Eur Respir J, 1993, 6; 803–810
Printed in UK - all rights reserved

European Respiratory Journal
ISSN 0903 - 1936

Differences in BAL fluid variables in interstitial lung diseases evaluated by discriminant analysis

M. Drent*, P.G.H. Mulder**, Sj.Sc. Wagenaar+, H.C. Hoogsteden++, H. van Velzen-Blad+, J.M.M. van den Bosch*

Differences in BAL fluid variables in interstitial lung diseases evaluated by discriminant analysis. M. Drent, P.G.H. Mulder, Sj.Sc. Wagenaar, H.C. Hoogsteden, H. van Velzen-Blad, J.M.M. van den Bosch. ©ERS Journals Ltd 1993.

ABSTRACT: The aim of this study was to investigate the possibility of distinguishing between patients with similarities in clinical presentation, suffering from three frequently occurring interstitial lung diseases, by means of discriminant analysis, using a number of selected variables derived from bronchoalveolar lavage fluid (BALF) analysis.

The study involved all 277 patients, who had an initial bronchoalveolar lavage (BAL) in the period 1980–1990. These patients belonged to the following diagnostic groups: sarcoidosis (n=193), subacute extrinsic allergic alveolitis (EAA) (n=39) and idiopathic pulmonary fibrosis (IPF) (n=45). Thirty healthy volunteers were used as controls. Cellular and non-cellular constituents of BALF were evaluated.

Variables, which could be used to discriminate among the three diagnostic groups were: yield of recovered BALF, total cell count, and percentages of alveolar macrophages, lymphocytes, polymorphonuclear neutrophils, eosinophils and plasma cells in BALF. When the set of data used to predict the membership of patients to diagnostic groups (test set) was the same as that in which the discriminant analysis was performed (learning set), 93% of the cases were correctly classified. This percentage decreased to 90%, however, when the test set was different from the learning set.

It is possible to discriminate among patients with sarcoidosis, EAA or IPF with these selected variables. It appears that bronchoalveolar lavage (BAL) is useful as an adjunct in concert with other diagnostic methods.

Eur Respir J., 1993, 6, 803–810.

Depts of *Pulmonary Diseases, +Pathology and †Immunology, St. Antonius Hospital, Nieuwegein, The Netherlands. Depts of **Epidemiology and Biostatistics and ++Pulmonary Diseases, Erasmus University, Rotterdam, The Netherlands.

Correspondence: J.M.M. van den Bosch
Dept of Pulmonary Diseases
St. Antonius Hospital
PO Box 2500
3430 EM Nieuwegein
The Netherlands

Keywords: Bronchoalveolar lavage
extrinsic allergic alveolitis
idiopathic pulmonary fibrosis
interstitial lung diseases
sarcoidosis
statistical analysis

Received: August 4 1992
Accepted after revision February 14 1993

This study was supported by a grant from Glaxo B.V., The Netherlands.

A biopsy of patients with interstitial lung diseases (ILD) is not always available, thus, other diagnostic features are needed. Bronchoalveolar lavage (BAL) has proved to be of diagnostic value for analysis of inflammatory processes in the lung, particularly in ILD [1–4]. These disorders usually show an alveolitis, characterized by an accumulation of inflammatory and immune effector cells within the interstitium and the alveolar structures [3–5]. The increase in different cell types and immunoglobulins (Igs) in BAL fluid (BALF) may vary among the various ILD [1–4]. In order to discriminate between the ILDs, characteristic BAL features should be considered, together with unique clinical parameters [6, 7].

Sarcoidosis patients usually show an accumulation of activated, proliferating T-lymphocytes in the BALF, although some cases may show normal values [8–10]. Also, in extrinsic allergic alveolitis (EAA) lymphocytes are increased in the BALF. However, the composition of lymphocyte subpopulations in BALF obtained from EAA patients differs from that observed in other diseases, as is usually characterized by a low CD4/CD8 ratio [11, 12]. In contrast, in sarcoidosis a high CD4/CD8 ratio is frequently found [9, 10]. Also, in EAA, a mild increase of polymorphonuclear/neutrophils (PMN) and mast cells can be found shortly after inhalation of antigen [11, 12]. Plasma cells (0.1–2%) have been observed in BALF of EAA patients [13, 14]. Idiopathic pulmonary fibrosis (IPF) has no specific diagnostic BAL features. Random increases of BALF lymphocytes, PMN and eosinophils occur in about two-thirds of IPF patients [3, 15, 16].

In BALF of sarcoidosis, EAA and IPF patients, in addition to changes in cellular constituents, Igs and Ig to albumin ratios are frequently increased [1, 3, 4].

In order to evaluate whether a number of selected variables derived from BALF analysis can distinguish between the three diagnostic groups, we have performed a discriminant analysis in patients with sarcoidosis, EAA and IPF.

Materials and Methods

Patients

Retrospectively, the initial BAL of patients with sarcoidosis, EAA and IPF were selected out of all BAL analyses (n=2,008) performed during a 10 yr period, between 1980 and 1990.

