ERJ Express. Published on August 9, 2007 as doi: 10.1183/09031936.00126306

Discrimination of exudative pleural effusions based on multiple biological parameters

Zoe D. Daniil[1], Elias Zintzaras[2], Theodoros Kiropoulos[1], Andriana Papaioannou[1], Agela Koutsokera[1], Alcibiadis Kastanis[2], Konstantinos I. Gourgoulianis[1]

[1]Department of Respiratory Medicine and [2]Department of Biomathematics, University of Thessaly School of Medicine, Larissa, Greece.

Zoe D. Daniil: zdaniil@med.uth.gr, Elias Zintzaras: zintza@med.uth.gr, Theodoros Kiropoulos: ktheod@med.uth.gr, Andriana Papaioannou: andriana78@panafonet.gr, Agela Koutsokera: ilkoy@otenet.gr, Alcibiadis Kastanis: disxaper@otenet.gr, Konstantinos I. Gourgoulianis: kgourg@med.uth.gr

*Address for Correspondence: Associate Professor Elias Zintzaras, The Institute for Clinical Research and Health Policy Studies, Tufts-New England Medical Center, Tufts University School of Medicine, 750 Washington Street, Tufts-NEMC #63, Boston, MA 02111, USA, Tel: +1-617-6367248, fax: +1-617-6368628. e-mail: zintza@med.uth.gr

Abstract

Background: Pleural effusion is a common complication of different diseases. Conventional methods are not always capable to establish the cause of pleural effusion and alternative tests are needed.

Aim: to explore the discrimination of pleural effusion groups: malignant, parapneumonic and tuberculous, based on a function of seven biological markers.

Methods: Adenosine deaminase (ADA), interferon-γ (INF-γ), C-reactive protein (CRP), carcinoembryonic antigen (CEA), interleukin (IL)-6, tumor necrosis factor (TNF)-α and vascular endothelial growth factor (VEGF) were measured in pleural fluid in: 45 patients with malignant (MPE), 15 with parapneumonic (PPE) and 12 with tuberculous (TBE) pleural effusion. ROC curve analysis, multinomial logit modeling and canonical variate analysis were applied to discriminate the pleural effusion groups.

Results: The three groups can be discriminated successfully using the measured markers. The most important parameters for discrimination were ADA and CRP. An individual with ADA more than 45 and CRP less than 4 is more likely to belong to TBE, whereas, one with ADA less than 40 and CRP more than 6 to PPE, and one with CRP less than 4 to MPE.

Conclusion: Combination of ADA and CRP might be sufficient to discriminate the three different groups of exudative pleural effusion, malignant, tuberculous and parapneumonic.

Background

Pleural effusion can occur as a complication of many different diseases. It is a common clinical problem and it has been estimated that there are over 800,000 cases per year in the USA [1]. The diagnosis of pleural effusions remains a controversial issue in terms of cost to both patients and the health care system. Conventional methods are not always capable to establish the cause of pleural effusion and alternative tests permitting a rapid and accurate diagnosis are greatly needed. Malignant disease involving the pleura and parapneumonic effusion are the leading causes of exudative pleural effusions. However, the diagnosis of tuberculous pleuritis should be also considered in any patient with an exudative pleural effusion.

A variety of biological markers have been proposed to facilitate the differential diagnosis among the above mentioned causes of pleural effusion, including pleural fluid concentrations of adenosine deaminase (ADA), interferon (INF)-γ, a variety of tumor marker and cytokines and C- reactive protein (CRP). Although, there is a large body of literature on these biological markers and their utility for the diagnosis of pleural effusion, so far, the diagnosis is usually based on each individual marker separately.

Aim of our study was to explore the discrimination of pleural effusion groups (malignant, parapneumonic and tuberculous pleural effusion) and to provide a classification-diagnostic rule based on a function of seven markers (parameters): ADA, INF-γ, CRP, carcinoembryonic antigen (CEA), interleukin (IL)-6, tumor necrosis factor (TNF)-α and vascular endothelial growth factor (VEGF).

Methods and Materials

Subjects

A total of 72 pleural fluid samples were prospectively collected from 72 patients at the Respiratory Medicine Department of the Medical School of the University of Thessaly in Larissa between January and June 2005. The study group included 51 men and 21 women with a mean age of 65 years. Forty-five patients (62%) had malignant pleural effusion (MPE), 15 patients (21%) had parapneumonic pleural effusion (PPE), and 12 patients (17%) had tuberculous pleural effusion (TBE). The study protocol was approved by the local ethics committee and all subjects gave their written informed consent.

Diagnostic criteria for pleural effusions

The determination of the etiology of the pleural effusions was based on the clinical presentation, the appropriate diagnostic tests, and the response to treatment for each patient. Accordingly, effusions were classified into the following groups defined by predetermined criteria: i) *Malignant pleural effusion (MPE):* Secondary to lung cancer (diagnosed by the demonstration of malignant cells at cytologic examination or in a biopsy specimen or histologically proven primary lung malignancy with the exclusion of any other cause of pleural effusion), and other malignancies (effusions that were clearly secondary to other malignancies with exclusion of other causes for the development of pleural effusions) (Table 1). ii) *Tuberculous pleural effusion (TBE):* the diagnosis was based either on the presence of positive stain or culture for Mycobacterium tuberculosis in the pleural fluid, sputum or pleural biopsy; or in the presence of typical caseating

granulomas in pleural biopsy. iii) *Parapneumonic pleural effusion (PPE)* were identified by the presence of pulmonary infections associated with acute febrile illness, pulmonary infiltrates, purulent sputum, and response to antibiotic treatment; identification of the organism in the pleural fluid; or the presence of empyema, associated with finding of franc pus in the pleural cavity.

