

J.-P. NAKACHE

P. LORENTE

J.-P. BENZÉCRI

J.-F. CHASTANG

**Aspects pronostiques et thérapeutiques de
l'infarctus myocardique aigu compliqué d'une
défaillance sévère de la pompe cardiaque.
Application des méthodes de discrimination**

Les cahiers de l'analyse des données, tome 2, n° 4 (1977),
p. 415-434

http://www.numdam.org/item?id=CAD_1977__2_4_415_0

© Les cahiers de l'analyse des données, Dunod, 1977, 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/>

ASPECTS PRONOSTIQUES ET THÉRAPEUTIQUES
DE L'INFARCTUS MYOCARDIQUE AIGU
COMPLIQUÉ D'UNE DÉFAILLANCE SÉVÈRE
DE LA POMPE CARDIAQUE.
APPLICATION DES MÉTHODES DE DISCRIMINATION
[AORTE]

par J.-P. Nakache, P. Lorente ⁽¹⁾
J.-P. Benzécri, J.-F. Chastang ⁽²⁾

1. Le problème thérapeutique

Depuis quelques années les malades atteints d'un infarctus myocardique aigu compliqué d'une défaillance sévère de la pompe cardiaque peuvent bénéficier de certaines techniques d'assistance. Jusqu'alors les thérapeutiques conventionnelles étaient impuissantes à améliorer ces états dramatiques où la destruction importante du muscle cardiaque conditionnait des perturbations circulatoires irréversibles. L'avènement des techniques d'assistance a permis d'envisager une amélioration du pronostic de ces cas gravissimes.

La technique la plus communément utilisée est la contre-pulsion par ballon intra-aortique (CPBIA) [1],[2]. Ce procédé consiste à introduire par l'artère fémorale jusque dans l'aorte thoracique descendante un ballon en polyuréthane de 24 cm de long et d'un diamètre inférieur à celui de l'aorte (2 cm). Ce ballon est relié à une pompe d'animation externe par l'intermédiaire d'un catheter de calibre 14 F (diamètre externe = 4 mm). C'est cette pompe qui permet d'assurer grâce à un système d'asservissement électronique basé sur l'électrocardiogramme du patient le gonflement et le dégonflement alternatif du ballon. L'ensemble du dispositif pneumatique contient de l'hélium et il est conçu de manière à propulser des volumes sanguins de l'ordre de 20 à 40 ml par cycle.

La synchronisation sur le cycle cardiaque grâce au signal électrocardiographique donne la séquence suivante :

. Pendant la phase de repos cardiaque (diastole) le gonflement du ballon chasse le sang d'une part vers la circulation artérielle périphérique et d'autre part vers le coeur (système artériel coronaire). L'effet favorable espéré est d'améliorer l'irrigation générale et donc l'apport en oxygène au niveau des tissus périphériques et du coeur lui-même. A la fin de la période de gonflement du ballon l'aorte s'est donc partiellement vidée de son contenu sanguin grâce à cette impulsion fournie par la pompe d'assistance (d'où le nom de contre-pulsion diastolique intra-aortique).

. Pendant la phase de contraction cardiaque (systole) le ballon reste dégonflé. Le ventricule gauche peut alors chasser le sang dans l'aorte dont l'impédance d'entrée a décru (puisque'elle s'est vidée en partie) : il en résulte une diminution du travail effectif du ventricule gauche et des besoins en oxygène du muscle cardiaque.

(1) Groupe de recherches INSERM U88 et U14 (Laboratoire du Professeur F. GREMY).

(2) Laboratoire de statistique. Université Pierre et Marie Curie

Augmentation de l'apport et diminution de la demande en O_2 peuvent contribuer à améliorer un état myocardique critique. Cependant l'institution d'une telle thérapeutique nécessite l'intervention d'une équipe médico-chirurgicale entraînée et d'un matériel important. On ne recourt donc à la contre-pulsion que si sans elle les chances de survie paraissent réduites. D'où un problème de décision que l'analyse des données est susceptible de résoudre.

La présente étude comporte deux phases. D'une part, analyse des dossiers de 101 patients reçus à l'unité de soins intensifs de l'hôpital Lariboisière alors que la contre-pulsion n'y était pas encore pratiquée (§ 2). D'autre part, application des conclusions de la première phase au traitement de 55 patients reçus dans une période où la contre-pulsion pouvait être appliquée (§ 5). En étudiant les 101 premiers dossiers par l'analyse des correspondances (§ 3) et par plusieurs méthodes de discrimination (§ 4), on a pu avec une certitude appréciable (10% seulement de cas douteux) discriminer entre les patients décédés dans les quatre semaines qui suivent l'infarctus, et ceux qui survivent au-delà de ce terme. Ainsi quand un patient nouveau se présente, la même règle de discrimination permet (d'après les examens auxquels celui-ci est soumis) de savoir (et cela par des calculs très simples) s'il est de ceux qui sans aide mécanique spéciale n'ont pas de chance de survie ; à un tel patient on applique, dans la mesure du possible, la contre-pulsion. L'expérience de 55 cas a montré le bien fondé de cette règle thérapeutique (§ 5).

2 Le tableau des données de l'étude de base

Considérons successivement l'ensemble des patients, les variables retenues, le codage de celles-ci.

2.1 Les patients : 101 patients présentant un infarctus myocardique aigu compliqué de défaillance cardiaque font l'objet de cette étude. Dès leur arrivée dans l'unité de Soins Intensifs de l'Hôpital Lariboisière, ils ont été soumis à des examens cliniques biologiques et hémodynamiques dans le cadre d'une surveillance continue.

Aucun d'entre eux n'a d'autre affection ou complication associée que l'atteinte de la pompe cardiaque : quant à la localisation de l'infarctus (ou mauvaise irrigation du muscle cardiaque), on distingue deux classes topographiques.

Infarctus du myocarde *antérieur* (52 IMA) ; et *postérieur* (49 IMP).

Sont dans l'étude, considérés comme décédés les malades (au nombre de 51) pour lesquels une issue fatale survient dans un délai de 4 semaines après le début de l'infarctus. Les cinquante patients dépassant ce cap sont considérés comme survivants.

2.2 Choix des variables : Vu le but poursuivi (une décision thérapeutique) les données cliniques et hémodynamiques analysées proviennent des examens initiaux effectués à l'entrée des patients : on a 19 variables quantitatives et 3 questions à deux modalités (sexe ; existence d'un infarctus dans les antécédents ; troubles de la conduction nécessitant un stimulateur temporaire). S'y ajoutent dans le tableau 1 qui illustre cet article, deux questions : survie - décès et localisation (IMA/IMP) non prises en compte dans les analyses ; car la première constitue justement la *variable à expliquer* tandis que la seconde apparaîtra d'elle-même sur le 2° axe issu de l'a. desc. (§ 3).

