Title: Predicting Mycobacterium tuberculosis in patients with community-acquired pneumonia

Authors: Rodrigo Cavallazzi MD\textsuperscript{1}, Timothy Wiemken PhD\textsuperscript{2}, Diana Christensen MD\textsuperscript{2}, Paula Peyrani MD\textsuperscript{2}, Francesco Blasi MD\textsuperscript{3}, Gur Levy MD\textsuperscript{4}, Stefano Aliberti\textsuperscript{5}, Robert Kelley PhD\textsuperscript{2}, Julio Ramirez MD\textsuperscript{2}, and the CAPO Investigators.

Affiliations:
1. Division of Pulmonary, Critical Care and Sleep Disorders Medicine, University of Louisville, Louisville, KY
2. Division of Infectious Diseases, University of Louisville, Louisville, KY
3. Dipartimento di Fisiopatologia Medico-Chirurgica e dei Trapianti, University of Milan, IRCCS Fondazione Cà Granda, Ospedale Maggiore Policlinico, Milan, Italy
4. Physician. Department of Pulmonology, Hospital Universitario de Caracas, Central University of Venezuela, Caracas, Venezuela
5. Dipartimento di Medicina Clinica e Prevenzione, University of Milan-Bicocca, San Gerardo Hospital, Via Pergolesi 33, Monza, Italy

Corresponding Author: Rodrigo Cavallazzi, MD
A3R40, Ambulatory Care Building
550 S. Jackson Street
Louisville, KY, 40202

Word Count:
Abstract: 217
Body: 3211

Key words: pneumonia, tuberculosis, epidemiology, public health, infection
Running Title: Mycobacterium tuberculosis in pneumonia
ABSTRACT

The 22 risk factors suggested by the Centers for Disease Control and Prevention (CDC) to suspect patients at risk for *Mycobacterium tuberculosis* (MTB) have not been evaluated in hospitalized patients with community-acquired pneumonia (CAP). Here, we evaluate which of the CDC risk factors best predict MTB in these patients. To our knowledge, this is the first time a score is developed assessing these risk factors.

This was a secondary analysis of 6,976 patients hospitalized with CAP enrolled in the CAPO International Cohort Study. Using Poisson regression, we selected the subset of risk factors that best predicted the presence of CAP due to MTB. This subset was compared to the CDC risk factors using ROC curve analysis.

Six risk factors were found to best predict CAP due to MTB: age <65, night sweats, hemoptysis, weight loss, MTB exposure, and upper lobe infiltrate. The area under the ROC curve for all CDC risk factors was 71%, and 89% for the subset of 6 risk factors.

The CDC-suggested risk factors are poor at predicting the presence of MTB in hospitalized patients with CAP. With a subset of 6 risk factors identified in this study, we developed a new score, which will improve our capacity to isolate patients at risk of CAP due to MTB at the time of hospitalization.
INTRODUCTION

Tuberculosis (TB) is the seventh leading cause of death in the world, and among infectious diseases it is second only to HIV.[1] Pulmonary TB classically presents as a chronic pneumonia with approximately four weeks between symptom onset to the first health consultation.[2] It may also present acutely in which it is essentially identical to the presentation of a patient with community-acquired pneumonia (CAP).[3]

*Mycobacterium tuberculosis* is identified as the etiologic agent for a small proportion of hospitalized patients with CAP. In fact, guidelines for managing CAP in several countries recognize *Mycobacterium tuberculosis* as a potential pathogen. [4-8]

Identifying and isolating hospitalized patients with pulmonary TB, before a final diagnosis is established, is essential in preventing nosocomial transmission of TB and outbreaks of multi-drug resistant TB.[9] To this end, institutional adherence to international standards of TB control is important. Alas, there is evidence that institutions often lack infection control strategy and adequate environmental measures to prevent transmission of TB.[10,11]
It remains the clinician responsibility to make sure that patients suspected of having TB undergo respiratory isolation.[12] To identify these patients, clinicians and infection preventionists rely primarily on twenty-two risk factors established for patients with chronic pneumonia caused by *Mycobacterium tuberculosis*. These factors were drafted by the United States Centers for Disease Control and Prevention (CDC).[13-15] However, there is little in the literature on the relevance and magnitude of these risk factors to predict CAP due to *Mycobacterium tuberculosis*.

