Eur J Cancer Clin Oncol, Vol. 20, No. 10, pp. 1243–1247, 1984.
Printed in Great Britain.

0277-5379/84$03.00+0.00

Discriminant Analysis of Hodgkin's Disease and non-Hodgkin's Lymphomas by Age and Serum Proteins

PAOLO MERUCCI,† FRANCO TAGGI,‡ ALESSANDRO MAROLLA,* MARIA R. ABBOLITO,* GAETANO VITELLI,* PAOLO MAROLLA,* LORETTA MELIFFI,* FLAVIA CHIAROTTI,‡ FRANCESCA MENNITI IPPOLITO‡ and FRANCO AMEGLIO*

**Laboratory of Clinical Pathology, Regina Elena Institute for Cancer Research, Rome, Italy and ‡Istituto Superiore di Sanità, Rome, Italy*

Abstract—*Immunonephelometric evaluations of 13 serum proteins were made in 71 patients with two types of lymphoproliferative diseases: Hodgkin's disease (32 patients) and non-Hodgkin's lymphomas (39 patients). The subjects were differentiated by discriminant analysis by means of age and three selected proteins: properdin factor B, IgM and ceruloplasmin. The results obtained permitted classification of 90% of the cases reported.*

INTRODUCTION

IN A general sense no diagnostic criteria may be obtained from the modifications of serum proteins for a screening of cancer patients in a given population [1–7]. As aspecific markers, many proteins have generally been studied both individually and grouped in various ways [2–5, 6, 8]. Their correlation with clinical patterns in cancer patients has also been examined in various types of malignancies [1, 2, 4, 7, 8].

With this type of monitoring more information can be obtained for Hodgkin's disease (HD) and for non-Hodgkin's lymphomas (NHL). Interest in these markers is not purely academic but has led to a deeper understanding of the clinical characteristics of these diseases [5, 6, 9–16].

In the present study a discriminant analysis was performed utilizing 15 different variables: age (at the time of onset of disease), sex and the following 13 serum proteins: albumin (ALB), transferrin (TRF), C'3, C'4, properdin factor B (PFB), IgG, IgM, IgA, ceruloplasmin (CER), α-acid glycoprotein (AAG) and haptoglobin (HPT).

The aim of the present report was to ascertain whether a statistical elaboration of the determined values of the above-mentioned variables could permit calculating, by means of a specific program inserted into the computer, a 'canonical variable' to verify a possible division of the patients into the two classes (HD and NHL), correctly differentiated with respect to histological diagnosis. Other authors have found this type of analysis useful to distinguish between metastatic cancer of the prostate and non-metastatic cancer on the basis of five variables [17]

MATERIALS AND METHODS

The study was carried out on two groups of patients suffering from lymphoproliferative diseases (HD and NHL): 32 patients with HD and 39 with NHL of both sexes and diagnosed according to histological examination. The staging of the patients is reported below, although this was not used in the statistical analysis:

HD patients: 17 patients with mixed cellularity: five stage II, five stage III, seven stage IV; six patients with nodular sclerosis: one in remission, one stage II, four stage IV; one patient with lymphocytic type, in remission; one patient with lymphocyte depletion, stage II; and seven non-classified patients.

NHL patients: 14 with reticular sarcomas: six stage II, one stage III, seven stage IV; seven patients with lymphomas of low-grade malignancy: three stage II, three stage IV and one stage III; 11 patients with lymphomas of high-grade malignancy: four stage II, two stage III, four stage IV and one in remission; one patient with chronic

Accepted 14 April 1984.
†To whom correspondence and requests for reprints should be addressed at: Istituto Regina Elena, Viale Regina Elena, 291, 00161 Rome, Italy.

lymphatic leukemia (CLL); one patient with an uncertain histology; and five patients with no ascertained histological grading.