Ce n'est pas le lieu d'enseigner l'hémodynamique à des statisticiens ; toutefois ceux-ci ayant étudié la physique, s'intéresseront aux quelques notes qui suivent :

CODE	VARIABLES	classe 1	classe 2	classe 3	classe 4	classe 5
M/F	SEXE	masculin	féminin			
AG	AGE (années)	57	58-69	70		
PM	INFARCTUS DANS LES ANTECEDENTS	absence	présence			
TS	DATE DE LA PREMIERE OBSERVATION PAR RAPPORT AU DEBUT DE L'INFARCTUS	1 jour	2 jours	3 jours		
M/I	LOCALISATION DE L'INFARCTUS	antérieur	postérieur			
HR	FREQUENCE CARDIAQUE (systoles / min)	82	84-96	99		
CJ	INDEX CARDIAQUE (l/min.m ⁻²)	1.23	1.24 - 1.56	1.58 - 1.92	1.93 - 2.40	2.41
SI	INDEX SYSTOLIQUE (m/systole.m ⁻²)	13.8	13.9 - 17.4	17.5 - 21.7	21.8 - 26.9	27.2
DA	PRESSON DIASTOLIQUE AORTIQUE (mmHg)	58	59 - 67	68		
PA	PRESSON AORTIQUE MOYENNE (mmHg)	69	70-81	82		
R	RESISTANCE ARTERIELLE SYSTEMIQUE (dyn.sec.cm ⁻⁵ .m ²)	2156	2172-2809	2846-3164	3173-3897	4031
LF	PRESSON DE REMPLISSAGE VG (mmHg)	14	15-17	18-20	21-24	25
MP	PRESSON ARTERIELLE PULMONAIRE MOYENNE (mmHg)	19	20-23	24-27	28-31	32
RF	PRESSON TEDIASTOLIQUE VD (mmHg)	7	8-11	12		
PR	RESISTANCE PULMONAIRE TOTALE (dyn.sec.cm ⁻⁵ .m ²)	797	807-1108	1131-1608	1610	
LS	TRAVAIL GAUCHE SYSTOLIQUE (g.m/m ²)	10.49	10.72-14.83	15.29-24.0	24.70	
RS	TRAVAIL DROIT SYSTOLIQUE (g.m/m ²)	2.74	2.82-4.15	4.16-6.15	6.22	
L'	PUISSANCE GAUCHE MINUTE (g.m/min.m ⁻²)	1022	1031-1325	1328-2098	2131	
RM	PUISSANCE DROITE MINUTE (g.m/min.m ⁻²)	242	246-387	388-564	573	
LT	INDEX GAUCHE (g.m/m ² .mmHg ⁻¹)	0.44	0.45-0.64	0.67-0.93	0.95-1.65	1.72
PT	INDEX DROIT (g.m/m ² .mmHg ⁻¹)	0.20	0.21-0.35	0.36-0.53	0.54-1.14	1.16
S/D	SURVIE/DECES	survivant	décédé			
CD	TRoubles DE CONDUCTION	absence	présence			
PV	INDEX DE GENERATION DE PRESSON VENTRICULAIRE DROITE (sens dimension)	0.84	1.05-1.62	1.83-2.88	3.00	

Tableau 1 : Liste des variables de l'étude, du code utilisé dans l'AFC et du découpage en classes

Le tableau 2 fournit les pourcentages d'inertie expliquée par les cinq premiers axes factoriels

Axes factoriels	Pourcentages	Pourcentages cumulés
F ₁	12.890	12.890
F ₂	8.180	21.070
F ₃	7.114	28.184
F ₄	5.066	33.250
F ₅	4.704	37.954

Tableau 2 : Pourcentages d'inertie expliqués par les 5 premiers facteurs de l'AFC

Le plan factoriel (F₁, F₂) explique 21,1% de l'inertie totale du nuage. Ce taux peut sembler relativement faible : d'une manière générale, l'analyse des tableaux sous codage disjonctif complet donne toujours de tels taux d'inertie ce qui fait mal juger de la part d'information extraite. Comme il est expliqué en [Bin. Mult.] § 2' Note, (Cahiers, Vol II n° 1 ; pp 59-60) voir aussi [Burins] Vol II n° 2 p 233) il convient en fait dans l'analyse de tels tableaux de considérer non les valeurs propres λ_α mais leur carré $\Lambda_\alpha = \lambda_\alpha^2$; les Λ_α correspondent à l'analyse du tableau de Burt (*loc. cit.*) qui donne à un changement d'échelle près les mêmes vecteurs, mais avec des pourcentages d'inertie différents, dont la considération se justifie en ce qu'ils s'accordent avec le cas limite d'un codage continu des variables ; ici on aura quant aux Λ_α les pourcentages suivants nettement plus favorables et répétons-le conformes aux faits (cf tableau 2').

Axes factoriels	Pourcentages	Pourcentages cumulés
F ₁	35.636	35.636
F ₂	14.350	49.986
F ₃	10.860	60.846
F ₄	5.507	66.353
F ₅	4.744	71.097

Tableau 2' : Pourcentages d'inertie calculés à partir du tableau de Burt

CI : Index cardiaque en litres par minute et par m^2 de surface corporelle ($l/mn.m^{-2}$). SI : Index systolique en $ml/systole.m^{-2} = CI / \text{fréquence cardiaque}$. R : Résistance artérielle périphérique (ou systémique) calculée suivant la loi de Poiseuille = $(MA - RF)mmHg \times 1332 \times 60/CIml/mn.m^{-2} = R \text{ dynes.sec. cm}^{-5}. m^2$.

PR : Résistance pulmonaire totale calculée suivant la même loi = $MP \text{ mm Hg} \times 1332 \times 60/CI \text{ ml/mn. } m^{-2} = PR \text{ dynes.sec. cm}^{-5}. m^2$.

LS : Index du travail systolique du ventricule gauche = $SI \text{ ml/syst. } m^{-2} \times (MA - LF)mmHg \times 1,055 \times 13,6 \times 10^{-3} = LS \text{ Kgm.m}^{-2}.10^3$

RS : Index du travail systolique du ventricule droit = $SI \text{ ml/syst.}m^{-2} \times (MA - LF) \text{ mn Hg} \times 1,055 \times 13,6 \times 10^{-3} = RS \text{ Kgm.m}^{-2}.10^3$

RS : Index du travail systolique du ventricule droit = $SI \text{ ml/syst.}m^{-2} \times (MP - RF)mn \text{ Hg} \times 1,055 \times 13,6 \times 10^{-3} = RS \text{ Kgm.m}^{-2}. 10^3$

LM : Index de puissance minute ventriculaire gauche = $LS \times \text{fréquence cardiaque} = LM \text{ Kgm.m}^{-2}.10^3/mn$.

LI : Index gauche = $LS/LF = LI \text{ Kgm.m}^{-2}. 10^3. \text{ mm Hg}^{-1}$

RI : Index droit = $RS/RF = RI \text{ Kgm.m}^{-2}.10^3. \text{ mm Hg}^{-1}$

RM = $RS \times \text{fréquence cardiaque} = RM \text{ Kgm. } m^{-2}. 10^3/mN$

RV = $(MP - RF)/RF = RV \text{ sans dimension}$.

