JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

D. SCHWARTZ

Table ronde sur l'informatique médicale

Journal de la société statistique de Paris, tome 111 (1970), p. 73-90
<http://www.numdam.org/item?id=JSFS_1970__111__73_0>

© Société de statistique de Paris, 1970, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) 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/

II

TABLE RONDE SUR L'INFORMATIQUE MÉDICALE

Personne n'ignore le développement extraordinaire qu'a pris l'informatique au cours de ces dernières années, et beaucoup savent qu'il en est de même dans le domaine particulier de la médecine. L'informatique médicale est la matière d'innombrables séminaires, symposiums, colloques ou tables rondes. Mais dans beaucoup de cas on mélange ce qui est fait, ce qu'on va faire, ce qui pourrait être fait, ce qu'on souhaite faire... Pour donner à notre table ronde un aspect original et concret, nous allons vous exposer ce qui a été fait ou ce qui est tout prêt d'être fait, en un même lieu — Villejuif —, autour d'un même ordinateur de grande puissance, par les équipes associées de l'Unité de Recherches statistiques de l'Institut national de la Santé et de la Recherche médicale, et de l'Institut Gustave-Roussy.

La médecine est un domaine très vaste : elle comporte, dans un centre hospitalo-universitaire notamment, les soins aux malades, la gestion de l'hôpital, l'enseignement, la recherche, la documentation. C'est à tous ces problèmes que se sont attaquées, sur le plan informatique, les équipes dont je vous ai parlé. C'est dire qu'un panorama complet de leurs activités recouvre un éventail très large. J'ai donc demandé à plusieurs personnes de venir exposer, l'une après l'autre, leur contribution à ce travail collectif. Pour ne pas dépasser le temps qui nous est imparti, chacun des orateurs devra vous dire en quelques minutes ce qui mériterait de longs développements. C'est là une difficile épreuve dont je les remercie par avance très vivement; et je leur passe maintenant la parole.

D. Schwartz

GESTION AUTOMATISÉE D'UN HOPITAL

Il n'est pas question ici de décrire ce que pourrait être idéalement la gestion automatisée et intégrée d'un hôpital, mais de vous dire brièvement où en est l'expérience que tente l'Institut Gustave-Roussy dans ce domaine. Il faut entendre ici le mot gestion dans son sens large : gestion des malades, des dossiers médicaux, des lits d'hospitalisation et non seulement de gestion financière.

Pourquoi *automatiser*, parce qu'il arrive un moment où le *volume* des données à traiter et la *complexité* des traitements sont tels qu'augmenter le personnel ne suffit plus à résoudre les problèmes. D'autre part, l'Institut désirait réaliser une gestion *intégrée*. Pour réaliser cet objectif un ordinateur était nécessaire; c'est l'UNIVAC 1107 de l'INSERM (Institut national de Statistique et Recherche médicale) auquel nous sommes reliés par un petit ordinateur l'UNIVAC 1005, qui sert à réaliser cette expérience.

En ce qui concerne l'analyse et la programmation, le Service informatique de l'Institut

a reçu une aide du constructeur et fait appel aux services de la S. E. M. A. (Société d'Étude et de Mathématiques appliquées).

1. GESTION FINANCIÈRE

La première application réalisée fut la *gestion du personnel*, qui non seulement calcule et imprime les bulletins de paie mais assure aussi bien d'autres fonctions : liste des personnes à convoquer pour visites médicales, fin de période d'essai, fin de stage, liste pour les élections, statistiques et états de fin d'année.

Dans le courant de l'année 1968 nous avons pris en charge la *gestion des stocks*, *matériel* et *alimentation*, et depuis janvier 1969 nous assurons la gestion des stocks *pharmaceutiques*.

L'enregistrement des travaux exécutés par les services généraux de l'Institut (atelier, service informatique, reproduction), des commandes de matériel, et des sorties de stock nous permet depuis janvier 1969 de fournir à chaque service l'état exact de ses dépenses, et en fonction de son budget, les crédits dont il dispose pour le mois suivant.

La *facturation d*es actes médicaux à la Sécurité sociale est depuis janvier 1968 faite chaque mois sur ordinateur à partir des données recueillies à la caisse.

