A decision tree for differentiating tuberculous from malignant pleural effusions

José M. Porcel a,*, Carmen Alemán b, Silvia Bielsa a, Javier Sarrapio b, Tomás Fernández de Sevilla b, Aureli Esquerda c

a Pleural Disease Unit, Department of Internal Medicine, Arnau de Vilanova University Hospital, IRBLLEIDA, Avda Alcalde Rovira Roure 80, 25198 Lleida, Spain
b Department of Internal Medicine, Hospital General Vall d’Hebron, Autonomous University, Barcelona, Spain
c Pleural Disease Unit, Department of Laboratory Medicine, Arnau de Vilanova University Hospital, IRBLLEIDA, Avda Alcalde Rovira Roure 80, 25198 Lleida, Spain

Received 5 September 2007; accepted 4 March 2008
Available online 24 June 2008

KEYWORDS
Decision tree; Algorithm; Pleural effusion; Tuberculosis; Cancer

Summary
Objective: To improve physicians’ ability to discriminate tuberculous from malignant pleural effusions through a simple clinical algorithm that avoids pleural biopsy.
Design: We retrospectively compared the clinical and pleural fluid features of 238 adults with pleural effusion who satisfied diagnostic criteria for tuberculosis (n = 64) or malignancy (n = 174) at one academic center (derivation cohort). Then, we built a decision tree model to predict tuberculosis using the C4.5 algorithm. The model was validated with an independent sample set from another center that included 74 tuberculous and 293 malignant effusions (validation cohort).
Results: Among 12 potential predictor variables, the classification tree analysis selected four discriminant parameters (age > 35 years, pleural fluid adenosine deaminase > 38 U/L, temperature > 37.8°C, and pleural fluid LDH > 320 U/L) from the derivation cohort. The generated flowchart had 92.2% sensitivity, 98.3% specificity, and an area under the ROC curve of 0.976 for diagnosing tuberculosis. The corresponding operating characteristics for the validation cohort were 85.1%, 96.9% and 0.958.
Conclusions: Applying a decision tree analysis that contains simple clinical and laboratory data can help in the differential diagnosis of tuberculous and malignant pleural effusions.

Introduction

Tuberculosis and cancer represent the two most frequent causes of exudative pleural effusions with predominantly mononuclear cells in pleural fluid.1,2 Unfortunately,
isolation of *Mycobacterium tuberculosis* in pleural fluid is difficult (as well as a late event) because tuberculous pleurisy is primarily an immunological process with a small number of tuberculous bacilli. The demonstration of granuloma in a biopsy specimen from the parietal pleura suggests tuberculous pleuritis. However, this procedure has been questioned because of the availability of pleural fluid surrogate markers, such as adenosine deaminase (ADA) and interferon-gamma, which are accurate enough in supporting a diagnosis of tuberculous pleuritis. On the other hand, a major obstacle in diagnosing malignant effusions is the presence of false negative cytological results in about 40% of cases. Needle biopsy of the pleura has a low sensitivity in detecting malignancy and thoracoscopy, which can be a definitive procedure for both tuberculosis and neoplasm, is invasive and not widely available.

Decision analysis techniques are a systematic approach to decision making in complex situations under conditions of uncertainty. A decision tree is a flowchart for modeling a decision analysis. It consists of a starting point (i.e., the clinical question to be addressed in the analysis) as well as branching points (i.e., points at which alternatives become possible). To our knowledge, no study has previously addressed the common clinical dilemma of differentiating tuberculosis from malignant pleural effusions through a decision tree analysis.

The present study uses classification tree analysis to develop a clinical algorithm for discriminating tuberculous from malignant pleural effusions. We hypothesized that a simple flowchart incorporating basic clinical and pleural fluid data can predict with high sensitivity and specificity the probability of tuberculosis, thus preventing in most cases the performance of invasive diagnostic procedures.

**Materials and methods**

**Patients and measurements**

The local ethics committee of the two participating centers approved this study and all subjects signed written informed consent. We retrospectively compared clinical (gender, age, temperature), radiological (effusion size and laterality), and pleural fluid (total leukocyte count, differential white cell count, glucose, protein, lactate dehydrogenase (LDH), ADA and pH) data from patients with tuberculous and malignant effusions. Using these 12 variables, we then conducted classification tree analysis to mathematically derive an algorithm that accurately discriminated between the two conditions.

