

CH. BASTIN

G. FLANDRIN

Typologie des lymphocytes et pathologie lymphocytaire (étude préliminaire)

Les cahiers de l'analyse des données, tome 5, n° 3 (1980),
p. 347-359

http://www.numdam.org/item?id=CAD_1980__5_3_347_0

© Les cahiers de l'analyse des données, Dunod, 1980, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

TYPOLOGIE DES LYMPHOCYTES
ET PATHOLOGIE LYMPHOCYTAIRE
(ÉTUDE PRÉLIMINAIRE)
[TYP. LYMPH.]

par Ch. Bastin ⁽¹⁾
et G. Flandrin ⁽²⁾

0 Analyse morphologique et traitement statistique

L'analyse morphologique des éléments figurés du sang sur frottis a été pendant longtemps et demeure pour une grande part, l'un des éléments fondamentaux du diagnostic de nombreuses maladies hématologiques. Malgré l'importante contribution que l'analyse morphologique a apporté dans l'étude des maladies du sang, elle a longtemps souffert de la subjectivité de cette analyse simple au microscope d'une part et des problèmes sémantiques de classification des cellules d'autre part.

Les possibilités actuelles d'analyse automatique d'image, à partir de données morphologiques obtenues par caméra de télévision asservie à un système de microscopie, a ouvert le champ à d'autres possibilités pour l'analyse morphologique, en permettant d'envisager de véritables études morphométriques. Parmi les premières applications de cette morphométrie figurent en hématologie, les divers appareils analyseurs d'images, destinés au décompte "différentiel" des éléments nucléés du sang (formule sanguine) (Utilisation d'un appareil automatique à formule sanguine dans un laboratoire spécialisé d'hématologie. G. Flandrin, T. Bouffette, M.T. Daniel, F. Valensi. *Nouv. Rev. Fr. Hématol.* 1978.19. p.273-288).

La disponibilité de l'un de ces systèmes au laboratoire d'hématologie de l'hôpital Saint-Louis, nous a permis de l'utiliser dans un but de morphométrie des cellules lymphoïdes normales et pathologiques. D'une analyse morphologique plus précise, on peut en effet, espérer un certain nombre de progrès :

- une meilleure reproductibilité de classifications,
- l'espoir de réaliser des diagnostics automatiques,
- et surtout la possibilité éventuelle d'affiner le sous-classement d'hémopathie en groupe de malades non encore établi et présentant un intérêt biologique ou clinique.

Les données numériques obtenues à travers cet analyseur d'images ont été traitées par C. Bastin, J.L. Chasson, J. Juan, M. O. Lebeaux, A. Lebossé, M. Simon, au laboratoire de statistique de l'université P. et M. Curie du professeur J.P. Benzécri.

(1) Assistante au laboratoire de statistique de l'université Pierre et Marie Curie.

(2) Professeur agrégé. Laboratoire central d'hématologie-cytologie. Hôpital Saint-Louis.

(*) Travail réalisé en partie sur contrat libre INSERM n° 80 20 09.

1 Matériel et méthodes

1.1 Matériel : L'étude a porté sur 100 éléments lymphoïdes du sang périphérique de chaque sujet. 50 sujets ont été explorés parmi lesquels 8 sujets normaux (témoins volontaires du laboratoire) 10 malades atteints de leucémie lymphoïde chronique, 32 sujets atteints de leucémie aiguë lymphoblastique. Ces leucémies aiguës lymphoblastiques sont séparées selon la classification F.A.B. en trois groupes distincts : (*Proposal for the classification of acute leukemia*. Bennett J.M., Catovsky D., Daniel M.T., Flandrin G., Galton D., Gralnick H.R., Sultan C.. Brit. J. Hemat. 1976. 33. p. 451-458).

- L₁ (petites cellules),

- L₂ (grandes cellules),

- L₃ (ou leucémie à cellules de Tumeur de Burkitt) (*Acute leukemia with Burkitt's tumor cells Study of six cases with special reference to lymphocytes surface marker*. Flandrin G., Brouet J.C., Daniel M.T., Preud'homme J.L. *Blood* 1975 45 p. 183-188.

1.2 Méthode

1.2.1 : L'analyseur d'image (LARC Corning Medical) utilise une chaîne de travail uniformisant les conditions d'étalement, de coloration et de lecture des images. Les "étalements" sont préparés par une centrifugeuse permettant d'obtenir un film monocellulaire uniforme, ce qui contraste avec les techniques traditionnelles d'étalement manuelles : la répartition des cellules est uniforme, toute distorsion et rupture cellulaire est éliminée. Enfin presque toute la surface de la lame est propre à l'examen.

La coloration est effectuée avec un colorant type Giemsa, par un appareil automatique contrôlant les conditions de température et de pH des réactions de coloration et permettant donc de contrôler sa stabilité. Les images TV sont obtenues à travers deux filtres colorés (filtre bleu et filtre jaune), dans le but d'appréhender les notions de couleur.

Précédant la saisie des données, un ajustement des résultats de la coloration par rapport à des données témoins, permet de corriger les petites variations des réactions colorées inhérentes à tout système de coloration panoptique de ce type.

Le programme d'analyse d'image fourni par le système permet notamment d'obtenir 8 données qui seront utilisées pour cette étude :

- n° 1 surface nucléaire
- n° 2 indice de forme [(périmètre nucléaire)²/surface nucléaire].
- n° 3 indice de "proximité" (rapport du nombre de points du cytoplasme dans le premier et le second anneau autour du noyau, sur le nombre de points du périmètre nucléaire).
- n° 4, n° 5 absorption du cytoplasme à travers le filtre bleu et le filtre jaune.
- n° 6 surface cytoplasmique.
- n° 7, n° 8 absorption du noyau à travers le filtre bleu et le filtre jaune.

