

F. TEKAÏA

Analyse statistique et recherche des facteurs pronostiqués dans la pathogénèse du sida

Les cahiers de l'analyse des données, tome 15, n° 3 (1990),
p. 257-260

http://www.numdam.org/item?id=CAD_1990__15_3_257_0

© Les cahiers de l'analyse des données, Dunod, 1990, 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/>

ANALYSE STATISTIQUE ET RECHERCHE DES FACTEURS PRONOSTIQUES DANS LA PATHOGÉNÈSE DU SIDA

[RECH. PRON. SIDA]

*F. TEKAÏA**

Le mercredi 20 Juin 1990 se tint, dans la Salle Chamberland de l'Institut PASTEUR, une réunion de réflexion consacrée à la Recherche de facteurs pronostiques dans l'évolution du SIDA. Une notable partie des articles du présent cahier a pris forme au cours de cette réunion pluri-disciplinaire. La note qui suit reprend, en le complétant, l'exposé d'introduction adressé par F. Tekaïa aux participants. Le lecteur en usera comme d'un fil d'Ariane pour suivre à l'aise le labyrinthe du cahier!

1 Étapes d'une collaboration

On trouve dans *CAD*, Vol. XIII, n°4, (1988), un article, déjà consacré au thème de notre réunion, et signé par cinq pastoriens:

Intérêt de l'antigénémie dans la définition des stades de l'infection par le virus VIH, [STADES INF. VIH], par F. Tekaïa, J. de Saint Martin, Ph. Sansonetti, F. Vuillier et J.M. Claverie.

Comme tous les travaux relatifs au SIDA que nous évoquerons ici, cet article est fondé sur la base de données CONSULT, laquelle donne accès aux dossiers des malades séropositifs suivis au service de consultation de l'Hôpital de l'Institut Pasteur (cf. §2). Au printemps de 1990, a été soumis pour publication dans *CAD* un deuxième travail:

Estimation du stade de l'infection par le VIH chez les sujets séropositifs, [STADES VIH], par F. Tekaïa, Ph. Sansonetti et J.M. Claverie.

Ce travail repose sur une suite d'analyses effectuées sur un tableau croisant un ensemble de consultations avec les modalités de 13 variables biologiques et cliniques qu'il vaut la peine d'énumérer:

(*) Institut Pasteur, Unité d'informatique scientifique, 25 rue du Docteur Roux, 75724, Paris CEDEX 15.

{Nombre de lymphocytes, de cellules T4, de cellules T8, de plaquettes; taux de β 2-microglobuline, d'IgG, d'IgM, d'IgA; antigénémie-p25; symptômes (signes fonctionnels, neurologiques, infections...); les classifications et stades selon le CDC et l'HIP}.

On a considéré, dans certaines analyses, jusqu'à 2084 consultations relatives à 1000 patients.

Le Pr. Benzécri ayant prêté toute son attention à ce travail, a demandé des éclaircissements, suggéré des compléments, repris par lui-même les données... et rédigé une note communiquée aux participants avant la réunion:

État du système immunitaire et histoire clinique chez les patients infectés par le virus VIH, [IMM. CLIN. VIH], par J.-P. Benzécri.

Comme cette note est écrite selon la conception de l'analyse des données propre à son auteur, il s'imposait de faire, avant tout débat, un exposé d'introduction; et cet exposé a été rédigé sous le titre:

Analyse des données biologiques et pathologie clinique, [ANA. CLIN.], par J.-P. Benzécri.

Suivit une discussion, où des médecins et biologistes, travaillant dans le domaine du SIDA (cf. §3), firent des propositions de recherches pour poursuivre l'étude longitudinale de la cohorte de patients, en recueillant éventuellement de nouvelles données. La discussion ayant été enregistrée, J.-P. et F. Benzécri en ont fait un compte rendu; soumis à la vérification des participants, mais rédigé par des statisticiens, à l'usage de statisticiens et au niveau de compétence de ceux-ci.

Compte rendu d'une discussion entre informaticiens et spécialistes du SIDA, après un exposé d'analyse des données, [DISCUSSION PASTEUR], par J.-P et F. Benzécri.

Il importe de dire que les rédacteurs de ce compte rendu ont mis au point définitivement leur texte en recourant à une excellente publication pastoriennne:

VIH et SIDA, in ADIP, lettre de l'association pour le développement de l'Institut Pasteur, n° 31, Juin 1990; ouvrage collectif de 116 pages.

Parmi les projets issus du débat, l'un des plus faciles à réaliser était d'extraire de la base CONSULT un tableau de données concernant les malades pour lesquels on dispose d'une suite de plusieurs consultations, afin de suivre sur les axes factoriels le développement de la maladie. Ce tableau a fait l'objet d'analyses, qui ont aussi tenu compte d'hypothèses suggérées par la lecture du rapport collectif que nous venons de citer. Et les premières conclusions de ces analyses ont été introduites dans le §3.3 de [ANA. CLIN.].

