

G. D. MAÏTI

G. CARGILL

Variation de l'acidité œsophagienne et des reflux gastro-œsophagiens en fonction de l'âge et des circonstances diagnostiques chez 2739 patients

Les cahiers de l'analyse des données, tome 22, n° 1 (1997), p. 83-94

http://www.numdam.org/item?id=CAD_1997__22_1_83_0

© Les cahiers de l'analyse des données, Dunod, 1997, 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/>

VARIATION DE L'ACIDITÉ ŒSOPHAGIENNE ET DES REFLUX GASTRO-ŒSOPHAGIENS EN FONCTION DE L'ÂGE ET DES CIRCONSTANCES DIAGNOSTIQUES CHEZ 2739 PATIENTS

[ACIDITÉ ŒSOPHAGIENNE]

G. D. MAÏTI*

G. CARGILL**

1 Origine et structure des données

Les médecins généralistes et spécialistes, particulièrement les gastro-entérologues, recourent aujourd'hui communément à des centres spécialisés en explorations fonctionnelles digestives pour un examen minutieux de l'acidité œsophagienne résultant de l'activité gastro-œsophagienne chez leurs patients. Dans le présent article, on considère les résultats d'examens diagnostiques, effectués dans un tel centre, tous suivant le même protocole. Les cas n'ont pas été choisis suivant quelque critère que ce soit: il s'agit de 2739 examens consécutifs, réalisés sur une période de plusieurs mois. Assurément, la répartition des âges et des pathologies n'est pas conforme à l'état sanitaire de la France; ni même, plus précisément, de la Région Parisienne, où opère le Centre. De plus, le recours du patient à un médecin, puis de celui-ci au Centre, résulte autant de dispositions personnelles que de la pathologie et de la physiologie objectives elles-mêmes. Cependant, la netteté des résultats que révèle notre analyse suffit, selon nous, à suggérer des conclusions positives et à écarter le projet d'un diagnostic purement physiologique.

1.1 Les variables et leur codage par modalités

De façon précise, le fichier traité donne, pour chaque examen, outre un numéro de dossier, 10 variables. Les deux premières, âge et diagnostic (ou plutôt: classe pathologique) sont à part. Les huit autres offrent les résultats d'une pH-métrie œsophagienne ayant duré trois heures.

Considérons, afin d'illustrer notre introduction, les cinq premières lignes du tableau de base, soumis au traitement statistique multidimensionnel. Selon le format usuel du logiciel MacSAIF, vient d'abord une ligne de titre. Puis un nombre, dont les trois derniers chiffres indiquent le nombre des variables; tandis que le nombre des milliers a été introduit ici pour spécifier au

(*) Docteur ès Science, Expert clinicien; société STATMATIC; Épinay sur Seine.

(**) Médecin, gastro entérologue; Paris 13.

RGO données d'activité gastro- sopathagienne (SyQ:DphM:Ph)
2739010

	path	âge	nrf	kay	nrg	dur	rg>5	rg>>	%ph4	%ph5
1	5	108	2	28	4	11	0	74	271	380
2	5	780	16	12	0	9	0	20	106	133
3	5	730	7	2	0	6	0	16	21	39
4	5	804	41	79	3	185	3	685	241	2424
5	5	359	6	493	2	1434	2	6540	7028	7974

programme de lecture le nombre des individus: 2739; lequel dépasse la borne commune de 1000. Viennent ensuite les sigles des variables, puis, ligne après ligne, les cas individuels. On rappelle que cette disposition par ligne, possible ici du fait du petit nombre des variables, n'est pas indispensable pour le traitement informatique, même si elle est commode pour la simple lecture des données. Ici, en tête de ligne, le numéro du dossier tient lieu de sigle pour chaque cas individuel; les dix nombres qui suivent étant les valeurs des variables.

La variable path, pathologie déclarée, admet huit modalités, chacune désignée par un chiffre de 1 à 8. Soit:

- 1: π ORL : ORL pure (212 cas) ;
- 2: π PUL : PULmonaire pure (360 cas) ;
- 3: π O+P : ORL + PULM (155 cas) ;
- 4: π OPG : ORL + PULM + GASTROentérologie (517 cas) ;
- 5: π Eso : ŒSOPHAGIENNE pure (1145 cas) ;
- 6: π Mal : Malaises (186 cas) ;
- 7: π r+M : Reflux + Malaises (78 cas) ;
- 8: π var : Pathologies Diverses (86 cas) ;

on notera qu'après chaque numéro, est placé un sigle formé de quatre lettres; ces sigles nous ont servi dans l'examen du fichier; mais il faut dire d'emblée que nos données montrent un lien si fort entre l'état œsophagien et l'âge, que la pathologie, considérée isolément n'a pas de caractère physiologique clair; ce qui nous a conduit à introduire des couples pathologie-âge (cf. *infra*, §2.2).

