



## Original Article

## The Combined Use of Pleural Fluid Parameters Improves Diagnostic Accuracy in Complicated Pleural Infection/Empyema<sup>☆</sup>



Lucía Ferreiro,<sup>a,\*</sup> Óscar Lado-Baleato,<sup>b</sup> Juan Suárez-Antelo,<sup>c</sup> María E. Toubes,<sup>c</sup> María E. San José,<sup>d</sup> Adriana Lama,<sup>c</sup> Nuria Rodríguez-Núñez,<sup>c</sup> José Manuel Álvarez-Dobaño,<sup>a</sup> Francisco J. González-Barcala,<sup>a</sup> Jorge Ricoy,<sup>c</sup> Francisco Gude,<sup>b</sup> Luis Valdés<sup>a</sup>

<sup>a</sup> Servicio de Neumología, Complejo Hospitalario Clínico Universitario de Santiago de Compostela, Grupo Interdisciplinar de Investigación en Neumología, Instituto de Investigaciones Sanitarias de Santiago (IDIS), Santiago de Compostela, A Coruña, Spain

<sup>b</sup> Unidad de Epidemiología Clínica, Complejo Hospitalario Clínico Universitario de Santiago de Compostela, Grupo Epidemiología de Enfermedades Frecuentes, Instituto de Investigaciones Sanitarias de Santiago (IDIS), Santiago de Compostela, A Coruña, Spain

<sup>c</sup> Servicio de Neumología, Complejo Hospitalario Clínico Universitario de Santiago de Compostela, Santiago de Compostela, A Coruña, Spain

<sup>d</sup> Servicio de Análisis Clínicos, Complejo Hospitalario Clínico Universitario de Santiago de Compostela, Grupo Interdisciplinar de Investigación en Neumología, Instituto de Investigaciones Sanitarias de Santiago (IDIS), España, Santiago de Compostela, A Coruña, Spain

## ARTICLE INFO

## Article history:

Received 3 January 2019

Accepted 28 February 2019

Available online 1 October 2019

## Keywords:

Pleural effusion

Pleural infection

Complicated pleural infection/empyema

Diagnosis

Inflammatory markers

Logistic regression analysis

## ABSTRACT

**Introduction:** Identifying infectious pleural effusions (IPE) that will progress to complicated infection or empyema is challenging. The purpose of this study was to determine whether a model based on multiple biochemical parameters in pleural fluid can predict which IPEs will produce empyema.

**Methods:** A prospective study was performed of all cases of IPEs treated in our unit. IPEs were classified as uncomplicated or complicated (empyema). Logistic regression was used to estimate the risk for complicated pleural infection (empyema). A predictive model was developed using biochemical parameters in pleural fluid. Discriminatory power (areas under the ROC curve), calibration, and diagnostic accuracy of the model were assessed.

**Results:** A total of 177 patients were included in the study (74 with uncomplicated infectious pleural effusion, and 103 with complicated pleural effusion/empyema). The area under the curve (AUC) for the model (pH, lactate dehydrogenase and interleukin 6) was 0.9783, which is significantly superior to the AUC of the individual biochemical parameters alone (0.921, 0.949, and 0.837, respectively;  $P < .001$  using all parameters). The rate of correct classification of infectious pleural effusions was 96% [170/177: 72/74 (97.3%) for uncomplicated and 98/103 (95.1%) for complicated effusion (empyema)].

**Conclusion:** The multiple-marker model showed better diagnostic performance for predicting complicated infectious pleural effusion (empyema) compared to individual parameters alone.

© 2019 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

### La combinación de la determinación de parámetros bioquímicos del líquido pleural mejora la predicción diagnóstica de la infección pleural complicada/empiema

## RESUMEN

**Introducción:** Predecir cuándo un derrame pleural infeccioso puede evolucionar hacia una infección complicada/empiema es difícil de establecer. Nuestro propósito es analizar si un modelo predictivo construido con parámetros bioquímicos del líquido pleural puede ayudar a identificar estos derrames.

**Métodos:** Se estudió de forma prospectiva a todos los pacientes diagnosticados de derrame pleural infeccioso y se clasificaron en no complicados y complicados/empiemas. Se realizó un análisis de regresión logística para estimar la probabilidad de infección pleural complicada/empiema. Con base en parámetros

## Palabras clave:

Derrame pleural

Infección pleural

Infección pleural complicada/empiema

Diagnóstico

<sup>☆</sup> Please cite this article as: Ferreiro L, Lado-Baleato Ó, Suárez-Antelo J, Toubes ME, San José ME, Lama A, et al. La combinación de la determinación de parámetros bioquímicos del líquido pleural mejora la predicción diagnóstica de la infección pleural complicada/empiema. Arch Bronconeumol. 2019;55:565–572.

\* Corresponding author.

E-mail address: [lferfer7@gmail.com](mailto:lferfer7@gmail.com) (L. Ferreiro).

Marcadores inflamatorios  
Análisis regresión logística

bioquímicos del líquido pleural, se construyó un modelo predictivo y se determinaron su discriminación (áreas bajo la curva ROC), calibración y precisión diagnóstica.

**Resultados:** Se incluyó a 177 pacientes (74 infecciones pleurales no complicadas y 103 complicadas/empiemas). El área bajo la curva del modelo construido (pH, lactato deshidrogenasa e interleucina 6) fue 0,9783, significativamente mejor que cualquiera de las variables bioquímicas utilizadas de forma individual (0,921, 0,949 y 0,837, respectivamente;  $p < 0,001$  usando todos los parámetros). La tasa de clasificación correcta fue del 96% de los derrames (170/177; 72/74 [97,3%] de los no complicados y 98/103 [95,1%] de los complicados/empiemas).

**Conclusión:** El modelo predictivo analizado tiene una buena rentabilidad para el diagnóstico de las infecciones pleurales complicadas/empiemas, superior a la de cualquiera de las variables individuales que lo componen.