Blood was collected from the fasting subjects by venopuncture. The sera were obtained after coagulation at room temperature and processed within 2–4 hr after the collection or stored in a refrigerator at −20°C until employed for analysis. Sera were defrosted only once and then discarded.

A Beckman ICS II nephelometer was used to determine the following variables: ALB, TRF, CER, C'3, C'4, PFB, HPT, AMG, AAT, IgA, IgG and IgM. All Beckman antisera were used. Results were expressed as mG/dl. Age was expressed in years.

Immunonephelometric methods offer a good correlation with values obtained using immunodiffusion techniques [7, 18, 19]. The variation coefficients found for every protein on ten repetitive tests are as follows: ALB: 0.69; TRF: 0.92; IgG: 0.93; IgA: 1.68; IgM: 2.67; C'3: 2.12; C'4: 1.03; PFB: 1.29; AMG: 1.53; HPT: 1.32; AAT: 1.15; CER: 0.80; and AAG: 1.4.

A statistical evaluation of the results was made with an IBM 4341 computer at the Istituto Superiore di Sanità of Rome with a package of BMDP programs (P7M Applied Program) from the University of California [20]. Evaluation of the final classification was made employing the Jackknifed method [20–22].

Data obtained were analyzed in multivariate terms by discriminant analysis.

RESULTS

Discriminant analysis

The values obtained for the 13 proteins in the two classes of patients were inserted into the computer together with two other variables: the sex of the patient and the time of disease onset. The results obtained by the computer according to the applied program, P7M, established that only four variables were able to discriminate between HD and NHL patients with a 90% rate of precision (Fig. 1).

The resulting variables were: age of disease onset, CER, PFB and IgM. Adding more variables to the program (such as HPT, AAG and AAT) [20] other than those previously selected, did not alter the results (88.6% for AAT, 89.0% for HPT and 90.1% for AAG). If the 'age' variable was not considered, the resulting percentage of corrected discrimination was reduced to 72%. Similar results (70%) were obtained removing the four selected variables. In this case the second choice was AAT and C'4. Removal of the 'age' variable primarily disturbs the NHL group, while the choice of AAT and C'4 creates confusion in the distribution in the HD group in particular.

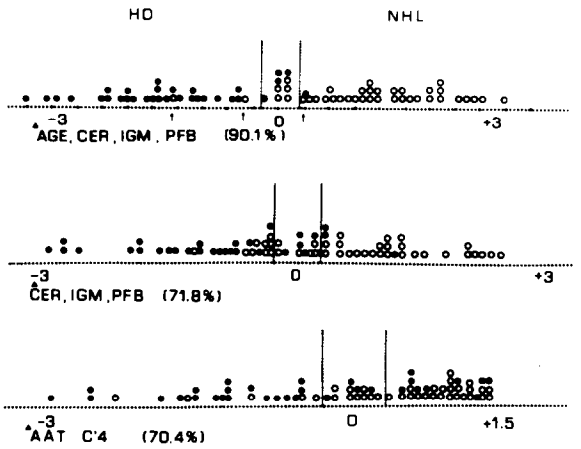

Fig. 1. Discriminant analysis. Distribution of Hodgkin's (HD) and non-Hodgkin's (NHL) patients according to values of the canonical variable (C.V.) calculated on the basis of 4, 3 or 2 variables. The variables selected are: (1) age at onset of disease, ceruloplasmin (CER), properdin factor B (PFB) and IgM (top); (2) CER, PFB and IgM (center); and (3) α_1-antitrypsin and C'4 (bottom). Dark circle: histologically diagnosed HD; white circle: histologically diagnosed NHL. Positive values: exact position of NHL; negative values: exact position of HD. Vertical lines: ±10% limit around the '0' equivalent to 50% probability of correct classification of patients. Numbers within parentheses indicate the percentage of patients correctly classified.

Table 1 presents the percentage of correct classification (according to the Jackknifed classification) of patients with HD and NHL by means of the following variables: age at disease onset, CER, PFB and IgM.