De parti pris, dans cette étude, seules les cellules lymphoïdes normales ou pathologiques sont mesurées ; celles-ci sont sélectionnées au microscope par l'observateur. La saisie des données pour 100 cellules prend approximativement 20 minutes.

1.2.2 Les données recueillies ont donc la structure suivante : à chaque sujet i (sain ou malade ; ou plus exactement état de sujet, car on aura éventuellement à considérer plusieurs prises de sang faites sur un même sujet), correspond un ensemble Ci de 100 cellules c ; et

pour chaque cellule c on a un ensemble de 8 mesures. La présente étude préliminaire porte sur un ensemble I de 50 sujets ; soit au total, un ensemble C de 4944 cellules lymphoïdes (C. Ly.) (4944 et non 5000 parce que le nombre de cellules par sujet n'est pas toujours exactement 100). Dans les analyses dont nous rendons compte ici, les mesures effectuées sur ces leucocytes ne sont pas traitées directement, mais codées sous forme logique, ainsi qu'on le fait ordinairement pour les données numériques hétérogènes. D'après l'histogramme des valeurs prises, l'intervalle de variation de chaque mesure, e.g. l'aire du cytoplasme est divisé en 5 classes (disons de très faible à très fort) et chaque cellule est décrite non par une ligne de 8 mesures, mais par une ligne, suivant un ensemble J de 40 modalités (8×5), avec par exemple, dans la case "aire du cytoplasme très faible" un 1 si la cellule rentre dans cette modalité et un 0 si non.

2 La chaîne des analyses effectuées

Avant d'exposer en détail les résultats de quelques unes des premières analyses décrites ci-dessous, nous décrirons l'enchaînement de ces analyses en anticipant sur leur exploitation future pour le diagnostic.

Une première analyse a porté sur un tableau $C \times J$, 4944×40 .

- C = ensemble des C. Ly.
- J = ensemble des 40 modalités descriptives des 8 variables.
- $k(c, j) = 1$ si la C. Ly. c entre dans la modalité j .
- $k(c, j) = 0$ sinon.

Comme la projection de ces quelques 5000 points est de peu d'intérêt, car peu lisible, on a adjoint à ce tableau, un tableau supplémentaire $I \times J$, 50×40 (dit des C. Ly. moyennes).

- I = ensemble des sujets.
- $k(i, j) = 1$ si la C. Ly. moyenne du sujet i entre dans la modalité j ; 0 sinon.

Ceci nous permet de représenter, par exemple dans le plan 1×2 les différents sujets. On constate une disposition satisfaisante de ces individus en particulier les points représentatifs des quelques sujets sains sont groupés et les cas de leucémie de Burkitt occupent une extrémité du nuage des individus. Afin d'examiner la diversité des formes des C. Ly. de chaque sujet, on a réalisé, pour chaque individu le graphique de la projection de ses 100 cellules sur les plans 1×2 et 1×3 .

Il apparaît alors des similitudes frappantes entre les sous-nuages des sujets relevant d'un même diagnostic : sujets sains, leucémies lymphoblastiques de type L_3 , de type L_2 . Donc la proximité observée sur les points moyens résulte d'une similitude plus profonde que nous appellerons similitude de *forme lymphocytaire*. En bref, à un type donné de pathologie correspond la présence ou l'absence des mêmes formes de C. Ly.

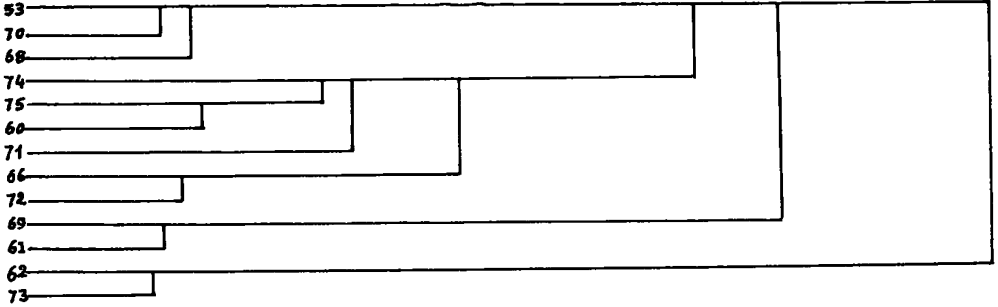
Après cette première analyse, on a entrepris de préciser statistiquement ce que l'on doit entendre par *forme lymphocytaire*. Pour cela on a effectué une *classification ascendante hiérarchique* (C.A.H.) sur l'ensemble des 4944 lymphocytes, décrits par les 8 premiers facteurs de l'analyse factorielle. Cette opération, qui est le maillon le plus coûteux de la chaîne de traitement, a permis de proposer, en coupant l'arbre de C.A.H. à différents niveaux (choisis d'après l'histogramme des niveaux des noeuds), trois partitions plus ou moins fines

de C (voir figure).

Q_1 = partition en 7 classes ;

Q_2 = partition en 13 classes ;

Q_3 = partition en 44 classes.



Sous-arbre de C.A.H. en 13 classes

Désormais l'ensemble des 50 observations individuelles, mis à la base de notre étude est décrit par un tableau $I \times Q$ (où Q peut être, au choix, Q_1 , Q_2 ou Q_3) ; avec : $k(i, q)$ = nombre de cellules

lymphoïdes de l'individu i (de I) rentrant dans la classe q (de Q). La ligne de ce tableau afférente au sujet i peut être appelée sa *formule lymphocytaire*. L'examen direct du tableau $I \times Q$ (au moins dans le cas de Q_1 et Q_2 ; c'est-à-dire avec 7 ou 13 classes), montre que certains diagnostics se caractérisent par l'absence ou la présence de certains types de cellules. Cela n'a rien d'étonnant dans la mesure où en effectuant la C.A.H., nous n'avons fait que réaliser rigoureusement, en exploitant au mieux toutes les données, ce que nous avions fait d'abord visuellement sur le plan 1×2 issu de l'analyse des cellules, c'est-à-dire que nous avons divisé le champ de variabilité possible des C. Ly. .