© 2019 SEPAR. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Introduction

Pleural infection (PI) is a common finding in clinical practice, as 57% of patients with pneumonia can develop infectious pleural effusion (IPE).<sup>1</sup> In the initial (exudative) phase, pleural fluid (PF) shows characteristics of neutrophilic exudate with no other distinguishing biochemical data; the yield from culture is generally low; and response to antibiotic treatment is usually favorable.<sup>2</sup> In recent years, however, increasing numbers of patients are developing complicated PI/empyema, especially children and the elderly.<sup>3,4</sup> These patients will require antibiotics and chest tube drainage, and occasionally, intrapleural fibrinolytics and surgical management.<sup>2</sup> As a result, hospitalization rates for empyema are higher,<sup>3</sup> hospital stays are longer,<sup>5</sup> and mortality rates can be as high as 30% in immunocompromised patients.<sup>3,5,6</sup>

The biochemical characteristics of PF in the fibrinopurulent and organizing phases of PI are totally different from those of the exudative phase, and placement of a chest tube is often clearly indicated.<sup>2</sup> However, no sufficiently reliable clinical, laboratory, or radiological predictors have yet been developed to identify patients with a simple PI that might evolve into a complicated PI/empyema.

The aim of this study was to assess if a predictive model based on combination of biomarkers of inflammatory response in PF can predict which patients with IPE will develop more complicated PI.

## Materials and Methods

### Patients

This was a prospective study of all patients with IPE who were seen in our department between June 2013 and December 2016. If a patient underwent several thoracenteses, the biochemical data of the first procedure were used.

### Definitions

IPE was diagnosed on the basis of lung infection associated with an acute clinical picture progressing with fever, pulmonary infiltrates, purulent sputum, and response to antibiotic treatment, or identification of an organism in PF.

PIs were considered uncomplicated if they resolved with antibiotic treatment alone, and complicated if the effusion required a chest drain for resolution; empyema was defined as pus in the pleural space or a positive PF culture.<sup>7</sup>

We excluded patients under the age of 18 years and those who had had previous empyema, major surgery in the 5 previous days, pneumonectomy on the infected side, those who were pregnant or breast-feeding, and patients with a life expectancy of less than 3 months.

### Samples

PF was obtained by ultrasound-guided thoracentesis before starting treatment and with the patient fasting. PF samples were sent for microbiology (Ziehl–Neelsen staining and culture in aerobic, anaerobic, and Löwenstein media), cytology and biochemistry studies. After samples were centrifuged at  $1500 \times g$  at  $4^\circ C$  for 15 min, the supernatants were processed within 2 h of extraction and subsequently stored at  $-80^\circ C$ . Interleukin 6 (IL-6) and C-reactive protein (CRP) were determined by IMMULITE® 2000 (Siemens Healthcare Diagnostics Products Ltd., Llanberis, Gwynedd LL55 4EL, United Kingdom). Results were expressed in pg/ml and mg/l, respectively, and were available in less than 1 h.

### Statistical Analysis

Continuous variables were expressed as mean $\pm$ SD or median (25th–75th percentiles), according to whether distributions were normal or non-normal. Qualitative variables were expressed as absolute frequency and percentage. Normal distribution of the variables was calculated using the Kolmogorov–Smirnov test. The Chi-squared test was used to compare categorical variables, and the Mann–Whitney *U* test for continuous variables. The following were estimated to determine the diagnostic yield of the parameters: sensitivity, specificity, positive predictive value, negative predictive value, positive and negative likelihood ratios, and diagnostic accuracy. Optimal cut-off points were obtained using the Youden approach.<sup>8</sup> ROC curves were constructed and areas under the curve (AUC)<sup>9</sup> were identified to determine discrimination capacity. These values were interpreted according to the following limits: 0.60–0.69: poor discriminatory value; 0.70–0.79: moderate discriminatory value; 0.80–0.89: good discriminatory value; and 0.90–1.00: excellent discriminatory value.

Logistic binary regression models were constructed, in which all potentially predictive variables were initially included, followed by the elimination of those which contributed little or nothing to the explanatory power of the model according to the likelihood ratio test ( $P < .05$ ). Odds ratios were estimated from the regression coefficients and their 95% confidence intervals. Different aspects of the performance of the model were studied, including discrimination, calibration, and diagnostic confidence.<sup>10,11</sup> Calibration was evaluated using the Brier score, and non-parametric estimates of the observed frequencies versus those predicted by the model were plotted. ROC curves and their corresponding AUC also were used to check the discrimination capacity of the models. We used the theoretical relationship between the probability of the IPE threshold and the relative value of the false positive and false negative results to determine the value of the models. For this purpose, the classification rule was defined as follows: each patient was classified as belonging to the infectious group if they presented predicted

probabilities equal or superior to the selected threshold, and negative otherwise. This means that individuals in whom the model predicted IPE but who did not have it would be false negatives, while cases in whom the model did not predict IPE and they had it were false positives. Bootstrap techniques were used to correct for possible overoptimism in both estimates of discrimination and patient classification. Finally, nomograms were built for the prediction models, based on the coefficients of the models obtained by logistic regression. Data were analyzed using the packages mgcv,<sup>12</sup> MASS,<sup>13</sup> rms,<sup>10</sup> OptimalCutpoints<sup>13</sup> and pROC,<sup>14</sup> all of which are freely available in R.<sup>15</sup>

Our study was approved by the Clinical Research Ethics Committee of the hospital (registry 2013/013).

## Results

A total of 813 thoracenteses were carried out during the study period; 107 effusions were excluded as follows: no definitive diagnosis (27), presence of exclusion criteria (33), data from repeated patients (21), possible double diagnosis (18), and incomplete data (8). Of the 706 patients who were eventually diagnosed, 177 were IPE. Of these, 74 were uncomplicated, 65 were complicated and 38 were empyema (total complicated PI/empyema, 103).

PF culture was positive in 34/177 patients (19.2%). *Streptococcus milleri* (*Streptococcus intermedius*, *constellatus*, and *mitis*) was the group most frequently cultured (13 patients; 7.3%), followed by *Streptococcus pneumoniae* (5 patients; 2.8%), *Streptococcus oralis* (4 cases; 2.3%), other *Streptococcus* types (*viridans*, group C beta-hemolytic, and *species*) (3; 1.7%), *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa* (2 cases each; 1.1%), nocardia, *Fusobacterium nucleatum*, and *Parvimonas micra* (1 case each; 0.6%). Fifty patients (28.2%) had received antibiotic treatment before thoracentesis was performed.