In this study, consecutive patients with sarcoidosis (n=193) at time of diagnosis were included. The patient group consisted of patients detected on routine chest X-ray film (n=37), patients with respiratory and general constitutional symptoms (n=110), and patients with erythema nodosum and/or arthralgia and hilar lymphadenopathy (*i.e.* Löfgren's syndrome; n=46). All patients presented with a stage I or II X-ray film; none with a stage III X-ray film. The diagnosis was histologically proven, by biopsy of mediastinal lymph nodes, transbronchial biopsy, open lung biopsy, or liver biopsy. BAL was performed when the sarcoidosis patients were admitted to the hospital, to establish the diagnosis, and before corticosteroids were given.

The diagnosis EAA was based on clinical information, chest X-ray film, pulmonary function tests, the presence of precipitins in peripheral blood, and the disappearance of symptoms after avoidance of the causative antigen or, in some cases, after a short treatment with corticosteroids. All EAA patients (n=39) had recent contact with the causative antigen, but not within the last 48 h before BAL.

Patients with IPF (n=45) commonly presented with an onset of breathlessness on exercise and non-productive cough and, sometimes, with constitutional symptoms. The diagnosis IPF was based on clinical information, chest X-ray film, pulmonary function tests, *i.e.* decrease of lung compliance and diffusion capacity for carbon monoxide, and hypoxaemia, especially on exercise, without hypercapnia. IPF was histologically proven by biopsy. The demonstrated alveolitis was characterized by an infiltration of mononuclear cells, interstitial pneumonitis and/or derangement of parenchymal structures, *i.e.* fibrosis. No patient received corticosteroid treatment, or other medication, either at the time of lavage or before.

The control group consisted of 30 healthy volunteers, without any pulmonary history, having normal chest X-ray film and lung function tests. Table 1 lists the characteristics of the groups studied.

Table 1. – Characteristics of controls and patients with interstitial lung diseases

	C	Sar	EAA	IPF
Patient n	30	193	39	45
Missing cases† n	0	3	1	1
Sex F/M	15/15	95/91	14/24	16/28
Age yrs*	33 (21–55)	35 (18–79)	50 (23–78)	60 (30–79)
NSm	15	140	340	27
Sm cig·day $^{-1}$**	15 (14.9±8.8)	44 (15.7±7.8)	4 (12.8±6.9)	17 (18.0±5.1)

n: number of cases; m: missing cases; *: mean with range in parenthesis; NSm: nonsmokers; Sm: smokers; **: mean±SD number of cigarettes a day; C: controls; Sar: sarcoidosis; EAA: Extrinsic allergic alveolitis; IPF: idiopathic pulmonary fibrosis. †: missing at least one discriminating variable and thus not used in analysis.

Bronchoalveolar lavage

BAL was performed, as reported previously, during fibreoptic bronchoscopy [11]. In short, the procedure is as follows. After premedication (atropine and sometimes diazepam or codeine), and local anaesthesia of the larynx and bronchial tree (tetracaine 0.5%), BAL was performed by standardized washing of the right middle lobe, with four aliquots of 50 ml sterile saline (0.9% NaCl) at room temperature. Simultaneously, peripheral blood samples were taken.

Recovered BALF was kept on ice in a siliconized specimen trap and was separated from cellular components by centrifugation (5 min, at 350×g). Supernatants were directly stored at -70°C after an additional centrifugation step (10 min, at 1,000×g). Cells were washed twice, counted, and suspended in minimal essential medium (MEM; Gibco, Grand Island, New York, USA) supplemented with 1% bovine serum albumin (BSA; Organon, Teknika, Boxtel, The Netherlands). Preparations of cell suspensions were made in a cytocentrifuge (Shandon). Cytospin slides of BAL cells were stained with May-Grünwald-Giemsa (MGG; Merck, Darmstadt, Germany) for cell differentiation. At least 1,000 cells were counted.

If more than 15% lymphocytes were present, T-cell (sub)populations were determined. Total T-cells and subpopulations were recognized by staining with monoclonal antibodies CD2 (OKT11), CD3 (OKT3), CD4 (OKT4) and CD8 (OKT8) (Ortho-pharmaceuticals, Beerse, Belgium). Identification of T-cells reacting with monoclonal antibodies was performed by means of a conventional indirect immunofluorescence technique using fluorescein isothiocyanate (FITC)-labelled goat-antimouse immunoglobuline-Ig (GAM, Nordic, Immunological Laboratories, Tilburg, The Netherlands and Central Laboratory of the Netherlands Red Cross Blood Transfusion Service (CLB), Amsterdam, The Netherlands) .

Albumin determinations were performed, according to a modification of the bromocresol purple (BCP) dye-binding method [17, 18]. In short, this method is as follows: in the presence of a solubilizing agent, BCP binds to albumin at a pH 4.9. The amount of albumin-BCP complex is directly proportional to the albumin concentration. The complex absorbs at 600 nm. Albumin concentrations in serum and BALF are expressed as $g \cdot l^{-1}$ and $mg \cdot l^{-1}$, respectively. Immunoglobulin concentrations, *i.e.* IgM, IgG and IgA in BALF were determined by an enzyme-linked-immunosorbent assay (ELISA) method; microtitre plates were coated with a rabbit anti-human-isotype antiserum (anti-IgM, (CLB, Amsterdam, The Netherlands), anti-IgG and anti-IgA (Dako, Glostrup, Denmark)). Bound Igs from BALF were visualized by using a horseradish peroxidase (HRP) labelled rabbit anti-human-immunoglobulin antiserum (with anti-IgA, anti-IgG, anti-IgM, anti-kappa, anti-lambda reactivity (Dako, Glostrup, Denmark)) and a chromogenic substrate ortho-phenyldiamine (OPD) (Baker, Chemicals BV, Deventer, The Netherlands). Ig concentrations in BALF were expressed in $mg \cdot l^{-1}$ using as a reference a commercial human standard serum, HOO-03 (CLB, Amsterdam, The Netherlands).