Mieux que par l'examen direct, on connaîtra la diversité des formules lymphocytaires en analysant le tableau de correspondance $I \times Q$. On obtient ainsi une représentation du nuage I , aisément interprétable en terme de formule lymphocytaire grâce à la représentation simultanée des ensembles I et Q (individus et classes de cellules). En outre, lorsque l'ensemble aura été élargi pour couvrir l'ensemble de la pathologie lymphocytaire (ce qui n'est pas fait avec 50 cas), on pourra effectuer une C.A.H. sur I , d'après le tableau $I \times Q$.

Reste l'exploitation des traitements ci-dessous décrits pour le diagnostic. Soumettre à la C.A.H. l'ensemble C de toutes les cellules est une opération coûteuse qu'on ne peut reprendre chaque fois qu'un cas nouveau se présente : il faut donc insérer dans chaque partition Q choisie une fois pour toutes (d'après un ensemble I de cas représentatif de l'ensemble de la pathologie lymphocytaire) les C. Ly. de tout sujet nouveau. Telle quelle la C.A.H., ne permet pas exactement cette insertion ; mais il existe une solution approchée dont nous avons vérifié qu'elle est pratiquement satisfaisante : rattacher chaque cellule c à la classe q du centre de laquelle elle est le plus proche. Ainsi à partir des résultats de mesure de 100 lymphocytes d'un sujet nouveau is , on pourra déterminer une formule lymphocytaire (c'est-à-dire le nombre $k(is, q)$ des C. Ly. rentrant dans chaque classe q), puis placer l'individu is sur les diagrammes d'analyse factorielle issus de l'analyse du tableau $I \times Q$. Donc effectuer un diagnostic et plus précisément suivre l'évolution des points figuratifs au cours du temps etc. . Tous ces calculs étant accessibles à un petit ordinateur.

3 Résultats

3.1 Analyse des cellules

3.1.1. Analyse du tableau $C \times J$, le tableau $I \times J$ étant adjoint en sup-

plémentaire : En regardant sur le plan 1×2 , la projection des différentes modalités des variables et leurs contributions relatives, on constate que l'axe 1 est un axe de taille des cellules ; les grandes cellules (c'est-à-dire ayant de grands noyaux, grands cytoplasmes et de surcroît une grande irrégularité de forme) se projettent avec des coordonnées négatives. L'axe 2 précise la morphologie des petites cellules. Les coordonnées négatives sont associées aux cellules à petit noyau (et donc ayant de forts coefficients d'absorption du noyau pour les filtres bleus et jaunes), les coordonnées positives aux cellules à petit cytoplasme (et donc faible paramètre 3) (figure 1 représentative des paramètres). Les grandes cellules ne sont pas différenciées par l'axe 2.

