*Journal of Pharmaceutical & Biomedical Analysis*
Vol. 8, No. 3, pp. 229–234, 1990
Printed in Great Britain

0731–7085/90 $3.00 + 0.00


# Chemometric study and analytical enzymatic methods for diagnosis of cholesterol gallstones

L. CAMPANELLA,* M. TOMASSETTI and G. VISCO

*Department of Chemistry, Rome University, "La Sapienza" Pz. Aldo Moro 5, Rome, Italy*

**Abstract**: The lithogenic index (IL) provides an estimate of cholesterol saturation in gallbladder bile and is of possible value for prediction of gallstone formation. A package for pattern recognition of analytical chemical data, known as "Parvus", was used to study the different values of IL obtained experimentally using common enzymic methods for cholesterol and bile salts and other analytical techniques for phospholipids. Ten patients were investigated and some interesting conclusions were drawn, both on the equivalence of various analytical methods for the determination of phospholipids and on the contribution of pattern recognition analysis to the diagnosis of gallstones.

**Keywords**: *Gallstones; lithogenic index; bile; enzyme analysis; pattern recognition analysis; chemometrics.*

## Introduction

A number of studies [1–9] have been carried out in this laboratory in search of a test for the diagnosis of gallstone disease. Both clinical interest and analytical problems are involved although the introduction of different diagnostic techniques [10] has partly diminished interest in this kind of test. Analytical and physico-chemical problems are still relevant, however, since they deal with the etiology of gallstones. In order to evaluate the IL as a test for the risk of gallstone formation [2, 8] concentrations of bile salts, phospholipids and cholesterol in the gallbladder bile have been determined in 10 patients selected for potential gallstone problems. The analytical problems of phospholipid determination in bile have also been studied. The choline-containing phospholipids have been measured by different enzymic methods [6, 9] and the total phosphorus content by a chemical spectrometric method [4]. The results were subjected to computer modelling using software for IL calculation [11] and a software package for pattern recognition analysis known as "Parvus" [12]. Useful observations have been obtained on the variation of the value of the lithogenic index as a function of the analytical methods adopted for determination of phospholipid concentration, and on the question whether these methods are effectively interchangeable for the calculation of IL.

## Experimental

### *Samples*

Bile samples were cholecystic gallbladder aspirates, supplied by the Second Medical Clinic of Rome University "La Sapienza". These were obtained from patients, aged 30–50 years, all of whom had liver disease and suspected cholesterol stones.

### *Chemical reagents and apparatus*

Enzymic-amperometric measurements for phospholipids containing choline were carried out by the flow apparatus. Reagents and enzymic immobilization methods were as previously described [6, 9] using a commercial oxygen probe in apparatus supplied by Instrumentation Laboratory (I.L. 213, Milan, Italy).

Choline oxidase was supplied by Sigma Chemical Co. (St. Louis, USA) and phospholipase D by Boehringer Biochemia (Mannheim, FRG). The enzymic colour tests for phospholipids [6, 9] were obtained from Poli S.p.A. (Milano, Italy; Cat. No. 3220), for cholesterol [1] from Boehringer (Mannheim, FRG; Cat. No. 124079) and for bile acids [2] from Nyegaard Co. (Oslo, Norway; Cat. No. 999955). The Fiske–Subbarow reagent for the Bartlett [13] spectrometric method for total phosphorus [4] was supplied by the Sigma Chemical Co. All other reagents were of analytical grade and supplied by Carlo Erba

* Author to whom correspondence should be addressed.

(Milano, Italy). Spectrometric measurements were performed with a Perkin–Elmer spectrophotometer model 320, with a 1-cm quartz cell by procedures described in previous papers [1, 2, 4, 6, 9, 13, 14].

*Software and hardware*

To calculate the lithogenic indices from the 70 experimental values for cholesterol (CH), bile salts (BS) and phospholipids (PL), a personal computer HP 86A was used with "Litogindex" program [11], a Basic advanced program ROM, two 360 kBytes floppy disks and a graphic printer HP 82905B. graphical representation of the "scores" and all the relative calculations were performed by the "Parvus" program [12] on an Apple IIe computer with at least 64 kBytes using a printer Apple ImageWriter II and superserial card. The histograms were produced by a "Lotus" V. 1A program of the Lotus Corporation [15] using an IBM XT3 with graphic printer Epson FX1000.