Maintenant que nous saisissons les dépenses et les recettes de l'Institut nous abordons les problèmes de *Trésorerie Comptabilité générale* avec l'espoir de faire en 1970 de la *comptabilité analytique* nous permettant de calculer les prix de revient des actes, ainsi qu'une attribution dynamique des crédits de fonctionnement tenant compte de l'activité des services.

2. Les problèmes de la GESTION DES MALADES ont été abordés en même temps que ceux de la gestion administrative mais leur réalisation en est beaucoup moins avancée car nous n'avons pas pu dans ce secteur bénéficier de l'expérience acquise par d'autre.

Notre premier travail a été depuis le 1er janvier 1968 de recueillir pour tous les nouveaux malades (8 500 en 1968) et pour tous les anciens malades (10 000 environ) qui se sont représentés, toutes les *données d'état civil* ou autre dont nous pouvions avoir besoin dans les diverses applications prévues : nom, prénom, adresse, n° de Sécurité sociale, mutuelle, nom du médecin envoyeur, etc. Nous tirons de cet enregistrement complété et rectifié tous les jours : des étiquettes auto-collantes, des fiches d'inscription à l'I. N. S. E. E. (organisme nous signalant les décès), etc. De plus une sélection de ces informations complétées par des données médicales recueillies par l'INSERM permet à cet organisme de nous fournir des statistiques.

Nous voulons faire converger vers cet enregistrement *les événements médicaux les plus importants* de la vie du malade dans un but de diffusion des informations et de recherche.

La *diffusion* n'a d'intérêt que si elle est rapide, ceci impose de recueillir l'information aussitôt l'événement réalisé et de préférence par son auteur. L'inventaire précodé dans lequel le médecin coche l'événement qu'il désire signaler nous a paru le mieux adapté à ce recueil rapide. La rigidité de ce système a été adoucie par la possibilité d'introduire en clair tout événement non inventorié. Le service O. R. L. (comité 1 et 2) note ainsi depuis un an pour chaque malade : le diagnostic (localisation, adénopathie, métastases), les décisions thérapeutiques, les interventions chirurgicales, etc. Les informations sont introduites codées, contrôlées en ordinateur, et éditées en clair. N'ayant pas la possibilité avec le matériel actuel d'interroger le fichier en temps réel, nous avons adopté le système suivant : chaque information dont la diffusion est importante (nommée I. D. A. : information à diffusion accélérée) déclenche l'impression de tous les événements antérieurs enregistrés pour ce malade, créant

ainsi un genre de résumé qui annule et remplace le précédent. Ce résumé dynamique est imprimé en autant d'exemplaires qu'il y a de destinataires. Lorsque cette expérience sera bien rodée, nous étendrons le système progressivement aux différents comités. Le laboratoire d'anatomo-pathologie a adopté pour l'enregistrement de ses analyses le système documentaire du « mot clé ». Depuis juin 1969, dans chaque compte rendu le médecin souligne quelques mots, qui sont perforés par la secrétaire du service, et introduit en clair en ordinateur. Un dictionnaire préétabli permet de rejeter les mots erronés et de traduire en code les autres.

La gestion des hospitalisations qui est en cours de démarrage devrait permettre d'améliorer le taux d'occupation des lits et par conséquent de raccourcir le temps d'attente d'hospitalisation des malades. Le rôle de l'ordinateur dans cette application est d'aider les chefs d'unités et d'hospitalisation à gérer les services en leur fournissant l'état d'occupation des lits, les prévisions de sortie et les listes d'attente avec les indications médicales essentielles leur permettant de classer ces demandes. Les auteurs des demandes sont journellement tenus au courant de la progression de leur demande : malade en n^{ieme} position sur la liste d'attente, malade convoqué pour telle date, malade hospitalisé.

Lorsque la décision de convoquer est prise par le chef d'unité d'hospitalisation les personnes intéressées en sont prévenues et la veille de l'arrivée du malade des fiches contenant les principaux renseignements sont distribuées à tous les secteurs : soins, diététiques, hôtellerie, secrétariat, etc., qui doivent prendre en charge ce malade.