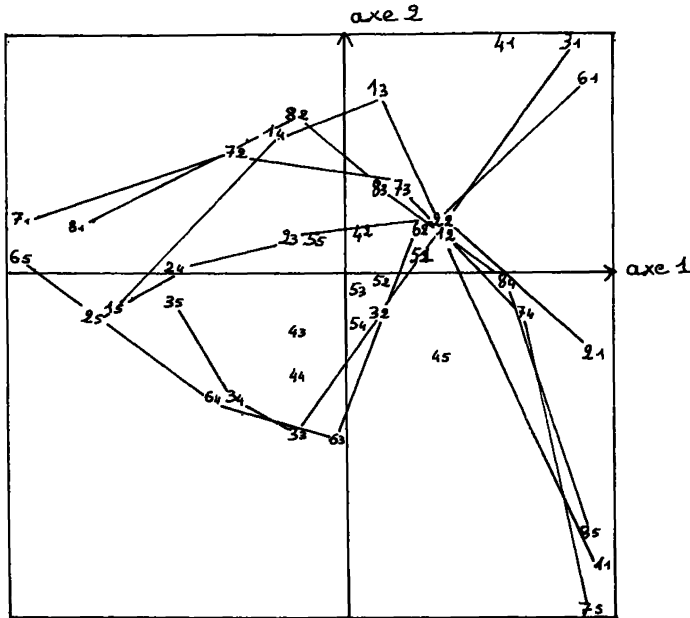
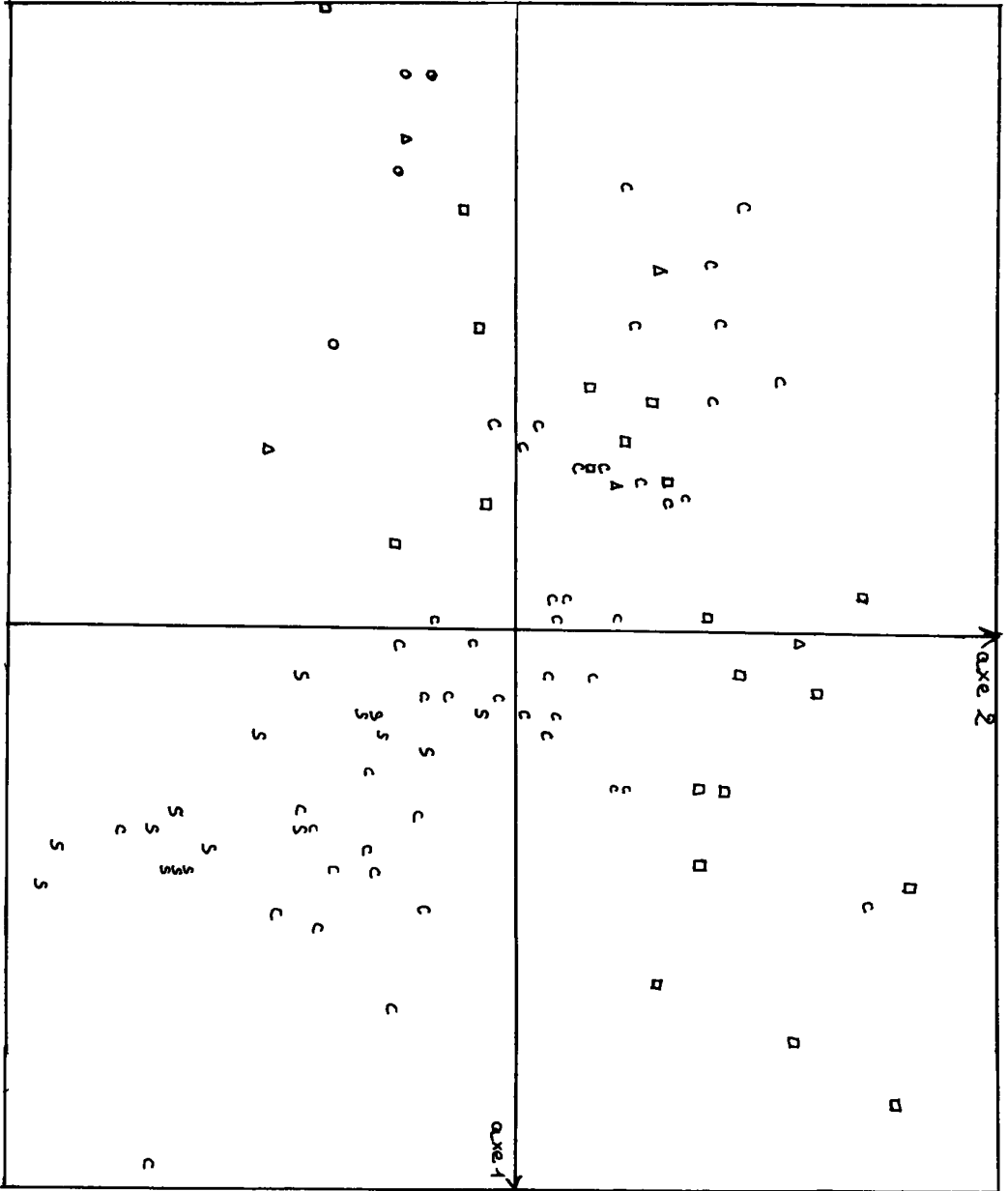


Figure 1 représentant les projections sur le plan 1×2 des 5 modalités des 8 variables.

Ces remarques permettent de préciser les regroupements d'individus observés, (voir figure 2). Les C. Ly. moyennes de sujets atteints de leucémie L_3 , se trouvent groupées tout à fait à gauche du graphique: ce sont de très grandes C. Ly.. En haut à droite, on trouve presque exclusivement des C. Ly. moyennes de sujets atteints de leucémie aiguë lymphoblastique de type L_1 : cellules assez petites, à petit cytoplasme et petit paramètre 3.



- C. Ty. moyenne d'un sujet
- S normal
- C atteint de LIC
- atteint de LAL de type L1
- Δ atteint de LAL de type L2
- o atteint de LAL de type L3

Figure 2 : Projection sur le plan 1 x 2 des C. Ty. moyennes de 98 sujets

En bas à droite, les C. Ly. moyennes des sujets sains sont très groupées et voisines de C. Ly. moyennes de sujets atteints de LLC ; de sont des C. Ly. petites, à petit noyau et forte absorption du noyau dans les filtres bleus et jaunes. En projection sur l'axe 1, les C. Ly. moyennes de sujets atteints de LAL de type L_2 sont assez peu dispersées ; ce sont des cellules de taille moyenne. Les C. Ly. moyennes de LLC, étant réparties dans tout le plan, on peut raisonnablement penser que cette maladie revêt des formes diverses. Une C.A.H. sur les C. Ly. moyennes nous a permis de retrouver ces résultats et de les préciser. Une C.A.H. sur les C. Ly. moyennes de sujets atteints de LLC permet de distinguer deux groupes de LLC : les malades ayant des C. Ly. à petit noyau et les autres.

3.1.2 *Représentation des sous-nuages Ci afférents à chaque sujet* : Sur le plan 1×2 de l'analyse de 4944 cellules on projette pour chaque individu ses 100 cellules lymphoïdes. Les C. Ly. moyennes des sujets sains sont très groupées dans le plan 1×2 (et d'ailleurs aussi dans les autres plans). Il est frappant de constater que les sous-nuages relatifs aux sujets sains ont tous à peu près la même forme, voir figure 3.

De même que les C. Ly. moyennes de LLC sont assez dispersées, les sous-nuages correspondants sont beaucoup plus étendus (voir fig. 3).

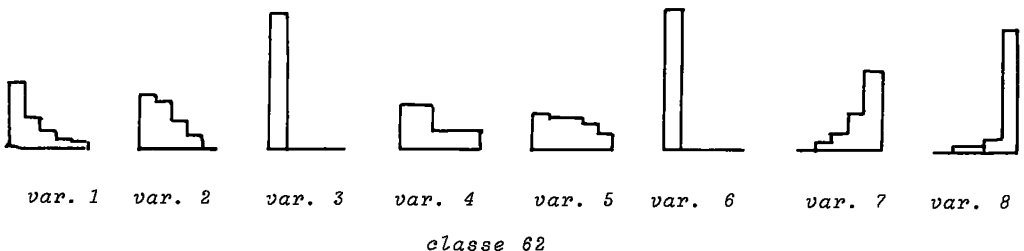
Pour ce qui est des LAL, on retrouve plusieurs types de nuages : les nuages relatifs aux leucémies L_3 , tous semblables, ceux relatifs aux leucémies L_2 ou aux leucémies L_1 (les plus variés) ; voir figure 3.