## Results and Discussion

The values obtained for CH and BS, in the bile samples are reported in Table 1. The concentrations of phospholipids have been determined both by enzymic-spectrometry, with or without blank subtraction, and by enzymic-amperometry [6]. Two sets of values were obtained spectrometrically, one using the Takayama's procedure [14] with reagent blank subtraction, and the other by a specially developed procedure [6] with subtraction of reagent and sample blanks. The latter technique seemed generally to provide better results [6, 9]. The enzymic-amperometric method measures free choline rather than the total of free choline and choline-containing phospholipids; thus, from the difference between the two values the concentration of the choline-containing phospholipids can be calculated [6, 9]. In the same samples total phosphorus was determined by Bartlett's spectrometric method [3, 13]; this measurement has been used to represent phospholipids concentration [2] before enzymic methods were well developed. All the results are summarized in Table 1.

The results for phospholipids concentration (by the five methods), for bile salts and for cholesterol were used to obtain five values for IL for each patient using the "Litogindex" program [11] (Table 2). All the IL values were examined using the "Parvus" package. In this the values are first normalized by the "normal" program but a particularly useful normalization of row, between zero and one, was carried out by the "Lotus" program [15] (Table 3A). Two of the five methods for phospholipids determination, yielded higher (method b), or lower (method e) values for IL for all patients, while the remaining methods (a, c, d) yielded intermediate values. A similar but reverse trend was shown from PL values in Table 1, "normalized" in the same way (Table 3B).

This observation seems to indicate that the different methods for the quantitative determination of phospholipids are not always equivalent when used for calculation of the IL value. This agrees with a comparison, by least-

**Table 1**
Experimental values in mmol $l^{-1}$ of: cholesterol (CH), bile salts (BS) and phospholipids (PL) for methods a–e

| Patient No. | CH | BS | PL (a) | PL (b) | PL (c) | PL (d) | PL (e) |
|---|---|---|---|---|---|---|---|
| 1 | 6.3 | 56.9 | 4.9 | 4.3 | 5.1 | 7.0 | 13.6 |
| 2 | 12.1 | 126.8 | 52.0 | 36.8 | 42.3 | 45.7 | 54.5 |
| 3 | 20.2 | 173.8 | 79.9 | 60.9 | 70.4 | 79.3 | 129.4 |
| 4 | 12.9 | 137.2 | 17.0 | 16.4 | 18.6 | 21.5 | 27.0 |
| 5 | 4.2 | 48.7 | 9.2 | 4.8 | 8.7 | 10.4 | 10.6 |
| 6 | 23.9 | 202.8 | 112.0 | 88.0 | 84.9 | 97.1 | 140.1 |
| 7 | 6.4 | 133.6 | 40.8 | 36.2 | 38.8 | 40.8 | 51.4 |
| 8 | 5.5 | 46.5 | 20.0 | 11.0 | 21.3 | 24.9 | 34.4 |
| 9 | 0.6 | 8.0 | 2.1 | 0.5 | 2.1 | 2.3 | 3.4 |
| 10 | 14.2 | 120.4 | 31.9 | 9.0 | 30.0 | 34.0 | 47.0 |

(a) Lecithin + choline by enzymic-amperometry [6].
(b) Lecithin only by enzymic-amperometry [6, 9].
(c) Lecithin + choline by enzymic-spectrometry with reagent and sample blank correction [6].
(d) Lecithin + choline by enzymic-spectrometry with reagent blank correction only [10].
(e) Total phosphorus by Bartlett's spectrometric method [3, 13].

**Table 2**
The fifty values of IL, calculated by the "Litogindex" program, from the data reported in Table 1, using PL values, reported in the corresponding (a), (b), (c), (d) or (e) columns of Table 1. (The CH and BS values are those reported in the first and second columns of Table 1)

| Patient No. | IL (a) | IL (b) | IL (c) | IL (d) | IL (e) |
|---|---|---|---|---|---|
| 1 | 3.21 | 3.44 | 3.15 | 2.45 | 1.54 |
| 2 | 0.79 | 0.99 | 0.90 | 0.86 | 0.77 |
| 3 | 0.85 | 0.99 | 0.91 | 0.85 | 0.69 |
| 4 | 1.88 | 1.92 | 1.77 | 1.52 | 1.29 |
| 5 | 1.56 | 2.37 | 1.62 | 1.42 | 1.40 |
| 6 | 0.76 | 0.86 | 0.87 | 0.82 | 0.71 |
| 7 | 0.50 | 0.54 | 0.51 | 0.50 | 0.43 |
| 8 | 1.10 | 1.76 | 1.07 | 1.00 | 0.89 |
| 9 | 1.57 | 6.77 | 1.57 | 1.42 | 1.10 |
| 10 | 1.28 | 3.06 | 1.34 | 1.23 | 0.99 |