Table 1 shows the characteristics of the study patients according to their classification as uncomplicated PI and complicated PI/empyema. The percentages of polymorphonuclear (PMN) leukocytes, lactate dehydrogenase (LDH), CRP, IL-6 and tumor necrosis factor alpha (TNF- $\alpha$ ) in PF increased significantly in line with the complexity of the effusion, while pH decreased. If the PI was complicated or empyema developed, significant increases were observed in the size of the PE, alcoholism, positive PF cultures, the need for chest tubes and treatment with fibrinolytic agents, and the presence of septa in the pleural space. No significant differences were found in CURB-65, PSI, length of hospital stay, need for ICU admission, and mortality. In contrast, heart failure was significantly more common in patients with uncomplicated PE than in those with complicated PE/empyema (9.5% vs. 1.9%,  $P=0.029$ ). Table 2 shows the yield of each PF parameter for the diagnosis of complicated PI/empyema. The parameters that showed the best AUC for diagnosis were LDH (0.949), pH (0.921), percentage of PMN (0.866), and IL-6 (0.837).

Fig. 1 shows the distribution of the concentrations of the main variables in PF for uncomplicated and complicated PI/empyema, and Table 2 shows the ROC AUC of all variables for the diagnosis of complicated PI/empyema.

We used logistic regression analysis to differentiate the uncomplicated PIs from complicated PI/empyema. The model included 3 variables: pH, LDH, and IL-6 in PF. The coefficients obtained in the regression analysis for the study model are shown in Table 3. The discrimination capacity obtained was an AUC of 0.9783, significantly better than any of the component variables individually ( $P<0.001$ ). The bootstrap-corrected AUC was slightly lower than the apparent AUC (0.9757), which also does not reflect any optimism. This model correctly classified 96% (170/177) of the effusions

(72/74 [97.3%] of uncomplicated PIs and 98/103 [95.1%] of the complicated PI/empyema cases).

The calibration plot of the constructed model is shown in Fig. 2. The calibration shows good concordance between the estimated probability and the observed frequency. Finally, Fig. 3 shows the constructed nomogram.

Table 4 shows the yield for different cut-off points of the predicted probability for the diagnosis of a complicated PI/empyema, obtained from the constructed model. In clinical practice, the calculation of the probability that the PI is complicated/empyema can be easily performed using an Excel calculator (Supplementary material [Annex], e-Figure 1).

Fig. 4 shows the distribution of the values of pH, IL-6, and LDH in PF, by uncomplicated PI (black lines) or complicated PI/empyema (red lines), and the optimal cut-off points.

The parameters of the model constructed to differentiate uncomplicated PI from complicated PI/empyema were used to build a three-dimensional graph that shows the relationship between the values of pH, LDH and IL-6 in PE (Fig. 5). Uncomplicated PI, represented by black circles, are grouped mainly in the area of the graph occupied by the cases with lower values of LDH and IL-6 and higher levels of pH.

Table 5 shows the errors that occur (false positives and false negatives) with different cut-off points for the probabilities obtained after applying the constructed model.

## Discussion

This study confirms that a predictive model built with a combination of biochemical variables in PF has excellent discriminatory value for predicting complicated PI/empyema, differentiating them from uncomplicated PI in 96% of cases.

Differentiating between uncomplicated PI and complicated PI/empyema has prognostic and therapeutic implications. In contrast to uncomplicated PI, complicated PI and empyema rarely resolve with antibiotics alone, and usually require chest tube drainage and occasionally intrapleural fibrinolytics or even surgical management. This generates a greater burden on health resources and more hospital admissions with extended stays, which can lead to increased morbidity and mortality. If the PF appears purulent, it is relatively simple to diagnose empyema, but fluids with a serous appearance can give a positive culture and need to be resolved by chest tube drainage.<sup>16</sup> It is important to establish early on which patients might have complicated PI/empyema, but to date no clinical or radiological data are available to identify these patients. pH, LDH, or glucose values in PF can help in this differentiation, but the margin of error is considerable.<sup>2</sup> Other authors have attempted to identify these PIs using multivariate analysis, and certain clinical (age, alcoholism, chest pain, tachycardia, use of intravenous drugs) or laboratory (albumin, sodium, platelets, CRP, leukocytes) characteristics have been reported to be helpful.<sup>7,17</sup>

The variables that best discriminated complicated PI/empyema from uncomplicated cases on an individual basis were LDH (AUC 0.949), pH (AUC 0.921), percentage of PMN (AUC 0.866), and IL-6 (AUC 0.837) (Table 2). Both LDH and pH are widely used for this purpose<sup>16,18</sup> and the yield obtained can be considered to be within the ranges observed in other series.<sup>19–22</sup> The percentage of neutrophils is often used in the diagnosis of IPE. Although this parameter is specific, it is not sensitive, since factors such as the stage of disease progression at which the thoracentesis is performed and previous use of antibiotics may influence the outcome.<sup>23</sup> However, it has given a good diagnostic yield in the diagnosis of complicated PI/empyema (cut-off point 61.1%; sensitivity 79.1%; specificity 85.5%; AUC 0.866), probably because all cases are at an advanced stage of infection (fibrinopurulent