3.2 *La classification des 4944 lymphocytes* : L'histogramme des indices de niveaux de la hiérarchie montre 3 ruptures : partitions en 7, 13 et 14 classes.

Nous avons étudié ces trois partitions Q_1 , Q_2 , Q_3 .

Pour chaque partition, nous avons construit le tableau $I \times Q$ des formules lymphocytaires. Et pour aider à l'interprétation des classes, nous avons compté pour chaque classe, le nombre de fois où apparaissent les différentes modalités des 8 variables. Du moins pour les partitions en 7 et 13 classes, ceci permet de décrire approximativement les cellules de chaque classe.

Voici les classes les plus particulières : les classes 62 et 73 sont des cellules à très petit cytoplasme et aussi très petit paramètre 3. Elles se retrouvent essentiellement chez des sujets atteints de LAL et n'apparaissent pour ainsi dire jamais chez des sujets sains.



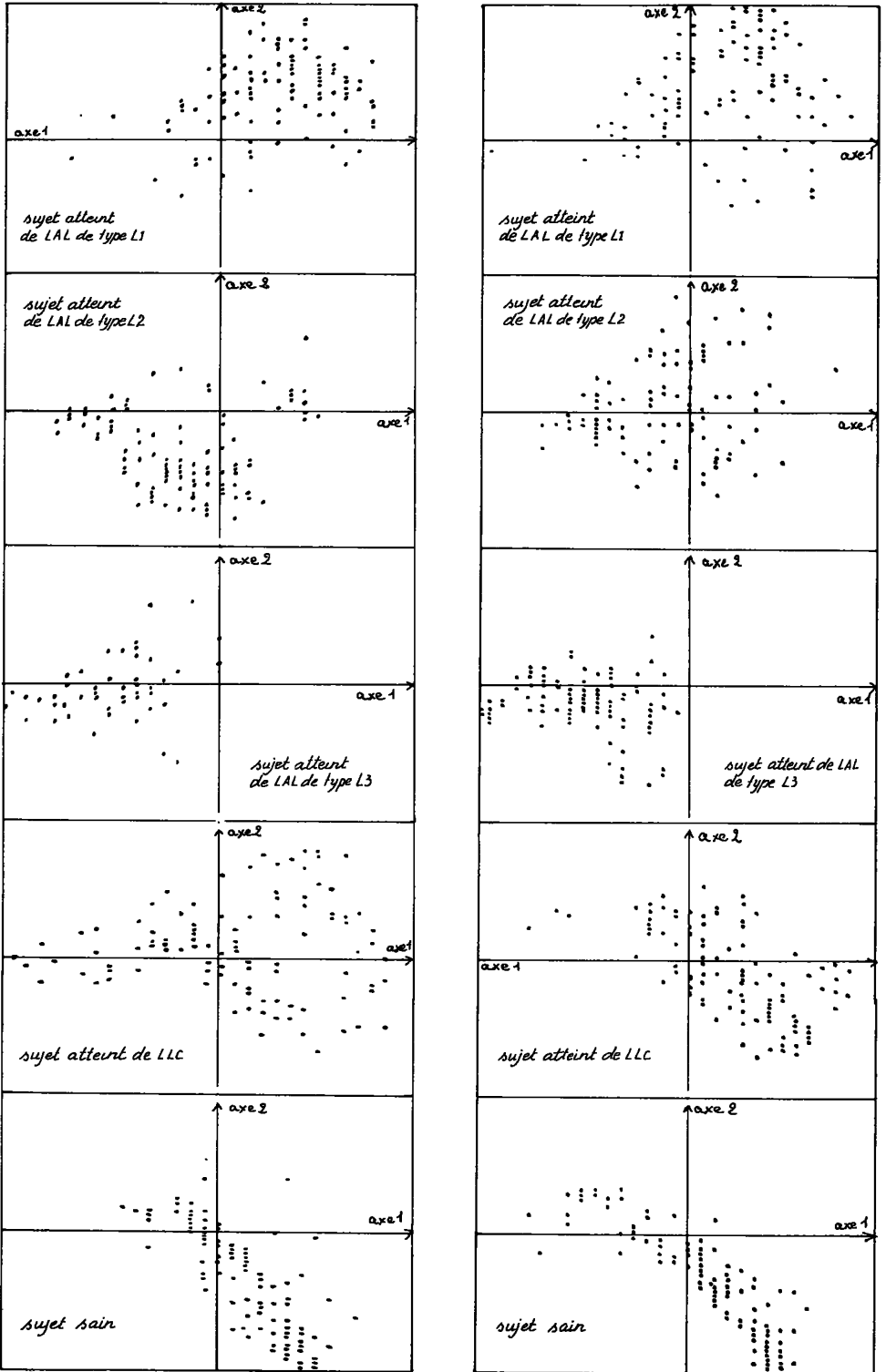
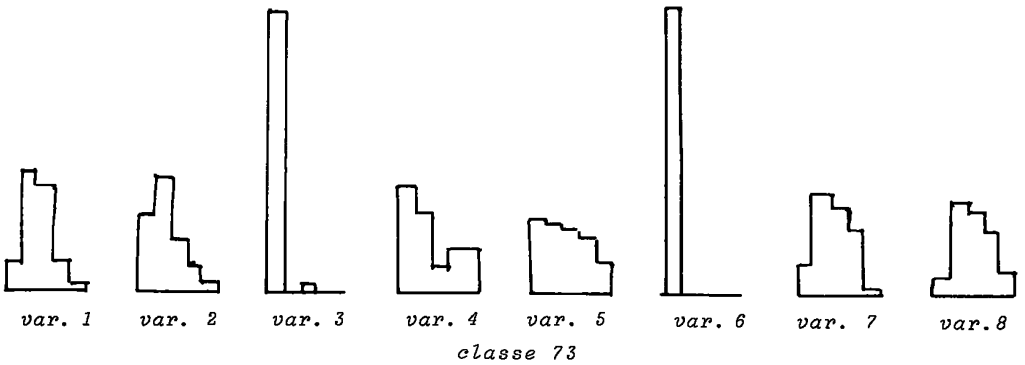
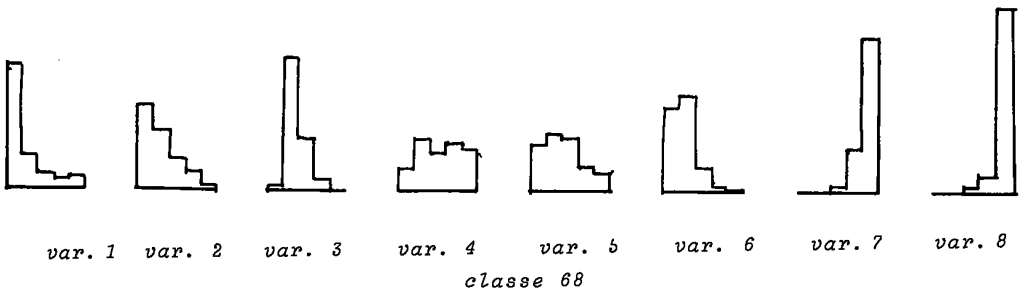
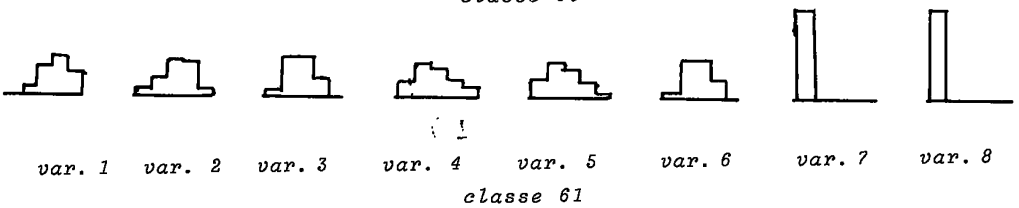
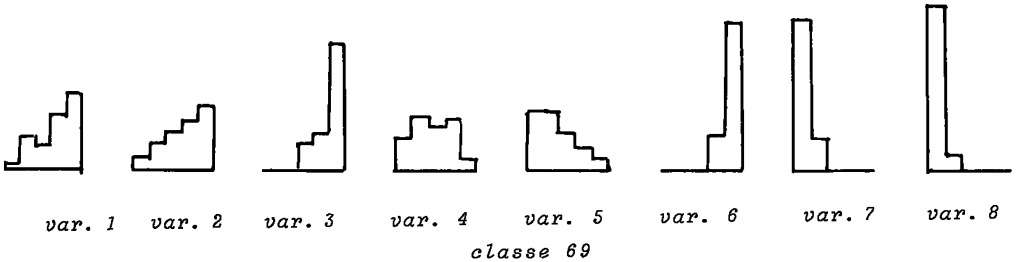