Sample collection and determination of ADA, CRP, INF-γ, IL-6, TNF-α, VEGF and CEA levels

The samples obtained by the first successful thoracentesis and before any treatment was initiated, immediately after the patients' admission. At the same time 10 mL of venous blood were obtained from each patient. Pleural fluid and blood samples were subsequently analyzed for total cell count, different cell count, glucose, total protein and LDH. All biochemical measurements were performed using standard commercially available methods (Olympus AU 600, Olympus Diagnostics Gmbh, Lismeehan, Ireland) while cell counts were obtained by manual microscopy. Cytologic examinations and both aerobic and anaerobic bacterial cultures were performed on all pleural effusions. Aliquots of pleural fluid and blood samples were immediately centrifuged at 1500 x g for 15 min at $4^{o}C$ and the supernatants were stored at $-80^{o}C$ until further analysis for the measurements of the above markers. ADA activity was measured by colorimetric method of Giusti. CRP measurements were performed by immunonephelometry with the Behring Nephelometer Analyzer II (BNII), using the N High Sensitivity kit [Dade Behring Gmbh, Germany)] The appropriate control and standard sera were provided by the same company, according to the manufacturer’s instructions. IL-6, VEGF and TNF-α levels

were measured with a commercially available enzyme-immunosorbent assay kit (Biosource Europe S.A.) according to the manufacturer's protocol. CEA levels were determined using an electrochemiluminescence immunoassay on Roche Modular E 170 analyzer (Roche Diagnostics Gmbh, Mannheim, Germany).

Statistical Methodology

The discriminating quality of each biological marker was evaluated independently using a receiver operator characteristics (ROC) curve analysis [2]. The ROC curve analysis was applied for two contrasts: PPE against TBE and MPE, and TBE against MPE and PPE. The quality of biological markers was assessed based on the area under the curve (AUC) metric. For each ROC curve, a cut-off point was determined as the value of the parameter that maximizes the sum of specificity and sensitivity, weighting equally their significance.

The discrimination between the three study groups (MPE, TBE, and PPE) based on all parameters simultaneously was investigated by fitting a multinomial logit model [3]. Multinomial Logistic Regression is useful for situations in which we want to be able to classify individuals based on values of a set of predictor variables (continuous and/or categorical). This type of regression is similar to logistic regression, but it is more general because the dependent variable is not restricted to two categories. In this study, we want to classify individuals into the three pleural effusion groups based on the measured markers/parameters, which are continuous. The overall significance of parameters was tested using a Likelihood Ratio test; parameters with significance level >0.05 were omitted from the model and then, a parsimonious model was used. This procedure

enables the easier interpretation of the results without a substantial loss of information. Then, the parameters of the parsimonious model can be used for classifying individuals to their groups. The logit model estimates the probabilities: p_1, p_2, and p_3 an individual to belong to groups MPE, TBE, and PPE, respectively. Then, an individual is classified to the group with the largest probability. The percentage of correct classified individuals into their groups was calculated, and the importance of the model was assessed based on the misclassified individuals. Detection rate (DR, sensitivity), false-positive rate (FPR, 1 minus specificity) and likelihood ratio (DR/FPR) were also considered for the significant parameters.

In order to provide a combination of parameters values for discriminating the three groups, parameters' values based on the cut-off points of the ROC curve analysis, were entered in the logit model and the respective probabilities p_i (i=1-3) were estimated. In addition, the three pleural effusion groups were discriminated using canonical variate analysis (CVA). CVA provides ordinations of patients on the basis of their biological marker measurement (parameters). CVA examines the separations among a set of groups of units. For this purpose, CVA seeks linear combinations of the *k* parameters, called canonical variates that have the greatest between-group variation relative to their within-group variability. The first and second canonical variates (denoted CV1 and CV2, respectively) are the eigenvectors of the $W^{-1/2}BW^{-1/2}$, where B is the between-group sums of squares and products (SSP) matrix, $W=\Sigma_{i=1-18}W_i$ and W_i is the within each group i SSP matrix. A two dimensional ordination from CVA (CV1 versus CV2), usually, accounts for most of the variation of the data [4, 5]. The statistical analysis was performed using SPSS v13, StatsDirect and routines in Compaq Visual Fortran90.

Results

The descriptive statistics for the seven parameters for each pleural effusion group are shown in Table 2. All parameters deviate from normality (except IL-6), but, after log-transformation most parameters looked quite normal with the exception of CEA. The data were analyzed raw since results can be interpreted more easily to clinicians and logit model for discriminant analysis may hold in the absence of normality [3], though any conclusions must be cautious. In addition, data transformation may mask the real effects of the parameters and their interactions [6]. Figure 1 shows the distribution of ADA and CRP after a log-transformation, which are typical of the data. Table 3 shows the Spearman's correlation coefficients (r) between the seven parameters. The nominal correlations between the parameters are generally low. However, the highest correlations were shown between TNF-α and ADA (r=0.62), CEA (r=-0.44), IL-6 (r=0.48), and between IL-6 and ADA (r=0.51), CRP (r=0.55).