The Community-Acquired Pneumonia Organization (CAPO) cohort study is a multicenter, international study of adult hospitalized patients with CAP.[16] The database for the study contains data on over 7,000 patients and includes all CDC risk factors. In addition, the etiology of CAP for each patient in the study is investigated during the initial workup. In some of these cases, *Mycobacterium tuberculosis* has been identified as the cause of CAP. Our teams’ access to this unique dataset provided us the opportunity to evaluate the role of CDC risk factors for TB in adult hospitalized patients with CAP both with and without TB.

We conducted this study with the following objectives: 1) to evaluate the role of the CDC TB risk factors in the identification of patients with CAP due to TB, and 2) to identify the subset of CDC TB risk factors that best predict the presence of CAP due to TB.
STUDY POPULATION AND METHODS

Study design

This was a secondary data analysis of the Community-Acquired Pneumonia Organization (CAPO) international cohort study database. Investigators for the study completed a case report form (CRF) for each case which was subsequently transferred to the University of Louisville Division of Infectious Diseases Clinical and Translational Research Support Center (CTRSC, http://www.ctrsc.net) through the Internet. Members of the CTRSC validated the cases as they were entered to ensure data quality. A sample of the CRF is available at the study website (www.caposite.com). Information on the CAPO project has been previously published. [16]

Patient Characteristics

The study population consisted of adult patients hospitalized with CAP from March 2001 through December 2011 from 113 hospitals in 33 countries. We reviewed medical records of patients with the diagnosis of CAP, which was established by the admitting physician. Patients enrolled in the study were ≥18 years old and met the criteria for diagnosis of CAP which includes a new pulmonary infiltrate (within 24 hours of admission) associated with at least one of the following: new or increased cough with/without sputum production, fever (> 37.8 °C or 100 °F) or hypothermia (< 35.6 °C or 96 °F), leukocytosis, left shift, or leukopenia. Pneumonia was considered to be community-acquired if patients had no history of hospitalization during the two weeks prior to admission. A case of CAP was considered due to Mycobacterium tuberculosis in
patients with positive respiratory cultures for *Mycobacterium tuberculosis*. There were no exclusion criteria.

**Measurements**

Of the 22 CDC risk factors for TB, we evaluated the following 20: 1) night sweats, 2) hemoptysis, 3) weight loss, 4) hoarseness, 5) HIV/AIDS, 6) history of positive PPD, 7) homeless, 8) alcohol or drug abuse, 9) health care worker, 10) prior history of TB, 11) age more than 65, 12) community living, 13) recent exposure to TB, 14) silicosis, 15) end stage renal disease (ESRD), 16) gastrectomy, 17) diabetes, 18) 10 % or less of ideal body weight, 19) other immunosuppressive state (cancer gastrointestinal tract, hematological disorders, long term cortisone use) and 20) intestinal bypass. We removed two risk factors (cough and pulmonary infiltrate) because these were part of the inclusion criteria to the study. All patients were admitted to the hospital with a working diagnosis of CAP and were initially treated empirically with antibiotics for bacterial CAP.

**Statistical Methods**

Baseline patient characteristics are presented as frequencies with percentages for categorical variables, and means with standard deviations for continuous variables, stratified based on TB status. We performed bivariate analysis of categorical variables using Chi-Squared or Fisher’s Exact Tests for categorical variables and the Mann-Whitney U Test for continuous variables.
We calculated the relative risk, sensitivity and specificity for the diagnosis of TB for each of the 20 CDC TB risk factors. We also calculated the sensitivity and specificity of having at least one of the CDC risk factors significantly associated with TB. Next, we calculated risk for CAP caused by TB for all patients according to the two following scores: 1) CDC TB Risk Score and 2) CAPO TB Risk Score.