Statistical methods

In order to distinguish between the three diagnostic groups, a discriminant analysis was performed, according to the following procedure. Each of the 277 patients in the total study group belongs to one, and only one, diagnostic group. Five of the cases (three sarcoidosis patients; one EAA patient and one IPF patient) had at least one missing discriminating variable, therefore 272 cases were used for the analysis. Thus, 190 patients belong to the sarcoidosis group, 38 to the EAA group and 44 to the IPF group (table 1). Hence, an arbitrary patient out of the total study group has a probability of 190/272=0.70, 38/272=0.14 and 44/272=0.16, respectively, of belonging to either one of the three diagnostic groups. These probabilities, which add up to 1 (as they should), are called "prior probabilities". If a set of predefined characteristics (so-called "explanatory variables") of a patient are known, then these characteristics can be involved in making these probabilities vary among patients. For example, if it is known that a patient is a smoker, then his probabilities may be different from the above prior probabilities, and so also from the probabilities of a non-smoker. The latter probabilities, which can be calculated if we know the smoking status of a patient, are the so-called "posterior probabilities". A statistical technique with which these posterior probabilities can be calculated from the prior probabilities and from the patients characteristics is called discriminant analysis [19].

By means of discriminant analysis, an allocation rule can be derived according to which patients **i** are allocated to one and only one diagnostic category **j** (**j**= 1,...,J) on the basis of a number of explanatory characteristics $x_{i1}, x_{i2}, \ldots, x_{in}$. Taking diagnostic category 1 (**j**=1) as a basis, the explanatory characteristics are summarized in J-1 linear discriminant functions $Y_{ij}(j=2,...,J)$ for each patient **i**: $Y_{ij}=\beta_{0j}+\beta_{1j}X_{i1}+\beta_{2j}X_{i2}+\ldots\beta_{nj}X_{in}$, where the β-coefficients are the same for all patients and are estimated by means of discriminant analysis. The estimated coefficients are presented in table 2.

The probability p_{ij} that patient i falls into diagnostic category **j** (j=1,...,J), given his discriminant function scores $Y_{ij}(j=2,\ldots,Y)$, equals:

$$P_{ij} = \frac{1}{1 + \sum_{j=2}^{J} \exp(Y_{ij})} \quad \text{for } j=1;$$

$$P_{ij} = \frac{\exp(Y_{ij})}{1 + \sum_{j=2}^{J} \exp(Y_{ij})} \quad \text{for } j>1,$$

so that it is guaranteed that $\Sigma_j p_{ij}=1$. Patient **i** is allocated to that diagnostic category **j** for which p_{ij} is maximum.

This analysis is based on the formula of Bayes [19], and on the assumption that in each of three diagnostic groups the explanatory variables have a multivariate Gaussian distribution with different means. If it is also assumed that the variances and covariances differ between the diagnostic groups, then also quadratic terms of the explanatory variables have to be included in the discriminant analysis. In this study, a number of variables (*e.g.* sex and smoking) are clearly non-Gaussian variables. However, discriminant analysis is known to be rather robust for deviations from the Gaussian distribution. The variables are selected into the analysis according to a stepwise procedure. The stepwise selection procedure used is as follows. At each step, the variable with the smallest Wilks' lambda was selected with the significance level used as a criterion for entry (p=0.05) and removal (p=0.10). Age, sex, smoking and yield are included in the analysis as standard personal characteristics, and because of their possible confounding with other variables in the analysis. As it is supposed that the discriminatory power of explanatory variables may depend on the smoking status of a patient, the interaction terms of all explanatory variables (and their quadratic terms) with smoking is also eligible for stepwise inclusion in the model. In order to test the goodness-of-fit of the discriminant analysis model, the predicted and the actual group

Table 2. – Variables eligible for discriminant analysis, with unstandardized canonical discriminant function coefficients mentioned only for those variables and higher order terms that were eventually selected in the discriminant functions*

Explanatory variable	Linear effect		Quadratic effect		Interaction linear and smoking*		Interaction quadratic and smoking*	
	F1	F2	F1	F2	F1	F2	F1	F2
Constant	-20.53985	-15.49421						
Age yrs	0.04000	0.00219					0.00036	0.00001
Sex female	-0.05678	-0.13240						
Smoking yes	0.65372	-0.13799						
BALF								
Yield (out/in)×100	0.01488	0.01539						
Cells $\times 10^4 \cdot ml^{-1}$	-1.76927	-0.21222						
AM %	0.21780	0.49864	0.00002	-0.00344				
PMN %			0.00421	0.00131				
Lym %	0.27903	-0.13799	-0.00070	0.00343				
Eos %			0.00451	0.00292				
MC %								
PC %	-1.28392	-2.54227	0.38701	0.62018				