**Table 3A**
IL values of Table 2, normalized from 0 to 1 by "row normalization"

| Patient No. | IL (a) | IL (b) | IL (c) | IL (d) | IL (e) |
|---|---|---|---|---|---|
| 1 | 0.88 | 1.00 | 0.85 | 0.48 | 0.00 |
| 2 | 0.09 | 1.00 | 0.59 | 0.41 | 0.00 |
| 3 | 0.53 | 1.00 | 0.73 | 0.53 | 0.00 |
| 4 | 0.94 | 1.00 | 0.76 | 0.37 | 0.00 |
| 5 | 0.16 | 1.00 | 0.23 | 0.02 | 0.00 |
| 6 | 0.31 | 0.94 | 1.00 | 0.69 | 0.00 |
| 7 | 0.64 | 1.00 | 0.73 | 0.64 | 0.00 |
| 8 | 0.24 | 1.00 | 0.21 | 0.13 | 0.00 |
| 9 | 0.08 | 1.00 | 0.08 | 0.06 | 0.00 |
| 10 | 0.14 | 1.00 | 0.17 | 0.12 | 0.00 |

**Table 3B**
PL values of Table 1, normalized from 0 to 1 by "row normalization"

| Patient No. | PL (a) | PL (b) | PL (c) | PL (d) | PL (e) |
|---|---|---|---|---|---|
| 1 | 0.06 | 0.00 | 0.09 | 0.29 | 1.00 |
| 2 | 0.86 | 0.00 | 0.31 | 0.50 | 1.00 |
| 3 | 0.28 | 0.00 | 0.14 | 0.27 | 1.00 |
| 4 | 0.06 | 0.00 | 0.21 | 0.48 | 1.00 |
| 5 | 0.76 | 0.00 | 0.67 | 0.97 | 1.00 |
| 6 | 0.49 | 0.06 | 0.00 | 0.22 | 1.00 |
| 7 | 0.30 | 0.00 | 0.17 | 0.30 | 1.00 |
| 8 | 0.38 | 0.00 | 0.44 | 0.59 | 1.00 |
| 9 | 0.55 | 0.00 | 0.55 | 0.62 | 1.00 |
| 10 | 0.60 | 0.00 | 0.55 | 0.66 | 1.00 |

squares fitting, of the PL values corresponding to two of the five methods (Table 1) for PL determination. On the other hand, the application of the Mann–Whitney [16] and Wilcoxon [17] tests to the PL values of Table 1 generally support the $H_0$ (null hypothesis) [17]. This point is important from a diagnostic point of view, and is worthy of further consideration; therefore the five IL values, corresponding to the five methods of analysis for phospholipids, were considered as "objects" in relation to the 10 patients as variables. The convergence of the five methods was estimated by the "KNN" program which performs a piecewise-linear classification of the objects, and for each object computes the Euclidian distance [17] from the others and recognizes the K Nearest [12]. Table 4 and the histograms of Fig. 1 summarize the findings in this kind of analysis. The IL values in columns (b) and (e), of Table 2, relative to the methods for PL(b) and PL(e) of Table 1, have greater Euclidian distance than the other three methods. This conclusion is confirmed by the relative projection of the "scores" of the "eigenvectors" with 98% of information (Fig. 2), obtained by the "Varvar" program [12]. The original values were normalised by "autoscaling" and then the generalized covariance matrix was carried out by the "Matcal" program [12]; finally the "loading" and the "scores" were calculated by the "Eigen" program [12]. In Fig. 2 it can be easily observed that a group of three points [correlated by the three methods (a), (c) and (d)] is compact, while the other two points [correlated by (b) and (e) methods], are separated from each other and are almost equidistant from the group of the other three. This is all in good agreement with the results by "KNN" method and with previous observations derived from "normalizations" in Table 3.

