>
</tr>
<tr>
<td>2003</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>0.75</td>
<td>103</td>
<td>39%</td>
<td>No information</td>
<td>47</td>
<td>2</td>
<td>19</td>
<td>15</td>
<td>0.76 (0.63–0.86)</td>
<td>0.90 (0.70–0.99)</td>
</tr>
<tr>
<td>2004</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>1.2</td>
<td>49</td>
<td>69%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>21</td>
<td>2</td>
<td>13</td>
<td>13</td>
<td>0.62 (0.44–0.78)</td>
<td>0.87 (0.60–0.98)</td>
</tr>
<tr>
<td>2006</td>
<td>Adult</td>
<td>Medical</td>
<td>ICU</td>
<td>PCT-Kryptor</td>
<td>1</td>
<td>76</td>
<td>47%</td>
<td>Septic shock</td>
<td>29</td>
<td>2</td>
<td>38</td>
<td>7</td>
<td>0.81 (0.64–0.92)</td>
<td>0.95 (0.83–0.99)</td>
</tr>
<tr>
<td>2006</td>
<td>Adult</td>
<td>Surgical</td>
<td>ICU</td>
<td>PCT-Kryptor</td>
<td>9.7</td>
<td>67</td>
<td>46%</td>
<td>Septic shock</td>
<td>28</td>
<td>9</td>
<td>27</td>
<td>3</td>
<td>0.90 (0.74–0.98)</td>
<td>0.75 (0.58–0.88)</td>
</tr>
<tr>
<td>2006</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>PICU</td>
<td>PCT-LIA</td>
<td>1</td>
<td>83</td>
<td>61%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>42</td>
<td>6</td>
<td>26</td>
<td>9</td>
<td>0.82 (0.69–0.92)</td>
<td>0.81 (0.64–0.93)</td>
</tr>
<tr>
<td>2003</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>1.6</td>
<td>51</td>
<td>39%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>16</td>
<td>8</td>
<td>23</td>
<td>4</td>
<td>0.80 (0.58–0.94)</td>
<td>0.74 (0.55–0.88)</td>
</tr>
<tr>
<td>2006</td>
<td>Adult</td>
<td>Medical</td>
<td>HW</td>
<td>PCT-Kryptor</td>
<td>1</td>
<td>93</td>
<td>80%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>56</td>
<td>9</td>
<td>10</td>
<td>18</td>
<td>0.76 (0.64–0.85)</td>
<td>0.52 (0.29–0.76)</td>
</tr>
<tr>
<td>2004</td>
<td>Adult</td>
<td>Medical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>0.6</td>
<td>76</td>
<td>62%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>39</td>
<td>9</td>
<td>20</td>
<td>8</td>
<td>0.83 (0.69–0.92)</td>
<td>0.69 (0.49–0.85)</td>
</tr>
<tr>
<td>2006</td>
<td>Paediatric</td>
<td>Medical and surgical</td>
<td>PICU</td>
<td>PCT-LIA</td>
<td>0.28</td>
<td>36</td>
<td>67%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>20</td>
<td>3</td>
<td>9</td>
<td>4</td>
<td>0.83 (0.63–0.95)</td>
<td>0.75 (0.43–0.95)</td>
</tr>
<tr>
<td>2001</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-LIA</td>
<td>1.1</td>
<td>78</td>
<td>77%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>58</td>
<td>4</td>
<td>14</td>
<td>4</td>
<td>0.97 (0.85–1.00)</td>
<td>0.78 (0.52–0.94)</td>
</tr>
<tr>
<td>2011</td>
<td>Adult</td>
<td>Medical</td>
<td>ICU</td>
<td>PCT-Kryptor</td>
<td>2.2</td>
<td>66</td>
<td>83%</td>
<td>Severe sepis and septic shock</td>
<td>31</td>
<td>0</td>
<td>11</td>
<td>24</td>
<td>0.56 (0.42–0.70)</td>
<td>1.00 (0.72–1.00)</td>
</tr>
<tr>
<td>2008</td>
<td>Adult</td>
<td>Surgical</td>
<td>–</td>
<td>PCT-LIA</td>
<td>1.1</td>
<td>63</td>
<td>65%</td>
<td>No information</td>
<td>34</td>
<td>5</td>
<td>17</td>
<td>7</td>
<td>0.83 (0.68–0.93)</td>
<td>0.77 (0.55–0.92)</td>
</tr>
<tr>
<td>2004</td>
<td>Adult</td>
<td>Medical</td>
<td>–</td>
<td>PCT-LIA</td>
<td>0.5</td>
<td>104</td>
<td>39%</td>
<td>No information</td>
<td>17</td>
<td>5</td>
<td>58</td>
<td>24</td>
<td>0.42 (0.28–0.58)</td>
<td>0.92 (0.82–0.97)</td>
</tr>
<tr>
<td>2007</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>HW and ED</td>
<td>PCT-Kryptor</td>
<td>0.25</td>
<td>151</td>
<td>64%</td>
<td>No information</td>
<td>77</td>
<td>23</td>
<td>32</td>
<td>19</td>
<td>0.80 (0.71–0.88)</td>
<td>0.58 (0.44–0.71)</td>
</tr>
<tr>
<td>2010</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-Q</td>
<td>0.5</td>
<td>114</td>
<td>63%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>53</td>
<td>5</td>
<td>37</td>
<td>19</td>
<td>0.74 (0.62–0.83)</td>
<td>0.88 (0.74–0.96)</td>
</tr>
<tr>
<td>2011</td>
<td>Adult</td>
<td>Medical and surgical</td>
<td>ICU</td>
<td>PCT-Kryptor</td>
<td>2</td>
<td>76</td>
<td>42%</td>
<td>Sepsis, severe sepis, and septic shock</td>
<td>31</td>
<td>9</td>
<td>35</td>
<td>1</td>
<td>0.97 (0.84–1.00)</td>
<td>0.80 (0.65–0.90)</td>
</tr>
</tbody>
</table>