**CDC TB Risk Score**

The CDC TB Risk Score is the name we designated for a score calculated by summing each of the CDC risk factors for each patient. This score ranged from zero points (no risk factors present) to 20 points (all risk factors present).

**CAPO TB Risk Score**

To better elucidate which of the 20 CDC TB risk factors were predictive of TB etiology in our patients, we used a multivariate Poisson regression model with robust error variance. We included in the model variables that were statistically significant (P ≤ 0.05) on bivariate analysis. In order to avoid colinearity in this model, we combined the following variables into one variable: 1) prior history of TB, 2) recent exposure to TB, and 3) history of positive PPD. Similarly, we also combined the weight loss and 10% or less of ideal body weight variables into one variable. In the final Poisson regression model, variables that were not statistically significant (P ≤ 0.05) were deleted.

The CAPO TB Risk Score utilized only the subset of variables selected in the final multivariate regression model. For the CAPO TB Risk Score, we also included the
infiltrate localization (upper lobe versus non-upper lobe) variable. The CAPO TB Risk Score was defined as the sum of each of the identified variables.

*Comparison of CDC TB Risk Score and CAPO TB Risk Score*

These two scores were then subjected to ROC curve analysis for comparison of their ability to predict CAP due to TB. Statistical differences in the ROC curves were calculated according to previously published methods.[17] Unless otherwise specified, we considered a P value ≤0.05 statistically significant. We performed statistical analyses with SAS v9.3 (SAS Institute, Cary, NC) and Stata v10 (StataCorp, College Station, TX).

**RESULTS**

*Demographic Information*

This study included 6,976 patients admitted to the hospital with criteria for the diagnosis of CAP. The mean age for the study population was 59.5 (SD: 17.5), and 4,206 (60.3%) were male. Of the total patients included in the study, 60 were diagnosed with CAP due to TB. Table 1 shows the demographic and clinical characteristics of the patients. Patients were distributed among four CAPO world regions: Asia/Africa/Australia (9 TB-CAP; 76 non-TB CAP), Europe (20 TB-CAP; 2,035 non-TB CAP); Latin America (19 TB-CAP; 1,837 non-TB CAP); and USA/Canada (12 TB CAP; 2,122 non-TB CAP).

*CDC TB Risk Score*
The following CDC risk factors had a positive significant association with a diagnosis of TB: 1) night sweats, 2) hemoptysis, 3) weight loss, 4) HIV/AIDS, 5) 10% or less of ideal body weight, 6) prior history of TB, 7) recent exposure to TB, and 8) history of positive PPD. Conversely, patients with age > 65 years and diabetes had a significantly lower risk of having TB. Thus, we present the accuracy and relative risk of these two variables with an inverse code, i.e., the association of TB with age < 65 years and absence of diabetes (see table 2).

At least one of the risk factors significantly associated with TB was present in 59 (98.3%) patients with TB and 6,256 (88.6%) patients without TB. The presence of at least one of the risk factors significantly associated with TB had a sensitivity of 98.3%, specificity of 11.4%, positive predictive value of 0.9%, and negative predictive value of 99.9% for the diagnosis of TB. Thus, in a patient without any of the CDC risk factors, TB is unlikely.

**CAPO TB Risk Score**

The CAPO TB Risk Score consisted of the following predictive variables: 1) night sweats, 2) hemoptysis, 3) combined weight loss/< 10% of ideal body weight, 4) combined prior history of TB/recent exposure to TB/history of positive PPD, and 5) upper lobe infiltrate localization. We excluded age < 65 years because it did not substantially change the model. Of the 60 patients with CAP due to TB, none had all six risk factors included in the CAPO TB Risk Score. Of the 6,916 patients with CAP not due to TB, 28% did not have any of the six risk factors included in the CAPO TB Risk Score.
The percentage of patients with CAP based on the number of CDC and the CAPO TB Risk Score risk factors is depicted in Figure 1.