If the significance of the values for IL (Table 2) is considered from a clinical point of view [1, 8, 18, 19] it is evident that for four patients, (Nos 2, 3, 6 and 7) IL values are in every case $<1$ and are independent of the analytical method for phospholipids measurement; for the other four patients, (Nos 1, 4, 5 and 9) IL values are in every case $\geqslant 1$.

This has clinical significance [2, 8]. Carey and Small [20] consider the first type of patient

**Table 4**
"Euclidian distances" from different "objects", corresponding to the series of five IL values, relative to the five methods for phospholipid determination, obtained by the "KNN" program

| | (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|---|
| (a) | 0.00 | 5.19 | 2.17 | 2.10 | 4.00 |
| (b) | 5.19 | 0.00 | 4.37 | 5.41 | 7.84 |
| (c) | 2.17 | 4.37 | 0.00 | 1.76 | 4.89 |
| (d) | 2.10 | 5.41 | 1.76 | 0.00 | 3.33 |
| (e) | 4.00 | 7.84 | 4.89 | 3.33 | 0.00 |

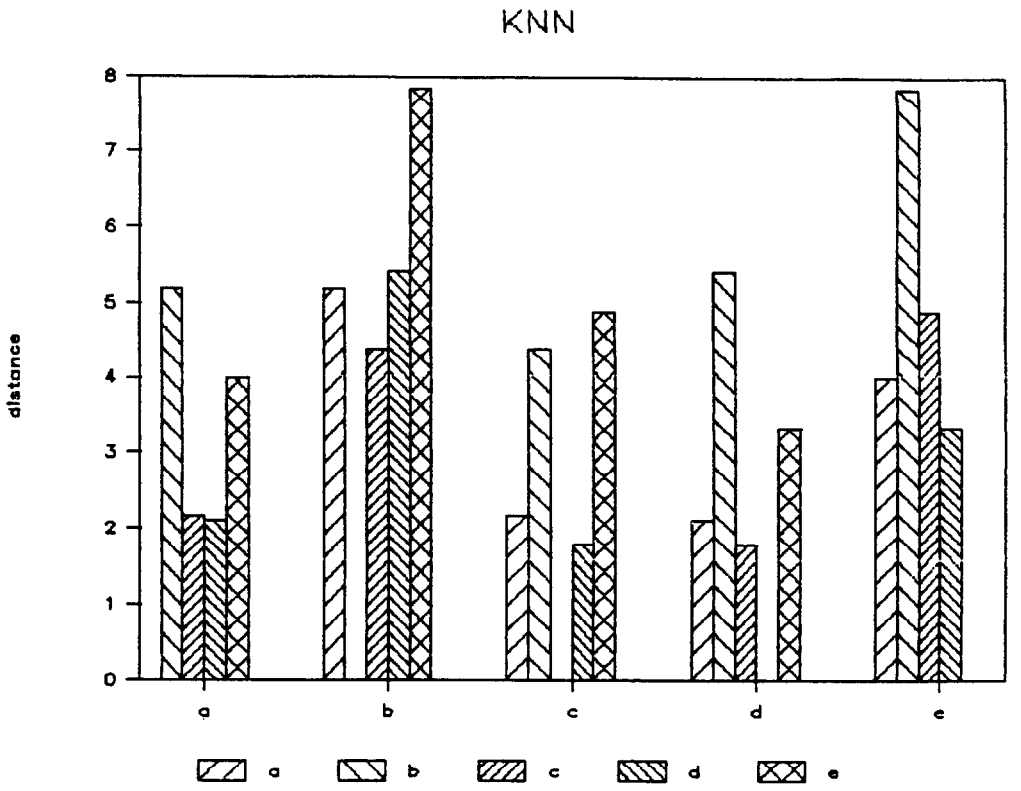


**Figure 1**
Histogram of "Euclidian distances" obtained by the values reported in Table 4.

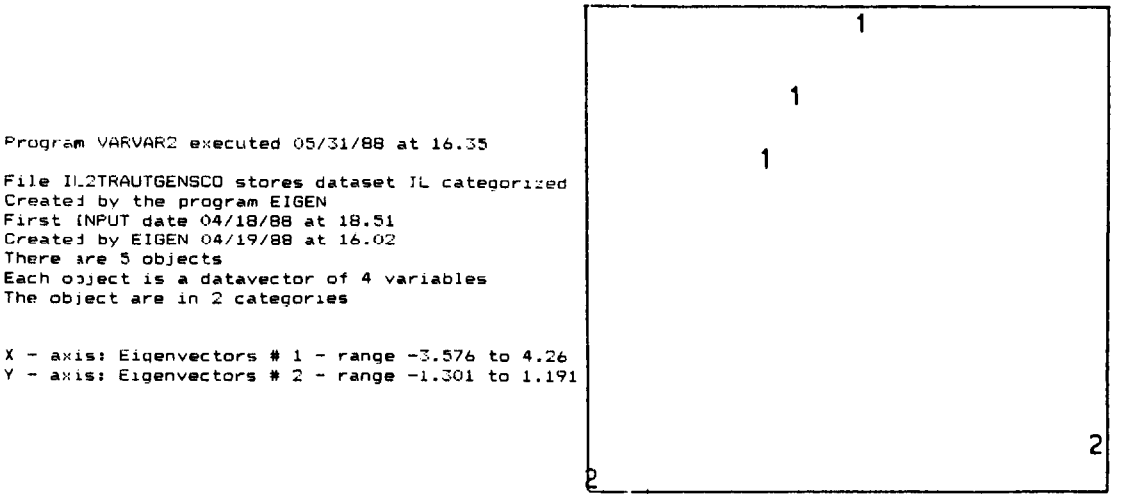