La variable âge est donnée en mois. Ainsi, pour le 1-er cas, l'âge est de 108 mois, soit: 9 ans; pour le 2-nd cas, 780 mois, soit: 65 ans; etc. Sur notre échantillon, la distribution de l'âge est bimodale: sur l'histogramme, (cf. §1.2) on voit une densité en forte décroissance, quasi exponentielle, de 0 mois (i.e. moins de 15 jours) jusqu'à \approx 10 ans (120 mois); puis s'étend une large distribution en cloche dont le maximum est vers 50 ans (600 mois). On a donc pris 5 modalités d'âge: {â β , â ∂ , âe, âM, âV} délimitées par les bornes {5, 30, 120, 600}; avec des sigles qui évoquent respectivement {Bébé, deux ans, enfant, Mûr, Vieux}; même si, e.g., âM s'étend de 10 à 50 ans. L'important est pour nous de suivre, à la fois, la forme de l'histogramme et ce que des analyses multipliées nous ont appris du lien entre âge et pathologie: e.g. π Mal offre 149 cas dans â β , 29 dans â ∂ ; et seulement 8 cas au-delà de 30 mois.

```

ph découpage des variables (SyQK:DphM:PhDcodx)
données d'observation
sigles et pivots; le nombre des var est 10
path a 8 µod
πORL πPul πO+P πOPG π so πMal πr+M πvar 1 2 3 4 5 6 7 8
âge a 5 µod
âβ âð âe âM âV 5 30 120 600 2000
nrf a 6 µod
nrf0 nrf1 nrf2 nrf3 nrf4 nrf5 4 8 16 32 64 999
kay a 5 µod
k0 k1 k2 k3 k4 5 25 75 200 10000
nrg a 4 µod
nrG0 nrG1 nrG2 nrG3 0 1 5 15
dur a 4 µod
Δ0 Δ1 Δ2 Δ3 0 15 30 60
rg>5 a 4 µod
rgL0 rgL1 rgL2 rgL3 0 1 2 5
rg>> a 6 µod
rg+0 rg+1 rg+2 rg+3 rg+4 rg+5 0 30 100 300 1000 100000
%pH4 a 5 µod
4pH0 4pH1 4pH2 4pH3 4pH4 25 150 500 1500 10000
%pH5 a 5 µod
5pH0 5pH1 5pH2 5pH3 5pH4 100 300 1000 3000 10000

```

La variable nrf, nombre de reflux, est de l'ordre de quelques unités ou de quelques dizaines. On y a distingué 6 modalités; avec des bornes en progression géométrique, selon ce que suggère l'histogramme.

L'indice de KAYE, note d'appréciation globale, est partagé en 6 modalités. L'histogramme montre, ici encore, une forte décroissance; les bornes sont telles que les modalités soient de poids équivalent (les effectifs extrêmes étant dans le rapport de 3 à 2).

Sous le sigle nrg sont dénombrés les reflux gastro-œsophagiens qui durent plus de 30 secondes, avec un $\text{pH} \leq 2$; i.e., très acide, l'acidité étant d'autant plus forte que le pH est faible. Sans être uniforme, la distribution de cette variable se prête à un codage barycentrique (i.e., en bref, continu) suivant 4 pivots. De façon précise, on a, pour chaque cas, après codage, un bloc de 4 nombres {nrG0, nrG1, nrG2, nrG3}; si nrg=0, le codage est {1, 0, 0, 0}; pour nrg=1, on a {0, 1, 0, 0}; entre 1 et 5, la séquence est {0, t, 1-t, 0}, t étant d'autant plus fort qu'on est plus près du pivot 1; entre 5 et 15, l'on a de même {0, 0, u, 1-u}, avec u décroissant de 1 à 0; enfin, au-delà du dernier pivot, 15, le codage donne {0, 0, 0, 1}. Ainsi, sans multiplier les modalités, on conserve au mieux l'information apportée par nrg.

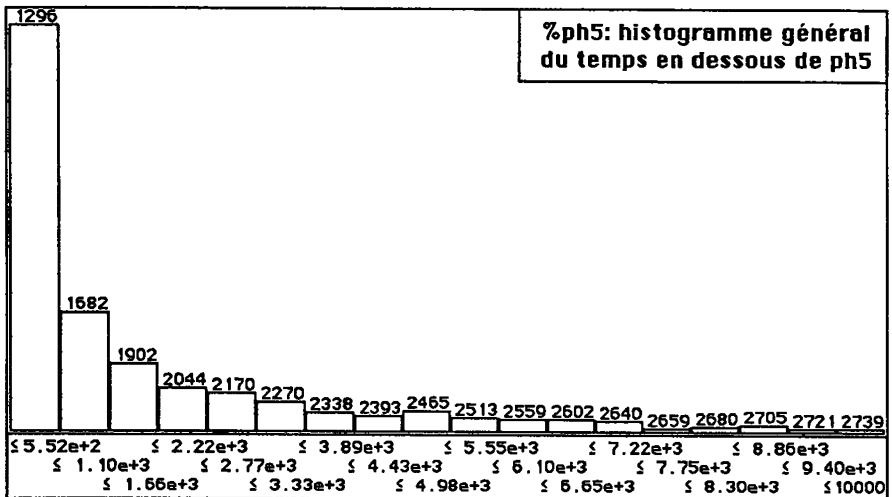
La variable, dur, durée moyenne des reflux, offre, comme nrf, un histogramme en décroissance quasi exponentielle, d'où le choix des bornes séparant les modalités.