Comparison of Scores to Predict CAP Due to TB

The area under the ROC curve of the CAPO TB Risk Score (0.89; 95% CI: 0.85 to 0.93; P < 0.001) was significantly higher as compared with a model that included the CDC TB Risk Score as the predictive variable (0.71; 95% CI: 0.64 to 0.78; P < 0.001). Figure 2 displays the receiver operating characteristic curves of the risk scores.

DISCUSSION

This study indicates that most of the CDC TB risk factors have low sensitivity for predicting TB in hospitalized patients with CAP. Using a subset of CDC TB risk factors, as well as the presence of an upper lobe infiltrate, we have developed a simpler and more accurate prediction score: the CAPO TB Risk Score. This large dataset allowed us to perform robust statistical analyses and provide new findings to the pre-existing literature. This new score could be used as a tool for healthcare practitioners at the bedside for predicting the risk of TB in hospitalized patients with CAP.

Establishing the factors that are associated with an increased risk of TB in hospitalized patients with CAP is important for several reasons. First, recognizing these factors in hospitalized patients should prompt respiratory isolation to prevent nosocomial transmission of TB. Outbreaks in health-care settings have been associated with delayed
A challenge for the prevention of nosocomial transmission of TB is that institutions do not always comply with international standards of TB control.[9-12] Second, patients with these factors should be scrutinized for *Mycobacterium tuberculosis* infection to allow a more timely diagnosis of TB. Third, if empirical antibiotics for CAP are to be administered these factors suggest preference should be given to antibiotics that do not decrease the yield of diagnostic tests for *Mycobacterium tuberculosis*. In this light, there is evidence that empiric treatment with fluoroquinolones leads to a delay in the onset of anti-tuberculosis treatment.[18,19]

When assessing the probability of TB in hospitalized patients with pneumonia, it is important to use a sensitive tool to minimize the consequences of missing a diagnosis of pulmonary TB. Applying the presence of at least one of the CDC TB risk factors as criteria for suspecting TB and initiating respiratory isolation, one would have missed only one case of TB in the cohort of almost 7,000 hospitalized patients with CAP included in our study. However, such a highly sensitive tool comes at the cost of low specificity with 88.6% of the patients without TB receiving respiratory isolation and workup for TB, which may be infeasible. Using Poisson regression with robust error variance, we created a risk score based on five variables with high discrimination value for diagnosing TB. This model compared favorably with a model that had the sum of CDC TB risk factors as the predictive value.
While only a small proportion of hospitalized patients presenting with CAP had a
diagnosis of TB in our study population (0.86%), this may not be the same in areas with
high incidence of TB. For instance, in a cohort of 346 patients hospitalized for CAP in
Malaysia, 17 (5%) patients had pulmonary TB, and *Mycobacterium tuberculosis* was the
fourth most frequently identified pathogen.[20]

Few other studies evaluated risk factors for predicting a diagnosis of TB in patients with
pneumonia. Gaeta et al performed a case-control study which included patients that
presented to an urban emergency department with a clinical picture consistent with
pneumonia. Their study included 50 patients with TB that were gender and age-matched
to 50 patients without TB. They found that the following factors were significantly
associated with a diagnosis of TB: HIV infection, injection drug use, recent PPD
conversion, prior history of TB, hemoptysis, and chest x-ray consistent with TB. They
subsequently used these risk factors as criteria for respiratory isolation in a validation
sample of 103 patients, 22 of whom had TB. They found that the use of any of the
criteria for the diagnosis of TB provided a sensitivity of 96%, specificity of 14%, positive
predictive value of 23%, and negative predictive value of 92% for diagnosing TB.[21]
Unlike our study, Gaeta et al did not attempt to perform multivariate analysis of the
variables associated with TB. This may explain some difference in the variables used in
our TB Risk Score and the criteria they used for respiratory isolation. For instance,
although HIV/AIDS was a significant risk factor for TB on univariate analysis in our
study, it was no longer significant on multivariate analysis.
In a study with a design similar to ours, Liam et al evaluated clinical features that can distinguish TB from non-TB in 346 patients hospitalized with CAP. Seventeen of these patients had TB. The emphasis in their study was on physiological variables. The following features were significantly associated with a diagnosis of TB: duration of symptoms of more than 2 weeks before hospital admission, night sweats, upper lobe involvement in the chest radiograph, total white blood cell count on admission $\leq 12 \times 10^9$/L, and lymphopenia. Patients with age $< 40$ years had lower risk of TB.[20] A distinguishing feature of their study is that half of their patients with TB had disease duration longer than 2 weeks as compared with only 7% of patients with non-tuberculosis pneumonia. In the Liam study, chronic duration was heavily weighted in the group with TB, making the comparison of CAP due to TB and non-TB less compelling.