The AUCs derived from ROC curve analysis for each parameter are depicted in Table 4. In discriminating TBE from MPE and PPE, ADA provided the largest AUC (0.94) with a cut-off point of 42.2, and in discriminating the PPE from MPE and TBE, CRP provided the largest AUC (0.92) with a cut-off point of 5.5.

Only parameters ADA (p=0.01) and CRP (p=0.05) were significant important to the simultaneous system after fitting the logit model. The fitting of the model, thought it is not optimal, it may be considered satisfactory as it can be seen from the half-normal plot of the residuals (i.e. the difference between observed and predicted probabilities) (Figure 2). Thus, a parsimonious model was fitted including only the parameters ADA ($p<0.01$)

and CRP ($p<0.01$). Table 5 shows the classified individuals based on the logit model: four individuals from MPE were classified as PPE, one individual from TBE was classified as PPE, and three individuals from PPE were classified as MPE (2 cases) or TBP. The overall proportion of misclassified individuals was 11%. The effect of adding markers to the significant parameters ADA and CRP in a logit model was explored (Table 6). However, combinations up to four parameters were considered in this model since our aim was to investigate the utility of the minimum possible parameters; in addition, false positive results were kept to minimum. Obviously, the more information is included in the logit model the less misclassification rate you are expected, however, the best trade-off between number of parameter and misclassification rate seems to be the model included the ADA and CRP parameters, though, inclusion of the INF-γ could be an alternative option. Table 7 shows the detection rate (DR, sensitivity), false-positive rate (FPR, 1 minus specificity) and likelihood ratio (DR/FPR) for considering ADA and CRP and combination with the rest parameters in the logit model. For ADA and CRP in the model, the likelihood ratio is greatest for TBE (likelihood ratio=54) suggesting that the two parameters can be much better at classifying TBE correctly while minimizing incorrect classification of the two other groups.

The respective estimated probabilities for classifying an individual to MPE, TBE, or PPE are the following: $p_1=e^{9.23-0.16(ADA)-0.62(CRP)}/D$, $p_2=e^{0.47+0.06(ADA)-0.70(CRP)}/D$, or $p_3=1/D$, where $D=1+e^{9.23-0.16(ADA)-0.62(CRP)}+e^{0.47+0.06(ADA)-0.70(CRP)}$. These probabilities predict the group membership of an individual. Combination of values below and above the cut-off points of the parameters CRP (5.5) and ADA (42.4) can be used to estimate the probabilities pi (i=1-3), and thus to provide a crude rule for diagnosis. Table 8 shows

the estimated values of p_i (i=1-3) for various combinations of ADA and CRP. Thus, an individual with ADA more than 45 and CRP less than 4 is more likely to belong to TBE, whereas, one with ADA less than 40 and CRP more than 6 belongs to PPE, and one with CRP less than 4 to MPE.

The CVA discriminated the three groups successfully. Figure 3 shows the two-dimensional ordination form CVA which counts for 71% of the total variation of the data. The first and second canonical variates (CV1 and CV2, respectively) were: CV1=(0.87)ADA+(0.16)CEA+(0.53)IL6+(0.41)INF+(0.52)VEGF+(0.26)CRP+(0.77)TNF and

CV1=(0.17)ADA+(0.02)CEA+(0.11)IL6+(0.58)INF+(0.05)VEGF+(0.90)CRP+(0.15)TNF, respectively. The first axis CV1 clearly separated the MPE from the TBE patients, while the second axis CV2 separated the MPE from PPE patients. In CV1 and CV2 equations, the largest coefficients were appeared for ADA and CRP, respectively, indicating that these parameters are the most important for discrimination relative to the others [4, 5]

Discussion

The determination of biological markers in pleural effusions has been proposed as an alternative, noninvasive way of establishing a diagnosis of pleural effusion. However, the use of these measurements in clinical practice remains controversial [7, 8]. In this prospective study, ADA, CRP, CEA, INF-γ, IL-6, TNF-α and VEGF were measured in pleural fluid obtained from patients with exudative pleural effusion. To our knowledge, this is the first study to attempt exploring the usage of the above-mentioned seven

parameters, simultaneously, in discriminating the three different causes of exudative pleural effusion (malignant, parapneumonic and tuberculous pleural effusion).

Pleural fluid ADA activity has been shown to be a valuable biochemical marker that has a high sensitivity and specificity for the diagnosis of tuberculosis (TB) [9], but its diagnostic usefulness depends on the local prevalence of TB, laboratory methodology, and population ethnicity. The only other parameter that has comparable sensitivity and possibly higher specificity is INF-γ [10]. However, the latter's high cost and relatively long reaction time precludes its routine use [11]. A meta-analysis showed that studies conducted in Europe had significantly better diagnostic performance than those from other regions [12]. ADA combines low cost, easy performance and high diagnostic efficiency for identifying TB pleurisy [13]. According to the present study, ADA provides the largest AUC (0.94) for the discrimination of TBE from MPE and PPE (cut-off point: 42.4) and this finding is in agreement with previous studies [9]. In addition, when the logit model was fitted, ADA was found significant. INF-γ provided also a large AUC (0.93) for the discrimination of TBE from MPE and PPE. However, the logit model failed to show a significant role of this finding.