The canonic variable (C.V.) obtained may be calculated as follows: C.V. = (0.00808 × IgM) − (0.04749 × PFB) − (0.04784 × CER) + (0.06035 × age) + 0.34113. The final negative value of the C.V. obtained corresponds to the HD group, while the positive value corresponds to the NHL group.

The posterior probability of an exact classification for each subject according to the algebraic value of the canonic variable is presented in Table 2.

DISCUSSION

The serum proteins studied may be biologically classified as follows: (a) index of nutrition (ALB);

Table 1. Jackknifed classification of correct discrimination between HD and NHL patients

Class	% correct	No. of NHL patients	No. of HD patients	Total No. of patients
NHL	94.9	37	2	39
HD	84.4	5	27	32
Total	90.1	42	29	71

HD: Hodgkin's disease; NHL: non-Hodgkin's lymphoma. Variables employed: age (at disease onset), CER, PFB and IgM.

Table 2. Posterior probability of correct classification

C.V. (absolute value)	HD (C.V.-) (%)	NHL (C.V.+) (%)
0.5	70.9	75.2
1.0	91.5	95.7
1.5	97.6	98.8
2.0	99.4	99.7
2.5	99.9	99.9

C.V.: canonical variable; HD: Hodgkin's disease patients; NHL: non-Hodgkin's lymphoma patients.

(b) molecules of transport (CER, TRF); (c) immunological defenses (IgG, IgM, IgA, C'3, C'4, PFB); and (d) acute-phase reactant proteins (APRPs, that is, HPT, AAT, AAG and AMG). This latter group may include many proteins already listed in the previous groups [9, 23–28].

A subdivision of this sort is not rigid in as much as, given the multifunctional aspects of the previously mentioned proteins, they may pertain to more than one single group.

The choice of the 13 proteinic variables studied was determined in part by the availability of methods and in part by the discrepancies of behavior in the two diseases. Although it was already known that some proteins in our sequence were correlated variates, we let the computer choose the most suitable ones.

We recently reported the most relevant variations of the 13 serum proteins observed simultaneously in a group of neoplastic patients [7, 18]. The data obtained were compared with those observed in a healthy population. The differences between the two were extremely significant ($P<0.001$) for CER, AAT, PFB, ALB, TRF and IgM and only partially significant for AMG ($P<0.05$).

The present investigation allowed us to examine a larger number of patients, homogeneous for the two classes of tumors, and to analyze our results in greater depth. The two groups of patients were histologically subdivided.

The statistical evaluation of data obtained led us to choose four of the 15 variables considered (age, CER, PFB and IgM) by means of which we found the highest statistical discrimination between HD and NHL patients (Fig. 1).

A 'canonical variable' differentiating HD patients from NHL patients may be calculated according to the algebraic formula on the basis of the algorithm indicated (see Results) [21, 22].

Regarding CER and PFB, it is not surprising that their coefficients are negative. In fact, these two proteins are often increased in the sera of HD patients and in this way give a final negative result in the C.V. (H.D. area). On the contrary, age and IgM, presenting lower values in HD patients, are indicated by positive values and their increase modifies the resulting C.V. towards the positive area (NHL). The positioning of the two groups of patients with respect to point zero is obviously completely conventional.

There is a partially definite separation between the two groups of subjects and the following may occur (Fig. 1): (a) a high probability of exact patient classification; (b) a low probability of exact patient classification; and (c) erroneously classified patients.

Three cases are extremely incorrectly placed (two NHL and one HD patient). This apparently incorrect placement is due to age, which does not correspond to the more frequent distribution observed in both diseases. Two NHL patients are 15 and 25 yr old respectively. In the case of the HD patient, the age is over 60 yr. All of these subjects present values of the three remaining variables coherent with those of the category to which they belong. Unclearly defined cases and those with different probabilities of being correctly classified (dispersion along the axis) indicate a lack of clinical homogeneity in the two groups.