While our study and that of other investigators [20,21] evaluated risk factors for a diagnosis of TB in a cohort of patients presenting with CAP, others have assessed clinical and radiological findings of tuberculous pneumonia (as opposed to cavitary TB) in a cohort of adult patients with TB. For instance, in a cohort of 2,228 patients with tuberculosis in Brazil, 59 (2%) patients presented with non-cavitary pulmonary consolidation, which occurred predominantly in the upper lobes. [22] In another study of 16 patients with tuberculous pneumonia, consolidations occurred more evenly in the lung lobes, and these patients had more fever, less hemoptysis, and shorter duration of disease as compared with patients with cavitary TB.[23] In TB patients from our study, only one had a cavitary lesion, and most lesions were located in the upper lobe. The predominance of non-cavitary lesions in our patient population perhaps supports the notion that the
absence of cavitation is more common when TB presents acutely as a result of primary infection, although this has been disputed by other investigators.[24,25] Using DNA fingerprinting technique, Geng et al established that the main determinant of TB radiographic appearance is the patient’s immune status. Time from infection was not a predictor of radiographic appearance in multivariate analysis.[25]

Age > 65 years and diabetes are factors that recognizably increase the risk of TB. However, in our study these factors were protective against it. The association in our study might be explained by residual confounding. Also, this finding should be interpreted in the context of the study population. Our study population included only patients presenting with CAP. Older age and presence of diabetes are themselves risk factors for CAP in general.[26,27] Thus, it is conceivable that age > 65 years and diabetes remain risk factors for TB but not when the comparison is made with patients with CAP due to other etiologies. Furthermore, one could ponder that TB may have a less florid presentation in patients with diabetes, and thus it may be less likely to present as CAP in these patients. Finally, the elderly may have a different presentation due to a higher rate of reactivation disease or altered immune status. In this context, it has been shown that the young are more prone to TB due to recently transmitted infection.[28]

Our study has a number of limitations. Because the study was retrospective and observational, clinicians did not use a structured research interview form when assessing their patients. This may have led to under or over estimation of particular risk factors. There is also the risk of selection bias if for instance some patients suspected of having
TB were initially excluded from the study by investigators. While our overall study sample size was large, the number of patients who had TB was moderate, making the assessment of the more infrequent risk factors imprecise. Finally, the workup for TB in hospitalized patients with CAP is variable, and therefore we cannot exclude the possibility that some of the patients had CAP due to TB, received empirical treatment with fluoroquinolones, and were then misclassified as having CAP not due to TB. 

Our study also has several strengths. It is the first study to examine the CDC TB risk factors in hospitalized patients with CAP. This study is also multicenter and multinational, making the findings more generalizable. It has an overall large sample size compared with other studies assessing risk factors for TB in patients with pneumonia, and it assesses more risk factors for TB than prior studies.[20,21]

The available literature on the value of risk factors for diagnosing TB in hospitalized patients is still scarce. Future studies are needed to prospectively evaluate the clinical practice value of the risk factors found to be significantly associated with TB in our study, and the newly built regression model.