TNF-α is a proinflammatory cytokine that is known to regulate growth and differentiation of a variety of immune cells. Increased levels of TNF-α have been found both in infectious [8, 14] and malignant pleural effusion [16]. Porcel et al. suggested that pleural TNF-α could be a biochemical marker for inflammation in patients with PPE [7]. Additionally, elevated levels of pleural TNF-α identified more reliably the subgroup of patients with no purulent-appearing PPE who required invasive management with tube thoracostomy than traditional fluid chemistries [7]. However, others have reported no

difference between TNF-α level in exudative pleural effusion of various etiologies [16]. In the present study we found that TNF-α is more increased in TBE but this finding was not significant.

IL-6 has long been regarded as a proinflammatory cytokine induced by lipopolysaccharide along with TNF-α and IL-1. IL-6 is often used as a marker for systemic activation of proinflammatory cytokines [17]. IL-6 has been found to be elevated in malignant pleural effusion, especially after pleurodesis [15]. However, conflicting results have been reported in distinguishing malignant from infectious pleural effusion by IL-6 levels [8]. According to our findings, IL-6 was increased in TBE and PPE but these findings were not significant after fitting the logit model.

CRP is an acute-phase protein widely used as a marker of inflammation and tissue injury [18]. CRP level has been studied in pleural fluid and has been found to be higher in benign than in malignant exudates [19]. Turay et al have shown that pleural fluid CRP levels > 30 mg/l have a high sensitivity (93.7%) and specificity (76.5%) and a positive predictive value of 98.4% for the diagnosis of parapneumonic effusion [20]. Our results showed that CRP provides the largest AUC (0.92) for discriminating PPE from MPE and TBE, and it was a significant parameter in discriminating the groups.

Several studies have confirmed the presence of high levels of VEGF in exudative, especially malignant and inflammatory effusions [21, 22]. Finally, the conclusion arising from the literature is that there is an overlap between the VEGF levels in the various groups of pleural effusion therefore; VEGF levels are unlikely to be useful diagnostically [21]. The present study has demonstrated that VEGF levels are increased in both

malignant and parapneumonic pleural effusions and that VEGF is not significant parameter for the discrimination between the three groups.

Analysis of tumor markers in serum has been applied to the diagnosis, prognosis and treatment of patients with lung cancer. Furthermore, pleural fluid from patients with malignant pleural effusion is known to contain detectable levels of tumor markers. CEA was one of the first markers to be evaluated in lung cancer [23]. Determination of CEA concentration can be used as a diagnostic tool for malignant pleural effusions since 40-70% of these fluids give positive CEA assays [24]. In accordance with previous reports CEA was found to be the best single tumor marker in pleural fluid [25, 26] However, although CEA has been studied the most and has been shown to be very specific, its sensitivity remains approximately 29 to 77% with variable cut-off values. Lee et al demonstrated that in cases of suspicious malignant effusion showing a negative cytology, particularly in the absence of a visible tumor and/or unsuitability for invasive procedures, the determination of tumor markers in the pleural fluid might be helpful as a complementary tool for the differential diagnosis of pleural effusion [27]. Finally, although an elevated level of CEA in pleural fluid is suggestive of malignancy, CEA can be elevated in 9% of pleurisy owing to benign diseases, especially in empyemas and in complicated parapneumonic effusions. Identifying the most frequent causes of false-positive results of CEA helps to correctly interpret the findings of this tumor marker [28]. In the present study, the CEA levels were highest in malignant pleural effusions. However, a positive significance of this parameter to the simultaneous system was not found.

In the present study the number of cases might be considered relatively low compared to the number of parameters used in discrimination. Although in general, small sample size tend to result in inefficient discrimination, the discrimination accuracy does not strictly depend on sample size but on data nature (distribution), variability of data, and the number of variables examined [29-31]. Nevertheless, in order to strengthen farther the utility of the multinomial logit model and the derived classification rule, CVA was applied to the data. Then, CVA verified the existence of three distinct groups and the discrimination value mainly of ADA and CRP, though, the importance most of the parameters in CV1 should also be considered. However, accumulation of more data will enable us to explore further the validity of the discrimination results by applied additional techniques such as classification trees [32]. When a small dataset is used to both estimate the model parameters and test the model, then, it is expected the results to be favourable. Ideally, the two independent datasets should be used: one to derive the parameters and one to test them. Alternatively, there are more complex methods for parameter estimation such as bootstrapping [33], but this estimation is beyond the scope of this article.

In conclusion, combinations of markers can be useful in the discrimination of exudative pleural effusion groups. Although, our analysis was focused on ADA and CRP the data presented here does not exclude the usefulness of one or more additional markers. Therefore, further and larger studies should not be focused only on ADA and CRP and then, the results should be synthesized to provide more evidence.

List of abbreviations used:

Malignant pleural effusion (MPE);

Parapneumonic pleural effusion (PPE);

Tuberculous pleural effusion (TBE);

adenosine deaminase (ADA);

interferon-γ (INF-γ);

C-reactive protein (CRP);

carcinoembryonic antigen (CEA);

interleukin-6 (IL-6);

tumor necrosis factor-α (TNF-α) and

vascular endothelial growth factor (VEGF).

Competing Interests

The authors declare that they have no competing interests.

Authors' contributions

ZD participated in the design of the study, collection of the clinical data and drafted the manuscript. EZ participated in the design of the study, performed the statistical analysis, leaded the interpretation the data and drafted the manuscript. TK carried out the immunoassays. AP participated in the collection of the data and helped to draft the manuscript. AK participated in the data analysis and interpretation of the data. KG conceived of the study, participated in its design and coordination. All authors read and approved the final manuscript.