The high percentage of patients correctly placed emphasizes the well-known biological differences between these diseases, shown by the diversity of the variables examined. On the other hand, it is known that the histological diagnosis, controlled in a large number of patients, revealed many incorrect evaluations and, consequently, an erroneous interpretation of the histological pictures [29].

The chosen variables have been extensively studied for many years and the majority of them belong to the APRPs [4, 23, 30].

Analysis of the variables

The increase in PFB and C'4, together with C'3 and other APRPs, may be considered an expression of the inflammatory conditions in neoplastic and lymphoproliferative diseases [25, 31, 32].

In HD patients IgM decreases because of synthesis rate impairment rather than because of an increase in catabolism. The decrease is related to an advancement of the clinical stage more than to radiant therapy [33–35].

Copper and CER increases in the sera in the lymphoproliferative diseases, together with many others [6, 36], represent one of the factors most studied and one of the most reliable parameters regarding HD activity [36].

This kind of research is also clinically important because the natural history of these two neoplastic diseases is not always clearly distinguished and the therapeutic strategy differs from one to another [9].

Discriminant analysis presents a new point of view of the problem and permits making better use of commonly employed variables. In the future, results that include a larger number of cases could be correlated to the histological stage. Prolonged observation of the C.V. values could be clinically important if the behavior of the parameter was synchronous to the progression or regression of the disease.

Acknowledgement—The authors wish to express their gratitude to Ms Gillian McIntire for revision of the text and preparation of the manuscript.

REFERENCES

1. Blatt J, Reaman G, Poplack DG. Biochemical markers in lymphoid malignancy. *N Engl J Med* 1980, **303**, 918–922.
2. Cooper EH. Multiparametric markers in the monitoring of cancer. *Recent Prog Cancer Res* 1979, **67**, 97–103.
3. Child JA, Cooper EH, Illingworth S, Worthy TS. Biochemical marker in Hodgkin's disease and non-Hodgkin's lymphoma. *Recent Results Cancer Res* 1977, **64**, 180–189.
4. Cooper EM, Stone J. Acute phase reactant proteins in cancer. *Adv Cancer Res* 1979, **67**, 97–103.
5. Kaplan HS. Hodgkin's disease: biology, treatment, prognosis. *Blood* 1981, **57**, 813–822.
6. Ray GR, Wolf PH, Kaplan HS. Value of laboratory indicators in Hodgkin's disease: preliminary results. *Natl Cancer Inst Monogr* 1973, **36**, 315–323.
7. Ameglio F, Abbolito MR, Vitelli G, Marolla P, Merucci P. Nephelometric determinations of 11 proteins in sera of cancer patients: clinical views. *J Exp Clin Cancer Res* 1982, **1**, 153–158.
8. Ward AM, Cooper EH, Turner R, Anderson JA, Neville AM. Acute phase reactant proteins profiles: an aid to monitoring large bowel cancer by CEA and serum enzymes. *Br J Cancer* 1977, **35**, 170–178.
9. Bonadonna G, Robustelli della Cuna G. *Manuale di Oncologia Medica.* Milan, Masson, 1980.
10. Kaplan HS. *Hodgkin's Disease.* Cambridge, MA, Harvard University Press, 1972.
11. Weltmann JE, Cunningham JK, Gellhorn A. The clinical picture of Hodgkin's disease. *Cancer Res* 1966, **26**, 1047–1060.
12. Arends T, Conrad E, Rundles RW. Serum proteins in Hodgkin's disease and malignant lymphoma. *Am J Med* 1954, **16**, 833–841.
13. Kaplan HS. *Hodgkin's Disease* (2nd Edn). Cambridge, MA, Harvard University Press, 1980.
14. Bjorkholm M, Holm G, Mellstedt H *et al.* Prognostic factors in Hodgkin's disease. *Scand J Haematol* 1978, **20**, 306–318.
15. Goulian M, Fahey JL. Abnormalities in serum proteins and protein bound hexose in Hodgkin's disease. *J Lab Clin Med* 1961, **57**, 408–419.
16. Kaplan HS. Hodgkin's disease: unfolding concepts concerning its nature, management and prognosis. *Cancer* 1980, **45**, 2439–2474.
17. Ward AM, Cooper EM, Houghton AL. Acute phase reactant proteins in prostate cancer. *Br J Urol* 1977, **49**, 411–418.
18. Ameglio F, Marolla P, Abbolito MR *et al.* Marcatori biologici aspecifici e malattia neoplastica. *Atti 32° Congr Ass Pat Clin, San Remo, 1982.* In press.
19. Peracino A, Marcovina S, Fenilli B. Immunonephelometric and immunoturbidometric techniques for plasma protein evaluation. *Ric Clin Lab* 1978 **8**, 125–134.
20. Dixon WJ, ed. *BMDP Manual.* Los Angeles, CA, University of California Pres, 1981, 519–536.
21. Overall JE, Kleitt JC. Applied multivariate analysis. In: *Use of Clinical Variates for Classification.* New York, McGraw Hill, 1978, Ch. 14.
22. Mosteuer F. The jacknifed. *Rev Inst Int Statist* 1971, **39**, 363–368.
23. Ritzman SE, Daniels JC. *Serum Protein Abnormalities. Diagnostic and Clinical Aspects.* Boston, MA, Little Brown & Co., 1975.
24. Hoffbrand BI. Hodgkin's disease and hypogammaglobulinemia. A rare association. *Br Med J* 1964, **1**, 1156–1158.
25. Rottino A, Levy AL. Behavior of total serum complement in Hodgkin's disease and other malignant lymphomas. *Blood* 1959, **14**, 246–254.
26. Muller-Eberhard HJ, Schreiber RD. Molecular biology and chemistry of the alternative pathway of complement. *Adv Immunol* 1980, **29**, 1–47.
27. Goldstein IM, Kaplan HB, Edelson HS, Weissmann G. Ceruloplasmin: an acute phase reactant that scavenges oxygen derived free-radicals. *Ann NY Acad Sci* 1982, **389**, 368–379.