**Table 1**  
Main Characteristics of Patients With Infectious Pleural Effusion.

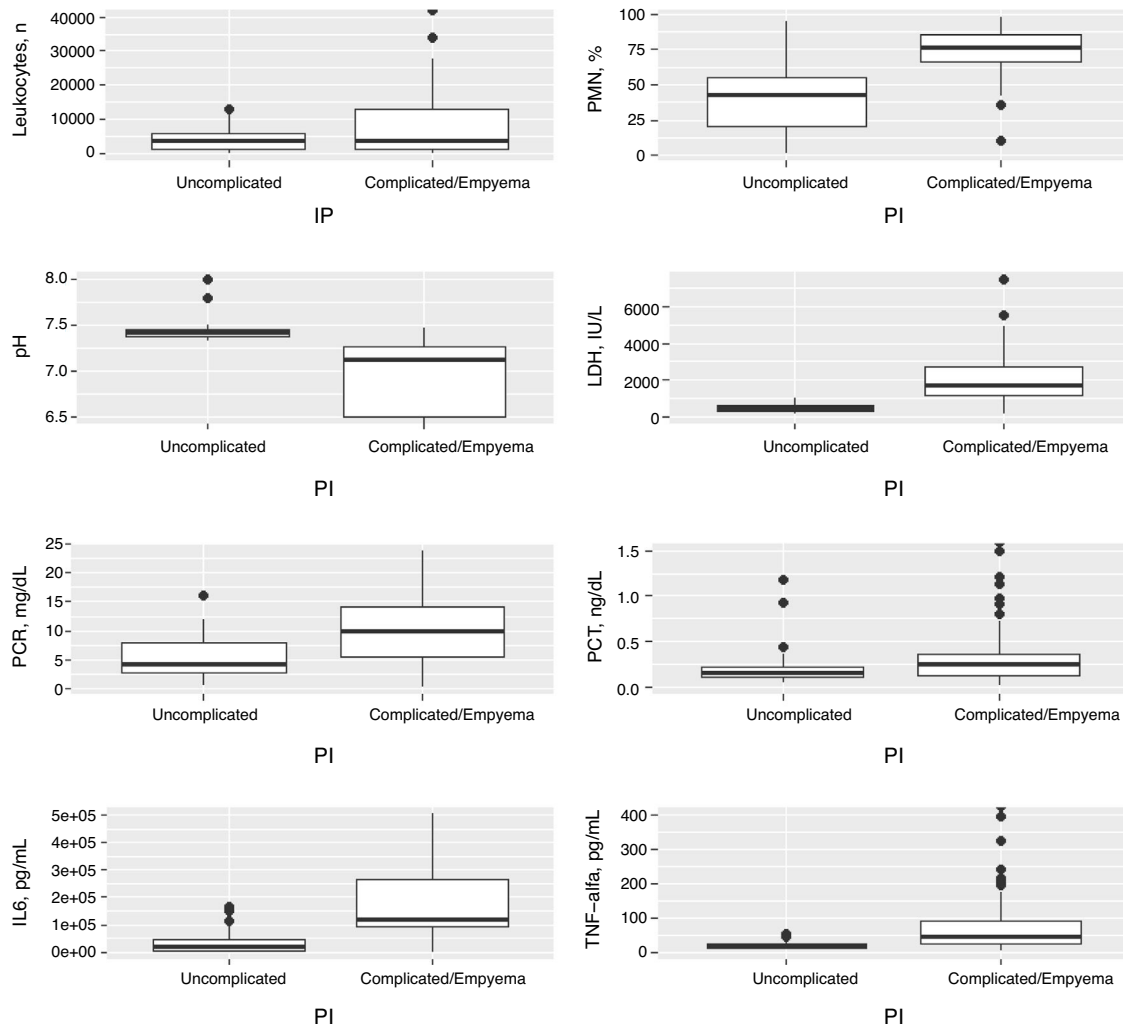
	Uncomplicated PI	Complicated PI/empyema	P
<i>n</i>	74	103	
Age, years	67.5 (49.8, 82.3)	59 (50, 76)	0.070
Men, <i>n</i> (%)	42 (56.8)	84 (81.6%)	0.000
Leukocytes in PF (cells/ $\mu$ l)	4010 (1752, 6688)	3760 (1213, 17,353)	0.317
PMN (%)	31 (15.5, 55.5)	76 (65, 85)	0.000
pH	7.42 (7.38, 7.49)	7.08 (6.52, 7.26)	0.000
LDH in PF (U/l)	395.5 (249.3, 592.5)	1958 (1127, 4749)	0.000
CRP in PF, mg/dl	3.9 (2.3, 6.3)	9 (5.1, 14.4)	0.000
PCT PF (ng/ml)	0.15 (0.09, 0.23)	0.23 (0.08, 0.37)	0.180
IL-6 in PF (pg/ml)	9807 (2593, 37,252)	100,000 (54,724, 226,400)	0.000
TNF- $\alpha$ in PF (pg/ml)	17.5 (11.6, 22.5)	45.4 (24.9, 101.3)	0.000
CURB-65, <i>n</i> (%)			0.583
1	38 (52.8)	62 (60.2)	
2	28 (38.9)	35 (34)	
3	6 (8.3)	6 (5.8)	
PSI class, <i>n</i> (%)			0.475
I	9 (12.5)	10 (9.7)	
II	13 (18.1)	22 (21.4)	
III	17 (23.6)	328 (31.1)	
IV	29 (40.3)	30 (29.1)	
V	4 (5.6)	9 (8.7)	
Effusion size, <i>n</i> (%)			0.000
<1/3	53 (71.6)	30 (29.1)	
>1/3 and <2/3	19 (25.7)	60 (58.3)	
>2/3	2 (2.7)	12 (12.6)	
Time since onset			
Early (<72 h)		83 (80.6)	
Late (>72 h)		20 (19.4%)	
APACHE II (mean score)	9 (6, 12)	8 (6, 13)	0.709
SIRS (mean score)	2 (1, 2)	2 (1, 3)	0.126
SEWS (mean score)	2 (1, 3)	2 (1, 4)	0.203
Chest tube, <i>n</i> (%)	2 (2.7)	90 (87.4)	0.000
Fibrinolytics, <i>n</i> (%)	0 (0)	76 (87.4)	0.001
Pleural fluid cultures, <i>n</i> (%)	0 (0)	34 (33)	0.000
Days of admission (median)	11 (7, 19)	13 (10, 20)	0.093
30-day mortality, <i>n</i> (%)	3 (4.1)	4 (3.9)	0.954
Smoker (PYI)	12 (1, 37.5)	16 (1, 30)	0.458
Alcohol, <i>n</i> (%)	6 (8.1)	23 (22.3)	0.009
COPD, <i>n</i> (%)	12 (16.2)	9 (8.7)	0.101
Chronic liver disease, <i>n</i> (%)	4 (5.4)	3 (2.9)	0.323
Heart failure, <i>n</i> (%)	7 (9.5)	2 (1.9)	0.029
Cerebrovascular accident, <i>n</i> (%)	7 (9.5)	6 (5.8)	0.265
Chronic renal failure, <i>n</i> (%)	5 (6.8)	6 (5.8)	0.678
Rheumatic disease, <i>n</i> (%)	4 (5.4)	4 (3.9)	0.448
Diabetes mellitus, <i>n</i> (%)	15 (20.3)	15 (14.6)	0.289
Time, <i>n</i> (%)			0.128
<30 (1)	71 (95.9)	93 (90.3)	
>30 days (2)	3 (4.1)	10 (9.7)	
Fever, <i>n</i> (%)	46 (62.2)	71 (68.9)	0.218
Prior antibiotic treatment, <i>n</i> (%)			0.484
No or 1 day	60 (81.1)	79 (76.7)	
$\geq 2$ days	14 (18.9)	24 (23.3)	
Free septa, <i>n</i> (%)	11 (3)	41 (64.1)	0.000
Septa, <i>n</i> (%)	2 (6.1)	35 (53.0)	0.000
UCL admission, <i>n</i> (%)	9 (12.2)	8 (7.8)	0.481