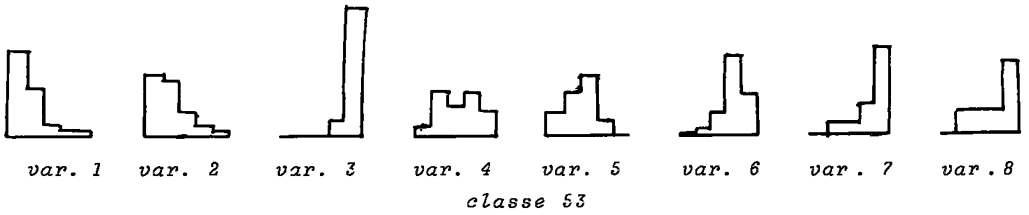
figure 3: projection sur le plan 1+2 des 100 C. Ly d'un même sujet



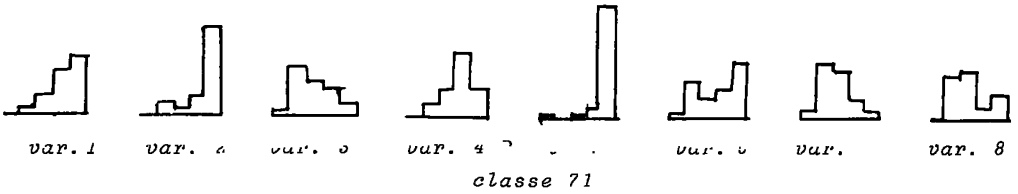
Les classes 68 et 53 sont des cellules dont le noyau est petit et fortement absorbé dans les filtres jaunes et bleus. Les cellules de la classe 53 sont des C. Ly. de sujets sains ; celles de la classe 68 sont des C. Ly. de sujets atteints de LLC et à un moindre degré de sujets sains.