L'automatisation des *Rendez-Vous* qui est en cours de programmation et devrait être en application fin 1970 permet :

— d'ordonnancer, pour un même malade, les séquences d'actes demandés, en fonction d'un certain nombre de critères;

— de fixer au malade la date et l'heure de son prochain rendez-vous, réduisant ainsi la durée d'attente;

— de planifier la charge des services réalisant les actes.

3. Il nous reste encore bien des problèmes à résoudre : *automatisation des laboratoires*, *gestion des repas*, recueil de tous les actes en vue de la *facturation*. Ce n'est que lorsque nous serons arrivés au bout de ces réalisations que nous pourrons enfin déduire la facturation de ce fichier des malades, et réunir ainsi les deux grandes chaînes administratives et médicales.

Autre « Intégration » à réaliser : pour des raisons de commodités se sont développés deux dictionnaires : celui de la documentation et celui du dossier médical, il nous faudra dans l'avenir les confronter et les réunir.

M. Renaud
Responsable du Service informatique
à l'Institut Gustave-Roussy

Pr Denoix et Renaud M. *Un exemple de logistique hospitalière à l'Institut Gustave-Roussy.* Informatique et Gestion, *10* : 45-51 (1969).

Renaud M. — *Études et Réalisations du Service informatique de l'Institut Gustave-Roussy.* Informatique médicale et hospitalière, Symposium de Toulouse : 495-509 (1968).

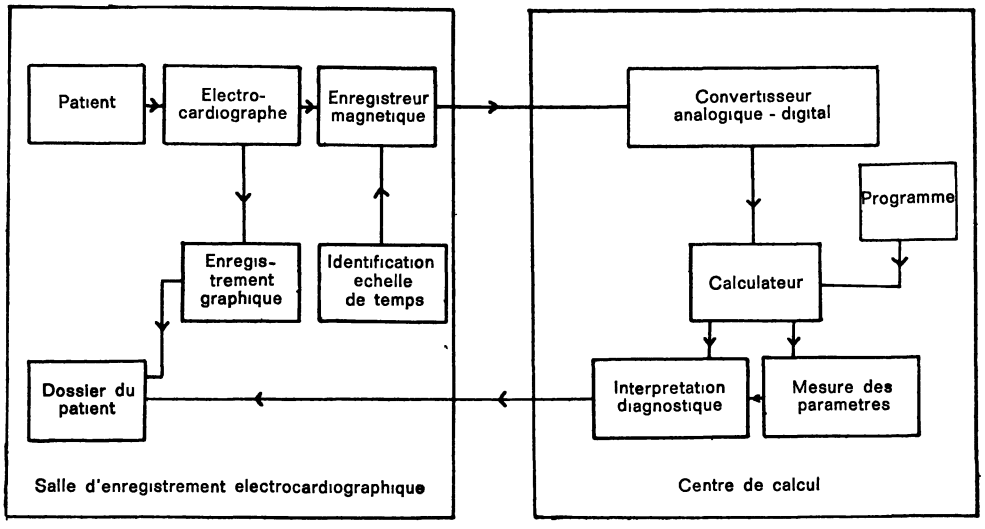
L'ANALYSE DES SIGNAUX BIOLOGIQUES

L'une des applications les plus fructueuses de l'informatique au domaine médical est sans doute l'analyse automatique des signaux biologiques. Les raisons en sont de deux types : les signaux électriques recueillis sur un sujet peuvent être enregistrés indépendamment de toute interprétation et contiennent donc une information objective qui mérite d'être traitée objectivement, d'autre part le calcul automatique seul permet d'analyser la masse de données contenues dans un signal complexe ou dans plusieurs signaux enregistrés en simultanéité.

Essentiellement, on peut distinguer deux grands domaines d'application :

— lorsque les progrès de la connaissance médicale permettent déjà à des spécialistes de tirer des renseignements substantiels sur l'état du sujet à partir de l'observation de tracés enregistrés par exemple sur papier graphique, il est possible d'envisager une *automatisation* de cette activité. Tel est le cas de l'électrocardiographie. Il existe des programmes permettant à un calculateur d'interpréter un électrocardiogramme à 12 dérivations scalaires en moins de deux minutes. Il est aisé d'imaginer les applications possibles de cette technique nouvelle qui ajoute à l'interprétation humaine un bénéfice de précision et de reproductibilité fort appréciables. L'application au dépistage et à la prévention des maladies cardiovasculaires par exemple semble pleine d'intérêt.