*: The only variable which had a significant interaction with smoking was age squared. F1: function 1; F2: function 2; AM: alveolar macrophages; PMN: polymorphonuclear neutropils; Lym: lymphocytes; Eos: eosinophils; MC: mast cells; PC: plasma cells.

membership were compared in a different testing set. Therefore, the total set was randomly and evenly split per diagnostic group in a testing and a learning set. The same variables as selected in the discriminant analysis on the total group were used to estimate the discriminant functions in the learning set. Next, these functions were used to compare predicted and actual group membership in the testing set.

Results

Tables 3–5 report the results of BALF cell and protein analyses of the controls and patients with sarcoidosis, EAA, or IPF. Since statistically significant differences were found among nonsmokers (NSm) and smokers (Sm) within the studied groups, these data are shown separately. The recovery or yield (out/in)×100% of BAL was 56±0.8% in the sarcoidosis group, 47±1.4% in the EAA group, 45±1.5% in the IPF group and 58+2.5% in the control group. These results were presented as mean±standard error of the mean. No differences were found between Sm and NSm.

The explanatory variables used in the analysis are listed in table 2. The explanatory variables eventually selected according to the stepwise procedure are those in table 2 with an explicitly mentioned coefficient. Indeed, two discriminant functions appear to be necessary for appropriately discriminating between three diagnostic groups. Function 1 discriminates the three groups for 62%, and function 2 for 38% (see also table 6 and figure 1). Variables used in both functions are total cell count, the percentage of alveolar macrophages (AMs), polymorphonuclear neutrophils (PMNs), lymphocytes, plasma cells and eosinophils.

Table 3. – Absolute number of cells in BALF of controls and patients with interstitial lung diseases

Study group		Am	PMN	Lym	Eos	MC	PC
C	NSm	8.7±4.0	0.15±0.12	1.2±0.9	0.03±0.05	0.006±0.01	0.0±0
	Sm	23.9±12.0	0.25±0.34	1.7±3.5	0.10±0.11	0.006±0.02	0.0±0
Sar	NSm	11.7±7.3	0.27±0.57	6.8±5.8	0.10±0.15	0.03±0.05	0.002±0.001
	Sm	23.2±17.0	0.34±0.36	7.7±9.6	0.16±0.22	0.07±0.13	0.0±0
EAA	NSm	12.1±7.5	1.71±2.00	26.3±18.6	1.01±1.52	0.40±0.51	0.24±0.41
	Sm	24.3±10.5	3.22±2.01	25.0±22.3	1.23±1.33	0.42±0.28	0.47±0.95
IPF	NSm	15.7±20.0	2.83±3.92	3.7±6.3	1.80±3.11	0.09±0.14	0.003±0.001
	Sm	38.1±25.6	6.91±4.00	1.7±1.9	3.04±5.70	0.19±0.40	0.001±0.001

Data are expressd as mean±SD absolute number of cells$\times 10^4 \cdot ml^{-1}$. BALF: bronchoalveolar lavage fluid. NSm: nonsmokers, Sm: smokers. For further abbreviations see legends to tables 1 and 2.

Table 4. – Differential cell count in BALF of controls and patients with interstitial lung diseases

Study group		Am	PMN	Lym	Eos	MC	PC
C	NSm	87.0±5.2	1.6±1.4	11.0±5.2	0.34±0.52	0.07±0.12	0.0±0
	Sm	91.6±8.8	1.2±1.6	6.8±8.6	0.41±0.45	0.02±0.06	0.0±0
Sar	NSm	63.2±17.8	1.7±5.4	34.3±17.9	0.55±0.73	0.16±0.31	0.001±0.001
	Sm	74.9±16.7	1.2±1.2	23.0±16.7	0.72±1.07	0.18±0.25	0.0±0
EAA	NSm	38.2±8.9	4.3±3.4	58.1±14.9	2.6±3.0	0.85±0.71	0.43±0.77
	Sm	57.3±11.2	5.8±2.4	40.8±15.0	3.3±4.7	0.90±0.66	0.45±0.90
IPF	NSm	67.8±19.4	11.7±10.1	13.8±15.7	6.4±9.3	0.31±0.35	0.001±0.003
	Sm	76.0±25.5	11.5±21.5	5.1±4.8	7.1±12.2	0.34±0.56	0.0±0

Data are expressd as mean±SD percentage of the total cell count. For abbreviations see legends to tables 1 and 2.