Sous ng>5, sont dénombrés les reflux durant plus de 5 minutes. La forme de l'histogramme se prête à un codage barycentrique, en tout analogue à celui adopté pour la variable nrg.

La variable $rg \gg$ mesure la durée du plus long reflux. Ici encore, l'allure exponentielle suggère le choix des bornes; et échappe au codage barycentrique, déséquilibré quand, même dans l'intervalle entre deux pivots successifs, la densité décroît très fortement.

Les deux dernières variables { $\%pH4$, $\%pH5$ }, donnent (multipliée par 10000) la fraction du temps d'examen où le pH œsophagien du patient est descendu respectivement en dessous de 5 ou de 4. Par exemple, rentrent dans la modalité 5pH4 les cas où le pH a été inférieur à 5 dans plus de 30% du temps de l'examen (lequel, on l'a dit, dure 3 heures).

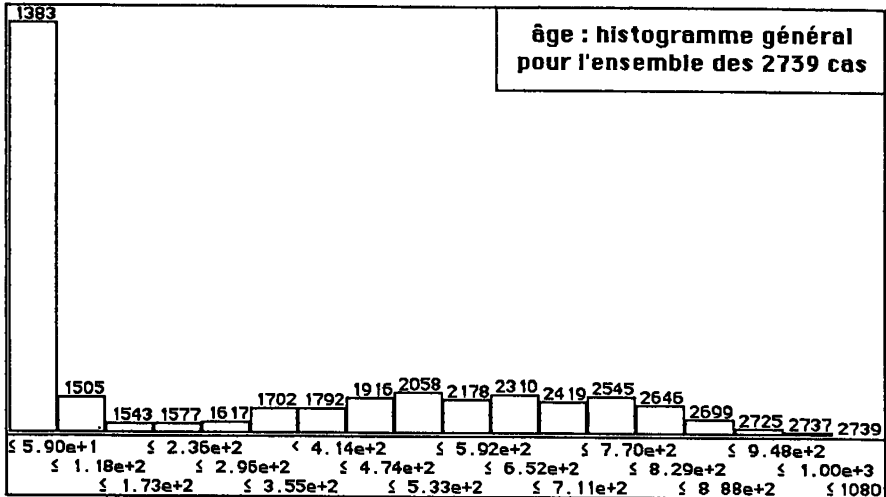
1.2 Exemples d'histogrammes



Pour saisir la distribution des variables, il convient d'en afficher de multiples histogrammes, en revenant encore à de semblables observations quand une analyse le suggère. On se bornera ici à quelques exemples sur lesquels le lecteur biologiste pourra s'initier à la lecture des diverses informations.

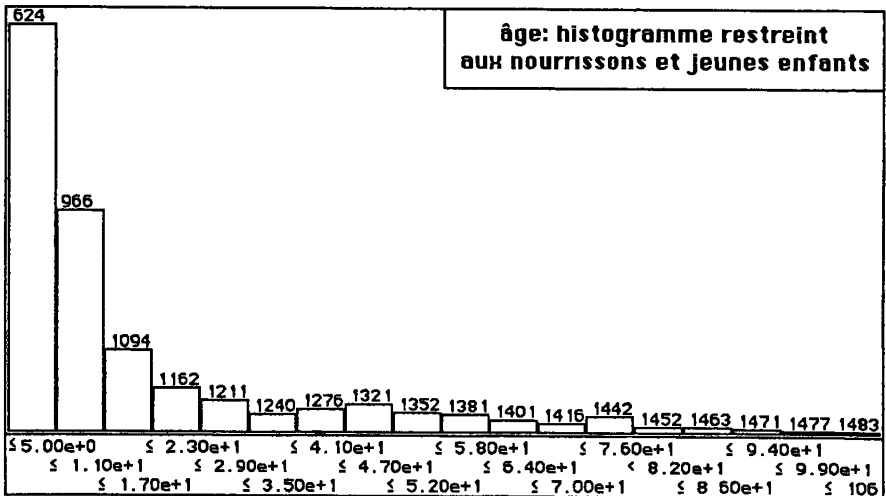
La variable $\%pH5$ offre un histogramme à décroissance quasi exponentielle. Le premier créneau, au haut duquel on lit '1296', comprend les 1296 valeurs les plus faibles (afférentes aux cas où l'acidité est le moins prononcée); soit de 0 à 552/10000, ou en core de 0 à 5,52% du temps en dessous du pH 5. Le créneau suivant, qui comprend les valeurs de rang 1297 à 1682, s'étend entre les bornes 5,52% et 11%; etc.