The sample used to develop the decision tree (derivation cohort) included all patients with demonstrated tuberculous and malignant effusions identified at the Vall d’Hebron University Hospital in Barcelona (Spain) from 1995 to 2006. Data collected from tuberculous and malignant effusion cases at the Arnau de Vilanova University Hospital in Lleida (Spain) during the same period of time were used to validate the decision tree (validation cohort).

A pleural effusion was categorized as malignant if malignant cells were demonstrated in pleural fluid or pleural biopsy. Tuberculous pleuritis was diagnosed if Ziehl–Neelsen stains or Lowenstein cultures of pleural fluid, sputum, or pleural biopsy tissue samples were positive or a pleural biopsy specimen showed granulomas in the parietal pleura. Fever was defined as a temperature $\geq 37.8$ °C, and pleural effusions were deemed to be large if they occupied two-thirds or more of the hemithorax.

Biochemical measurements on pleural fluid samples were carried out on discrete analyzers (Hitachi models 717, 917 or Modular DP, Roche Diagnostics, Mannheim, Germany) using standardized photometric methods. Specifically, pleural ADA activity and pH were assessed with an automated ultraviolet kinetic test (Roche Diagnostics, Barcelona, Spain) and a blood gas machine, respectively. White blood cells were manually counted in a Thoma chamber.

**Statistical analysis**

Continuous data are presented as medians (quartiles) according to their non-parametric distribution. We used Fisher’s exact test or Mann–Whitney *U* test for group comparisons of categorical and continuous variables, respectively.

Discrimination of tuberculous and malignant effusions was accomplished using the C4.5 method for decision tree analysis implemented through Weka 3.4 statistical software (The University of Waikato, New Zealand). Classification trees discriminate between outcome classes (e.g., tuberculosis vs. malignancy) by first searching the range of each potential predictor (e.g., a given pleural fluid parameter) and finding the split that maximizes the likelihood of the given data set. Within each resulting subset (or node), the algorithm again searches the range of each variable to choose the optimal split. This process is continued until all observations are perfectly discriminated, or the sample size within a given node is too small to divide further (e.g., *n* = 5 or less). The final output of the resulting classification tree is a graphical display of decision criteria for each split as well as the resulting predicted probabilities of being a given case across the final splits (i.e., terminal nodes). A 10-fold internal cross-validation analysis was performed as an initial evaluation of the test error of the algorithm. Briefly, this process involved splitting up the data set into 10 random segments and using nine of them for training and the 10th as a test set for the algorithm. Of note, for using the C4.5 decision tree, data must be available for all variables in the model.

Finally, we evaluated the performance of the decision algorithm with respect to sensitivity, specificity, likelihood ratios and area under the receiver operating characteristic curve (AUC), both in the derivation and the validation cohorts. *P* values $< 0.05$ were considered statistically significant. Calculations other than the decision tree were performed with a statistical software package (SPSS version 11.5; Chicago, IL, USA).

**Results**

A total of 64 tuberculous effusions and 174 malignant effusions formed the learning set (derivation cohort), which generated the decision algorithm. The decision tree was
then tested using collected data from 74 new tuberculous effusions and 293 malignant effusions (validation cohort). Previously, 19 and 25 cases from the derivation and validation cohorts, respectively, were excluded from the analysis, because data were not available on all variables used in the C4.5 decision tree.

Table 1 summarizes the methods used for diagnosing tuberculous and malignant effusions. The malignant pleural effusions in the derivation cohort included lung (71), unknown primary (21), breast (18), hematologic (16), mesothelioma (14), gynecologic (11), gastrointestinal (8), and other tumors (15). Likewise, primary sites of the tumor in the validation cohort were as follows: lung (98), breast (61), hematologic (31), unknown primary (30), gynecologic (30), mesothelioma (12), gastrointestinal (8), and others (23).

Characteristics of the tuberculous and malignant effusions used to create and test the decision tree are shown in Tables 2 and 3, respectively. Except for the laterality of pleural effusions, both cohorts were homogeneous.

Fig. 1 displays the decision tree produced by the C4.5 method to discriminate tuberculous from malignant effusions. The classification tree identified the following predictors of tuberculous pleurisy in order of importance: age < 35 years, pleural fluid ADA > 38 U/L, presence of fever and pleural fluid LDH > 320 U/L. In the derivation set group, the tree had a sensitivity of 92.2% (95% CI, 85.6–98.8%), a specificity of 98.3% (95% CI, 96.3–100%), and an AUC of 0.976 (95% CI, 0.946–1) for identifying effusions of tuberculous origin. Similarly, the decision tree performed with 85.1% sensitivity (95% CI, 77–93.2%), 96.9% specificity (95% CI, 95–98.9%), and AUC of 0.958 (95% CI, 0.929–0.987) in the validation cohort. In this latter group, only 9 of 293 (3%) malignant effusions (six lymphomas, one lung adenocarcinoma, one breast cancer, and one cancer of unknown origin) would have been misclassified as tuberculosis.