2.3 Codage des données : Les informations principales sont fournies par un ensemble hétérogène de variables exprimées dans des unités physiques différentes. Il a donc été nécessaire de procéder à une homogénéisation des variables qui a conduit à la construction d'un tableau sous forme disjonctive complète [3]. Au vu de son histogramme, l'intervalle de variation de chacune des variables physiques est divisé en classes d'effectifs approximativement égaux, en tenant compte de la signification physiopathologique des bornes. Par exemple, pour la variable Pression artérielle pulmonaire moyenne, on a distingué cinq classes dont les bornes sont en millimètres de mercure (mmHg) : 19,5 ; 23,5 ; 27,5 ; 31,5 ; ainsi un sujet dont l' PAPM est 22 mmHg, aura dans la col. PA2 (car il entre dans la deuxième classe 19,5 < 22 < 23,5) et 0 dans les autres colonnes. (PA1, PA3, PA4, PA5) afférentes à la variable PAPM. Pour la variable date de l'observation DT, on a trois modalités (examiné pendant la première journée, la deuxième journée, la troisième et plus) ; etc. L'ensemble des bornes choisies est donné dans le tableau des variables (Tableau 1).

Au § 3, on analyse le tableau $101 \times (83 + 4) = (101 \text{ patients}) \times [83 \text{ modalités (des variables principales) + 4 modalités (des 2 variables supplémentaires : localisation et issue)}]$. Les diverses méthodes de discrimination essayées au § 4, (classification hiérarchique [4] ; agrégation autour de centres variables (E. Diday [5]) ; hyperplan médiateur ; méthode de plus proches voisins [6]), reposent aussi sur le même tableau 101×83 ; la distance entre individus (lignes) étant la distance du χ^2 ; fait exception la méthode d'affectation barycentrique du § 4.5, laquelle par des calculs très simples (ne requérant pas d'ordinateur) effectue l'analyse d'un tableau de correspondance à deux lignes et 83 colonnes auquel les individus sont adjoints en lignes supplémentaires. Enfin au § 5 les dossiers des 55 malades hospitalisés après l'étude des 101 dossiers de base, sont comme ceux-ci, codés sur une ligne de zéro et de un suivant les 83 modalités des variables principales.

3. Etude de l'échantillon de base par l'analyse des correspondances (AFC) [7]

Le tableau disjonctif complet ainsi obtenu a été analysé par l'AFC. Les paramètres à expliquer de l'étude - I (infarctus postérieur), A (infarctus antérieur), S (survie) et D (décès) - ont été placés en variables supplémentaires sur les axes déterminés par les seules variables supplémentaires et par les patients.

3.1 Interprétation du 1° axe factoriel

3.1.1 Contributions

Le tableau 3 représente les contributions relatives des 87 paramètres binaires sur l'axe F_1 .

Les contributions les plus importantes sont relatives aux variables : puissance gauche minute (LM), travail gauche systolique (LS), fréquence cardiaque (CI), index cardiaque (SI), résistance pulmonaire totale (PR), index gauche (LI), qui sont les paramètres de fonction ventriculaire gauche.

Paramètres Code	. Modalités				
	1	2	3	4	5
MF	0.05	0.05			
AG	0.21	0.03	0.07		
PM	0.07	0.07			
TS	0.19	0.04	0.07		
HR	0.04	0.00	0.07		
CI	0.33	0.12	0.01	0.13	0.44
SI	0.40	0.09	0.00	0.14	0.38
DA	0.01	0.01	0.05		
MA	0.10	0.00	0.09		
R	0.12	0.05	0.00	0.03	0.19
LF	0.21	0.01	0.03	0.01	0.06
MP	0.11	0.01	0.03	0.00	0.04
RF	0.24	0.05	0.08		
PR	0.39	0.08	0.05	0.47	
LS	0.47	0.05	0.03	0.55	
RS	0.14	0.06	0.00	0.33	
LM	0.34	0.14	0.04	0.57	
RM	0.06	0.07	0.00	0.20	
LI	0.40	0.04	0.02	0.11	0.39
RI	0.07	0.11	0.03	0.07	0.22
CD	0.17	0.17			
RV	0.00	0.04	0.02	0.11	
AI	0.00	0.00			
SD	0.62	0.62			

Tableau 3 : Contributions relatives du premier axe factoriel aux modalités des paramètres (e.g. la contribution à AG1 est de 21%)

Note : Pour comprendre le terme de fonction ventriculaire gauche, on se souviendra du schéma général de la circulation sanguine. Le sang cyanosé (bleu) fait retour à l'auricule droit par les veines caves il gagne le ventricule droit qui chasse le sang par l'artère pulmonaire dans les poumons. Le sang hématosé (rouge) revient par les veines pulmonaires dans l'auricule gauche, gagne le ventricule gauche, d'où il est chassé dans la circulation générale par l'artère aorte.

C'est la fonction du ventricule gauche, fonction très lourde dont on conçoit qu'elle domine le pronostic des patients.

3.1.2 Lignes polygonales

En reliant entre elles les modalités binaires correspondant à ces paramètres de fonction ventriculaire gauche, on obtient des lignes polygonales qui suivent assez régulièrement la direction de F_1 avec les valeurs les plus basses à gauche et les plus élevées à droite (Figure 1) ; tel est le cas par exemple pour les quatre modalités du paramètre LS (travail systolique ventriculaire gauche) qui sont rappelons-le (cf tableau 1) :

- LS1 correspondant aux valeurs les plus basses du travail gauche ($LS < 10.49$)
- LS2 correspondant à des valeurs un peu moins basses du travail gauche ($10.72 \leq LS < 14.83$)
- LS3 correspond à des valeurs un peu plus élevées du travail gauche ($15.29 \leq LS < 24.0$)
- LS4 correspondant à des valeurs élevées du travail gauche ($LS \geq 24.70$) (à titre indicatif l'intervalle des valeurs normales de LS est : $63 \pm 12 \text{ Kg} \cdot 10^{-3} / \text{m}^2$; on voit que les moins atteints des patients sont encore très en dessous de cette norme).

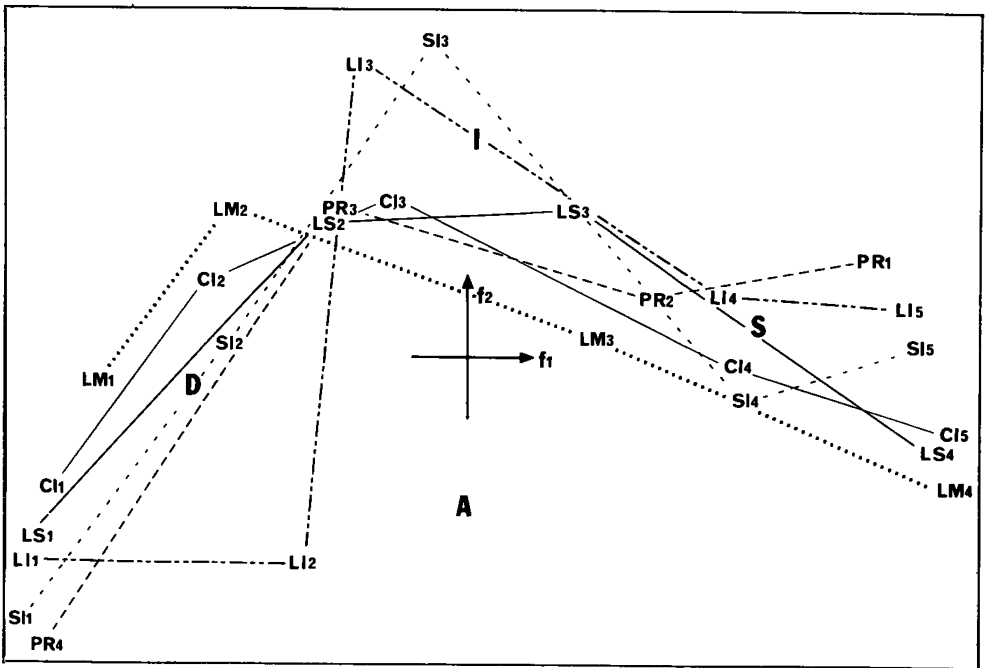


Figure 1 : Lignes polygonales des paramètres les plus liés à l'axe F_1

Ces lignes polygonales permettent de mettre en évidence des relations non linéaires entre les paramètres initiaux, de tester le découpage en classes et de valider *a posteriori* les données.