APACHE II: Acute Physiology and Chronic Health Evaluation II; CRP: C-reactive protein; ICU: intensive care unit; IL-6: interleukin 6; LDH: lactate dehydrogenase; PCT: procalcitonin; PI: pleural infection; SEWS: Standardized Early Warning Score; SIRS: Systemic Inflammatory Response Syndrome; TNF- $\alpha$ : tumor necrosis factor-alpha.

**Table 2**  
Sensitivity, Specificity, Positive and Negative Predictive Values and AUC for the Diagnosis of Complicated Pleural Infection/Empyema.

Parameter	Cut-off	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	PLR	NLR
Leukocytes in PF	12,690 cells/ $\mu$ l	0.545 (0.458, 0.633)	31.3	95.8	90.9	51.1	7.5	0.72
PMN (%)	61%	0.866 (0.807, 0.924)	79.1	85.5	87.8	75.6	5.46	0.24
PF pH	7.35	0.921 (0.881, 0.961)	85.5	84.8	81.4	88.2	5.81	0.18
LDH in PF (U/l)	1004 IU/l	0.949 (0.913, 0.984)	85.6	100	100	84.1	-	0.14
CRP in PF	6.7 mg/dl	0.766 (0.692, 0.839)	67.4	78.5	81	63.9	3.14	0.41
PCT in PF	0.25 ng/ml	0.574 (0.467, 0.682)	48.6	79.5	79.5	48.6	2.37	0.64
IL-6 in PF	64,560 pg/ml	0.837 (0.772, 0.901)	73.7	92.9	93.3	72.5	10.5	0.28
TNF- $\alpha$ in PF	25 pg/ml	0.808 (0.741, 0.875)	75.5	83.3	85.5	72.2	4.53	0.29

AUC: area under the ROC curve; CRP: C-reactive protein; LDH: lactate dehydrogenase; IL-6: interleukin 6; NLR: negative likelihood ratio; NPV: negative predictive value PCT: procalcitonin; PLR: positive likelihood ratio; PPV: positive predictive value; TNF- $\alpha$ : tumor necrosis factor-alpha.



**Fig. 1.** Distribution of concentrations in the number of leukocytes, polymorphonuclear leukocytes (%), pH, LDH, CRP, PCT, IL-6 and TNF- $\alpha$  in pleural fluid in the uncomplicated and complicated pleural infections and empyemas. CRP: C-reactive protein; IL-6: interleukin 6; LDH, lactate dehydrogenase; PCT: procalcitonin; PF: pleural fluid; PMN: polymorphonuclear leukocytes; TNF- $\alpha$ : tumor necrosis factor-alpha.

or organizing). Very little experience with IL-6 is available. It has only been used in 1 study and only to diagnose IPE.<sup>24</sup> IL-6 is an immune system mediator, and serum levels may be raised in some autoimmune diseases, infections, and sepsis, and in the rejection of organ transplants. AUC has a good discriminatory value (0.837), higher than the above-mentioned study (0.70), and has higher sensitivity and specificity (73.7% and 92.9%, respectively). In contrast, the discriminatory value of CRP was moderate (AUC 0.766). This could mean that once a certain level of inflammation is reached in PF, IL-6 values did not continue to rise in the same proportion as the infection. As suggested by previous studies,<sup>25–27</sup> the discriminatory power of procalcitonin was very poor (AUC 0.574).

The presence or absence of complicated IP/empyema was considered for prediction of the outcome. In this context, the logistic regression model is the most widely used statistical technique for binary medical outcomes.<sup>10</sup> When developing a prediction model, the aim is to quantify the accuracy of the model (model yield), so calibration, discrimination, diagnostic accuracy, and validity of the predictions were performed for new patients. The model constructed to differentiate both PIs uses the variables pH, LDH, and IL-6. The probability of the patient presenting complicated PI/empyema rises as values of LDH and IL-6 increase, and pH decreases. The AUC of the model, after bootstrapping, is slightly

**Table 3**

Logistic Regression Model for the Diagnosis of Complicated Pleural Infections/Empyema.

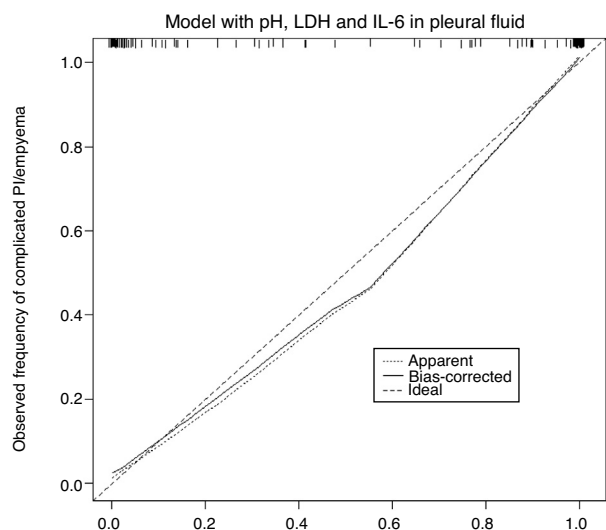
	Coefficients (SE)	P	OR (95% CI)
Intercept	62.69 (29.86)	0.024	
pH	-11.45 (3.81)	0.002	1.06 $\times 10^{-5}$ (6.08 $\times 10^{-9}$ , 0.018)
LDH in PF (IU/l) log	5.10 (1.24)	0.000	164.02 (14.43, 1863.85)
IL-6 in PF (pg/ml) log	1.52 (0.62)	0.014	4.57 (1.35, 15.41)

R2 = 0.843, Brier = 0.052, AUC = 0.978 (0.956, 1).