**Figure 2**
Graphical representation of the "scores" of eigenvectors with 99.8% of the total information. Five objects (methods) and 10 variables (patients); after normalization with autoscaling and obtaining the generalized covariance matrix. Points marked by the index "1" relate to methods (a), (c) and (d) (Table 1) while those marked by the index "2" relate to methods (b) and (e).

to be normal and predict gallstone formation for the second group. From Table 2 it will be observed that for two patients, (Nos 8 and 10) the calculated IL values are $<1$ or $\geq 1$, according to the analytical methods used for the determination of phospholipids concentration.

It was decided to investigate how a package for pattern recognition analysis ("Parvus") contributed to a decisional procedure of this kind. With this aim, eight of the 10 patients in Table 2 were divided into two classes, sick (IL $\geq 1$) and healthy (IL $<1$). The eight patients were

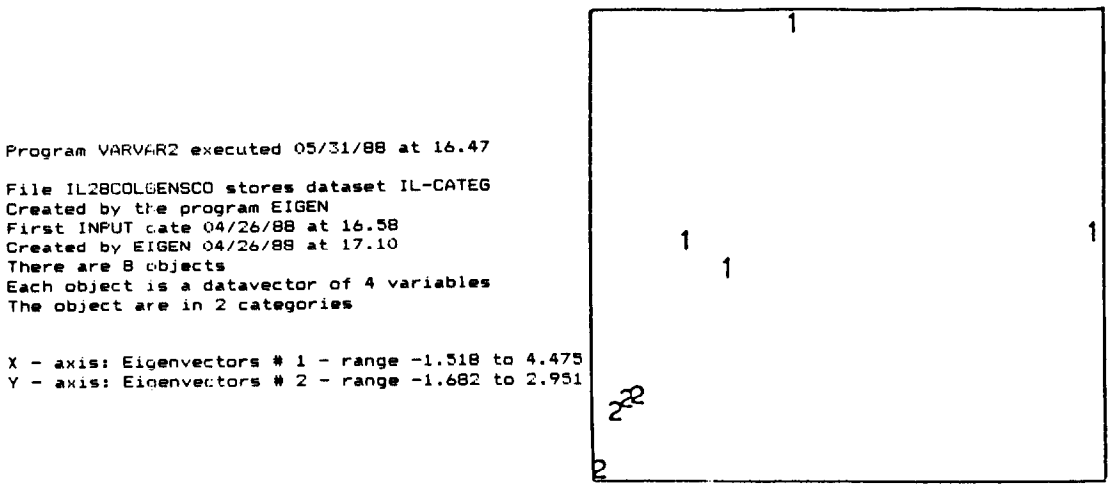


**Figure 3**
Graphical representation of the "scores" of eigenvectors with 99.6%, of the total information. Eight objects (patients) and five variables (methods), after normalization by column centring, categorization (sick and healthy) and obtaining the generalized covariance matrix. Points marked by index "1" relate to sick patients which those marked by the index "2" relate to healthy patients.