(Continues on next page)
We used the MIDAS module‡ for STATA (version 12) for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression. We used Proc for the bivariate summary receiver operating curve analysis and to calculate κ statistics. We used Proc GLIMMIX in SAS (version 9.3) to do the metaregression. Graphs were produced with the MIDAS module and the GLIMMIX in SAS (version 9.3) to do the metaregression.

Role of the funding source
The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
Our database search retrieved 3487 articles. After reviewing the titles and abstracts, we excluded 3321. After a full text review we excluded a further 136, leaving 30 studies for inclusion (figure 1). Because in one study investigators reported diagnostic accuracy separately for medical and surgical patients, the study was divided into two parts, thus we analysed 31 datasets. Search of the reference lists of the identified articles and previous systematic reviews,‡‡‡ did not identify any more relevant articles.

The table shows the main study characteristics. 3244 critically ill patients were included in the analysis, of whom 1863 (57%) had sepsis and 1381 (43%) had systemic inflammatory response syndrome of non-infectious origin. 21 of 30 studies reported classification of severity of illness (sepsis, severe sepsis, or septic shock). Of 1173 patients, 499 (42%) had severe sepsis, and 440 (38%) had septic shock. The prevalence of sepsis among studies ranged between 34% and 88% (mean 60%). Only four studies were done in a paediatric setting, whereas 27 investigated adult patients (table). Sites of infection—eg, lung, abdomen, bloodstream, urinary tract—varied. The source of infection (community-acquired or nosocomial) also differed between studies.

Most studies were done in intensive care units, four of them in a paediatric intensive care unit, and most (20 of 30) were done in Europe (table). The cutoff for procalcitonin concentration differed substantially between studies (median 1·1 ng/mL, IQR 0·5–2·0).

Most studies (17 of 30) used a quantitative manual procalcitonin assay for diagnosis of sepsis (table). The appendix shows assay characteristics, the methodological quality of the included studies according to the Quality Assessment of Diagnostic Accuracy Studies.
checklist, how the studies scored on each item, and how the items were assessed. We omitted item 12 of the checklist (clinical data) because the index test is fully automated and no further clinical data are needed to interpret the test results.

The inter-rater reliability for assessment of quality items was 0·59 (p<0·0001). Overall, the methodological quality was moderate. None of the studies fulfilled all of the items, but all studies fulfilled at least four items. 22 studies (73%) met at least 50% of the items.24,26–28,30–36,39,41,43–45,47–52 Items 3 (reference standard), 5 (partial verification bias), 6 (differential verification bias), and 14 (withdrawals) were fulfilled by all studies. Reports of test review bias (item 10) and uninterpretable results (item 13) were poor (appendix).

We identified publication bias by Deeks’ regression test of asymmetry (t=4·12; p<0·0005; appendix).

Pooled sensitivity was 0·77 (95% CI 0·72–0·81) and pooled specificity was 0·79 (95% CI 0·74–0·84; figure 2).

The area under the receiver operating characteristic curve was 0·85 (95% CI 0·81–0·88; figure 3). Substantial heterogeneity exists among the studies (overall I² for bivariate model 96%, 95% CI 94–99). We recorded no evidence of a threshold effect (tested with the STATA MIDAS module). The proportion of heterogeneity probably caused by different cutoffs was small (0·05). To identify the source of heterogeneity, we did metaregression analyses.