In conclusion, our findings indicate that using the CDC risk factors for TB as a tool to isolate patients hospitalized with CAP will identify most patients with CAP due to TB, but will unnecessarily isolate an unacceptably large number of patients. The newly built CAPO TB Risk Score has high accuracy for the diagnosis of TB and is a promising tool to help clinicians in the decision making of isolating hospitalized patients with CAP at risk for TB.
REFERENCES


15. Introduction to the Core Curriculum on Tuberculosis: What the Clinician should Know. CDC, 2011.


Table 1. Demographic and clinical characteristics of the patients.

<table>
<thead>
<tr>
<th></th>
<th>TB CAP, n= 60</th>
<th>Non-TB CAP, n= 6916</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>40.7 (16.9)</td>
<td>64.7 (19.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Male gender, number (%)</td>
<td>35 (58.3)</td>
<td>4171 (60.3)</td>
<td>0.76</td>
</tr>
<tr>
<td>Cavitary lesion, number (%)</td>
<td>1 (1.7)</td>
<td>24 (0.34)</td>
<td>0.19</td>
</tr>
<tr>
<td>Upper lobe infiltrate</td>
<td>35 (58.3)</td>
<td>1485 (21.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Pleural effusion, number (%)</td>
<td>10 (16.7)</td>
<td>1490 (21.1)</td>
<td>0.4</td>
</tr>
</tbody>
</table>

TB, tuberculosis; CAP, community-acquired pneumonia.
Table 2. Accuracy and relative risk of the factors significantly associated with TB presenting as pneumonia

<table>
<thead>
<tr>
<th>Factor</th>
<th>TB CAP</th>
<th>Non-TB CAP</th>
<th>Sens</th>
<th>Spec</th>
<th>RR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &lt; 65 yo</td>
<td>55</td>
<td>3051</td>
<td>91.7</td>
<td>55.9</td>
<td>13.7 (5.5 to 34.2)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Night sweats</td>
<td>18 (30%)</td>
<td>298</td>
<td>30.0</td>
<td>95.7</td>
<td>9.2 (5.4 to 15.9)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Hemoptysis</td>
<td>12 (20%)</td>
<td>291</td>
<td>20.0</td>
<td>95.8</td>
<td>5.5 (3.0 to 10.2)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>15 (25%)</td>
<td>436</td>
<td>25.0</td>
<td>93.7</td>
<td>4.8 (2.7 to 8.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Absence of Diabetes</td>
<td>56</td>
<td>5721</td>
<td>93.3</td>
<td>17.3</td>
<td>2.9 (1.1 to 8.0)</td>
<td>0.02</td>
</tr>
<tr>
<td>10 % or less of ideal body weight</td>
<td>12 (20%)</td>
<td>251</td>
<td>20.0</td>
<td>96.4</td>
<td>6.4 (3.4 to 11.9)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prior history of TB</td>
<td>5 (8.3%)</td>
<td>228</td>
<td>8.3</td>
<td>96.7</td>
<td>2.6 (1.1 to 6.5)</td>
<td>0.049</td>
</tr>
<tr>
<td>Recent exposure to</td>
<td>4 (6.7%)</td>
<td>32</td>
<td>6.7</td>
<td>99.5</td>
<td>13.7 (5.3 to 36)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>TB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>History of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive PPD</td>
<td>3 (5%)</td>
<td>88</td>
<td>5</td>
<td>98.7</td>
<td>4.0 (1.3 to 12.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.3%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight loss</td>
<td>27 (45%)</td>
<td>517</td>
<td>45</td>
<td>92.5</td>
<td>9.7 (5.9 to 16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.5%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TB, tuberculosis; CAP, community-acquired pneumonia; Sens, sensitivity; Spec, specificity; RR, relative risk.
Figure 1. Percentage of patients with CAP based on the number of CDC and the CAPO TB Risk Score risk factors.
Figure 2. Receiver operating characteristic curves of 2 risk scores to predict tuberculosis in hospitalized patients with community-acquired pneumonia