Table 5. – Proteins in BALF of controls and patients with interstitial lung diseases

Study group		Lalb $mg \cdot l^{-1}$	Lalb/s-alb ×100	IgM/Lalb	IgG/Lalb	IgA/Lalb
C	NSM	60.5±25.6	0.16±0.05	0.01±0.00	0.11±0.08	0.05±0.04
	Sm	68.2±46.9	0.14±0.04	0.0±0	0.21±0.27	0.02±0.01
Sar	NSm	150.9±198.8	0.39±0.57	0.01±0.01	0.42±0.29	0.10±0.14
	Sm	144.6±150.6	0.37±0.39	0.01±0.01	0.41±0.42	0.07±0.07
EAA	NSm	185.4±152.3	0.49±0.40	0.10±0.12	1.62±1.86	0.49±0.86
	Sm	171.5±163.8	0.44±0.36	0.04±0.05	0.77±0.37	0.16±0.11
IPF	NSm	141.3±98.9	0.37±0.28	0.01±0.01	0.40±0.19	0.13±0.10
	Sm	81.5±38.9	0.22±0.11	0.0±0	0.35±0.30	0.15±0.16

Data are expressd as mean±SD. Lalb: lavage albumin; s-alb: serum albumin; IgM, IgG and IgA: immunoglobulins M, G and A. For further abbreviations see legends to table 1 and 3.

Table 6. – Canonical discriminant functions

Fcn	Eigen-value	Percentage of variance	Cumulative percentage	Canonical correlation	After fcn	Wilks' lambda	Chi-squared	DF	Significance
1*	2.7879	62.14	62.14	0.8579	0	0.0978	584.588	28	0.0000
2*	1.6982	37.86	100.00	0.7933	1	0.3706	249.939	13	0.0000

*: marks the two canonical discriminant functions (DF) remaining in the analysis; fcn = function.

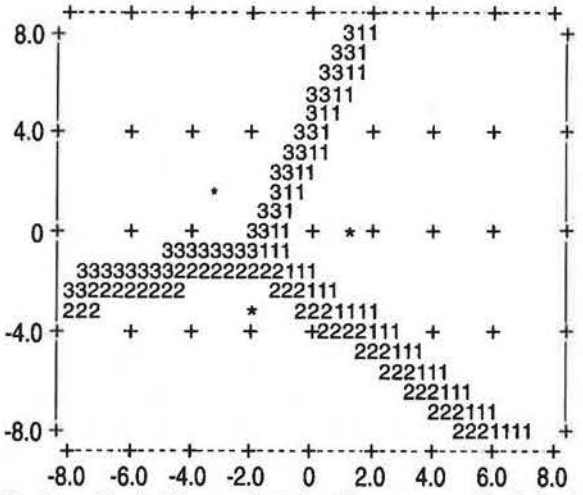

Fig. 1. – Territorial map: clustering (*=group centroid) of the three groups in the plane scanned by the discriminant functions. The symbols 1 represent sarcoidosis patients, 2 extrinsic allergic alveolitis patients and 3 idiopathic pulmonary fibrosis patients; Across: function 1, down: function 2.

The classification results for all cases used in the analysis are shown in table 7. These results are obtained by applying the allocation rule as described above. The percentage of patients correctly classified in all patients with a given actual diagnosis, called the "diagnostic effectiveness" [21] is 100((188+28+36)/272)=92.6. The diagnostic effectivity for sarcoidosis is 100(188/190)=98.9%, for EAA 100(28/38)=73.7% and IPF 100(36/44)=81.8% (table 7). The predicted value of a classification can be calculated as the probability that a patient actually belongs to the predicted group. For the prediction "sarcoidosis", the predicted value (PV^+) equals 100(188/204)=92.2%, and for the prediction "EAA" and "IPF" these values are 100(28/29)=96.6% and 100(36/39)=92.3%, respectively.

The three respective prior probabilities that an arbitrary patient (without using any additional information) actually belongs to a diagnostic group are for sarcoidosis 100(190/272)=69.9%, for EAA 100(38/272)=14.0% and for IPF 100(44/272)=16.2%.

The specificity, *i.e.* the probability of the prediction "non-sarcoidosis" in the group without sarcoidosis is 100((28+1+1+36)/(38+44))=80.5%, for "non-EAA" 100((188+2+7+36)/(190+44))=99.6% and for "non-IPF" 100((188+0+9+28)/(190+38))=98.7%.

The predicted value of the negative result, *i.e.* the predicted value of the group with "non-sarcoidosis" (PV^-) equals 100(28+1+1+36)/(29+39)=97.1%, for "non-EAA 100(188+7+2+36)/(204+39)=95.9%, and for "non-IPF" 100(188+0+9+28)/(204-29)=96.6% (table 7).

The results of the goodness-of-fit test were as follows: in the learning set the discriminant functions were estimated again, yielding a correct classification in 96.2% of the cases. Next, these functions were applied to the testing set in order to predict group membership, yielding 90.1% correctly classified cases.

A separate analysis was performed in a subset of 196 patients, whose BALF was assayed for Igs. In this subgroup, group membership is now correctly predicted in 28 out of 36 EAA patients (77.8%, data not shown). In contrast to the original discriminant analysis without including the Igs, the diagnostic effectiveness was 73.7% (table 7). In the original analysis, 10 EAA patients were incorrectly predicted. Inclusion of Igs in the discriminant analysis corrects the prediction in two out of these 10 cases. No such correction was found in IPF patients. In contrast, inclusion of T-cell subpopulations and CD4/CD8 ratio in the analysis did not result in a better prediction because of too many missing data.