To compare medical with surgical patients we did a stratified bivariate regression analysis. We obtained data from 13 studies (nine provided data for medical patients and four provided data for surgical patients). The diagnostic accuracy in surgical patients was higher than that in medical patients as measured by the area under the summary receiver operating characteristic curve (0·83 [95% CI 0·80–0·86] vs 0·79 [0·75–0·83]; not tested for significance). We also compared adult with paediatric...

Figure 2: Sensitivity and specificity of procalcitonin assay for diagnosis of sepsis
patients (0.85 [0.82–0.88] vs 0.85 [0.81–0.88]; not tested for significance). Analysis of the other covariates yielded no significant results (data not shown). Thus, the heterogeneity could not be explained by metaregression analysis.

Discussion
Procalcitonin can differentiate effectively between sepsis and systemic inflammatory response syndrome of non-infectious origin. Previously, two meta-analyses have investigated the diagnostic accuracy of procalcitonin in critically ill patients, with conflicting results.10,11

In a meta-analysis from 2006, including studies published between April, 1996, and October, 2004, Uzzan and colleagues11 reported that the summary receiver operating characteristics curve for procalcitonin was better than for C-reactive protein for identification of sepsis. However, the investigators restricted the population to surgery or trauma patients. Therefore, no conclusion can be drawn for patients other than surgical. Furthermore, the researchers did not assess the heterogeneity of patients from different settings, with different sites of infection, or other study-specific covariates.

In a meta-analysis from 2007, including 18 studies published between April, 1996, and November, 2005, Tang and colleagues10 concluded that procalcitonin is not able to discriminate between sepsis and systemic inflammatory response syndrome. The diagnostic accuracy of procalcitonin was low; mean sensitivity and specificity were both 71% (95% CI 67–76) and the area under the summary receiver operator characteristic curve was 0.78 (95% CI 0.73–83). However, their findings were heavily biased because of their selection criteria. First, studies were excluded that had sites of infection typical in sepsis, such as abdominal sepsis, pancreatitis, or meningitis. Second, studies that assessed the ability of procalcitonin to diagnose septic shock were excluded. Because progression of sepsis to septic shock is associated with an increase in procalcitonin concentration,1 exclusion of patients with septic shock could reduce the overall estimate of diagnostic accuracy. To prevent systematic bias, we included all eligible studies that investigated the diagnostic capacity of procalcitonin in the continuum from sepsis to severe sepsis and to septic shock. Third, they included studies that assessed patients who did not have systemic inflammatory response syndrome or who were not critically ill, which might cause underestimation of diagnostic accuracy.

Accordingly, 23 studies included in the previous meta-analyses10,11 were excluded from our systematic review because 13 included healthy controls or patients who did not have systemic inflammatory response syndrome in the control group,14–44 and seven did not provide clear definitions for the target condition or included patients who had infection without systemic inflammatory response syndrome and thus were not in accordance with our selection criteria.44–50 Furthermore, four studies had insufficient information to construct the 2 × 2 contingency table.79–83 One investigated the predictive value of procalcitonin for tumour necrosis factor α and interleukin 6 concentrations.28 Another did multiple measurements in several patients and one study investigated the prognostic value of procalcitonin for infection after cardiac surgery.76

Furthermore, the meta-analysis of Tang and colleagues10 has substantial shortcomings in its quantitative data analysis. It summarised pairs of sensitivity and specificity into a single measure of diagnostic accuracy. Thus, important information is missing. To retain the two-dimensional character, we used the bivariate mixed-effects regression model.

Our meta-analysis has several limitations. First, we detected substantial heterogeneity between studies but none of the study characteristics were responsible for the majority of this heterogeneity. The studies differ in several ways—eg, methodological quality, patients’ clinical spectrum, admission category, and procalcitonin assay used. Thus, further unrecorded differences between the studies probably contribute to the heterogeneity. Use of a more homogenous population would solve this difficulty, but would cause selection bias.

Second, a reliable test of infection is still absent, so observational studies are biased by the choice of...
gold standard. According to our inclusion criteria, the presence of infection had to be microbiologically confirmed or at least clinically suspected. All included studies fulfilled this requirement (appendix), but most did not provide much detailed information about how infection was proved. Nevertheless, depending on previous antibiotic treatment, bacteraemia occurs in only about 30% of patients with sepsis. Additionally, absence of standardisation of clinical and radiological findings could cause interobserver variability, which could lead to false-negative or false-positive judgments about the patient’s medical condition. We only included studies that had a well defined reference standard for sepsis. Nevertheless, we do not know definitively whether all patients with infection were identified as such.

