

Diagnostic sur profils protéiques inflammatoires. Approche linguistique par sous-ensembles flous

R. Bartolin, V. Bonniol, G. Bouvenot, C. Arnaud, E. Sanchez

Service de Médecine Interne, Hôtel-Dieu, 13224 Marseille Cedex 1, France. IRTHM, Hôtel-Dieu, 13224 Marseille Cedex 1, France. INSERM U.21, Hôpital P.-Brousse, 94807 Villejuif, France. Laboratoire Central, Hôtel-Dieu, 13224 Marseille Cedex 1, France. Service de Biomathématiques, Faculté de Médecine, 13385 Marseille Cedex 5, France

RESUME

Un système original de classification linguistique sur des profils protéiques inflammatoires a été mis au point. Des expressions du type : "légèrement augmenté ou augmenté", qualifiant des variations protéiques, ont été représentées par des sous-ensembles flous. Des méthodes de confrontation entre le modèle et les données du patient permettent d'apprécier le degré de compatibilité entre le profil biologique de celui-ci et l'éventail des syndromes étudiés. Le système propose au biologiste, pour chaque bilan, un diagnostic final constitué de un ou plusieurs syndromes ordonnés.

MOTS CLEFS

Classification linguistique, protéines de l'inflammation, aide au diagnostic, sous ensembles flous

INTRODUCTION

Les syndromes inflammatoires biologiques demeurent toujours un sujet important de réflexion en médecine interne à cause de leur fréquence et de leurs implications diagnostiques et thérapeutiques. Suite aux recherches effectuées par P.GIRAUDET et la Société Française de Biologie Clinique, le concept de profil protéique inflammatoire est actuellement bien admis. Les méthodes de classification étiopathogénique, fonction des taux plasmatiques des protéines, présentent cependant quelques particularités : un patient peut répondre simultanément à plusieurs états et il peut donc être difficile de lui attribuer un diagnostic sans hésiter, le modèle décisionnel n'étant habituellement pas de nature probabiliste. Les raisonnements automatiques issus de la théorie des sous-ensembles flous et les approches linguistiques semblent mieux adaptés à la résolution de ces problèmes.

MATERIELS

Echantillon

Il comporte 163 adultes des deux sexes, hospitalisés au CHU de Marseille durant le premier semestre 1987, chez lesquels les prélèvements ont été effectués le matin à jeun et à distance de toute thérapeutique anti-inflammatoire ou à tropisme hépatorenal. Il est composé de trois groupes: 130 sujets inflammatoires (VS > 40 mm) atteints d'affections bien définies (critères habituels), 11 sujets non inflammatoires (VS < 10 mm) sièges de pathologies frontières perturbant le profil protéique et 22 sujets normaux (VS < 10 mm et absence de toute pathologie organique).

Biochimie

Les dosages protéiques ont été effectués sur sérum par néphélométrie cinétique (BNA Behring), au sein du laboratoire central de l'Hotel Dieu de Marseille.

Informatique

Le logiciel, utilisable en conversationnel, a été réalisé sur MINI 6 Bull dans le service informatique de la faculté de Médecine de Marseille. Il comporte 40 sous-programmes écrites en FORTRAN 77, soit environ 3500 lignes d'instructions.

METHODES

Méthodologie propre à l'application

Le modèle de référence comprend 11 groupes de syndromes : le groupe normal, 2 groupes de maladies infectieuses (Bactérienne et Virale), 2 groupes frontières de syndromes non inflammatoires (Hémolyse intra-veineuse pure et Insuffisance rénale glomérulaire) et 6 groupes de syndromes inflammatoires (Vascularite, Syndrome néphrotique, Glomérulo-néphrite aiguë, Hémolyse avec inflammation, Connectivites, Lupus individualisé). Les protéines dosées sont au nombre de 5 : Fraction C3 du complément (FC3), α -1-Anti-Trypsine (A1AT), Orosomucoïde (Om), Haptoglobine (Hp) et Protéine C Réactive (PCR). Selon l'usage [Frot et coll., 1984], les valeurs des taux protéiques sont exprimés de façon relative au taux de référence de notre laboratoire. L'influence du sexe et de l'âge est prise en compte dans le quotient de normalisation calculé à partir d'une régression polynomiale [Bartolin, 1987].

Méthodologie Mathématique

La théorie des Sous-Ensembles Flous (S.E.F.) permet de pondérer par degrés la notion d'appartenance ou de compatibilité [Zadeh, 1965] et ainsi de traiter automatiquement par ordinateur des cas frontières ainsi que les nombreuses situations dans lesquelles les seuils sont trop rigides et n'ont pas de réalité physiopathologique. Les résultats numériques des dosages protéiques sont transformés en nombres flous; de même (fig.1) les concepts linguistiques traduisant la connaissance médicale sont représentés par des S.E.F. [Bartolin et coll., submitted]. La comparaison modèle-patient est effectuée par des mesures de compatibilité [Bartolin et coll., 1988] et l'ensemble, non flou, des diagnostics est obtenu par application de la fonction "pouvoir séparateur" [Dujet et coll., 1989].