De plus les contributions importantes des variables supplémentaires S (survie) et D (décès) à l'axe F_1 permettent d'affirmer la valeur pronostique de cet axe quant à la survie au bout de 4 semaines ; interprétation pleinement et définitivement confirmée par l'examen détaillé du nuage des individus.

3.2 Interprétation du 2° axe factoriel

3.2.1 Contributions

Le tableau 5 fournit les contributions relatives des variables à l'axe F_2 . Les plus importantes sont relatives aux variables suivantes : pression de remplissage ventriculaire droit (RF), travail droit systolique (RS), puissance droite minute (RM), index droit (RI), index de génération de pression ventriculaire droite (RV), fréquence cardiaque (HR), pression artérielle moyenne (MP). Ce sont les paramètres importants de fonction ventriculaire droite.

Paramètres Code	Classes				
	1	2	3	4	5
MF	0.02	0.02			
AG	0.02	0.01	0.01		
PM	0.02	0.02			
TS	0.02	0.01	0.01		
HR	0.18	0.00	0.22		
CI	0.03	0.01	0.05	0.00	0.01
SI	0.14	0.00	0.20	0.00	0.00
DA	0.06	0.01	0.14		
MA	0.05	0.04	0.17		
R	0.01	0.00	0.01	0.00	0.07
LF	0.09	0.11	0.00	0.03	0.24
MP	0.19	0.03	0.01	0.04	0.29
RF	0.02	0.08	0.24		
PR	0.03	0.01	0.05	0.22	
LS	0.09	0.04	0.05	0.02	
RS	0.34	0.02	0.01	0.16	
LM	0.00	0.05	0.00	0.05	
RM	0.44	0.00	0.04	0.18	
LI	0.08	0.08	0.16	0.01	0.00
RI	0.47	0.01	0.01	0.08	0.03
CD	0.05	0.05			
RV	0.53	0.00	0.21	0.08	
AI	0.25	0.25			
SD	0.01	0.01			

Tableau 5 : Contributions relatives du deuxième axe factoriel aux paramètres

Note : Quant aux termes fonctions ventriculaires droites, nous renvoyons le lecteur à l'explication donnée au § 3.1 pour la fonction ventriculaire gauche. On notera que le travail du ventricule droit étant moins lourd que celui du ventricule gauche, elle n'apparaît pas nettement sur l'axe décès-survie. On a bien du côté SURVIE (F_1 positif) les caractéristiques d'une fonction droite relativement bonne, mais les modalités les plus défavorables s'éloignent dans la direction du 2° axe positif.

3.2.2 Lignes polygonales

Reliant entre elles les classes de ces variables, on obtient (Fig. 2) des lignes polygonales suivant la direction de F_2 . On remarque pourtant que les valeurs les plus basses de ces indices sont très bien corrélées avec F_2 à la partie supérieure du graphe alors que les valeurs les plus élevées de ces indices tendent à se diriger dans la direction de F_1 vers la zone SURVIE. En conséquence l'axe F_2 sépare les infarctus antérieurs et postérieurs moins nettement que F_1 ne sépare les survivants et les décédés.

De plus, comme le montre le tableau 5, F_2 est bien corrélé avec les variables supplémentaires A (infarctus antérieur) et I (Infarctus postérieur) ce qui force la caractérisation de l'axe F_2 comme axe topographique.

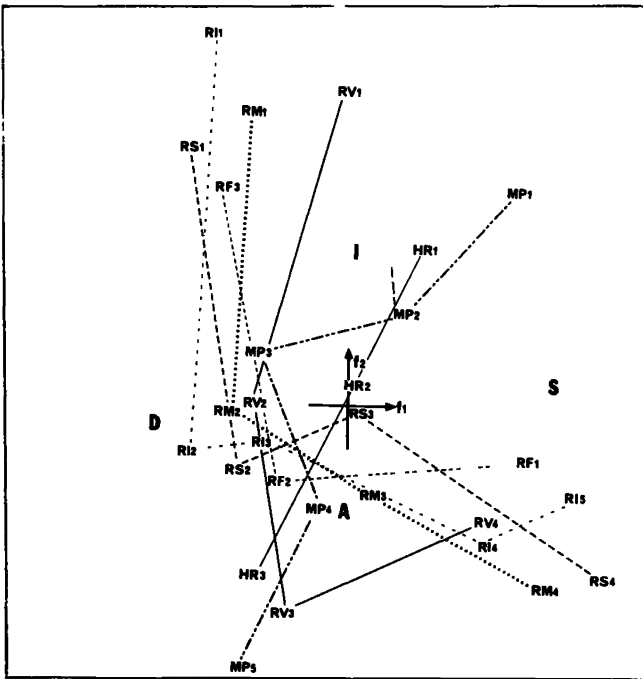


Figure 2 : Lignes polygonales les plus liées à l'axe F_2

3.3 Interprétation du plan factoriel (F_1, F_2)

La figure 3 représente la projection sur le plan (F_1, F_2) de l'ensemble des patients et des paramètres (parmi les 83) qui sont bien représentés sur ce plan. Les patients sont identifiés suivant la survie et la localisation de l'infarctus. L'examen de cette figure montre que les survivants et les décédés occupent deux zones différentes. Ces zones sont mieux séparées à la partie inférieure de la figure où prédominent les IMA. Au contraire, à la partie supérieure où les IMP sont les plus nombreux, ces deux zones se recouvrent nettement. La prédominance des IMP, en haut, et des IMA, en bas, est statistiquement significative. En délimitant les 2 zones SURVIE et DECES par les pointillés on obtient:

dans la zone SURVIE, 43 survivants et 2 décédés (soit $43/50 = 86\%$ de survivants bien classés)

dans la zone décès, 42 survivants et 2 décédés (soit $42/51 = 82\%$ de décédés bien classés)

dans l'intersection des 2 zones, 5 survivants et 7 décédés (soit $12/101 = 12\%$ de mal classés au total)