28. Waldmann T, Trier T, Fallon JH. Albumin metabolism in patients with lymphomas. *J Clin Invest* 1963, **42**, 171–178.
29. Miller THP, Byrne GE, Jones SE. Mistaken clinical and pathologic diagnosis of Hodgkin's disease: a Southwest Oncology Group study. *Cancer Treat Rep* 1982, **66**, 645–651.
30. Hobbs JR. Tumor markers in monitoring cancer. *Ric Clin Lab* 1978, **8**, 29–37.
31. Desforges JF, Rutherford CJ, Piro A. Hodgkin's disease. *N Engl J Med* 1979, **301**, 1212–1222.
32. Lichtenfeld JL, Wiernike PH, Mardiney MR Jr, Zarco RM. Abnormalities of complement and its component in patients with acute leukemia, Hodgkin's disease and sarcoma. *Cancer Res* 1976, **36**, 3678–3680.
33. Goldman JM, Hobbs JR. The immunoglobulins in Hodgkin's disease. *Immunology* 1967, **13**, 421–431.
34. Berth WF, Wochner RD, Waldmann TA. Metabolism of human macroglobulins. *J Clin Invest* 1964, **43**, 1036–1040.
35. Anomah Ngu V, McFarlane H, Osunkoya BO, Udeozo IOK. Immunoglobulins in Burkitt's lymphoma. *Lancet* 1966, **ii**, 414.
36. Halliwell B, Gutteridge JMC. Ceruloplasmin and superoxide radicals. *Lancet* 1982, **11**, 556.

After submitting this paper for publication, another 22 cases (7 HD and 15 NHL) were observed. The rate of exact classification performed by C.V. was the same as that previously found (of 22 patients, 20 were correctly positioned).