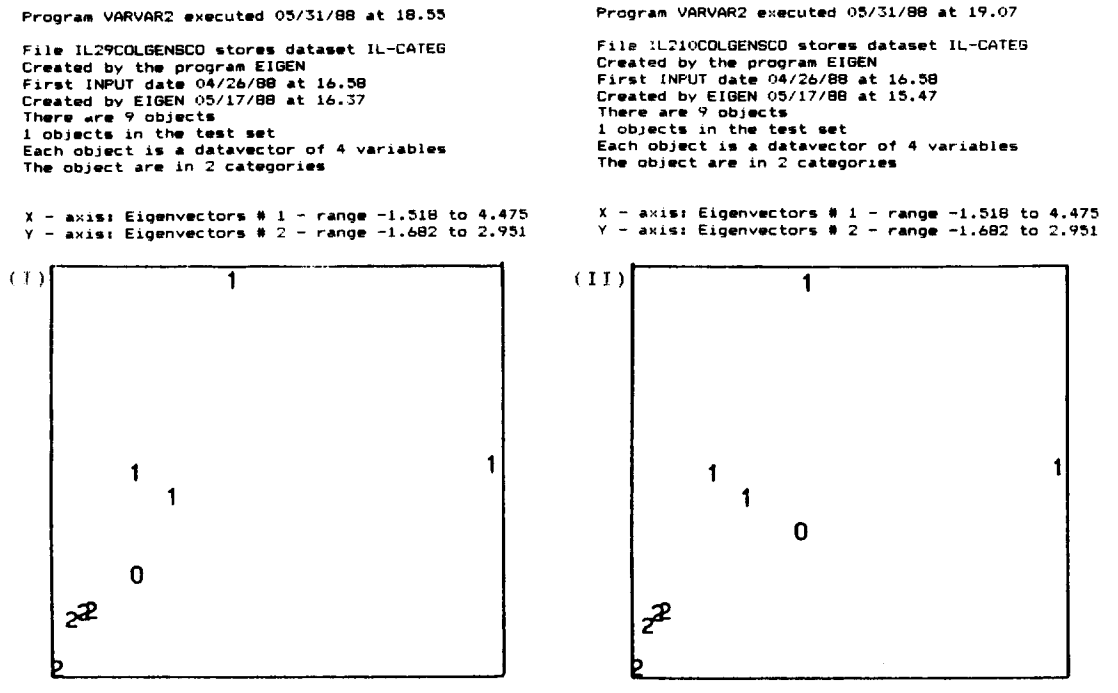


**Figure 4**
(I) Graphical representation of the "scores" of eigenvectors with 99.6%, of the total information. Nine objects (patients), one of whom (No. 8 in Table 1), is placed in the "test set" and five variables (methods). (II) The same representations, with 99.6% of the total information; again nine objects (patients), but in place of patient No. 8 of Table 1, patient No. 10 is considered in the "test set". The calculation procedure and the indices are the same as Fig. 3 while the index zero is assigned to the patient placed in the "test set".

considered as objects, and the IL values, from the five methods of phospholipid determination, as variables. An acceptable "separation" of the scores of the two eigenvectors (to which most of the total information is associated) was obtained by using the programs "Normal", "Matcal" and "Eigen" [12], (Fig. 3). Then the possibility was investigated of assigning the two patients considered above (Nos 8 and 10) to the right class, according to the IL values found. The values for IL calculated for the two patients have been placed separately in a "test set"; it was observed whether the corresponding values of the scores in the eigenvalues projection of the two eigenvectors with greater information fell in the zone of the healthy or that of the sick patients. It is evident from Fig. 4 that patient No. 10 is easily assignable to the sick patient zone, while patient No. 8 cannot be classified, i.e. no accurate diagnosis is possible. For this patient it is clear that the "IL test" is strongly affected by the choice of the experimental method of analysis of phospholipids.

## Conclusions

On the basis of the results obtained by the application of the package "Parvus" for pattern recognition of analytical chemical data, three conclusions can be drawn.

Firstly, not all analytical methods for the determination of phospholipids in bile are equivalent and exchangeable for the determination of IL values.