La variable âge a , comme on l'a dit, une distribution complexe. Un premier histogramme, global, en donne une vue d'ensemble; où l'on reconnaît



la décroissance exponentielle initiale, suivie d'une cloche unimodale très large.

Afin de saisir exactement la distribution des nourrissons et des jeunes enfants, on considérera un histogramme partiel restreint aux 1483 cas des sujets les plus jeunes. La décroissance exponentielle stricte se poursuit jusque vers 30 mois; ensuite s'étalent les cas des jeunes enfants; vers 9 ou 10 ans, la densité est minima; et c'est d'après cela qu'on a séparé les enfants, âe, des sujets dits Mûrs ou Vieux: âM, âV.



2 Analyse multidimensionnelle des données physiologiques

2.1 Le tableau de BURT: physiologie, pathologie, âge

Avec 10 variables découpées, au total, en 52 modalités, on construit un tableau de BURT, kB , tableau carré symétrique croisant avec lui-même l'ensemble J des modalités. En bref, $kB(j, j')$ est le nombre des cas dans les modalités j et j' . (Plus exactement, dans le cas présent, certaines variables ayant été découpées en classes, il s'agit d'un tableau de BURT généralisé; cf. [BURT. COD. BARY.]).

De ce tableau, il s'impose ici de faire l'analyse avec, en principal, le sous-tableau carré croisant avec lui-même l'ensemble des 39 modalités afférentes aux huit variables issues de l'examen physiologique; les 5 modalités de l'âge et les 8 modalités de la pathologie restent en supplément. Le tableau des individus, après codage, peut lui-même être adjoint en supplément à l'analyse (indifféremment, ici, comme lignes ou comme colonnes).

L'analyse du carré 39×39 produit un 1-er axe d'activité gastro-œsophagienne; sur lequel les modalités de chacune des 8 variables physiologiques s'ordonnent depuis le niveau inférieur d'acidité œsophagienne et d'activité gastro-œsophagienne ($F1 < 0$) jusqu'au niveau le plus fort ($F1 > 0$). L'axe 2 est lié à l'axe 1 par un fort effet GUTTMAN; les deux axes rendant compte ensemble de près de 70% de l'inertie du nuage. Pour chacune des variables, on a, dans le plan (1, 2), un schéma en croissant: sur l'axe 2, les modalités moyennes, ($F2 < 0$), s'opposent aux modalités extrêmes ($F > 0$).

D'une part, ce résultat atteste l'excellente cohérence des données; d'autre part, le plan (1, 2) offre un cadre parfait où considérer les rapports entre individus ou groupes d'individus; et, d'abord, les modalités supplémentaires d'âge, \hat{a} , et de pathologie, π , dont chacune n'est autre que le centre de gravité du groupe des cas individuels qu'elle comprend.

Sur l'axe 1 des acidités, l'ensemble des modalités π se range dans l'ordre suivant, de ($F1 < 0$) à ($F1 > 0$):

$$\{\pi_{Mal}, \pi_{Pul}, \pi_{var}, \pi_{r+M}, \pi_{O+P}, \pi_{ORL}, \pi_{OPG}, \pi_{Eso}\};$$

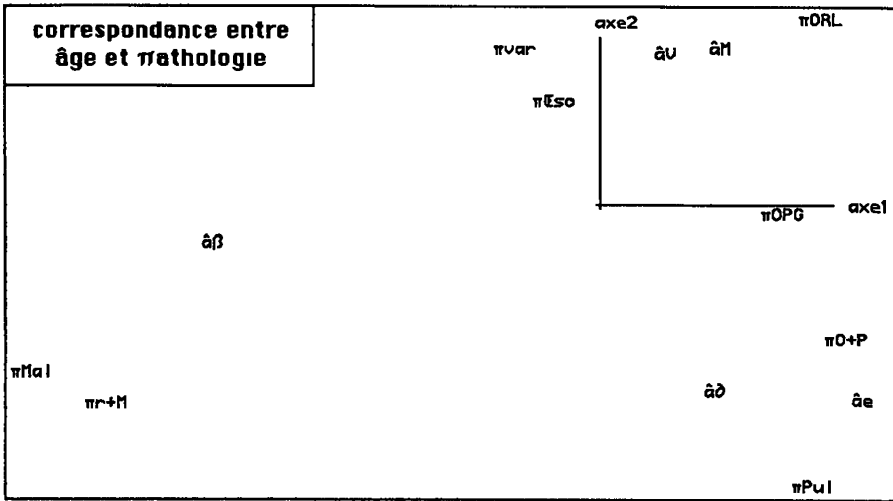
Il est normal que l'acidité maxima échoie à la pathologie œsophagienne, π_{Eso} ; suivie de π_{OPG} , seule pathologie mixte où rentre π_{Eso} . Toutes les autres pathologies sont sur ($F1 < 0$); l'acidité minima allant avec π_{Mal} .