References

1. Marel M, Zrustova M, Stasny B, Light RW **The incidence of pleural effusion in a well-defined region.** Epidemiologic study in central Bohemia. *Chest* 1993; 104:1486-9.

2. Zintzaras E, Germenis AE. **Diagnostic performance of antibodies against tissue transglutaminase for the diagnosis of coeliac disease**: meta-analysis. *Clin Vacc Immun.* 2006; 13:15-28.

3. Jobson J.D. **Applied Multivariate Data Analysis, Springer**, New York, 1992.

4. Zintzaras E, N.P. Brown, A. Kowald. **Growing a classification tree using the apparent misclassification rate**. *Computer Applications in the Biosciences* 1994; 10: 263-271.

5. Zintzaras E, J. Margaritopoulos., J. Tsitsipis. **Statistical tree classification of aphids based on morphological characters**. *Comp Electron Agric* 1999; 24: 165-175.

6. Zintzaras E. **Statistical assessment of the synergistic effect between a vaccine and a drug.** *J Pharm Pharmacol* 1992; 44: 873-874.

7. Porcel JM, Vives M, Esquerda A. **Tumor Necrosis Factor-α in Pleural Fluid. A Marker of Complicated Parapneumonic Effusions**. *Chest* 2004; 125:160-164.

8. Yokoyama A, Kohno N, Fujino S, Abe M Ishida O Hiwada K. **Soluble interleukin-6 receptor levels in pleural effusions.** *Respir Med* 1996; 90: 329-332

9. Perez-Rodriguez E, Walton IJP, Hernandez JJS, Pallaris E, Rubi J, Castro DJ, Diaz Nuevo G. **ADA1/ADAp ratio in pleural tuberculosis**: an excellent diagnostic parameter in pleural fluid. *Respir Med* 1999; 93:816- 21.

10. Villegas MV, Labrada LA, Saravia MG. **Evaluation of polymerase chain reaction, adenosine deaminase, and interferongamma in pleural fluid for differential diagnosis of pleural tuberculosis**. *Chest* 2000; 118:1355- 64.

11. Ferrer J. **Pleural tuberculosis**. *Eur Respir J* 1997; 10:942- 7.

12. Goto M, Noguchi Y, Koyama H, Hira K, Shimbo T, Fukui T. **Diagnostic value of adenosine deaminase in tuberculous pleural effusion**: a meta-analysis. *Ann Clin Biochem* 2003; 40:374- 81.

13. Chen ML, Yu WC, Lam CW, Au KM, Kong FY, Chan A. **Diagnostic value of pleural fluid adenosine deaminase activity in tuberculous pleurisy**. *Clinica Chimica Acta* 2004; 341:101-107.

14. Barnes PF, Fong SJ, Brennan PJ, Twomey PE, Mazumder A, Modlin RL **Local production of tumor necrosis factor and IFN-α in tuberculous pleuritis**. *J Immunol* 1990; 145:149-154.

15. Agrenius V, Gustafsson LE, Widstrom O. **Tumour necrosis factor-α and nitric oxide, determined as nitrite, in malignant pleural effusion**. *Respir Med* 1994; 88:743-748.

16. Soderblom T, Nyberg P, Teppo AM, Klockars M, Riska H, Petterson T. **Pleural fluid interferon-γ and tumour necrosis factor-α in tuberculous and rheumatoid pleurisy**. *Eur Respir J* 1996; 9:1652-1655.

17. Opal SM, DePalo VA. **Anti-inflammatory cytokines**. *Chest* 2000; 117:1162-1172.

18. Okamura JM, Miyagi JM, Terada K, Hokama Y. **Potential clinical applications of C-reactive protein**. *J Clin Lab Anal* 1990; 4:231-235.

19. Castano Vidriales JL, Amores Antequera C. **Use of pleural fluid C-reactive protein in laboratory diagnosis of pleural effusions**. *Eur J Med* 1992; 1:201-207.

20. Yilmaz Turay U, Yildirim Z, Turkoz Y, Biber C, Erdogan Y, Keyf AI, Ugurman F, Ayaz A, Ergun P, Harputluoglu M. **Use of pleural fluid C-reactive protein in diagnosis of pleural effusions** *Respir Med*. 2000; 94:432-435.

21. Chang D-S, Rodriguez RM, Perkett EA, Rogers J, Bienvenu G, Lappalainen U, Light RW. **Vascular endothelial growth factor in pleural fluid**. *Chest* 1999; 116:760-765.

22. Thickett DR, Armstrong L, Millar AB. **Vascular endothelial growth factor (VEGF) in inflammatory and malignant pleural effusions**. *Thorax* 1999; 54:707-710.

23. Niklinski J, Furman M. **Clinical tumor markers in lung cancer**. *Eur J Cancer Prev.* 1995; 4:129-138.

24. Sculier JP, Body JJ, Jacobowitz D, Fruhling J. **Value of CEA determination in biological fluids and tissues**. Eur J Cancer Clin Oncol 1987; 23:1091-1093.

25. Ferrer J, Villarino MA, Encabo G, Felip E, Bermejo B, Vila S, Orriols R. **Diagnostic utility of CYFRA 21–1, carcinoembryonic antigen, CA 125, neuron specific enolase, and squamous cell antigen level determinations in the serum and pleural fluid of patients with pleural effusions**. *Cancer* 1999; 86:1488–1495.