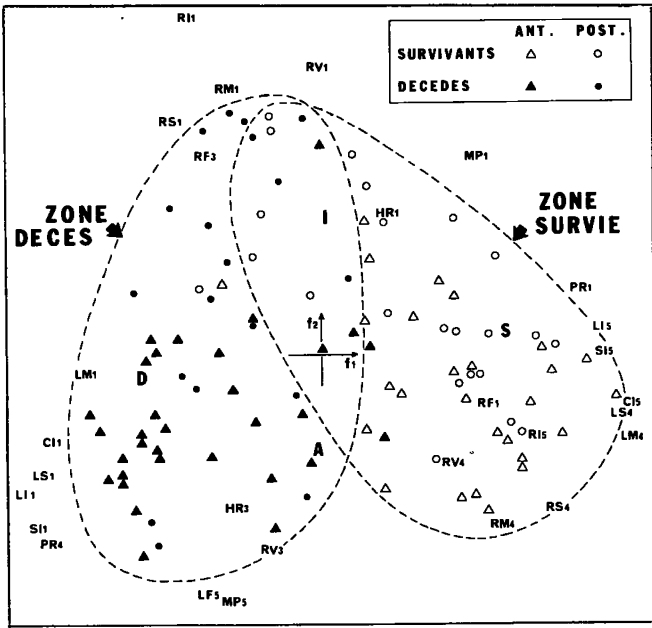


Figure 3 : Projection sur (F_1, F_2) des patients et des paramètres (parmi les 83) bien représentés sur ce plan

3.4 Interaction entre localisation et pronostic de l'issue

Afin de mettre en évidence une éventuelle interaction entre la localisation (apparue sur l'axe 2) et la discrimination entre survivants et décédés faite sur l'axe 1 on a introduit en éléments supplémentaires dans le plan (F_1, F_2), quatre nouvelles variables SA, SI, DA, DI qui sont les combinaisons des variables supplémentaires précédentes A (infarctus antérieur), I (infarctus postérieur), D (décès) et S (survie). On a donc posé :

SA = 1 si S = 1 et A = 1 (survivant ayant un infarctus antérieur)

SA = 0 sinon

SI = 1 si S = 1 et I = 1 (survivant ayant un infarctus postérieur)

SI = 0 sinon

DA = 1 si D = 1 et A = 1 (décédé ayant un infarctus antérieur)

DA = 0 sinon

DI = 1 si D = 1 et I = 1 (décédé ayant un infarctus postérieur)

DI = 0 sinon

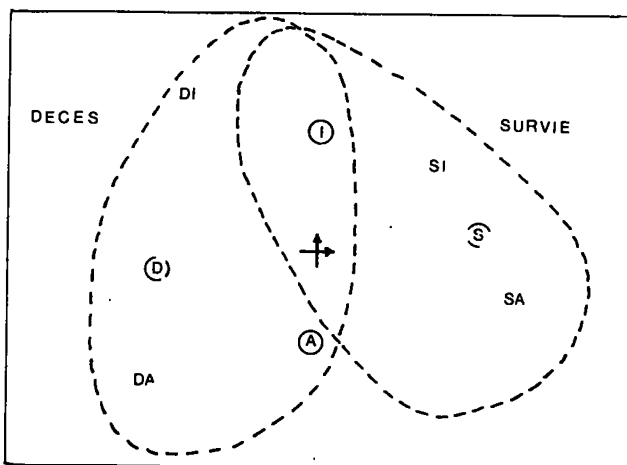


Figure 4 : Projection des 8 variables supplémentaires sur (F_1, F_2)

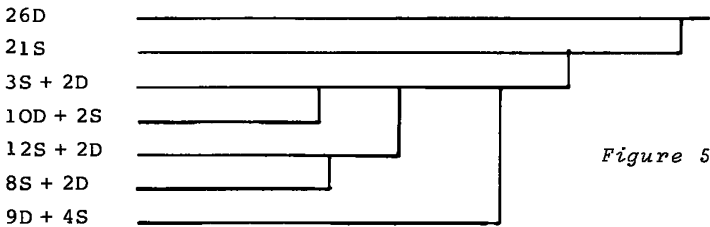
La figure 4 représente la position des 8 variables supplémentaires dans le plan 1×2 , on voit une interaction entre localisation et discrimination. La discrimination entre décédés et survivants est plus nette pour les patients ayant eu un infarctus antérieur que pour ceux ayant eu un infarctus postérieur.

Note : Infarctus postérieur, correspond on le voit à la partie positive de l'axe 2, c'est-à-dire à une déficience particulièrement marquée de la fonction droite dont on a expliqué plus haut le rôle. Du point de vue anatomique la correspondance entre le degré d'altération de la fonction droite et la localisation de l'infarctus s'explique par la vascularisation coronaire.

L'infarctus postérieur correspond, en général à une atteinte de la coronaire droite qui irrigue préférentiellement le ventricule droit et la paroi postérieure du cœur, tandis que l'infarctus antérieur est en général la conséquence d'une obstruction de la coronaire gauche qui assure la distribution sanguine essentiellement au ventricule gauche et n'apporte pratiquement rien au ventricule droit.

4. Classification et discrimination : Sur le plan 1×2 issu de l'analyse du tableau de correspondance 101×83 , l'ensemble des survivants (S) et celui des décédés (D) se séparent assez bien, mais empiètent suivant une zone comprenant environ 12 individus. En bref les diverses méthodes de discrimination qu'on a essayées, placent une cloison dans cette zone, séparant ainsi les individus en deux classes (S') et (D') qui coïncident avec (S) et (D) à quelques erreurs près ; lesquelles, plus ou moins nombreuses suivant les méthodes affectent presque exclusivement l'un ou l'autre des cas douteux qu'a montrés l'analyse factorielle. Le tableau 6 illustrant ce § 4 donne pour chacun des individus l'issue réelle ($S = 0$; $D = \bullet$) et l'affectation proposée suivant 5 méthodes différentes. De plus, on reproduit le plan 1×2 de l'analyse factorielle, (Figure 8) où sont identifiés par leur numéro et leur issue réelle les sujets mal classés par au moins une méthode. Après ce bilan global, énumérons les diverses méthodes appliquées.

4.1 Classification ascendante hiérarchique : Les dossiers individuels ont été classés, par agrégation suivant la variance en prenant pour distance euclidienne celle du χ^2 . Voici les classes les plus hautes de la hiérarchie obtenue, résumées par leur composition en S et D :



on a deux classes pures (26D, 21S) totalisant environ la moitié des sujets ; trois classes où l'une des issues prédomine nettement (10 + 2 ; 12 + 2 ; 8 + 2) ; deux mixtes (3 + 2 ; 9 + 4). Etendue à un plus grand nombre de cas, une telle classification pourrait fournir une typologie affinée de l'infarctus aigu ; mais elle n'a pas paru immédiatement utilisable pour discriminer parmi les patients nouveaux ceux pour lesquels on doit tenter une contre-pulsion par ballon (cf §§ 1 et 5). Nous ne donnerons donc pas le détail de la classification ascendante hiérarchique.

4.2 Agrégation autour des centres variables : On applique la méthode des nuées dynamiques de E. Diday en demandant une fois deux classes ; une autre 4 classes. (Pour un exposé de la méthode, cf E. Diday [5] ; pour une brève esquisse cf e.g. dans ces *Cahiers* Vol 2 n° 2 [8])

4.2.1 Agrégation en deux classes : Les deux classes S' et D' obtenues s'accordent bien avec les classes S et D définies par les issues réelles : il n'y a que 18 dossiers mal classés ; soit un pourcentage de bien classés dépassant 82%.

4.2.2 Agrégation en quatre classes : L'une des classes ne contient que des décédés ; une autre ne comprend que des survivants à une exception près ; il y a deux classes mixtes (20S + 6D) et (14D + 10S) ; en attribuant à chaque cas l'issue qui prédomine dans sa classe, on a 17 erreurs de pronostic, ce qui diffère peu du § 3.2.1.