Pour les âges, l'ordre est:

$$\{\hat{a}_e, \hat{a}_\partial, \hat{a}_\beta, \hat{a}_M, \hat{a}_V\};$$

seuls les âges $\{\hat{a}_M, \hat{a}_V\}$ (dix ans et au-delà) sont sur ($F1 > 0$), vers les fortes acidités; pour les nourrissons et jeunes enfants, l'acidité, toujours relativement faible, l'est le plus pour ce que nous avons noté \hat{a}_e , de 30 mois à 12 abs. On ne

s'étonnera pas que les sujets de {âM, âV} ne soient généralement soumis à un examen physiologique que pour des acidités relativement plus fortes que celles qui, chez des sujets très jeunes, suffisent à produire des effets difficilement supportés. D'ailleurs, il est dans l'ordre que les malaises, qui, ainsi qu'on l'a dit, sont surtout le fait de nourrissons et manquent presque après 10 ans, soient sur ($F1 < 0$), avec {âβ, âð, âe}; donc opposés à {âM, âV}.



Considérons, pour plus de précision, le sous-tableau de BURT croisant les pathologies avec les modalités de l'âge. Sur le tableau, comme sur le plan (1, 2), on voit πMal lié avec âβ et opposé à {âM, âV}. De plus, πEso lié à {âM, âV} et aussi à âβ, s'oppose à {âe, âð}; par quoi on expliquera que l'acidité minima soit pour âe et non pour âβ; ce, malgré le taux élevé de πMal propre à âβ.

2739 cas: path x âge
 trace : 3.855e-1
 rang : 1 2 3 4
 lambda: 2497 1217 138 3 e-4
 taux : 6477 3156 359 9 e-4
 cumul : 6477 9632 9991 10000 e-4

2739 cas: path x âge	5	âβ	âð	âe	âM	âV
πORL	8	23	28	110	43	
πPul	31	158	90	47	34	
πO+P	5	56	32	32	30	
πOPG	42	132	82	154	107	
πso	307	167	55	320	296	
πMal	149	29	3	2	3	
πr+M	57	17	2	2	0	
πvar	25	9	2	21	29	

Mais si l'on y regarde de plus près, on voit que, pour les cas de πEso compris dans les âges âβ et âe, la moyenne de F1 est négative (faible acidité); tandis que, comme cela semblerait devoir être la règle générale, elle est positive (forte acidité) pour les âges {âM, âV}. L'âge n'est pas seulement lié à l'activité œsophagienne par le biais d'une distribution variable des pathologies; il module également cette activité au sein d'une pathologie donnée.

2739 cas: tableau de BURT: en pr., le carré des 39 mod. physiologiques
 trace : 1.297e+0
 rang : 1 2 3 4 5 6 7 8 9 10
 lambda : 6044 2910 1380 668 443 278 201 158 156 139 e-4
 taux : 4660 2244 1064 515 341 214 155 122 120 107 e-4
 cumul : 4660 6904 7968 8484 8825 9039 9194 9316 9436 9544 e-4