Table 7. – Classification results from the patients with interstitial lung diseases

	Predicted group membership n				Percentage			
Actual group	Sar	EAA	IPF	Total	Spec	DE	PV+	PV-
Sar	188	0	2	190	80.5	98.9	92.2	97.1
EAA	9	28	1	38	99.6	73.7	96.6	95.9
IPF	7	1	36	44	98.7	81.8	92.3	96.6
Total	201	29	39	272				

Spec: specificity; DE: diagnostic effectiveness (*i.e.* sensitivity); PV+: positive predicted value; PV-: negative predicted value. For further abbreviations see legend to table 1.

Discussion

The aim of this study was to investigate whether a number of selected variables derived from routine BALF analysis, made it possible to discriminate between three disorders belonging to the ILD group, all of which show striking similarities in their clinical presentation. To this end, a discriminant analysis was used to distinguish between sarcoidosis, EAA and IPF.

To be of diagnostic value, BAL should be performed using a standard procedure and BALF recovery, *i.e.* yield, should be as comparable as possible. The recovery is, among others, related to BAL procedure and pulmonary function values, so standardization among varying diseases is very difficult. The recovery was included in the analysis, so the effects of the other variables in the model are adjusted for yield. In our studied patient population, the recovery was higher in sarcoidosis patients as compared to EAA and IPF patients. With regard to characteristic changes in cellular components of BALF in various ILD, our results confirm those described by others, showing high absolute and relative numbers of lymphocytes in the sarcoidosis and EAA groups [1–4], high numbers of plasma cells [13, 14] and mast cells [11, 12] in the EAA group, and an increased number of PMN and eosinophils in the IPF group [15, 16]. Interestingly, by including the BALF-cell profile only, we have found a high percentage of cases classified correctly (92.7%) and, as such, a very high overall diagnostic effectiveness. This percentage may be somewhat too optimistic as we found the diagnostic effectiveness decreasing from 96.2 to 90.1% when comparing the learning with the testing set.

In previous studies, immunological parameters have been included in discriminant analysis [6, 21] of the different ILD. Thus, BERTORELLI *et al.* [6] regard T-lymphocyte subpopulations, PMN and eosinophils in BALF, immune complex determination and gallium-67 lung scanning to possess the most important discriminant capacity. However, our study is not quite comparable with the above-mentioned study, because it included other variables and non-BALF parameters, and the use of different inclusion criteria. The EAA population of BERTORELLI *et al.* [6] was recently exposed to the causative antigen, and the authors do not state whether patients who had a BAL after provocation were also included. This is very important, because of the reported influence of the timing of BAL related to the last antigen exposure on the results of BALF analyses [11, 12]. It is also our experience, that the BALF profile after provocation differs from BAL at other times. Moreover, no correlation was made between the smoking status of a patient and the BALF analysis results.

In our study, we found that a little improvement of the diagnostic effectiveness for the EAA group can be obtained by including Igs in the discriminant analysis (77.8%) in addition to the BALF cell-profile. We found increased ratios of IgM, IgG and IgA to albumin in BALF in EAA patients. These increased ratios are supposed to be the result of the immune stimulation in the lung by inhalation of a causative antigen [22]. In contrast, BERTORELLI *et al.* [6] only described IgM in lung biopsy material both in EAA and IPF patients, but not in BALF. Our results suggest that the IgM ratio to albumin in BALF mainly differentiates EAA from the other two diagnostic groups. This was also reported by REYNOLDS *et al.* [22]. The IgG and IgA ratios to albumin are BALF is only high in the nonsmoking EAA patients, compared to the two other groups. BERTORELLI *et al.* [6] did not include the patient smoking status as a variable, this may be due to the conflicting data. REYNOLDS *et al.* [22] recently also reported the negative influence of smoking on total and functional lung Igs, and local immune response. In this study, and in a follow-up study (data not shown), we noticed that the demonstration of plasma cells in BALF is highly suggestive for the diagnosis of EAA [14]. Local production of Igs by plasma cells has been suggested and could be an explanation of the increased ratios of Igs to BALF albumin, especially in EAA patients [23]. Plasma cells and Igs are, therefore, useful in discriminating EAA from other ILD [13, 14].

Previously, other BALF variables have been selected to discriminate sarcoidosis from other lung diseases (COSTABEL *et al.* [21]). They also used a discriminant analysis. However, they selected other BALF variables, *i.e.* percentage of lymphocytes, the CD4/CD8 ratio and Leu7+ natural killer cells. They also concluded that the determination of multiple BALF variables may be of diagnostic help in sarcoidosis. The diagnosis of sarcoidosis can be made by the determination of BALF lymphocyte subpopulations, and may avoid the need for more invasive biopsy procedures [21]. Our results confirm this hypothesis. Moreover, our first discrimination was based on cell differentiation only, without including the CD4/CD8 ratio. However, our data are not comparable with those of BERTORELLI *et al.* [6] and COSTABEL *et al.* [22], because of a different study protocol. It has been reported [23–26] that the total cell count and the absolute and relative number of AMs are increased in BALF of current smokers. In addition, the percentage lymphocytes is decreased. These differences in the BALF profile suggest a modification of the inflammatory reactions in the lungs, due to smoking [27].