4.3 Méthode des plus proches voisins : On détermine pour chaque individu, une boule centrée au point qui le représente et contenant les k plus proches voisins : l'affectation de l'individu à une issue se fait en suivant la majorité de ces k voisins. La méthode peut être appliquée avec différentes formules de distance, le nombre k étant e.g. 5, 10, ou 20 selon l'effectif de l'échantillon disponible. Ici on a utilisé uniquement la distance du χ^2 (bien qu'il soit souvent avantageux de substituer à celle-ci une expression approchée calculée d'après les facteurs :

$$d^2(i, i') \approx \sum \{ (F_\alpha(i) - F_\alpha(i'))^2 \mid \alpha = 1, \dots, r \},$$

avec r variant e.g. de 2 à 6 ; pour un programme avec sa notice : cf dans ce cahier, M. O. Lebeaux [6] ; pour un exemple d'application, cf J.C. Turlot [9]. On a essayé pour k les deux valeurs 5 et 10.

4.3.1 Affectation d'après 5 voisins : Il n'y a que 16 individus mal classés.

4.3.2 Affectation d'après 10 voisins : Il y a 12 individus mal classés ; de plus 5 individus ne sont pas classés, car ils comportent 5 voisins S et 5 voisins D. Notons au passage que l'affectation majoritaire peut se trouver en défaut parce que $k = 10$ est un nombre pair.

4.4 Hyperplan discriminant : De nombreux auteurs ont abouti par des voies diverses à cette méthode dont nous donnerons ici une explication géométrique succincte. Le nuage I des individus est représenté dans un espace vectoriel : on munit cet espace de la métrique d'inertie du nuage, c'est-à-dire de la métrique euclidienne pour laquelle le nuage a tous ses moments d'inertie égaux (à l'inertie d'une sphère) ; I étant partagé en deux classes (ici S et D) on prend pour cloison un hyperplan perpendiculaire au segment joignant les centres de gravité de ces classes. Pratiquement, l'espace de la représentation sera l'espace engendré par les premiers axes factoriels (e.g. 2, 3 ou 4 axes) ; pour rendre le nuage sphérique il suffit d'amener les coordonnées à avoir même variance ; i.e. de substituer aux $F_\alpha(i)$ les $\lambda_\alpha^{-1/2} F_\alpha(i) = \varphi_\alpha(i)$ d'où la formule de distance :

$$d^2(i, i') = \sum \{ \lambda_\alpha^{-1} (F_\alpha(i) - F_\alpha(i'))^2 \mid \alpha = 1, \dots, r \}.$$

Dans la présente étude, on s'est borné à placer une cloison rectiligne dans le plan des axes 1 et 2 ; le taux d'erreur est le plus faible enregistré jusqu'ici : soit 11% (Figure 6)

Remarque : Pour faciliter au médecin l'usage de l'analyse discriminante (§5), on s'est borné ici à placer une cloison dans un plan (sans rechercher une meilleure discrimination dans un espace de dimension plus élevée) : pour cela, les calculs ne sont pas indispensables : il suffit de tâtonner en déplaçant une règle dans le plan jusqu'à obtenir une partition en deux classes S' et D' s'accordant le mieux avec les classes S et D.

4.5 Affectation barycentrique : Plaçons sur un axe deux points S et D figurant, les deux issues ; puis chacune des modalités en un point j du segment SD tel que $|jS|/|jD| = k(D, j)/k(S, j)$ (où l'on a noté $k(D, j)$ le nombre des décédés possédant la modalité j ; et de même $k(S, j)$ pour les survivants) ; puis enfin chaque individu au barycentre (point moyen) des modalités j qu'il possède. La représentation unidimensionnelle ainsi obtenue au prix de calculs minimales aboutit à une discrimination à peu près équivalente à celle du § 4.4 : dix erreurs seulement comme l'atteste l'histogramme ci-dessous. (Figure 7)

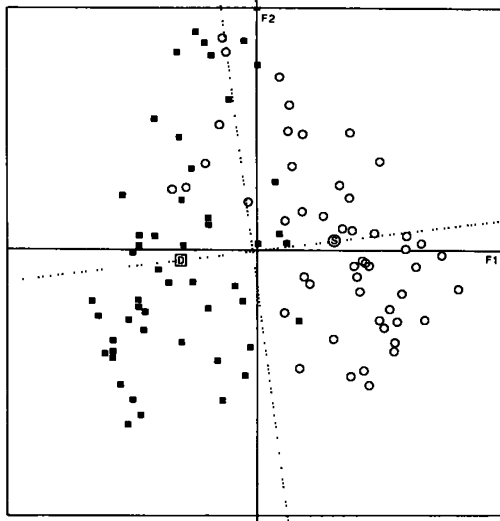


Figure 6 : Cloison rectiligne dans le plan (F_1, F_2) où le nuage a été rendu sphérique

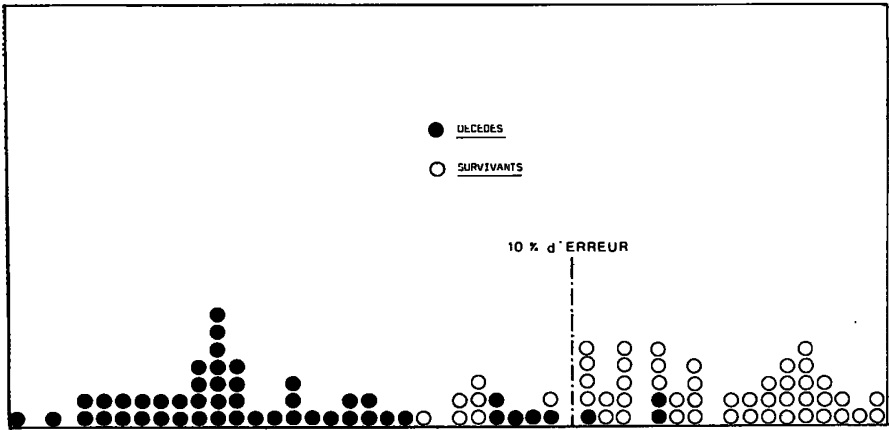


Figure 7 : Histogramme de la représentation barycentrique (Echantillon de base)