Third, implementation of some studies was reported poorly, especially with regard to uninterpretable results and test review bias (appendix). To minimise resultant bias and to ensure more homogeneity, investigators should use the Standards for Reporting of Diagnostic Accuracy checklist and also consider using the Quality Assessment of Diagnostic Accuracy Studies checklist.11

Fourth, we detected publication bias. Studies with desirable results are more likely to be published, which can lead to an overestimation of overall diagnostic accuracy. To solve this problem, we looked again for further studies by searching the databases and reference lists of primary studies, but could not identify additional relevant articles. Finally, we only included studies written in German, English, or French, which might have affected our findings.

The cutoffs that separated patients who had sepsis from those who did not varied greatly between studies. Some had a cutoff that led to the most favourable results for diagnostic accuracy. Others gave sensitivity and specificity at different thresholds. The difficulty is that the cutoffs were not subsequently validated. The values of diagnostic accuracy are correlated negatively with each other. To change the cutoff means changing sensitivity at the cost of specificity or vice versa. False-negative results leading to denial of treatment could be fatal in sepsis.6 However, to prevent the development of antibiotic resistance, and increased side-effects and costs, critically ill patients without bacterial infection should be identified correctly. Thus, a rational threshold is needed. We recommend different phases in testing diagnostic accuracy. First, investigators should examine the validity of procalcitonin in a selected group of patients to find a rational cutoff. Second, to ascertain diagnostic value in everyday clinical practice, the established cutoff has to be validated in a diagnostic controlled trial.

The most important feature of a biomarker is its potential to change clinical decision making. In recent years, cutoffs between 0·1 and 0·5 ng/mL have been calculated in patients with lower respiratory tract infections.7 Our meta-analysis provides important information for critically ill patients, for whom diagnostic decision making is of utmost importance. The median cutoff of the studies included was 1·1 ng/mL (IQR 0·5–2·0). The absence of a clinical threshold effect suggests that a cutoff of between 1·0 and 2·0 ng/mL is helpful for discrimination of patients with sepsis from other inflammatory conditions, in accordance with recommendations.80

Likelihood ratios and post-test probabilities are also relevant for clinicians. They provide information about the likelihood that a patient with a positive or negative test actually has sepsis or not. In our study, both likelihood ratio and post-test probability were moderate (figure 4). A positive likelihood ratio of 4 implies that a person with disease is four-times more likely to have a positive test result than is a healthy person. Given a pretest probability of 20%, the post-test probability for a positive test result is 48% (figure 4). Likewise a negative likelihood ratio of 0·29 reduces the post-test probability to 7% for a negative test result. However, these likelihood ratios are calculated from dichotomised data. The result of the procalcitonin test is either positive or negative. The disadvantage of making data dichotomous is that
useful information is lost. Patients with a high procalcitonin concentration are more likely to have sepsis than are patients with a low procalcitonin concentration. To provide more precise information about the reliability of the test, we suggest calculating likelihood ratios based on multiple cutoffs.

As our results show, procalcitonin is not a perfect marker for diagnosis of sepsis, but an ideal marker does not exist. Sepsis is a pathophysiological process rather than a specific syndrome and is too complex to be described by a single measure. Nevertheless, procalcitonin is one of the most promising parameters. Several other mediators and molecules of the host response to infection—C-reactive protein, soluble TREM1, interleukin 6, interleukin 8, and soluble PLAUR—have been investigated, but with no outstanding result.

In conclusion, procalcitonin is a helpful marker for diagnosis of sepsis in critically ill patients. However, it cannot be recommended as the single definitive test for sepsis diagnosis, but rather it must be interpreted in context with information from careful medical history, physical examination, and when feasible, microbiological assessment. Moreover, continuing re-evaluation during the course of disease is advisable.

Contributors

CW had the idea for and designed the study, searched the scientific literature, collected, analysed, and interpreted data, and wrote and critically revised the report. AP searched the scientific literature, collected data, and drafted and critically revised the report. FMB had the idea for and designed the study, interpreted data, drafted and critically revised the report, supervised the study, and gave administrative, technical, and material support. PS had the idea for and designed the study, statistically analysed and interpreted the data, drafted and critically revised the report, supervised the study, and gave administrative, technical, and material support.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

The work was supported by the German Centre for Sepsis Control & Care, funded by the Ministry of Education and Research (grant number 01EO1002), the Deutsche Forschungsgemeinschaft (Schl 3-1), Thuringian Ministry for Education, Science and Culture (ProExcellence; PE 108-2), the Thuringian Foundation for Technology, Innovation and Research, and the German Sepsis Society.

References

Articles


41 Meynaar IA, Droog W, Batstra M, Vreede R, Herbrink P. In critically ill patients, serum procalcitonin is more useful in differentiating between sepsis and SIRS than CRP, IL-6, or LBP. Crit Care Res Pract 2011; 2011: 594659.