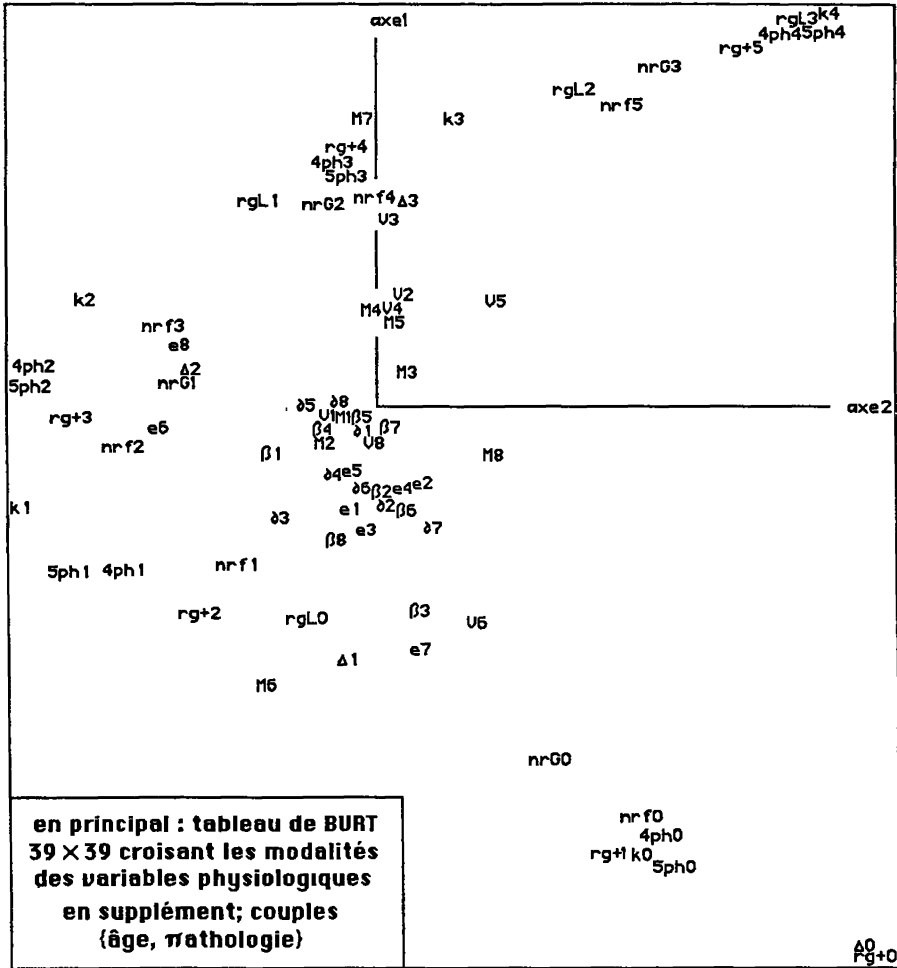
2.2 Combinaisons entre âge et pathologie

Il y a, *a priori*, 40 couples (\hat{a} , π); que l'on peut désigner par un sigle formé de deux caractères: la lettre propre à la classe d'âge étant suivie par le n° de la pathologie. Dans chacun de ces couples, rentre un nombre plus ou moins grand de cas. Comme on peut le voir sur le tableau de croisement $\pi_{ath} \times \hat{a}$, seul V7 est vide: il n'y a pas de sujet de plus de 50 ans, affecté de $\pi_7 + M = \pi_7$, conjonction de Malaise avec reflux.

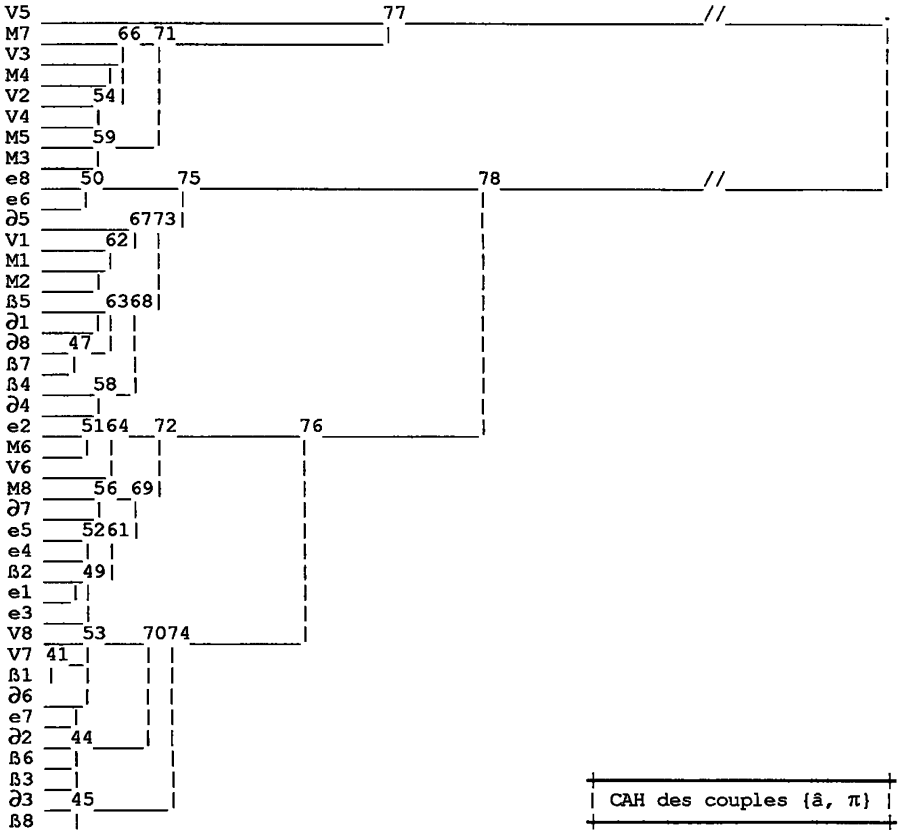
Ainsi qu'on l'a dit au début du précédent §, à chaque cas individuel est associée une ligne codant ce cas suivant l'ensemble des 52 modalités: en bref, il y a '1' pour la modalité de chaque variable où rentre le cas; et zéro ailleurs; pour une variable codée barycentriquement (i.e., cf. §1.1, pour nrg ou rg>5) la masse unité est partagée entre deux modalités consécutives. Donc, à chaque couple $\{\hat{a}, \pi\}$, correspond un bloc de lignes; dont le cumul peut lui-même être adjoint en supplément à l'analyse du tableau de BURT; sa place n'étant autre que la moyenne des cas individuels.