CI: confidence interval; IL-6: interleukin 6; LDH: lactate dehydrogenase; PF: pleural fluid; OR: odds ratio; SE: standard error.

lower than the apparent AUC (0.9757 vs. 0.9783), showing that there is no overoptimism in the estimate. It is interesting that the constructed model does not use percentage of PMNs as a variable, the AUC of which is higher than that of IL-6. This may be because the mechanism by which it rises is similar to that of any of the variables chosen and, therefore, its determination does not offer any advantage in increasing the performance of the model.

The calibration plot (Fig. 2) shows excellent concordance between the predicted probabilities and the observed frequencies, demonstrating the robustness of the model. The nomogram represents the weight of each of the selected variables and estimates the probability that a PI will be complicated PI/empyema, according to



**Fig. 2.** Calibration plot of the model constructed for the diagnosis of complicated pleural effusions/empyema. IL-6: interleukin 6; LDH: lactate dehydrogenase; PI: pleural infection.

the total score obtained from all the variables (Fig. 3, bottom).

Misdiagnosing an uncomplicated PI as complicated PI will subject the patient to unnecessary invasive testing. However, the opposite situation will lead to delays in diagnosis and, as such, to a potential increase in morbidity and mortality. Perhaps, therefore, the best strategy for classifying these patients would be to obtain the lowest number of false negatives, that is to say, to try to avoid complicated PIs being mistakenly classified as uncomplicated, in order to avoid the complications mentioned above. For each predicted probability, a certain yield

is obtained. A predicted probability of 60% seems to offer the greatest yield, since this is the point with the least possibility of committing errors when establishing the diagnosis of IPE (Table 5).

Our study has some limitations. First, the number of IPEs was low (177); second, patients were recruited in a single center. Although bootstrap techniques have been used to correct overoptimism, an external validation must be carried out in other centers. Finally, the calculation of the estimated probability of having a complicated PI/empyema is complex, but the construction of an Excel calculator, available in [supplementary material](#), will facilitate the practical implementation of this approach by clinicians.

In summary, our predictive model has a higher yield than that of any of the component variables taken individually for identifying the complicated PI/empyema. Although further studies are needed to confirm these findings, the results obtained appear to suggest that it could be applied in clinical practice.

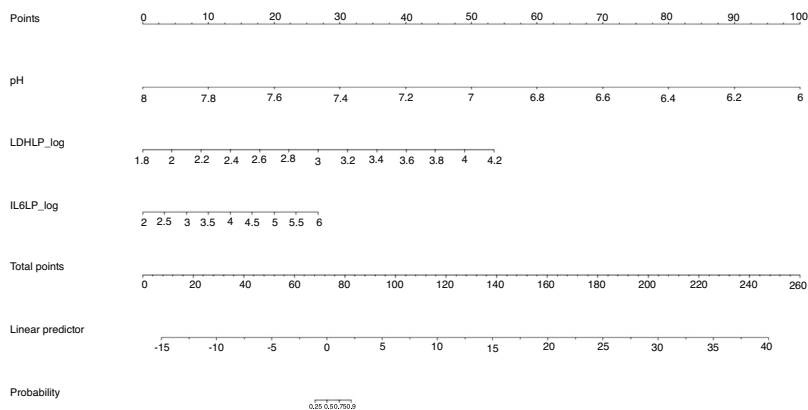
### Authorship

*Lucía Ferreira.* Study conception and design: Author. Data collection. Data analysis and interpretation. Critical review of the article. Final approval of the manuscript.

*Óscar Lado-Baleato.* Statistical analysis. Critical review of the article. Final approval of the manuscript.

*Juan Suárez-Antelo.* Co-author. Data collection. Data analysis and interpretation. Critical review of the article. Final approval of the manuscript.

*María E. Toubes.* Co-author. Data collection. Data analysis and interpretation. Critical review of the article. Final approval of the manuscript.

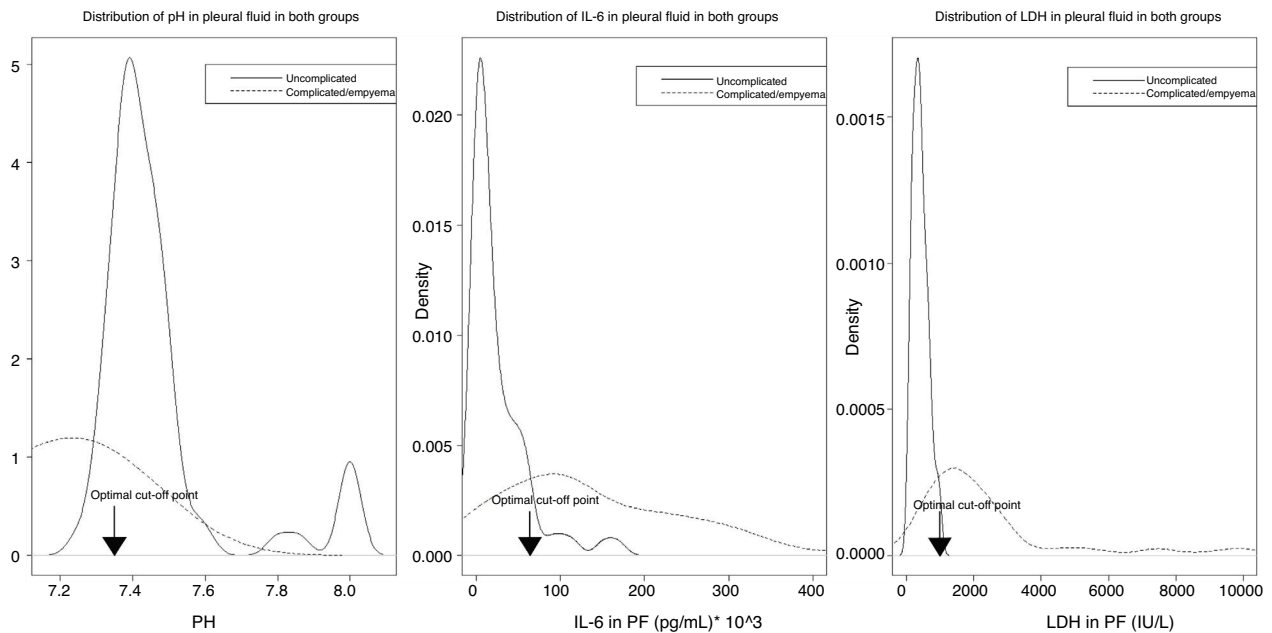


**Fig. 3.** Nomograms of the model constructed for the diagnosis of complicated pleural effusions/empyema. IL-6: interleukin 6; LDH: lactate dehydrogenase; PF: pleural fluid.