— L'interprétation de nombreux signaux de structure plus complexe traduisant en particulier des phénomènes aléatoires comme l'électro-encéphalogramme, est encore à l'heure actuelle trop rudimentaire pour qu'une analyse de routine puisse présenter un intérêt pratique. L'ordinateur doit être considéré dans ce cas comme un outil *de recherche*. Le calcul automatique permet alors d'extraire des signaux étudiés des paramètres que la seule observation ne permet pas d'appréhender (corrélations, spectres, reconnaissance de « patterns »...) et qui peuvent



se révéler fort importants soit pour la connaissance des mécanismes physiologiques, soit dans un but diagnostique.

En conclusion l'analyse des signaux biologiques par ordinateur est avant tout une technique de recherche, pouvant dans certains cas privilégiés donner lieu à des applications de routine.

A titre d'exemple et dans le but de préciser le type de méthodes utilisées pour l'analyse de l'électrocardiogramme, il est utile de décrire schématiquement le système étudié à la clinique cardiologique de l'hôpital Boucicaut. Ce schéma permet de distinguer clairement trois étapes :

— *une étape analogique :* enregistrement simultané en modulation de fréquence de 3 dérivations électrocardiographiques, d'une base de temps et de signaux d'identification, chaque enregistrement étant effectué sur une voie séparée de l'enregistreur magnétique;

— *la conversion analogique numérique :* elle est effectuée en synchronisation avec la base de temps à la fréquence de 500 points par seconde avec une précision de 12 bits. Les signaux d'identification sont reconnus et codés numériquement. Le résultat de la conversion est écrit sur bande magnétique numérique compatible avec le calculateur (3 000 points par dérivation sont conservés);

— *une étape numérique :* le programme réalise trois series d'opérations :
- le prétraitement (lissage, détection des complexes rapides...),
- la reconnaissance des ondes (repérage des différentes ondes des complexes rapides, de l'onde T et de l'onde P, détection d'arythmies, édition des durées et amplitudes moyennes des ondes après élimination des complexes « aberrants »...),
- le « diagnostic » (à partir des mesures ainsi obtenues, un algorithme simulant le raisonnement du cardiologue permet d'éditer les conclusions de l'interprétation des tracés).

Les programmes sont écrits en FORTRAN IV. Le test de l'ensemble du système est actuellement effectué sur un grand nombre de tracés normaux ou pathologiques.

P. Ducimetière

UTILISATION DES CALCULATEURS POUR L'AIDE AU DIAGNOSTIC

Les réalisations de l'informatique dans le domaine du diagnostic médical restent jusqu'à présent à mi-chemin entre la recherche et l'application en pratique médicale courante. L'idée d'utiliser les possibilités des calculateurs pour reproduire la démarche diagnostique ne s'est d'ailleurs pas imposée comme un besoin fondamental mais plutôt comme une voie de recherche originale. Les premiers résultats, publiés dès 1960, étaient encourageants mais les sujets abordés restent jusqu'à ce jour très étroits (cardiopathies congénitales, pathologie thyroïdienne, tumeurs osseuses) et d'un intérêt limité pour le praticien non spécialisé.

C'est à une pathologie assez fréquente, en particulier dans les services de pneumologie (10 % environ des hospitalisations) que nous sommes intéressés, la découverte d'une opacité arrondie unique intrathoracique lors d'un examen radiologique (examen systématique dans près de 30 % des cas) entraînant des discussions diagnostiques parfois difficiles.

Le problème posé par la recherche de l'étiologie de tels types d'images présente deux aspects très souvent indissociables :

— l'image observée est-elle bien une néoformation pulmonaire;

— quelle en est la maladie causale.

On peut évoquer trente diagnostics du cancer bronchique à l'image séquelle de fracture de côtes dont les fréquentes respectives sont évidemment très différentes.