Discussion

Based on readily available clinical and pleural fluid data, we developed a very simple and accurate decision model for the differential diagnosis of tuberculous and malignant effusion, a key dilemma when confronted with a lymphocytic pleural exudate. Although the gold standard for the diagnosis of tuberculous pleurisy remains microbiological (i.e., demonstration of tubercle bacilli in the pleural fluid, sputum or pleural biopsy specimen) or histological (i.e., demonstration of pleural granulomas), the proper diagnostic approach is still debated. Central to the controversy is the role of needle biopsy of the pleura vs. the use of pleural fluid ADA or, less commonly, interferon-gamma measurements alone.9

Avoidance of closed pleural biopsy is clinically attractive. In this regard, few studies have evaluated the utility of clinical and laboratory data in diagnosing tuberculous pleural effusion on the basis of mathematical predictive models.10–13 In an early study from Spain, 47 variables...
obtained from the medical records of 78 patients with tuberculous and 111 with non-tuberculous effusions were entered in a stepwise discriminant analysis. The most powerful predictor of tuberculous effusion was a discriminant function (formula) that included the patient’s age (years), tuberculin skin test (mm of induration at 48 h), nant function (formula) that included the patient’s age < 35 years (two points), temperature > 37.8 °C (two points), and pleural red blood cell count < 6000 cell/mm³) performed the best, in a logistic regression model, to discriminate 104 tuberculosis from 111 non-tuberculous effusions (AUC 0.991). External validation of the results from this study was lacking. In addition, a prediction rule using a complicated fractional polynomial equation to distinguish tuberculosis from malignancy, as in the above-mentioned studies, would require automation to be applied in the clinical setting. A third study devised a multivariate scoring system to discriminate pleural effusions caused by tuberculosis from those caused by malignancy. Four parameters were selected and scored by the model as predictive of tuberculosis, namely pleural fluid ADA > 40 U/L (five points), age ≤ 35 years (two points), temperature ≥ 37.8 °C (two points), and pleural red blood cell count ≤ 5 × 10⁹/L (two points). Summated scores of ≥5 yielded excellent discriminating characteristics (sensitivity 95%, specificity 94%, AUC 0.987). The study was flawed by the inclusion of a significant number of probable rather than certain cases in both the tuberculous and malignant groups. This study, also, was not validated. Lastly, a recent prospective study compared seven biological markers in the pleural fluid of 45 patients with malignant, 15 with parapneumonic and 12 with tuberculous pleural effusions. The application of a multinomial logit model revealed ADA and C-reactive protein as the most important parameters for discriminating between groups. Specifically, the combination of pleural fluid ADA > 42.4 U/L and C-reactive protein < 5.5 mg/dL resulted in the correct classification of 91.7% of tuberculous effusions. Limitations of this study were the small sample size and the absence of lymphomatous effusions among the malignant group.

The present investigation supports some previous findings, particularly the diagnostic utility of the ADA measurement as a marker of tuberculous pleurisy. The inclusion of other parameters in the classification tree, such as age or fever, is logical. It is accepted that tuberculosis tends to occur in younger patients and that most of them are febrile as compared with patients with a malignancy. Paradoxically, a high pleural LDH level was selected in the last tree node as indicative of malignant effusion, despite the greater mean concentrations of this enzyme in tuberculosis effusions.

**Table 3** Patients’ characteristics within the validation cohort

<table>
<thead>
<tr>
<th>Clinical data</th>
<th>Tuberculous effusions (n = 74)</th>
<th>Malignant effusions (n = 293)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>33 (25–44.3)</td>
<td>70 (60.5–77)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Male gender</td>
<td>55/74 (74.3%)</td>
<td>146/293 (49.8%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fever</td>
<td>55/74 (74.3%)</td>
<td>21/293 (7.2%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Large effusion</td>
<td>3/71 (4.2%)</td>
<td>50/250 (17.1%)</td>
<td>0.006</td>
</tr>
<tr>
<td>Unilateral effusion</td>
<td>68/71 (95.8%)</td>
<td>231/274 (84.3%)</td>
<td>0.010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pleural fluid data</th>
<th>Tuberculous effusions</th>
<th>Malignant effusions</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukocytes, ×10⁹/L</td>
<td>1.59 (0.54–3.27)</td>
<td>0.88 (0.40–1.71)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>88 (71–95)</td>
<td>84 (61–94)</td>
<td>0.166</td>
</tr>
<tr>
<td>Glucose, mg/dL</td>
<td>77 (55–95)</td>
<td>108 (86–133)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Protein, g/L</td>
<td>53.1 (47.6–57.3)</td>
<td>43.7 (38.2–49)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LDH, U/L</td>
<td>911 (675–1703)</td>
<td>545 (340–1016)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ADA, U/L</td>
<td>63.9 (50.4–91.2)</td>
<td>16.0 (10.2–23.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>pH</td>
<td>7.36 (7.30–7.43)</td>
<td>7.41 (7.34–7.47)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