Les classes 69 et 61 sont des cellules dont le noyau absorbe très peu à travers les filtres jaunes et bleus (faibles paramètres 7 et 8). Les cellules de la classe 69 ont un grand cytoplasme et fort paramètre 3. Ces cellules se retrouvent essentiellement chez des sujets atteints de LAL (presque exclusivement pour la classe 61)

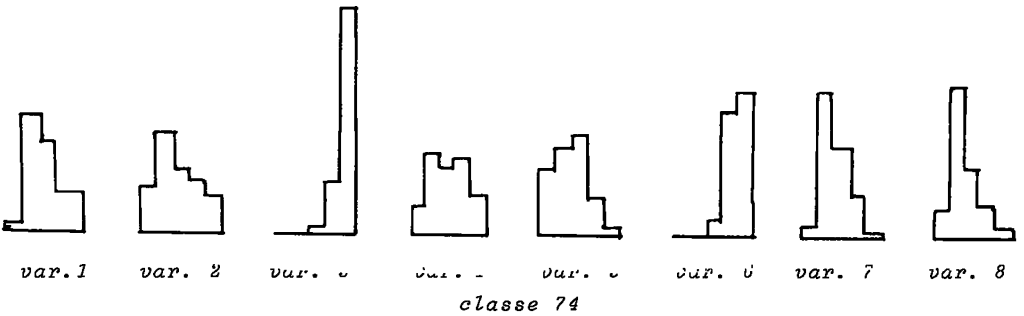




La classe 71 regroupe des C. Ly. à grand noyau, grand paramètre 2 et forte absorption du cytoplasme dans les filtres jaunes et bleus. Ces cellules se retrouvent presque uniquement chez des sujets atteints de LAL. et parmi eux surtout, chez des sujets atteints de leu-



La classe 74 regroupe des cellules à grand cytoplasme et se trouve peu chez les sujets atteints de LAL.



3.3 Formules lymphocytaires

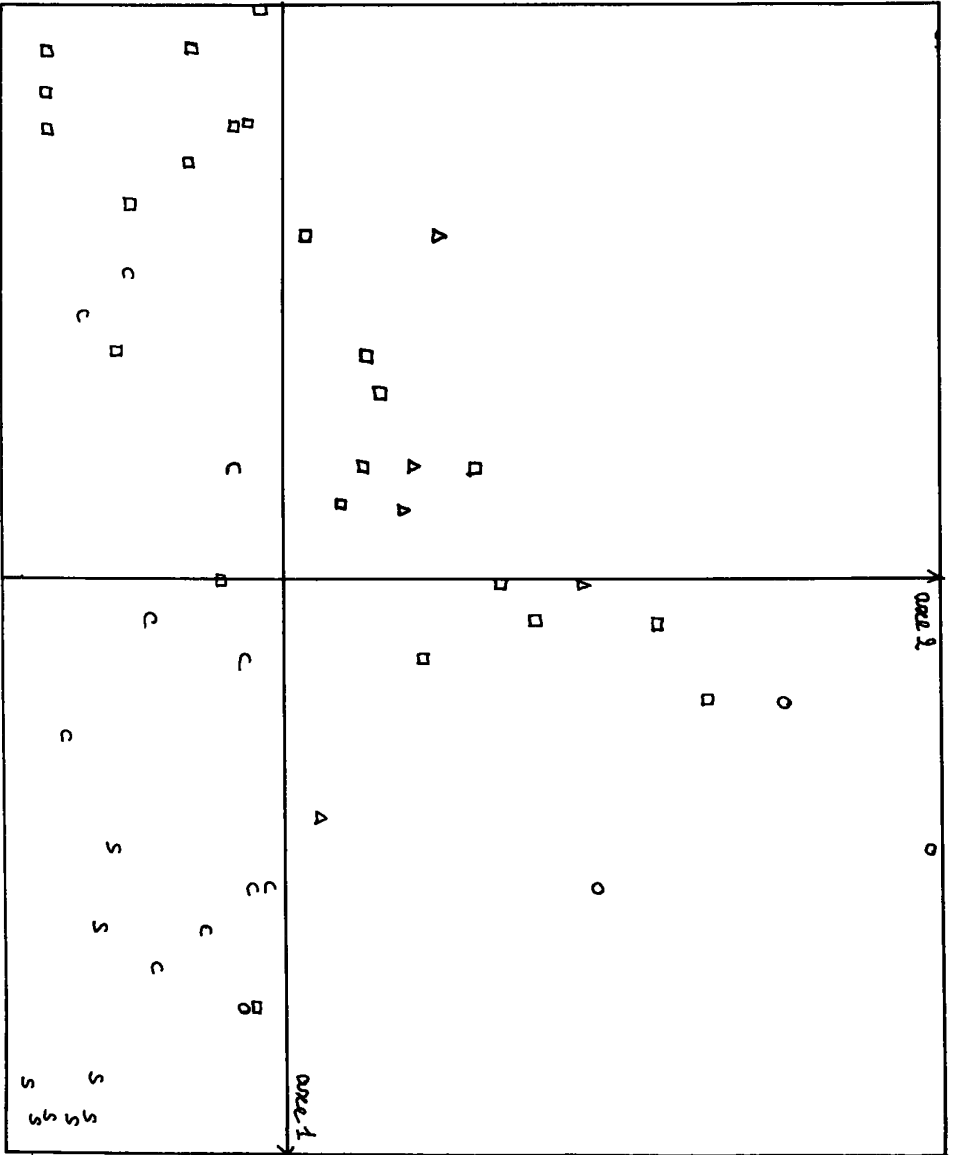
3.3.1 Analyse du tableau des formules : Nous avons fait l'analyse des correspondances des tableaux $I \times Q$ donnant pour chaque individu sa formule lymphocytaire.

L'analyse du tableau $I \times Q_2$ (partition en 13 classes), donne des résultats plus fins que celles du tableau $I \times Q_1$ (partition en 7 classes); par contre, le choix de la partition Q_3 en 44 classes, sur cet échantillon assez faible d'individus, ne donne pas de résultats sensiblement différents.