Recently, increased mast cells in BALF in sarcoidosis patients were reported. Mast cells were related to a more active sarcoidosis by BJERMER *et al.* [28], who did not differentiate between smokers and nonsmokers. This is in contrast with the results of VALEYRE *et al.* [29], who reported a lower incidence of sarcoidosis, with less severe symptoms, among smokers with a tendency to high numbers of mast cells in BALF. One has to take into account the fact that there is a profound effect of smoking on the number of mast cells in BALF.

This study reports only BALF features with regard to the presentation of the disease and not to the prognosis or the possible response to any therapy. Follow-up studies are needed to investigate the value of BAL to evaluate disease improvement or progression, and the possible influence of therapy on BALF profile.

Moreover, the differences between other disorders with similar clinical presentation to sarcoidosis should be included in the analysis to confirm the discriminant

analysis. For instance, tuberculosis, important in the differential diagnosis of sarcoidosis, has a characteristic BALF profile, different from sarcoidosis [30–33]. Pulmonary Hodgkin's disease, also important in the differential diagnosis of sarcoidosis, has characteristic diagnostic cytological features. Mononuclear Reed-Sternberg cells can be identified in BALF and fine needle aspiration [34]. COSTABEL *et al.* [35] suggested that BAL should be considered as a non-invasive diagnostic approach in cases of pulmonary shadowing associated with malignant haematological disorders. We also found that malign lymphomas, *i.e.* non-Hodgkin's lymphoma and Hodgkin's lymphoma, and tuberculosis have features in BALF profile which enables these disorders to be differentiated from other ILD, such as sarcoidosis (data to be published).

In the discriminant analysis, a number of 14 explanatory variables have been used, derived from 10 different variables measured in the studied patients. From the results so far, we can already conclude that, with a number of selected variables, it is possible to discriminate among patients with sarcoidosis, EAA or IPF. Furthermore, we demonstrated that there is an association between BALF profile and smoking. Therefore, smoking was included as a confounding variable in the discriminant analysis. BAL could serve as an adjunct in concert with other methods for establishing the diagnosis, especially with regard to differentiating between disorders with similar clinical presentation to the studied ILD.

Certainly, attempts still have to be made to develop a more general discriminant model, based on polytomous logistic regression analysis. Currently, we are developing a testing model, which can be used to predict the diagnosis of an arbitrary patient, using information provided only from BALF analysis.

Acknowledgements: The authors gratefully acknowledge M. Diamant for her advice and critical comments during the preparation of this manuscript and M. Donckerwolcke-Bogaert, M. Kolijn-Couwenberg and E. Tuenter for their laboratory work.