Secondly, the greatest differences between results are observed when methods (b) or (e) are adopted. An explanation is based on what each method actually measures. Method (e) measures total phosphorus rather than phospholipids concentration [4, 13], whereas method (b) measures choline-containing phospholipids [9]. For practical purposes this is lecithin [8, 20], the major choline-containing phospholipid in bile [2, 8]). The other three methods generally yield the sum of free choline and choline-containing phospholipids and produce similar IL values.

Thirdly, the application of pattern recognition analysis is helpful in the classification of patients as healthy or sick. In the case of patient Nos 8 and 10, for whom the IL index is equivocal, it is probable that without the "Parvus" package, both would be classified as sick since IL values in both cases are generally $\geq 1$ (see rows 8 and 10 of Table 2). The results from using "Parvus" show, however, that this decision is right only in the case of patient No. 10, and wrong for patient No. 8. Perhaps in these cases only a doubtful judgement can be formed from knowledge of the IL value only and further diagnostic criteria are required, for example ultrasound [10].

*Acknowledgements* — This work was supported by the National Research Council (CNR) of Italy. The authors wish to thank Professor M. Forina of Genova University who developed and made available the software package "Parvus".

## References

[1] U. Biader Ceipidor, R. Curini, G. D'Ascenzo, M. Tomassetti and A. Alessandrini, *Giorn. It. Chim. Clin.* **4**, 355–369 (1979).
[2] U. Biader Ceipidor, R. Curini, G. D'Ascenzo, M. Tomassetti, A. Alessandrini and C. Montesani, *Giorn. It. Chim. Clin.* **5**, 127–151 (1983).
[3] L. Campanella, M. Tomassetti, G. D'Ascenzo, G. De Angelis, R. Morabito and L. Sorrentino, *J. Pharm. Biomed. Anal.* **1**, 163–167 (1983).
[4] M. Tomassetti, L. Campanella, A.M. Salvi, G. D'Ascenzo and R. Curini, *J. Pharm. Biomed. Anal.* **2**, 417–424 (1984).
[5] M. Mascini, M. Tomassetti and M. Iannello, *Clin. Chim. Acta* **132**, 7–15 (1983).
[6] L. Campanella, M. Mascini, G. Palleschi and M. Tomassetti, *Clin. Chim. Acta* **151**, 71–83 (1985).
[7] L. Campanella, M. Tomassetti and M. Cordatore, *J. Pharm. Biomed. Anal.* **4**, 155–162 (1986).
[8] U. Biader Ceipidor and M. Tomassetti, *Clin. Lab.* **9**, 1–14 (1985).
[9] L. Campanella, M. Tomassetti and M.P. Sammartino, *Analyst* **113**, 77–80 (1988); and 113, 994 (1989).
[10] J. Tabrisky, R.R. Lindstrom, M.W. Heiman, J. Castagna and D. Sarti, *Gastroenterology* **68**, 1246 (1975).
[11] L. Campanella, G. Visco and M. Tomassetti, *Biochim. Clin.* **8**, 1980–1986 (1984).
[12] "School of elaboration of chemical-analytical data by software package Parvus", S.C.I. Istituto di Analisi e Tecnologie Farmaceutiche ed Alimentari, 3, Genova, 22–26, October (1984).
[13] G.R. Bartlett, *J. Biol. Chem.* **234**, 466–468 (1959).
[14] M. Takayama, S. Ithoh, T. Nagasaki and I. Tanimizu, *Clin. Chim. Acta* **79**, 93–98 (1977).
[15] Lotus Development Corporation, *Lotus* 1-2-3 *User Manual*, Lotus Corporation, Cambridge, MA, USA.
[16] S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York (1956).
[17] D.L. Massart, A. Dijkstra and L. Kaufman, *Evaluation and Optimization of Laboratory Methods and Analytical Procedures*. Elsevier, Oxford (1978).
[18] W.H. Admirand and D.M. Small, *J. Clin. Invest.* **47**, 1043–1052 (1968).
[19] A.L. Metzger, S. Heymsfield and S.M. Grundy, *Gastroenterology* **62**, 499–501 (1972).
[20] C. Carey and D.M. Small, *Clin. Invest.* **61**, 998–1026 (1978).

[*Received for review 15 August 1988; revised manuscript received 17 July 1989; re-typed manuscript received 4 December 1989*]