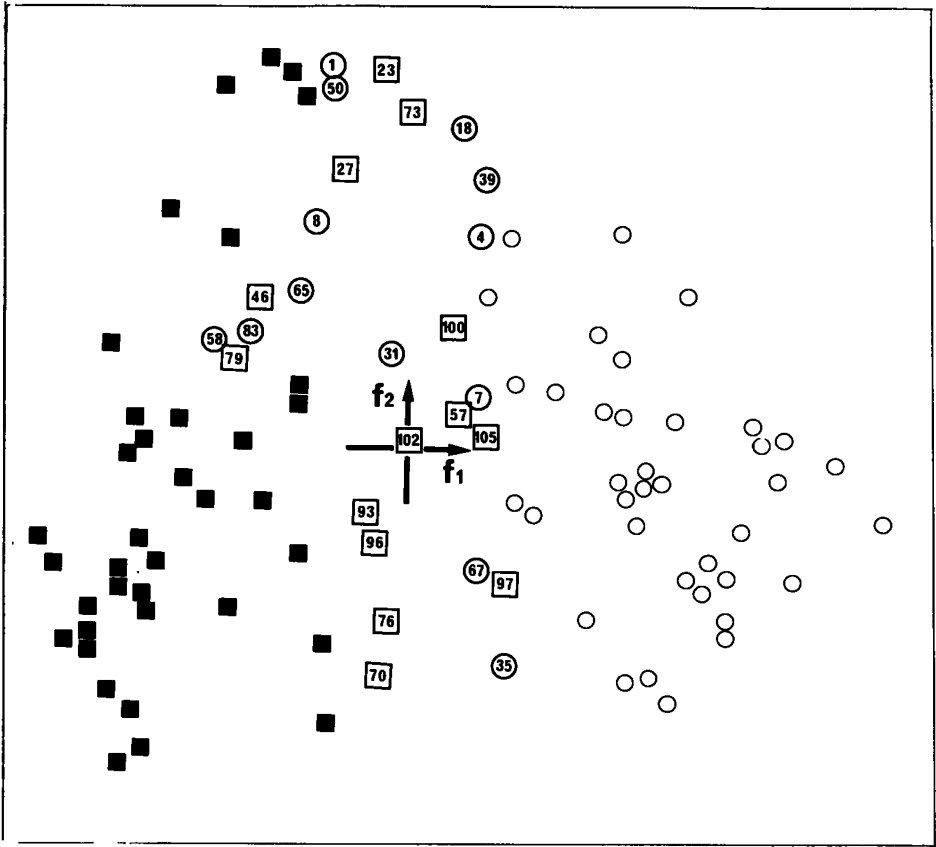


Figure 8 : Identification, par leur numéro et leur issue réelle (■ décédé et O survivant) des patients mal classés par au moins une méthode.

Les erreurs de classement de chacun de ces individus sont fournies dans le tableau 6.

Les patients bien classés par toutes les méthodes sont identifiés par un ■ pour les décédés et un O pour les survivants.

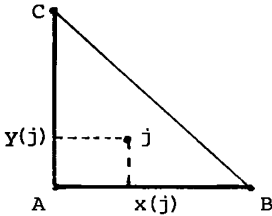
N° DU PATIENT	1	4	7	8	18	23	27	31	35	39	46	50	57	58	65	67	70	73	76	79	83	93	96	97	100	102	105
ISSUE REELLE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.2.1.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.2.2.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.3.1.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.3.2.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.4.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ISSUE 4.5.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Tableau 6 : Résultats des méthodes de discrimination ; on a noté 0 : décédé ; 0 : survivant ; les résultats des 38 sujets 0 et deux des 36 sujets 0 toujours bien classés n'ont pas été reportés ici : seules sont recensées les erreurs de discrimination situées dans le plan 1 x 2 ci-dessus issu de l'analyse des correspondances

- 4.2. Méthode des nuées dynamiques (2 classes ou 4 classes)
- 4.3. Méthode des plus proches voisins (5 voisins ou 10 voisins)
- 4.4. Hyperplan discriminant
- 4.5. Méthode d'affectation barycentrique

En fait, comme il est expliqué ailleurs dans ce cahier (cf § 3.3 p.38) cette représentation barycentrique est une analyse des correspondances : le tableau principal (2 x 83) à 2 lignes S et D (ligne des $k(S, j)$ et ligne des $k(D, j)$) ne fournit qu'un axe, l'axe SD, sur lequel chaque individu i est projeté en élément supplémentaire d'après la ligne des $k(i, j)$. Les calculs barycentriques effectués coïncident au facteur $\lambda^{-1/2}$ avec la formule de transition de l'analyse des correspondances.

Dans le cas d'une discrimination entre trois classes (e.g. A, B, C) on peut encore sans ordinateur, construire une représentation plane : il suffit de poser



$$x(j) = k(B, j) / (k(A, j) + k(B, j) + k(C, j))$$

$$y(j) = k(C, j) / (k(A, j) + k(B, j) + k(C, j))$$

$$x(i) = \sum \{x(j) k(i, j) / k(i) \mid j \in J\};$$

$$y(i) = \sum \{y(j) k(i, j) / k(i) \mid j \in J\};$$

où comme plus haut, $k(X, j) = \sum \{k(i, j) \mid i \in X\}$; i.e. dans le cas de données en 0 et 1, $k(X, j)$ est le nombre des individus de la classe X possédant le caractère j : les calculs sont le plus simple pour un tableau booléen mais valent pour tout tableau.

Au-delà de 3 classes, l'idée de placer les individus en éléments supplémentaires sur l'analyse du tableau décrivant les classes reste utile : mais sa mise en oeuvre requiert un ordinateur !

5 Pronostic et décision thérapeutique

5.1 Choix d'une méthode de discrimination : Rappelons le problème médical posé au § 1 : devant un cas nouveau d'infarctus aigu compliqué de défaillance cardiaque, reconnaître rapidement si le patient n'a que de faibles chances de survivre plus de quatre semaines ; et dans ce cas tenter de procurer au patient une assistance circulatoire par ballon intra-aortique. Il s'agit d'une décision thérapeutique fondée sur un pronostic, c'est-à-dire en terme statistique, sur une discrimination. Après l'analyse du § 3, on a essayé au § 4 diverses méthodes de discrimination : d'une part ces méthodes ont eu des taux de réussite inégaux, d'autre part, toutes ne sont pas susceptibles de fournir facilement un pronostic pour un cas nouveau. C'est pourquoi on retiendra les traitements mathématiques suivants :

a) Adjoindre le cas nouveau comme élément supplémentaire s aux résultats de l'analyse des 101 dossiers de base : le calcul des facteurs $F_\alpha(s)$ se faisant sans ordinateur par la formule usuelle :

$$F_\alpha(s) = (\lambda_\alpha^{-1/2} / k(s)) \sum \{k(s, j) G_\alpha(j) \mid j \in J\},$$

dont l'application est particulièrement facile car les $k(s, j)$ valent 0 ou 1 (1 si s possède la modalité j ; et 0 sinon) et de plus $k(s)$ vaut 22 (nombre des questions ; donc des modalités que s possède, en sorte que l'on a qu'à effectuer une addition de 22 nombres puis une multiplication.

b) Considérer dans le plan 1×2 le point $s = (F_1(s), F_2(s))$ parmi le nuage des 101 points de base ; et plus précisément placer s par rapport à la cloison mise en place au § 3.4. Il ne serait pas difficile de placer s e.g. par rapport à une cloison dans l'espace des 3 premiers axes, d'après le signe d'une combinaison linéaire $A_1 F_1(s) + A_2 F_2(s) + A_3 F_3(s)$

(les coefficients A_1, A_2, A_3 définissent la direction de la cloison) : mais puisqu'il s'agit d'aider un médecin à prendre une décision, il est préférable que celui-ci accède aussi simplement que possible, à une vue directe de l'ensemble des informations.

c) Placer simplement le cas nouveau par affectation barycentrique sur l'axe SD obtenu au § 4.5 : cet axe portant l'histogramme des 101 dossiers de base le cas nouveau y est situé par rapport à l'ensemble des 101 cas antérieurs.

d) Pour raisonner par analogie comme le font d'ordinaire les médecins, il est souhaitable de confronter le cas nouveau à ses 5 ou 10 plus proches voisins parmi les dossiers de base : la recherche des plus proches voisins suivant la distance du χ^2 (ou dans l'espace e.g. des 5 premiers facteurs) est fort simple, mais requiert un ordinateur ; cependant la recherche des plus proches voisins dans le plan 1×2 se fait sans calcul et suffit pratiquement. Eventuellement, on tiendra compte du facteur F_3 pour choisir parmi les voisins de s dans (F_1, F_2) ceux qui en sont le plus proches dans l'espace des 3 premiers axes ; mais l'examen direct des dossiers complets de quelques patients intéressera sans doute davantage le cardiologue, que toute recherche mathématique du point le plus proche.