A partir des probabilités à priori des affections pouvant être à l'origine de semblables images radiologiques, grâce au concours d'examens complémentaires, le médecin va étape par étape, cerner un diagnostic rarement évident d'emblée.

Nous nous sommes proposés d'établir un programme entre le médecin et un ordinateur susceptible de l'aider dans cette démarche.

Les premiers renseignements fournis par le médecin au calculateur ont trait à l'analyse de l'image radiologique.

La méthode utilisée suppose la connaissance d'un certain nombre de données : probabilités à priori des différentes étiologies dans la population à laquelle appartient le malade, probabilités conditionnelles des signes pris en compte.

Le recueil de ces informations est très souvent difficile; dans certains cas on dispose de statistiques correctement établies; dans d'autres il faut s'en tenir aux chiffres que des médecins très compétents considèrent comme vraisemblables, à la lumière de leur expérience. On peut alors en utilisant la méthode de Bayes calculer les probabilités des différents diagnostics que l'on peut évoquer devant une image donnée.

Une méthode semblable permet de traiter les informations obtenues par l'examen clinique. Mais déjà à ce niveau le calculateur suggère des questions, dont la priorité varie selon les résultats de l'analyse des signes radiologiques.

Il est très exceptionnel que les seuls renseignements cliniques et radiologiques suffisent pour décider d'un diagnostic avec certitude. Le plus souvent le médecin doit faire appel à des investigations complémentaires. Il dispose d'un assez large éventail d'examens (15 environ). A chacune de ces techniques on peut associer un coût complexe devant résumer plusieurs facteurs : prix de l'examen, contraintes plus ou moins désagréables qu'il impose au malade, risque de complication éventuelle lié à cet examen (un examen qui permet le diagnostic de la maladie A peut être dangereux si l'anomalie thoracique est due à l'affection B).

Il importe donc de déterminer en face de chaque cas une démarche optima permettant d'affirmer le diagnostic. Le calculateur devrait suggérer au médecin pour chaque malade une liste d'examens complémentaires, en fixant l'ordre dans lequel il est intéressant de les pratiquer.

Cette progression doit conduire soit au diagnostic sûr soit à deux ou trois diagnostics affectés de leurs probabilités, si l'affirmation ne peut être qu'histologique et nécessite une thoracotomie.

Un tel programme de formalisation du diagnostic des opacités pulmonaires arrondies nous paraît devoir intéresser, non seulement le médecin généraliste, mais aussi le spécialiste. Il est vraisemblable qu'il permettrait de plus une meilleure connaissance statistique de ce domaine de la pathologie pulmonaire.

Les résultats ultérieurs pourraient donc être de plus en plus sûrs et précis et faciliter le dépistage de masse des affections de l'arbre respiratoire et le traitement précoce des tumeurs bronchiques malignes.

Mme ALPEROVITCH

UTILISATION DES ORDINATEURS EN RADIOTHÉRAPIE

Les progrès de la radiothérapie au cours des deux dernières décennies ont été dus d'abord à l'introduction des hautes énergies puis à une rigueur de plus en plus grande des conditions de traitement. Cet effort d'accroissement de la précision des techniques nécessite une analyse de chacune des étapes de l'établissement du plan de traitement et de sa réalisation: repérage de la tumeur par rapport à des repères fixes, choix des conditions balistiques d'irradiation et calcul des doses, détermination des paramètres géométriques du traitement, réalisation de l'irradiation. Chacune de ces étapes peut bénéficier des progrès technologiques récents, notamment en informatique, et des méthodes d'automatisation susceptibles d'accroître la reproductibilité des séances de traitement et d'éliminer les causes humaines d'erreur ou d'inexactitude.

Les méthodes de dosimétrie permettent d'obtenir en laboratoire une précision supérieure à 1 % sur la détermination de la dose. Malheureusement, leur application au malade soulève de nombreuses difficultés.

Il faut d'abord déterminer dans chaque cas particulier les contours extérieurs du malade, la position des différents organes et évaluer leur densité.