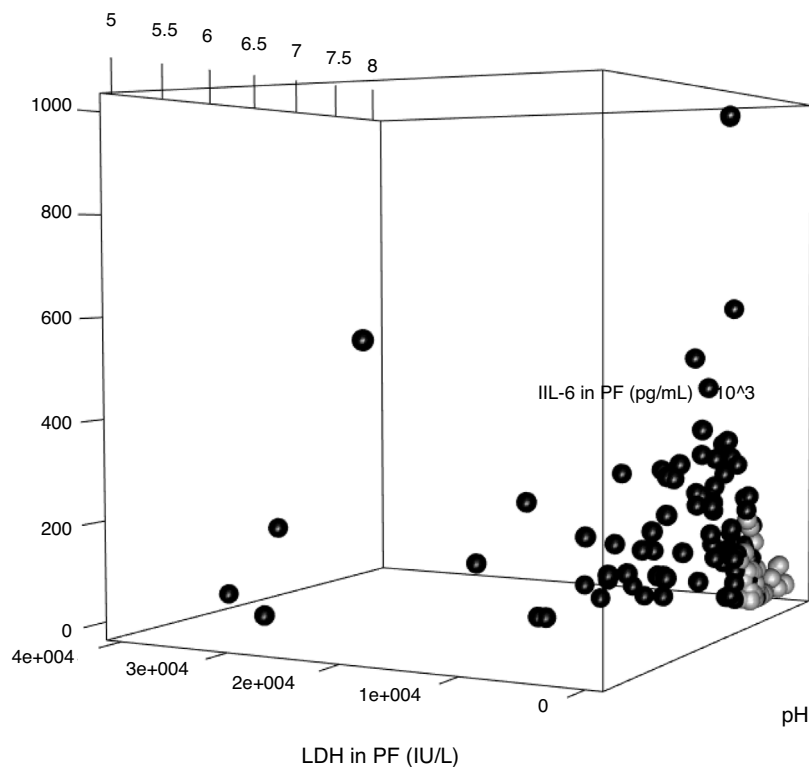
**Table 4**  
Sensitivity, Specificity, Positive and Negative Predictive Values, and Positive and Negative Likelihood Ratio to Differentiate an Uncomplicated Pleural Infection From a Complicated Pleural Infection/Empyema Using the Constructed Model.

Probability (%)	Sensitivity (CI)	Specificity (CI)	PPV (CI)	NPV (CI)	PLR (CI)	NLR (CI)
10	98.9 (93.9, 99.9)	68.1 (55.8, 78.8)	80 (70.5, 99.3)	97.9 (89.1, 98.7)	3.1 (2.2, 4.4)	0.01 (0.00, 0.1)
20	98.9 (93.9, 99.9)	75.3 (63.5, 84.9)	83.9 (74.8, 99.5)	98.1 (90.1, 98.9)	4 (2.6, 6.1)	0.01 (0.00, 0.1)
30	94.4 (87.5, 98.1)	82.6 (71.5, 90.6)	87.6 (78.9, 95.7)	91.9 (82.4, 95.9)	5.4 (3.2, 9.1)	0.06 (0.02, 0.1)
40	93.3 (86.1, 97.5)	86.9 (76.6, 93.8)	90.3 (82.1, 96.3)	90.9 (81.5, 95.8)	7.1 (3.9, 13.2)	0.07 (0.03, 0.16)
50	93.3 (86.1, 97.5)	89.8 (80.2, 95.8)	92.3 (84.5, 97.1)	91.1 (81.9, 96.4)	9.2 (4.5, 18.6)	0.07 (0.03, 0.16)
60	93.3 (86.1, 97.5)	97.1 (89.9, 99.6)	97.6 (91.7, 99.1)	91.7 (83.1, 98.9)	32.2 (8.2, 126.3)	0.06 (0.03, 0.14)
70	87.8 (79.1, 93.7)	98.5 (92.1, 99.9)	98.7 (93.2, 99.3)	86.1 (76.6, 99.6)	60.5 (8.6, 424.5)	0.12 (0.07, 0.21)
80	83.3 (74, 90.3)	98.5 (92.1, 99.9)	98.6 (92.8, 99.2)	81.9 (72.1, 99.4)	57.5 (8.2, 403.3)	0.16 (0.1, 0.26)
90	78.9 (69, 86.7)	100 (94.7, NaN)	100 (94, 100)	78.4 (68.4, NaN)	Inf (NaN, Inf)	0.21 (0.14, 0.31)

CI, confidence interval; Inf: infinite; PLR: positive likelihood ratio; NLR: negative likelihood ratio; NaN: not a number; NPV: negative predictive value; PPV: positive predictive value.



**Fig. 4.** Distribution of values of pH, IL-6, and LDH in PF, according to uncomplicated PI (solid line) or complicated PI/empyema (dashed line), and the optimal cut-off points.



**Fig. 5.** Three-dimensional graph showing the relationship between the selected variables (pH, LDH and IL-6 in pleural fluid) in the construction of the model to differentiate non-complicated pleural infections (black circles) from complicated pleural infections/empyema (gray circles).

*Esther San José.* Determination of study parameters in pleural fluid. Co-author. Data analysis and interpretation. Final approval of the manuscript.

*Adriana Lama.* Co-author. Data collection. Critical review of the article. Final approval of the manuscript.

*Nuria Rodríguez-Núñez.* Co-author. Data collection. Critical review of the article. Final approval of the manuscript.

*José Manuel Álvarez Dobaño.* Co-author. Data collection. Critical review of the article. Final approval of the manuscript.

*Francisco Javier González Barcala.* Co-author. Data collection. Critical review of the article. Final approval of the manuscript.

*Jorge Ricoy.* Co-author. Data collection. Critical review of the article. Final approval of the manuscript.