5.2 Application à 55 cas nouveaux : Pour décider de la thérapeutique, les médecins ont simplement considéré la place du cas nouveau dans le plan 1×2 . La figure résume les résultats de 55 cas :

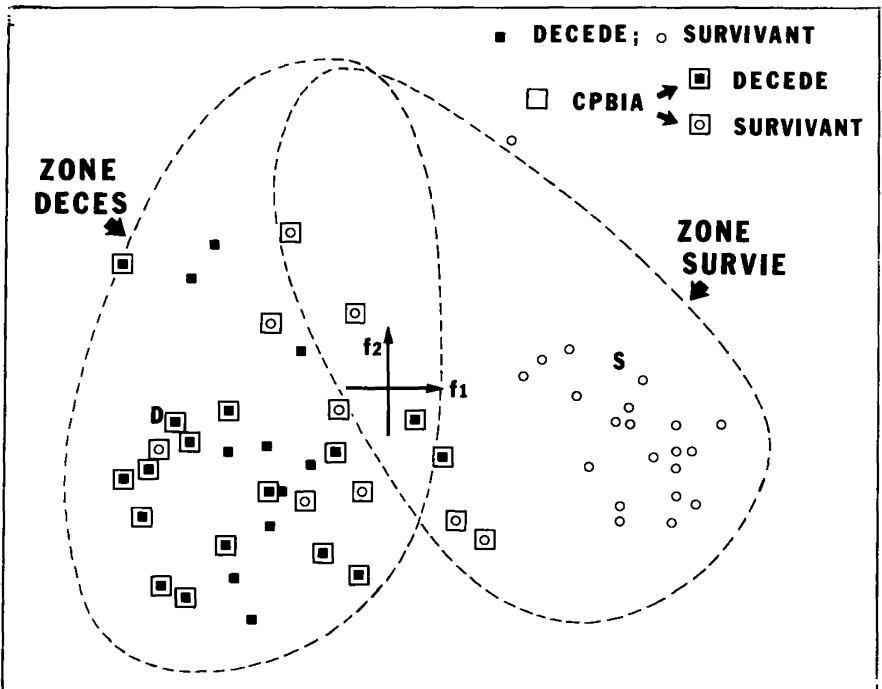


Figure 9 : Décision et issue thérapeutique à la 4^o semaine des 55 cas nouveaux sur le plan 1×2 de l'AFC des 101 dossiers de l'échantillon de base.

21 de ceux-ci se situent franchement dans la zone de survie : ils reçoivent le traitement médical classique et tous survivent (sigle O sur le graphe)

36 cas tombent dans la zone de décès ou à proximité de celle-ci ; une indication d'assistance circulatoire par contre-pulsion intra-aortique est alors posée ; et 25 patients peuvent en bénéficier. Parmi ceux-ci, 9 passent le cap de la phase aiguë et sont donc considérés comme survivants (sigle \square) sur le graphe) ; les 16 autres décèdent avant 4 semaines (sigle \blacksquare). Décèdent également les 10 sujets qui n'ont pu bénéficier de l'assistance circulatoire indiquée dans leur cas (sigle \blacksquare)

A posteriori on peut également considérer ce qu'aurait apporté à la décision thérapeutique la méthode d'affectation barycentrique (§ 4.5). Comme l'examen du plan 1 x 2, cette méthode (laquelle, rappelons-le, ne requiert aucun calcul sur ordinateur, non seulement pour les cas nouveaux, mais aussi pour le traitement des données de base) permet de distinguer les patients pour lesquels le pronostic de survie à court terme est certain, de ceux qui sans assistance circulatoire mécanique n'ont aucune chance de doubler le cap (ou en ont très peu) et doivent donc, dans la mesure du possible bénéficier de la contre-pulsion intra-aortique.

5.3 Conclusions

Quant à la thérapeutique : l'efficacité de la contre-pulsion par ballon intra-aortique (dans le traitement de la défaillance cardiaque à la phase aiguë de l'infarctus du myocarde, est confirmée et précisée. Il importe toutefois de répéter que nos statistiques ne concernent que la phase aiguë. L'évolution ultérieure est grevée d'une mortalité importante : sur les 9 patients ayant passé le cap des 4 semaines grâce à une assistance circulatoire, 5 décèdent secondairement dans un délai de un à plusieurs mois. Cette mortalité différée paraît due à un état général très altéré du réseau coronarien, au-dessus de tout recours chirurgical.

Quant aux méthodes multidimensionnelles : elles ont aidé le praticien dans le pronostic et la décision thérapeutique, en donnant des multiples informations recueillies ordinairement, une présentation géométrique simple se prêtant bien à la comparaison d'un cas nouveau aux cas antérieurs ; et cela moyennant des calculs faciles (toujours par l'insertion d'un cas nouveau : cf § 5.1 a,b,d : et même avec l'affectation barycentrique - § 5.1 c - pour une saisie globale de 100 dossiers)

BIBLIOGRAPHIE

- [1] DUNKMAN W.B., LEINBACH R.C. and al.
Clinical and hemodynamic results of intra-aortic balloon pumping and surgery for cardiogenic shock. CIRCULATION 46 : 465-477, 1972
- [2] BARDET J., MASQUET C., and al.
Clinical and hemodynamic results of intra-aortic balloon counterpulsation and surgery for cardiogenic shock. American Heart Journal : 93 : 280-290, 1977
- [3] LEBART L.
L'orientation du dépouillement de certaines enquêtes par l'analyse des correspondances multiples. Consommation - Annales du CREDOC, n° 2, 1975, DUNOD ed.
- [4] LEBART L.
On the BENZECRI'S method for finding eigenvector by stochastic approximation. Proceeding on COMPSTAT. Physica Verlag, Wien 202-211
- [5] DIDAY E.
Une nouvelle méthode de classification automatique et reconnaissance des formes : la méthode des nuées dynamiques. Revue de Statistique Appliquée, 1970, XIX, n° 2
- [6] LEBEAUX M.O.
Notice sur l'utilisation du programme [POUBEL] . Ce cahier pp 467-481
- [7] BENZECRI J.P.
L'Analyse des Données : Analyse des correspondances, tome II, DUNOD 1973, 2° édition 1976
- [8] BENZECRI J.P.
Actualité : l'analyse des données en archéologie préhistorique [PREHISTOIRE]. Cahiers de l'analyse des données pp 193-214
- [9] TURLOT J.C.
Sur la taxinomie du genre *Equus*. Cahiers de l'analyse des données (à paraître)