Aussi examinerons-nous en détail les résultats de l'analyse du tableau $I \times Q_2$ associé à la partition en 13 classes (voir figure 4).

Dans tous les plans, les projections des sujets sains sont très groupées et associées surtout à la classe 53.

Dans le plan 1×2 , les LLC sont en bas du graphique, assez nettement séparées des sujets atteints de leucémie lymphoblastique. La projection dans ce même plan des sujets atteints de leucémie aiguë lymphoblastique permet de distinguer nettement :



- S sujets sains
- C sujets atteints de LLC
- sujets atteints de LAL de type L1
- △ sujets atteints de LAL de type L2
- sujets atteints de LAL de type L3

Figure 4 : Analyse du tableau des formules lymphocytaires. Projection sur le plan 1 x 2 des 50 sujets.

- les sujets atteints de leucémie aiguë du type L_3 (en haut à droite) fortement associés à la classe 71.

- les sujets qui se projettent en bas à gauche du graphique sont atteints de leucémie aiguë du type L_1 exclusivement. Ils sont associés aux classes 62 et 73.

- tous les sujets atteints de leucémie aiguë de type L_2 sont entre ces deux groupes, assez mêlés avec des sujets de type L_1 .

3.3.2 *Réaffectation des cellules à une classe* : A partir des 13 classes obtenues par la classification automatique, nous avons pris leur centre de gravité et réaffecté toutes les cellules à la classe dont le centre est le plus proche. Afin de comparer cette partition à la précédente, nous avons fait l'analyse du tableau $I \times Q'$ des formules lymphocytaires obtenues après réaffectation. La carte obtenue est presque identique. Ce qui montre la validité de ce mode d'affectation d'une cellule à une classe et permet ainsi de classer de nouvelles C. Ly. par un procédé opérationnel sur des petits ordinateurs.

4 *Discussion*

Au terme de cette étude préliminaire, on peut déjà tirer un certain nombre de conclusions ayant un intérêt hématologique.

4.1 *Certaines données déjà reconnues* par la simple analyse au microscope se trouvent reproduites. Parmi elles, la claire distinction entre les populations des leucémies lymphocytaires chroniques et celle des leucémies aiguës. Parmi les leucémies aiguës lymphoblastiques, la subdivision cytologique en classe (L_1, L_2, L_3) trouve une certaine confirmation. Les patients L_3 ont une place très particulière avec notamment une forte association à la classe 71. Il n'est pas inintéressant de constater que l'individualisation de cette classe L_3 est récente (1976), et l'affirmation de son identité, d'abord fort contestée, n'a pu s'affirmer qu'à travers des preuves non morphologiques, comme l'existence en particulier d'une même anomalie cytogénétique (Berger R., Bernheim A., Flandrin G., Daniel M.T., Schaison G., Brouet J.C., Bernard J. Translocation (t(8 ; 14) dans la leucémie lymphoblastique de type Burkitt. Nouv. Pres. Med. 1979, 8, 181-183).

La distinction des groupes de leucémie aiguë L_1 et L_2 est moins claire, mais ce fait ne traduit peut être que l'incertitude du classement subjectif de ces deux catégories (Bennett J.M., Catovsky D., Daniel M.T., Flandrin G., Galton D.A.G., Gralnick H.R., Sultan C., *The morphological classification of acute lymphoblastic leukaemia - concordance among observers and clinical correlations*. Brit. J. Haemat. 1980)

A travers ce type d'étude morphologique, apparaît sans doute à ce niveau, la mise en cause de certaines classifications subjectives et la possibilité d'introduction de critères de sous-classifications plus reproductibles.

4.2 *Certaines constatations* sont plus originales et éventuellement non encore prévues par la simple analyse subjective. Le regroupement très étroit des profils lymphocytaires de sujets anormaux est certainement la constatation la plus frappante. Cette notion n'est, en

effet, pas évidente dans les traités classiques d'Hématologie et échappe certainement à l'analyse subjective. Ces limites étroites du profil lymphocytaire des sujets normaux, laissent espérer un abord morphométrique intéressant de discrètes modifications de ce profil, dans des états réactionnels ou inflammatoires, dont les subtiles changements échappent peut-être à l'analyse subjective.

4.3 L'autre constatation assez originale est l'affirmation de l'extrême polymorphisme des profils lymphocytaires individuels des différents malades atteints de leucémie lymphocytaire chronique. Cette question commence seulement, à être travaillée (Binet J.L., Vaugier G., Dighiero G., d'Athis P., Charron D. *Investigation of a new parameter in chronic lymphocytic leukemia : the percentage of large peripheral lymphocytes determined by the Hemalog D.* Amer. J. of Med. 1977, 63, 683-688). (Patterson K.G., Goldstone A.H., Richards J.D.M., Cawley J.C. *Studies of cell size in lymphoproliferative disorders.* Brit. J. Haemat. 1980, 46, 337-338) et devrait trouver son intérêt pratique dans la recherche des corrélations cliniques et pronostiques qui ne sont pas encore établies.

4.4 Cette étude préliminaire donne des résultats tout à fait encourageants. Et du point de vue statistique il serait très souhaitable de pouvoir disposer d'un échantillon plus important d'individus d'une part afin de confirmer et de préciser les constatations ci-dessus, d'autre part afin d'étudier dans ce cadre l'évolution de la pathologie des différents sujets notamment pour les leucémies lymphoïdes chroniques.

Peut-être aussi une étude plus vaste et plus approfondie permettra-t-elle de proposer de nouveaux classements à l'intérieur des LAL et des LLC.