*Francisco Gude.* Statistical analysis. Data analysis and interpretation. Critical review of the article. Final approval of the manuscript.

*Luis Valdés.* Principal author. Study conception and design. Data analysis and interpretation. Critical review of the article. Final approval of the manuscript.

**Table 5**  
False Positive and False Negative Results of the Model in the Diagnosis of Complicated Pleural Infections/Empyema at Different Percentages of Probability.

Probability (%)	False Positives	False Negatives
10	16	1
20	11	3
30	9	3
40	6	5
50	3	5
60	2	5
70	1	6
80	0	11
90	0	16
99	0	24

## Funding

This study was funded by a grant from the Spanish Society of Pulmonology and Thoracic Surgery (SEPAR 126/2012).

## Conflict of Interests

The authors state that they have no conflict of interests.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.arbr.2019.08.003.

## References

- Dean NC, Griffith PP, Sorensen JS, McCauley L, Jones BE, Lee YC. Pleural effusions at first ED encounter predict worse clinical outcomes in patients with pneumonia. *Chest*. 2016;149:1509–15.
- Ferreiro L, Porcel JM, Bielsa S, Toubes ME, Álvarez-Dobaño JM, Valdés L. Management of parapneumonic effusions. *Expert Rev Respir Med*. 2018;12:521–35.
- Grijalva CG, Zhu Y, Nuorti JP, Griffin MR. Emergence of parapneumonic empyema in the USA. *Thorax*. 2011;66:663–8.
- Givan DC, Eigen H. Common pleural effusions in children. *Clin Chest Med*. 1998;19:363–71.
- Maskell NA, Davies CWH, Nunn AJ, Hedley EL, Gleeson FV, Miller R, et al. U.K. controlled trial of intrapleural streptokinase for pleural infection. *N Engl J Med*. 2005;352:865–74 [Erratum *N Engl J Med*. 2005;352(20):2146].
- Rahman NM, Maskell NA, West A, Teoh R, Arnold A, Mackinlay C, et al. Intrapleural use of tissue plasminogen activator and DNase in pleural infection. *N Engl J Med*. 2011;365:518–26.
- Chalmers JD, Singanayagam A, Murray MP, Scally C, Fawzi A, Hill AT. Risk factors for complicated parapneumonic effusion and empyema on presentation to hospital with community-acquired pneumonia. *Thorax*. 2009;64:592–7.
- López-Ratón M, Rodríguez-Alvarez MX, Cadarso-Suárez C, Gude-Sampedro F. OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *J Stat Softw*. 2014;61:1–36.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under 2 or more correlated receiver operator characteristics curves: a nonparametric approach. *Biometrics*. 1998;44:837–45.
- Harrell FE. Regression modeling strategies: with applications to linear models logistic regression, and survival analysis. New York: Springer-Verlag; 2001.
- Ewout S. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.
- Wood SN. Generalized additive models: an introduction with R. 2nd ed. Boca Raton (Florida): Chapman and Hall/CRC; 2017.
- Venables WN, Ripley BD. Modern applied statistics with S. 4th ed. New York: Springer; 2002.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sánchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
- R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org> [accessed 4.03.2015].
- Davies HE, Davies RJO, Davies CWH, BTS Pleural Disease Guideline Group. Management of pleural infection in adults: British thoracic society pleural disease guideline 2010. *Thorax*. 2010;65 Suppl. 2:ii41–53.
- Falguera M, Carratalà J, Bielsa S, García-Vidal C, Ruiz-González A, Chica I, et al. Predictive factors, microbiology and outcome of patients with parapneumonic effusion. *Eur Respir J*. 2011;38:1173–9.
- Colice GL, Curtis A, Deslauriers J, Heffner J, Light R, Littenberg B, et al. Medical and surgical treatment of parapneumonic effusions. An evidence-based guideline. *Chest*. 2000;118:1158–71.
- Porcel JM, Galindo C, Esquerda A, Trujillano J, Ruiz-González A, Falguera M, et al. Pleural fluid interleukin-8 and C-reactive protein for discriminating complicated nonpurulent from uncomplicated parapneumonic effusions. *Respirology*. 2008;13:58–62.
- Porcel JM, Vives M, Cao G, Bielsa S, Ruiz-González A, Martínez-Iribarren A, et al. Biomarkers of infection for the differential diagnosis of pleural effusions. *Eur Respir J*. 2009;34:1383–9.
- Porcel JM, Bielsa S, Esquerda A, Ruiz-González A, Falguera M. Pleural fluid C reactive protein contributes to the diagnosis and assessment of severity of parapneumonic effusions. *Eur J Intern Med*. 2012;23:447–50.
- San José ME, Ferreiro L, Soneira ME, González-Barcala FJ, Vázquez MC, Golpe A, et al. Utility of measurement of interleukin-18 and interleukin-8 in the diagnosis of complicated parapneumonic pleural effusions. *Am J Clin Pathol*. 2014;142:467–73.
- Ferreiro L, Pereiro T, San José E, Toubes ME, Suárez-Antelo J, Álvarez-Dobaño JM, et al. Behavior of nucleated cells in various types of pleural effusion. *Rev Clin Esp*. 2017;217:136–43.
- San José ME, Valdés L, González-Barcala FJ, Vizcaino L, Garrido M, Sanmartín A, et al. Diagnostic value of proinflammatory interleukins in parapneumonic effusions. *Am J Clin Pathol*. 2010;133:884–91.
- San José ME, Valdés L, Vizcaíno LH, Mora T, Pose A, Soneira E, et al. Procalcitonin C-reactive protein, and cell counts in the diagnosis of parapneumonic pleural effusions. *J Investig Med*. 2010;58:971–6.
- Zou MX, Zhou RR, Wu WJ, Zhang NJ, Liu WE, Fan XG. The use of pleural fluid procalcitonin and C-reactive protein in the diagnosis of parapneumonic pleural effusions: a systemic review and meta-analysis. *Am J Emerg Med*. 2012;30:1907–14.
- Dixon G, Lama-Lopez A, Bintcliffe OJ, Morley AJ, Hooper CE, Maskell NA. The role of serum procalcitonin in establishing the diagnosis and prognosis of pleural infection. *Respir Res*. 2017;18:30.