26. Villena V, Lopez-Encuentra A, Echave-Sustaeta J, Martin-Escribano P, Ortuno-de-Solo B, Estenoz-Alfano J. **Diagnostic value of CA 72–4, carcinoembryonic antigen, CA 15–3, and CA 19–9 assay in pleural fluid**: a study of 207 patients. *Cancer* 1996; 78:736–740.

27. Lee JH, Chang JH. **Diagnostic Utility of Serum and Pleural Fluid Carcinoembryonic Antigen, Neuron-Specific Enolase, and Cytokeratin 19 Fragments in Patients With Effusions From Primary Lung Cancer**. *Chest* 2005; 128:2298-2303.

28. Garcia-Pachon E, Padilla-Navas I, Dosda D, Miralles-Llopis A. **Elevated Level of Carcinoembryonic Antigen in Nonmalignant Pleural Effusions**. *Chest* 1997; 111:643-647.

29. Raudys S, Pikelis V. **On dimensionality, sample size, classification error and complexity of classification algorithm in pattern recognition**. *Trans. Pattern Anal. Mach. Intell.* 1980; 2:242–252.

30. Wharton SW. **An analysis of the effects of sample size on classification performance of a histogram based cluster analysis procedure**. *Pattern Recognition*. 1984; 17: 239-244.

31. Boulesteix AL, Tutz G, Strimmer K. **A CART-based approach to discover emerging patterns in microarray data**. *Bioinformatics* 2003; 18: 2465-2472.

32. Zintzaras E, N.P. Brown, A. Kowald**, Non-parametric classification of secondary protein structures**. *Comp. Biol. Med*. 2005; 36: 145-156.

33. Zintzaras E, P. Bouka, A. Kowald, **Biometrical evaluation of bioequivalence trials using a bootstrap individual direct curve comparison method**. *Eur J Drug Metab Pharmacokinet.* 2002; 27: 11-16.

FIGURE LEGEND

Figure 1. Distribution of: ADA (a) and CRP (b) after a log-transformation.

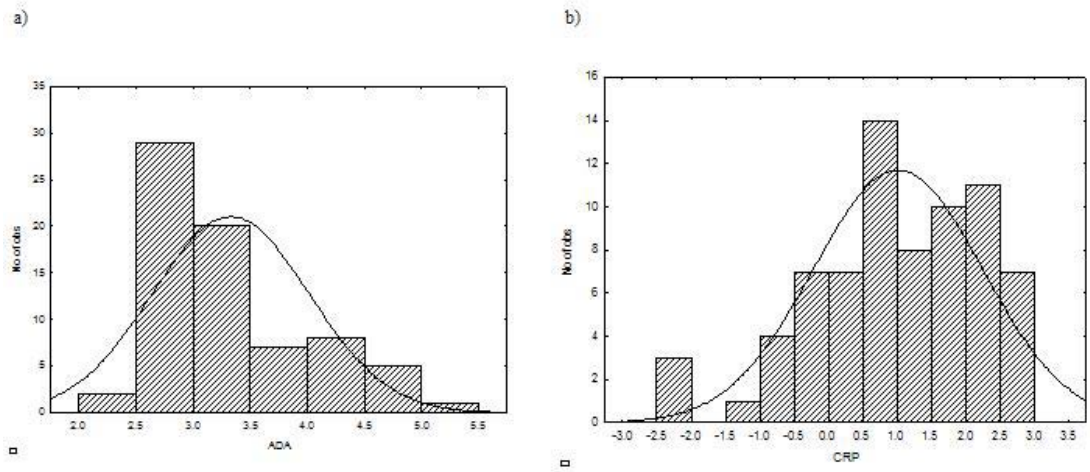

Figure 2. Half-normal plot of the residuals derived by fitting the multinomial logit model

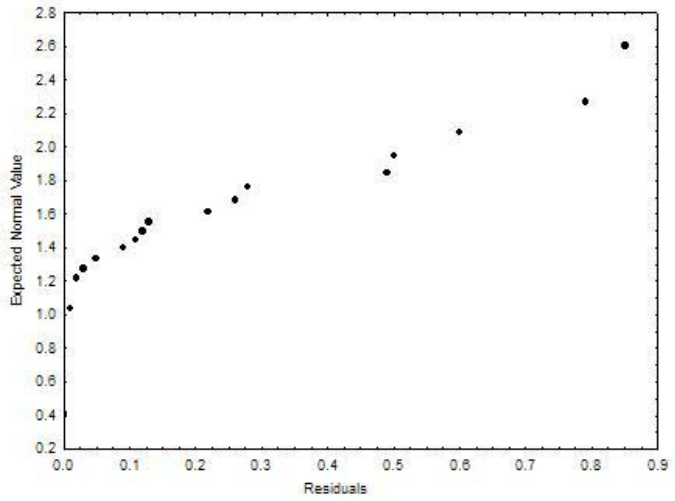

Figure 3. The two-dimensional ordination form CVA. The first and second canonical variates (CV1 and CV2, respectively) are shown.