Reste à construire ces lignes de cumul. On peut écrire un programme spécial dans un langage universel tel que PASCAL. Jouer avec les instructions propres à un tableur. Nous avons procédé autrement, en utilisant des programmes du logiciel MacSAIF: d'une part 'coupli', (programme recomposant des lignes par découpage de segments: cf. [ÉTAT MACSAIF]); d'autre part un programme de tri alphabétique d'une liste de chaînes de caractères. En bref, par 'coupli', on place d'abord, en tête de chaque ligne du tableau des données, le mot de 4 caractères donnant la valeur de l'âge (colonne âge). Un tri alphabétique donne alors les lignes des 2739 sujets, rangées par âge croissant; le mot initial (âge copié) étant supprimé (en utilisant une deuxième fois coupli), on peut séparer les cinq blocs de cas afférents aux cinq classes d'âge (de $\hat{a}\beta$ à $\hat{a}V$). Il vaut la peine de construire pour chacun des tableaux constitués par ces blocs, un tableau de BURT, découpé suivant les bornes mêmes déjà prises pour l'analyse globale. En particulier, la ligne afférente à la pathologie π , dans le tableau de BURT propre à l'âge \hat{a} est exactement le cumul des lignes propres au couple $\{\hat{a}, \pi\}$. On aurait pu encore recopier en tête de ligne l'âge suivi du numéro de diagnostic; mais les tableaux de BURT partiels ont leur intérêt propre; sur lequel on n'insistera pas.

C'est ainsi qu'on a construit le plan (1, 2) publié ici; où figure, d'une part, le nuage des 39 modalités principales des 8 variables physiologiques; et, d'autre part, en supplément, les 39 couples $\{\hat{a}, \pi\}$ non vides.



Sur ce plan, le lecteur pourra d'abord, pour chacune des 8 variables physiologiques, vérifier que ses modalités s'ordonnent suivant l'effet GUTTMAN annoncé au §2.1. Il tentera ensuite de trouver un ordre global dans le nuage des couples $\{\hat{\alpha}, \pi\}$. Certes, à un âge donné, la pathologie 5 (Eso) dépasse quant à F1 (axe vertical, acidité) la pathologie 2 (Pulmo). Et, pour une pathologie donnée, l'acidité croît ordinairement avec l'âge: mais pour π_5 , l'acidité est minima à l'âge $\hat{\alpha}_e$ (de 30 mois à 10 ans); les points $\{\beta_2, \partial_2, e_2\}$ sont groupés, M2 et V2 s'en séparant vers $(F1 < 0)$. Et il faut se garder de rien conclure de la très faible acidité de M6 et de V6; qui comptent 2 ou 3 cas.



La CAH des couples confirme que ceux-ci ne se rangent ni par âge ni par diagnostic. Il y a une branche 77 qui ne comprend que des couples en $\hat{a}M$ et $\hat{a}V$ (âges au-delà de 10 ans); mais 7 couples propres à ces âges ne sont pas dans cette branche; notamment {M1, M2, V1}, dont le poids est notable et qui se projettent sur ($F1 < 0$: acidité moyenne ou faible); mais, il est vrai, près de l'origine.

Il apparaît donc impossible de faire le diagnostic de la pathologie d'après les seules variables physiologiques mesurées lors de l'examen de pH-métrie.

En effet, considérons, par exemple, les deux couples B6 (149 nourrissons atteints de Malaises) et D2 (158 cas de très jeunes enfants présentant une pathologie pulmonaire pure. Dans le plan (1, 2), B6 voisine D2 et la CAH les agrège à un très bas niveau: ceci suggère, ainsi qu'on peut le vérifier, que les sous-nuages de cas afférents à ces deux couples se recouvrent en partie; d'où,

entre des cas de πMal (à l'âge $\hat{\alpha}\beta$) et des cas de πPul (à l'âge $\hat{\alpha}\beta$), une similitude qui interdit de distinguer entre ces deux diagnostics.

La seule voie encore ouverte est de faire le diagnostic au sein d'une seule tranche d'âge; ou, ce qui revient au même, en adjoignant l'âge aux huit variables décrivant l'acidité œsophagienne.

3 Essais de diagnostic par l'analyse discriminante, pour une tranche d'âge

Sous le nom d'analyse discriminante, la statistique multidimensionnelle propose, entre autres méthodes, celle de substituer aux notions initiales de cas individuel et de diagnostic, des points dans l'espace euclidien engendré par les axes issus de l'analyse factorielle. De façon précise, si I est l'ensemble des cas et J , l'ensemble des diagnostics, on attribuera à l'individu se projetant suivant le point i , le diagnostic afférent à celui des points j qui est le plus proche de i . La méthode offre des variantes selon le nombre des axes factoriels retenus pour l'espace au sein duquel sont calculées les distances.