LDH, lactate dehydrogenase; and ADA, adenosine deaminase.

**Figure 1** A decision tree for tuberculous and malignant effusion discrimination. Inside the boxes is the number of tuberculous (TB) and malignant (M) effusions from the derivation group. PF, pleural fluid; ADA, adenosine deaminase; and LDH, lactate dehydrogenase.
than in malignancy. Probably, the weight of this variable is negligible because 94% of effusions in both cohorts had already been classified before arriving at this point in the tree. Our classification tree performed as well as logistic regression models in identifying tuberculosis, but is much easier to interpret for clinical use. Unlike previous studies, the current proposed four-variable algorithm model was validated using an independent sample of patients, thereby demonstrating a very reliable discriminating ability (sensitivity 85.1%, specificity 96.9%, AUC 0.958), leading to the correct classification (accuracy) of 94.6% of patients. These figures are comparative to those obtained from the combination of histology and culture in pleural biopsy specimens for diagnosing tuberculous pleurisy.

Notably, lymphoma was the most frequently misclassified (6/31, 19%) tumor type after fitting the decision tree. This was not an unexpected finding, since hematologic malignancies have higher mean pleural fluid ADA levels than any other neoplasm invading the pleural surfaces.

This study has some limitations. First, the decision tree is valid only for the differential diagnosis of tuberculous and malignant effusions. However, other non-tuberculous pleural effusions, such as those secondary to heart failure or pneumonia, are easily diagnosed on clinical grounds. Second, our study was retrospective, but the potential impact of this limitation is minimal because the decision tree includes only objective clinical characteristics and laboratory parameters. Furthermore, we used strict criteria to define the outcome variable (tuberculosis vs. malignancy) to minimize misclassification bias. Finally, the prevalence of both tuberculosis and malignant effusions will affect the performance of the rule. The latest available data (2005) shows that the prevalence of tuberculosis in Spain is 22/100,000 inhabitants/year, while in the United States it is 3.4/100,000 inhabitants/year, far from the 15% of all pleural effusions submitted to thoracentesis. Although an evaluation of the tree performance in areas with low-pretest probability of tuberculosis is desirable, such a study is difficult to accomplish. However, even in those cases, ADA will still retain its high negative predictive value.

The general disadvantage for diagnosing pleural tuberculosis on the basis of pleural fluid biochemistries (i.e., ADA) is the failure to provide antituberculous susceptibility data, given concerns for drug-resistant forms of the disease. In Catalonia, the geographical region of Spain where the two participant centers are located, only 24 of 735 (3.2%) M. tuberculosis isolates exhibited any first-line drug resistance in tests performed in 2004. In the United States, a recent epidemiological survey found that 9.9% of patients with pleural tuberculosis had isolates that were resistant to at least one first-line drug, although multidrug resistance (resistance to both isoniazid and rifampin) was only 1%. In areas with low drug-resistance rates, the diagnosis of tuberculous pleurisy based solely upon pleural fluid ADA measurement seems reasonable.

In conclusion, the generated decision tree retains intuitive appeal and can easily be applied by the clinician. Our findings can be briefly summarized: a young patient with fever and a lymphocytic pleural exudate with high ADA activity and negative cytological studies has a tuberculous pleurisy, until proven otherwise. Accordingly, we advocate the initiation of an antituberculous therapy in this particular setting. Like all prediction rules, this one should be used with appropriate clinical judgment. In addition to pleural fluid cultures, sputum induction should also be obtained, when possible. Yet, invasive diagnostic procedures, such as pleural biopsy, could be reserved for cases unresponsive to treatment. The proposed clinical rule may be adequate for areas with a prevalence of tuberculosis comparable or greater to that reported in Spain, but less tenable if a lower disease prevalence or higher drug-resistant rates exist.

**Conflict of interest**

The authors have no conflict of interest.

**References**