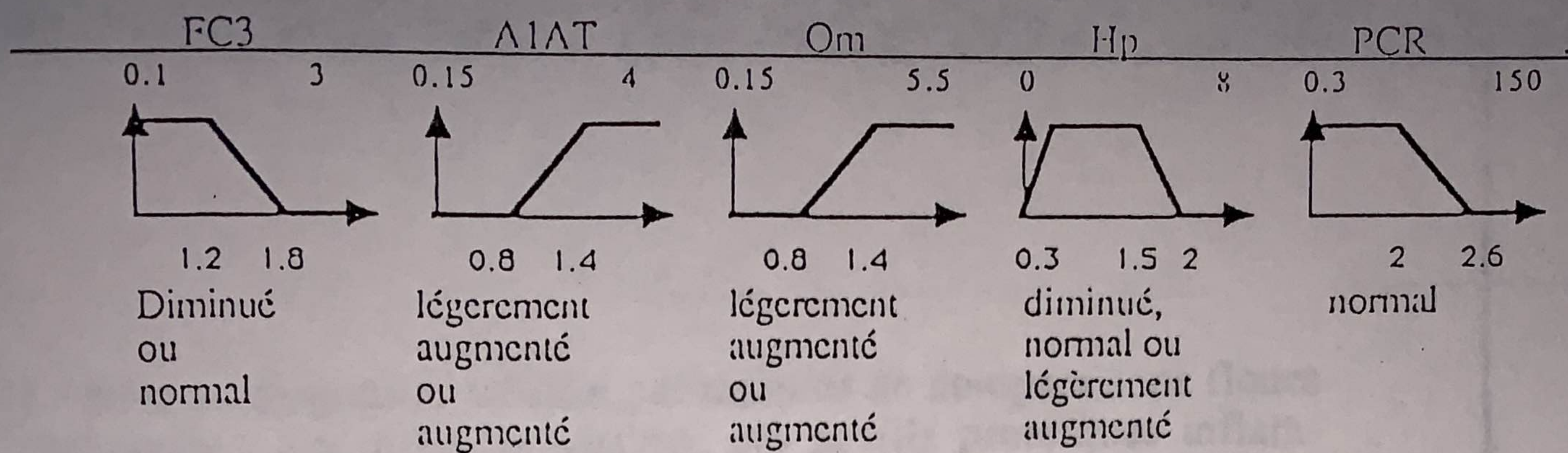


fig. 1 : Caractérisation du syndrome Lupus
Concepts linguistiques et traductions en S.E.F.

RESULTATS

Le système permet d'attribuer un ou plusieurs diagnostics (pour les formes mixtes) à chaque bilan biologique. Ces diagnostics, classés par ordre d'attribution préférentiel permettent aux biologistes et aux cliniciens d'évoquer l'étiologie correspondant aux syndromes sous-jacents tout en leur laissant la possibilité de recourir à des informations cliniques ou biologiques complémentaires pour confirmer l'hypothèse retenue. 56% de bonnes affectations ont été obtenues dont 39% de patients classés dans leur groupe d'origine et 17% en affectation mixte comprenant leur groupe d'origine. 6% des résultats ont été discordants tandis que 38% de sujets intermédiaires ont été mal affectés, ce qui peut s'expliquer par la physiopathogénie : le système effectue une classification en syndromes (définis par le profil des 5 protéines) et non pas en étiologies.

DISCUSSION

L'absence de spécificité des protéines retenues pour les syndromes étudiés fut un écueil lors de la modélisation, ce qui explique, en partie, la concordance relative obtenue. L'importance du taux d'affectations mixtes est un indicateur du chevauchement des sous ensembles flous. Enfin, en cas de mauvaise affectation, il n'y a pas toujours contradiction entre le résultat du classement et le diagnostic clinique initial : à la même maladie peuvent correspondre plusieurs profils protéiques (donc plusieurs syndromes différents) en fonction du stade évolutif ou de la survenue de complication (Lupus surinfecté : augmentation de la PCR et donc classement en Infection Bactérienne). Pour pallier ces difficultés, plusieurs solutions ont été envisagées : la pondération des sous ensembles flous [Bartolin et coll, 1988] et l'étude de nouveaux opérateurs [Sanchez et coll, 1988].

CONCLUSION

Un système de classification automatique basé sur la reconnaissance des profils protéiques et issu d'une approche linguistique présente un intérêt certain, à la fois pour le biologiste et pour le clinicien ; la méthodologie mathématique que nous avons utilisée permet une grande souplesse de modélisation et l'expression de résultats pondérés ou mixtes plus proche de la réalité physiopathogénique.

REFERENCES

- Bartolin R. (1987) : Aide au diagnostic médical par mesures de comparaisons floues et pouvoir séparateur : Approche linguistique des profils protéiques inflammatoires biologiques. Thèse d'Etat en biologie humaine, Marseille, France.
- Bartolin R., Bonniol V. and Sanchez E. (1988) : Inflammatory Protein Variations : Medical Knowledge Representation and Approximate Reasoning. In Uncertainty and Intelligent Systems, eds Bouchon B., Saitta L. and Yager R.R., Springer-Verlag, Berlin-Heidelberg, pp 306-313.
- Bartolin R., Bonniol V. and Sanchez E. : Protein Variations and Inflammatory Syndromes : A linguistic approach. Submitted to Methods of Information in Medicine.
- Dujet C., Sanchez E. and Bartolin R. (to appear in 1989) : Fuzzy sets and partitions. A new approach applied to medical decision assistance : inflammatory syndromes. Int. J. for A.I. in Medicine.
- Frot J.C., Hofmann H., Muller F., Benazet M.F. et Giraudet P. (1984) : Le concept de profil protéique. Ann. Biol. Clin. 42, 1-8.
- Sanchez E., Bartolin R. and Bonniol V. (1988) : Aggregation procedures in a biomedical problem on inflammatory syndromes. Int. Workshop on fuzzy systems applications Iisuka and Fukuoka (Japan), august 22-24, proceedings, pp 75-76.
- Zadeh L.A. (1965) : Fuzzy sets. Information and control 8, 338-353.

SUMMARY

Inflammatory protein variations are studied in a linguistic model in which the expressions are treated as labels of fuzzy sets. Protein values, measured over patients, are matched to the corresponding fuzzy sets by means of compatibility measures. Non fuzzy diagnostic groups are finally assigned to patients, after aggregation and separation.