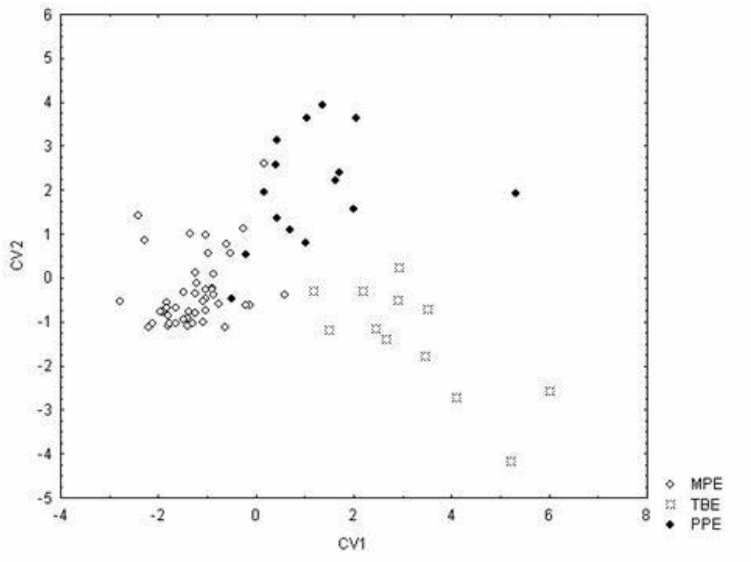
CV2
CV1
MPE
TBE
PPE

Table 1 Etiology of malignant pleural effusions

	n=45
Lung cancer	31
Ovary cancer	1
Renal cell carcinoma	1
Breast cancer	7
Prostate cancer	2
Gastric cancer	3

Table 2 Descriptive statistics (n, mean±SD, median (range)) of the seven parameters for each group of pleural effusion.

	MPE	**TBE**	**PPE**
	n=45	n=12	n=15
ADA	19.73±6.19 18.29(10.55-36.68	78.36±19.46 80.11(48.03-111.3)	51.61±43.10 31.47 (16.18-149.7)
CRP	2.76±2.80 1.9(0.1-12.7)	4.27±2.49 3.2(1.5-8.3)	10.94±4.30 11.6(1.8-17.1)
CEA	227.16±412.33 12.89(0.48-1433)	1.12±0.97 0.93(0.31-4)	2.23±2.88 1.53(0.59-12.4)
IL-6	8494.66±7754.58 4620(731-30220)	23622.75±8316.62 22615.5(10321-43819)	24340.93±9930.04 23819(9800-45315)
INF-γ	1.28±0.90 1.21(0.13-4.03)	23.61±23.12 14.8(0.74-63.6)	2.68±3.85 1.7(0.3-15.7)
VEGF	1259.04±1258.84 919(93-5886)	370.59±226.16 419(77.17-748)	1519.33±1025.04 1124(300-4015)
TNF-α	45.70±35.63 35.2(11.2-182)	261.89±276.92 211(53.9-1099)	293.19±669.69 115.2(27.7-2678)

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion; ADA: adenosine deaminase; CRP: C-reactive protein; CEA: carcinoembryonic antigen; IL-6: interleukin-6; INF-γ: interferon-γ; VEGF: vascular endothelial growth factor; TNF-α: tumor necrosis factor-α.

Table 3 Spearman's correlation coefficients between the seven parameters.

	ADA	**CEA**	**IL-6**	**IFN-γ**	**VEGF**	**CRP**	**TNF-α**
ADA	1.00						
CEA	-0.33	1.00					
IL-6	0.51	-0.39	1.00				
IFN-γ	0.38	-0.37	0.19	1.00			
VEGF	-0.18	0.23	0.16	-0.28	1.00		
CRP	0.29	-0.24	0.55	0.17	0.29	1.00	
TNF-α	0.62	-0.44	0.48	0.44	-0.18	0.38	1.00

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion; ADA: adenosine deaminase; CRP: C-reactive protein; CEA: carcinoembryonic antigen; IL-6: interleukin-6; INF-γ: interferon-γ; VEGF: vascular endothelial growth factor; TNF-α: tumor necrosis factor-α.

Table 4 AUC values produced from the ROC curve analysis

	AUC (95% ci)	
	TBE vs MPE and PPE	**PPE vs MPE and TBE**
ADA	0.94 (0.89-1)	0.70 (0.55-0.84)
CRP	0.57 (0.44-0.70)	0.92 (0.84-1)
CEA	0.17 (0.06-0.27)	0.34 (0.22-0.46)
IL-6	0.81 (0.70-0.91)	0.82 (0.72-0.92)
IFN-γ	0.93 (0.82-1)	0.51 (0.36-0.67)
VEGF	0.14 (0.05-0.24)	0.69 (0.55-0.83)
TNF-α	0.89 (0.81-0.97)	0.70 (0.56-0.84)

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion; ADA: adenosine deaminase; CRP: C-reactive protein; CEA: carcinoembryonic antigen; IL-6: interleukin-6; INF-γ: interferon-γ; VEGF: vascular endothelial growth factor; TNF-α: tumor necrosis factor-α.

Table 5 Confusion matrix for multinomial logit model predictions.

	Predicted classification			
Actual classification	MPE	TBE	PPE	Percent correct (95% confidence interval)
MPE	43	0	2	95.6% (90.8-100%)
TBE	0	11	1	91.7% (85.3-98.1%)
PPE	4	1	10	66.7% (55.8-77.6%)
Overall percentage	65.3%	16.7%	18.1%	88.9% (81.6-96.2%)

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion.

Table 6 Percentage of correctly classified individuals for each parameter and the effect of adding markers to ADA and CRP in the logit model. (For combinations ADA, CRP, IL-6, IFN-γ; ADA, CRP, IFN-γ, TNF-α; ADA, CRP, IFN-γ, CEA; ADA, CRP, IFN-γ, VEGF, the model cannot be fitted either the maximum likelihood estimates do not exist or some parameter estimates are infinite).