References

1. Klech H, Hutter C. – Clinical guidelines and indications for bronchoalveolar lavage (BAL): Report of the European Society of Pneumology Task Group on BAL. *Eur Respir J* 1990; 3: 937–974.
2. The BAL Co-operative Group Steering Committee. – Bronchoalveolar lavage constituents in healthy individuals, idiopathic pulmonary fibrosis, and selected comparison groups. *Am Rev Respir Dis* 1990; 141: 169–202.
3. Daniele RP, Elias JA, Epstein PE, Rossman MD. – Bronchoalveolar lavage: role in the pathogenesis, diagnosis and management of interstitial lung disease. *Ann Intern Med* 1985; 102: 93–108.
4. Reynolds SP, Jones KP, Eduards JH, Davies BH. – Immunoregulatory proteins in bronchoalveolar lavage fluid. A comparative analysis of pigeon breeders' disease, sarcoidosis and idiopathic pulmonary fibrosis. *Sarcoidosis* 1989; 6: 125–134.
5. Dugas M, Wallaert B, Tonnel A-B, Voisin C. – From subclinical alveolitis to granulomatosis. *Chest* 1989; 96: 931–933.
6. Bertorelli G, Pesci A, Consigli GF, *et al.* – Evaluation of some immunological parameters in interstitial lung disease by discriminant analysis. *Respiration* 1988; 54: 23–29.
7. Cantin A, Bégin R, Drapeau G, Rola-Pleszczynski M. – Features of bronchoalveolar lavage differentiating hypersensitivity pneumonitis and pulmonary sarcoidosis at time of initial presentation. *Clin Invest Med* 1984; 7: 89–94.
8. Thomas PD, Hunninghake GW. – Current concepts of the pathogenesis of sarcoidosis. *Am Rev Respir Dis* 1985; 135: 747–760.
9. Meyer KC, Kaminski MJ, Calhoun WJ, Auerbach R. – Studies of bronchoalveolar lavage cells and fluids in pulmonary sarcoidosis. *Am Rev Respir Dis* 1989; 140: 1446–1449.
10. Ward K, O'Conner C, Odlum C, Fitzgerald XM. – Prognostic value of bronchoalveolar lavage in sarcoidosis: the critical influence of disease presentation. *Thorax* 1989; 44: 6–12.
11. Bosch van den JMM, Heye C, Wagenaar SjSc, Velzen-Blad van HCW. – Bronchoalveolar lavage in extrinsic alveolitis. *Respiration* 1986; 49: 45–51.
12. Costabel U. – The alveolitis of hypersensitivity pneumonitis. *Eur Respir J* 1988; 1: 5–9.
13. Costabel U, Bross KJ, Guzman J, Matthys H. – Plasmazellen und Lymhocytensubpopulationen in der bronchoalveolären Lavage bei exogen-allergischer Alveolitis. *Prax Klin Pneumol* 1985; 39: 925–926.
14. Drent M, Velzen-Blad van H, Diamant M, *et al.* – Differential diagnostic value of plasma cells in bronchoalveolar lavage fluid. *Chest* 1993, (in press).
15. Dunnill MS. – Pulmonary fibrosis. *Histopathol* 1990; 16: 321–329.
16. Panos RJ, Mortenson RL, Niccoli SA, King TE. – Clinical deterioration in patients with idiopathic pulmonary fibrosis: causes and assessment. *Am J Med* 1990; 88: 396–404.
17. Carter P. – Ultramicroestimation of human serum albumin: binding of the cationic dye, 5,5'-dibromo-o-cresolsulfonphthalein. *Microchem J* 1970; 15: 531–539.
18. Louderback A, Measley A, Taylor NA. – A new dye-binder technic using bromocresol purple for determination of albumin in serum. *Clin Chem* 1968; 14: 793–794.
19. Cornfield J. – Discriminant functions. *Review Int Stat Inst* 1967; 35: 142–153.
20. Plomteux G. – Multivariate analysis of an enzymic profile for the differential diagnosis of viral hepatitis. *Clin Chem* 1980; 26/13: 1897–1899.
21. Costabel U, Zaiss A, Wagner DJ, *et al.* – Value of bronchoalveolar lavage lymphocyte subpopulations for the diagnosis of sarcoidosis. *In*: Grassi C, Rizzato G, Pozzi E, eds. Sarcoidosis and other granulomatous disorders. Amsterdam, Elsevier Science Publishers BV, 1988; pp. 429–432.
22. Reynolds SP, Edwards JH, Jones KP, Davies BH. – Immunoglobulin and antibody levels in bronchoalveolar lavage fluid from symptomatic and asymptomatic pigeon breeders. *Clin Exp Immunol* 1991; 86: 278–285.
23. Huninghake GW, Crystal RG. – Cigarette smoking and lung destruction: accumulation of neutrophils in the lungs of cigarette smokers. *Am Rev Respir Dis* 1983; 128: 833–838.
24. Tollerud DJ, Clark JW, Brown LM, *et al.* – The effect of cigarette smoking on T-cell subsets: a population based survey of healthy caucasians. *Am Rev Respir Dis* 1989; 139: 1446–1451.
25. Valberg PA, Jensen WA, Rose RM. – Cell organelle motions in bronchoalveolar lavage macrophages from smokers and nonsmokers. *Am Rev Respir Dis* 1990; 141: 1272–1279.
26. Hoogsteden HC, Hal van PTW, Wijkhuis JM, *et al.*

– Expression of the CD11/CD18 cell surface adhesion glycoprotein family on alveolar macrophages in smokers and nonsmokers. *Chest* 1991; 100: 1567–1571.
27. Hughes DA, Haslam PL. – Effect of smoking on the lipid composition of lung lining fluid and relationship between immunostimulatory lipids, inflammatory cells and foamy macrophages in extrinsic allergic alveolitis. *Eur Respir J* 1990; 3: 1128–1139.
28. Bjermer L, Rosenhall L, Ångström T, Hallgren R. – Predictive value of bronchoalveolar lavage cell analysis in sarcoidosis. *Thorax* 1988; 43: 284–288.
29. Valeyre D, Soler P, Clerici C, *et al.* – Smoking and pulmonary sarcoidosis: effect of cigarette smoking on prevalence, clinical manifestations, alveolitis, and the evaluation of the disease. *Thorax* 1988; 43: 516–524.
30. Harf R, Frobert Y, Boit N, *et al.* – Bronchoalveolar lavage findings in localised pulmonary tuberculosis. *Rev Pneumol Clin* 1985; 41: 101–105.
31. Raja A, Baughman RP, Daniel TM. – The detection by immunoassay of antibody to mycobacterial antigens and mycobacterial antigens in bronchoalveolar lavage fluid from patients with tuberculosis and control subjects. *Chest* 1988; 94: 133–137.
32. Drent M, Velzen-Blad van H, Mulder PGH, Bosch van den JMM. – Diagnostic value of bronchoalveolar lavage in non-immune compromised patients with suspected pulmonary and lymph node tuberculosis. II. BAL fluid analyses compared with a normal control group and a sarcoidosis patient group. *Eur Respir Rev* 1991: 15 (Abstract).
33. Baugman RB, Dohn MN, Loudon RG, Frame PT. – Bronchoscopy with bronchoalveolar lavage in tuberculosis and fungal infections. *Chest* 1991; 99: 92–97.
34. Flint A, Kumar NB, Naylor B. – Pulmonary Hodgkin's disease. Diagnosis by fine needle aspiration. *Acta Cytologica* 1988; 32: 221–225.
35. Costabel U, Bross KJ, Matthys H. – Diagnosis by bronchoalveolar lavage of cause of pulmonary infiltrates in haematological malignancies. *Br Med J* 1985; 290: 1041.