2 La base de données des consultants du SIDA: CONSULT

Il importe de décrire la base de donnée CONSULT; qui seule permet au statisticien de percevoir le réel; et offre, en retour, au médecin, au biologiste, l'occasion d'un survol; suivi, espérons-le, d'une féconde synthèse. La base présente les données cliniques et biologiques des dossiers des patients séropositifs suivis au service de consultation de l'Hôpital de l'Institut Pasteur. Chaque dossier est constitué de huit modules dont chacun comprend des données homogènes, groupées en vue de définir un objectif médical.

2.1 "IDENT": identification

Ce module comprend le numéro de dossier du patient, son état civil et son adresse; ainsi que les coordonnées du médecin traitant et celles de l'Hôpital qui a envoyé le patient à l'Institut Pasteur.

2.2 "ANAMNÈSE"

Vise à préciser dans quelles conditions le patient a contracté le VIH. Le module comporte donc des rubriques relatives à divers facteurs de risque: transmission sexuelle; risques médicaux; risque géographique si le sujet a vécu dans une région de forte endémie du SIDA.

2.3 "CLINIQUE"

Caractérise le stade de l'infection par le VIH, de la primo-infection jusqu'au SIDA déclaré, selon la classification, établie par le *Center for Diseases Control* d'Atlanta (CDC) et subdivisée à l'Hôpital de l'Institut Pasteur (HIP). Comprend des rubriques sur le type d'infection VIH ou autres; sur les symptômes neurologiques ou fonctionnels; ainsi que les résultats des épreuves d'allergie cutanée à la tuberculine et à la candidine.

2.4 "BIOIMM": biologie & immunologie

Rassemble les paramètres hématologiques, biochimiques, sérologiques et microbiologiques évalués lors de chaque consultation.

2.5 "IMMUNO": immunologie

Donne l'état immunitaire du patient à chaque consultation, suivant les rubriques suivantes: sérologie (VIH ou Hépatite B); recherche d'antigène VIH circulant; bilan détaillé de l'ensemble des populations lymphocitaires.

2.6 "RADIO": radiologie

Enregistre, s'il y a lieu, le compte rendu d'un examen radiologique, ou autre examen d'imagerie médicale, effectué à une date donnée.

2.7 "PEXPLO": protocole d'exploration

Suit, par périodes, les moyens d'exploration auxquels le patient a été soumis au cours de l'infection par le VIH.

2.8 “PTHÉRAP”: protocole thérapeutique

Note les périodes pendant lesquelles le patient a bénéficié d'un traitement par un médicament expérimental.

On voit qu'à l'exception des deux premiers, “IDENT” et “ANAMNÈSE”, chaque module contient des données évolutives relevées suivant une suite de consultations repérées par un numéro et une date. C'est par les numéros de dossiers et ceux de consultations que les informations afférentes à un même patient sont reliées entre elles; la date d'une consultation et celles des examens biologiques correspondant n'étant généralement pas les mêmes.

La base CONSULT, qui comprend à ce jour les résultats de quelque 3600 consultations afférentes à 1200 patients, est gérée par le Système de Gestion de Base de Données Relationnelles (SGBDR) “ORACLE”. Utilisée quotidiennement pendant la consultation, la base est aussi particulièrement apte à servir à des analyses statistiques dont on attend beaucoup pour le progrès de la connaissance du SIDA. Les médecins souhaitent que soient développés des outils d'aide à la décision, notamment pour déterminer le stade de l'infection et proposer un pronostique.

3 Le collège des participants à la réunion

Nous donnons sous ce titre seulement le signalement des protagonistes du débat rapporté dans [DISCUSSION SIDA]. Nos remerciements vont aussi aux pastoriens qui, en diverses circonstances, ont répondu à nos questions ou nous ont aidé de leurs suggestions; notamment à Fr. VVUILLIER et à G. DIGHIRO.

Jean-Paul BENZÉCRI: Professeur de statistique à l'Université Pierre et Marie CURIE.

Jean Michel CLAVERIE: Immunologiste, chef de l'Unité d'Informatique Scientifique à l'Institut Pasteur.

Yves RIVIÈRE: Chef de laboratoire à l'Unité d'Oncologie Virale.

Philippe SANSONETTI: Chef du Service des consultations à l'Hôpital de l'Institut Pasteur.

Fredj TEKAÏA: Statisticien à l'Unité d'Informatique Scientifique de l'Institut Pasteur; responsable de l'Analyse des Données et du développement des Bases de Données Biomédicales Relationnelles.

Jean-Louis VIRELIZIER: Chef de l'Unité d'Immunologie Virale à l'Institut Pasteur.

Cruz AYERBE: Médecin Stagiaire à l'Institut Pasteur.

Josè ALCAMI: Médecin Stagiaire à l'Institut Pasteur.