	Percent correctly classified			
	MPE	**TBE**	**PPE**	**overall**
ADA	97.8	75	20	77.8
CRP	95.6	0	73.3	75
IL-6	91.1	0	53.3	68.1
IFN-γ	100	75	13.3	77.8
TNF-α	93.3	0	46.7	68.1
CEA	88.9	41.7	0	62.5
VEGF	88.9	41.7	0	62.5
ADA, CRP	95.6	91.7	66.7	88.9
ADA, CRP, IL-6	95.6	91.7	80	91.7
ADA, CRP, IFN-γ	97.8	91.7	80	93.1
ADA, CRP, TNF-α	97.8	91.7	73.3	91.7
ADA, CRP, CEA	93.3	91.7	80	90.3
ADA, CRP, VEGF	95.6	91.7	80	91.7
ADA, CRP, IL-6, TNF-α	97.8	91.7	80	93.1
ADA, CRP, IL-6, CEA	97.8	91.7	86.7	94.4
ADA, CRP, IL-6, VEGF	97.8	91.7	86.7	94.4
ADA, CRP, TNF-α, CEA	93.3	91.7	80	90.3
ADA, CRP, TNF-α, VEGF	97.8	91.7	80	93.1
ADA, CRP, CEA, VEGF	97.8	91.7	93.3	95.8

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion; ADA: adenosine deaminase; CRP: C-reactive protein; IL-6: interleukin-6; INF-γ: interferon-γ; TNF-α: tumor necrosis factor-α.

Table 7 Detection rate (DR, sensitivity), false-positive rate (FPR, 1 minus specificity) and likelihood ratio (LR=DR/FPR) for considering the effect of adding markers to ADA and CRP in the logit model

	Group	DR	FPR	LR
ADA, CRP	MPE	96% (43/45)	14.8 (4/27)	6.5
	TBE	92% (11/12)	1.7% (1/60)	54
	PPE	67% (10/15)	5.3% (3/57)	12.6
ADA, CRP, IL-6	MPE	96% (43/45)	7.4% (2/27)	12.9
	TBE	92% (11/12)	1.7% (1/60)	55
	PPE	80% (12/15)	5.3% (3/57)	15.2
ADA, CRP, IFN-γ	MPE	98% (44/45)	11.1% (3/27)	8.8
	TBE	92% (11/12)	0% (0/60)	na
	PPE	80% (12/15)	3.5% (2/57)	22.8
ADA, CRP, TNF-α	MPE	98% (44/45)	11.1% (3/27)	8.8
	TBE	92% (11/12)	1.7% (1/60)	55
	PPE	73% (11/15)	3.5% (2/57)	20.9
ADA, CRP, CEA	MPE	93% (42/45)	11.1% (3/27)	8.4
	TBE	92% (11/12)	1.7% (1/60)	55
	PPE	80% (12/15)	5.3% (3/57)	15.2
ADA, CRP, VEGF	MPE	96% (43/45)	11% (3/27)	8.6
	TBE	92% ((11/12)	2% (1/60)	55
	PPE	80% (12/15)	4% (2/57)	22.8
ADA, CRP, IL-6, TNF-α	MPE	98% (44/45)	7.4% (2/27)	13.2
	TBE	92% (11/12)	1.7% (1/60)	55
	PPE	80% (12/15)	3.5% (2/57)	22.8
ADA, CRP, IL-6, CEA	MPE	98% (44/45)	7.4% (2/27)	13.2
	TBE	92% (11/12)	0% (0/60)	na
	PPE	87% (13/15)	3.5% (2/57)	24.7
ADA, CRP, IL-6, VEGF	MPE	98% (44/45)	7% (2/27)	13.2
	TBE	92% (11/12)	0% (0/60)	na
	PPE	87% (13/15)	4% (2/57)	24.7
ADA, CRP, TNF-α, CEA	MPE	93% (42/45)	11.1% (3/27)	8.4
	TBE	92% (11/12)	1.7% (1/60)	55
	PPE	80% (12/15)	5.3% (3/57)	15.2
ADA, CRP, TNF-α, VEGF	MPE	98% (44/45)	11% (3/27)	8.8
	TBE	92% (11/12)	0% (0/60)	na
	PPE	80% (12/15)	4% (2/57)	22.8
ADA, CRP, CEA, VEGF	MPE	98% (44/45)	4% (1/27)	26.4
	TBE	92% (11/12)	0% (0/60)	na
	PPE	93% (14/15)	4% (2/57)	26.6

na: non-applicable

Table 8 Estimated values of p_i (i=1-3) for various combinations of ADA and CRP. p1, p2 and p3 denote the probabilities an individual to belong to MPE, TBE and PPE, respectively.

ADA	CRP	p1 (MPE)	p2 (TBE)	p3 (PPE)
45	4.0	0.19	0.51	0.31
149.7*	0.1$	0.00	1.00	0.00
40	6.0	0.23	0.17	0.59
10.6$	17.1*	0.04	0.00	0.96
40	4.0	0.38	0.34	0.28
10.6$	0.1$	1.00	0.00	0.00

TBE: tuberculous pleural effusion; MPE: malignant pleural effusion; PPE: parapneumonic pleural effusion; ADA: adenosine deaminase; CRP: C-reactive protein.
*maximum observed value, $minimum observed value