Dans le cas présent, on a déjà vu que l'analyse écarte le projet de traiter l'ensemble I des 2739 cas de tous âges. Reste donc à tenter la discrimination pour le sous-ensemble des cas rentrant dans une classe d'âge donnée.

La tranche, $\hat{\alpha}V$, des sujets les plus âgés apparaît, *a priori*, comme définissant des conditions physiologiques plus stables que celles de nos autres modalités d'âge, marquées par le passage rapide de la naissance à l'adolescence. D'où le problème géométrique suivant, substitué à celui du diagnostic. Pour chacun des points de l'ensemble IV des 542 cas de la tranche d'âge $\hat{\alpha}V$, déterminer le couple, $V_x = \{V, \pi_x\}$, dont il est le plus proche.

Plus précisément, on se restreint aux 5 couples $\{V1, V2, V3, V4, V5\}$; car, dans $V6$ et $V7$, le nombre des cas est, respectivement, 3 et 0; et $V8$, (qui compte 29 cas), n'est qu'une absence de diagnostic ('divers'). Et on a procédé dans l'espace engendré par l'ensemble des 31 axes issus de l'analyse; ainsi qu'en projection dans l'espace engendré par les dix premiers axes. Les deux affectations ne différant entre elles que pour un petit nombre de sujets, on ne présentera que les résultats obtenus avec dix axes.

Le programme 'discr'i' produit un listage qui donne les affectations des individus successifs; puis des bilans, pour chaque diagnostic, des individus qui y ont été affectés; ceux-ci, ayant été rangés par blocs de cas rentrant dans la même pathologie, avec les numéros de dossier dans l'ordre croissant, il est facile de dénombrer les affectations exactes. (En bref, le fichier des données, afférentes aux sujets de $\hat{\alpha}V$, a été rangé par une méthode toute semblable à celle adoptée ci-dessus pour le groupement des classes d'âges).

Dur:W:PhtB ; nombre de facteurs utilisés = 10

affectation des 1V aux V π

(320-> V1) (361-> V1) (525-> V1) (526-> V1) (646-> V1) (715-> V3)
 (916-> V4) (1070-> V1) (1146-> V2) (1215-> V5) (1247-> V1) (1346-> V5)
 (1611-> V3) (1614-> V3) (1638-> V1) (1700-> V1) (1769-> V4) (1777-> V5) ...

Le tableau partiel ci-dessus donne les affectations obtenues pour les 18 premiers sujets rentrant dans le couple V1: il y a 8 affectations exactes; (V1 -> V1). En tout, le bilan est:

diagnostic véritable :	V1	V2	V3	V4	V5
nombre de cas :	43	34	30	107	296
affectations exactes :	25	9	14	1	102
affectations erronées:	18	25	16	106	194

On voit que la discrimination n'est aucunement satisfaisante.

Dans un autre essai, on a considéré, en particulier, les couples V2 et V5 (i.e.: pathologies Pulmonaire et Gastro-œsophagienne, au-delà de 50 ans). L'affectation se faisant à deux centres seulement, {V2, V5}, on a trouvé 18 affectations exactes pour les 34 cas que compte V2; et 104, pour les 296 cas de V5. Ce qui n'est pas plus satisfaisant.

4 Perspectives et conclusions

Il suffira de reprendre ici ce qu'annonçait l'introduction.

À nos yeux, la présente étude a pour premier mérite d'avoir montré la parfaite cohérence des données de pH-métrie, où l'analyse reconnaît une notion unique: l'activité gastro-œsophagienne. L'attention a été, d'autre part, appelée sur la forte corrélation, existant chez nos patients, entre cette activité et l'âge. Mais, même en se restreignant à une tranche d'âge, il semble impossible de fonder un diagnostic sur nos seules données physiologiques.

Reste qu'un ensemble de patients ne saurait représenter fidèlement le niveau de reflux gastro-œsophagien de la population générale, au sein de laquelle les patients ne se distinguent, d'ailleurs, par aucun critère précis. D'où l'intérêt d'une enquête épidémiologique de pH-métrie qui serait faite à l'occasion de visites médicales de routine.

Références bibliographiques

A. EL OUADRANI : "Généralisation du tableau de BURT et de l'analyse de ses sous-tableaux dans le cas d'un codage barycentrique"; [BURT. COD. BARY.]; in *CAD*, Vol.XIX, n°2; pp.229-246; (1994);

J.-P. & F. BENZÉCRI : "État des modifications apportées aux versions du logiciel MacSAIF"; [ÉTAT MacSAIF]; in *CAD*, Vol.XVIII, n°3; pp.321-344; (1993).