Il est alors nécessaire de calculer en fonction de ces différents paramètres la distribution de la dose correspondant à chaque faisceau d'irradiation ou à chaque source radioactive implantée. Jusqu'à présent la complexité avait conduit à mettre au point des méthodes empiriques. Ainsi en radiothérapie interstitielle, où on se contente le plus souvent d'évaluer l'ordre de grandeur de la dose délivrée. En radiothérapie externe, des abaques permettent de calculer la distribution de la dose à condition que tous les axes des faisceaux soient situés dans le même plan, que les paramètres géométriques ne prennent que certaines valeurs préétablies et que soient négligées les corrections complexes comme celles dues à l'hétérogénéité des tissus. Ce calcul demande à un technicien entraîné une demi-heure à deux heures par malade suivant la complexité du plan de traitement et il ne peut être fait que dans le plan central.

Ces difficultés privent les radiothérapeutes de renseignements essentiels sur la distribution des doses en tous les points du volume irradié.

En effet, le calcul mathématique de la dose en un point quelconque est possible mais comporte un tel nombre d'opérations élémentaires qu'il est irréalisable par les méthodes classiques. De plus le volume intéressant est généralement grand et il serait nécessaire d'effectuer le calcul plusieurs dizaines de milliers de fois.

Les radiothérapeutes hésitent d'autre part à utiliser certaines techniques de traitement du fait de la complexité des calculs qu'elles entraînent. Ainsi le radiothérapeute se limite le plus souvent à des faisceaux d'irradiation dont les axes sont coplanaires; de même le curiethérapeute n'accepte qu'avec réticence de modifier la forme ou la dimension des sources radioactives qu'il utilise par manque d'informations précises et faciles à obtenir sur les variations de la répartition des doses qui s'en suivent.

La complexité des calculs et le grand nombre de points pour lesquels le même processus de calcul doit être utilisé exigent l'emploi d'un ordinateur, d'autant plus que le radiothérapeute désire disposer du résultat dans un délai très bref (moins de 24 heures et si possible une heure) afin de juger de la valeur de son plan de traitement avant le début de celui-ci.

Enfin, il est possible de faire effectuer à l'ordinateur à priori plusieurs calculs correspondant à des plans de traitement légèrement différents pour permettre au radiothérapeute de choisir en connaissance de cause le plan de traitement le plus approprié.

Ces calculs peuvent être faits dans une première étape en faisant varier successivement chacun des paramètres définissant l'irradiation. Cette variation des paramètres peut d'ailleurs être commandée par l'ordinateur lui-même à partir du plan type d'irradiation.

Dans une étape ultérieure, le choix du « meilleur » plan de traitement peut être fait par l'ordinateur à partir de critères de choix établis par les radiothérapeutes après l'étude mathématique d'un programme d'optimisation.

Ces calculs de type scientifique ne peuvent être entrepris qu'avec un ordinateur d'une taille suffisante et suffisamment rapide.

Les temps de calcul sur l'ordinateur UNIVAC 1107 de l'INSERM sont de l'ordre de 2 à 5 mn par plan de calcul suivant la complexité des données. Un traceur de courbes relié à l'ordinateur permet de présenter directement les résultats sous forme de courbes.

Mme Dutreix

INFORMATIQUE ET ÉPIDÉMIOLOGIE

Il est habituel de diviser la recherche médicale en 4 chapitres qui sont : le diagnostic, l'étiologie (étude des « causes » des maladies), le pronostic et la thérapeutique. Bien entendu, il y a de très nombreuses façons d'aborder les problèmes qui se posent dans chacun de ces chapitres : on peut travailler sur l'homme ou sur l'animal; à partir d'expériences ou de l'observation, etc. Noux voudrions parler ici de la voie *épidémiologique*.

Qu'est-ce donc que l'épidémiologie?

Au sens moderne du terme on peut dire que c'est l'étude de la répartition d'une maladie en fonction du temps, de l'espace ou de tout autre critère. Cette définition est suffisamment générale pour qu'on puisse y faire entrer les quatre chapitres précédents : en effet, dire que le tabac joue un rôle étiologique dans le cancer de la vessie, c'est dire que la répartition des malades atteints de cette affection n'est pas la même chez les fumeurs et les non-fumeurs, etc.

Ces considérations générales étant faites, il est sans doute plus intéressant pour montrer la nécessité de l'informatique dans la recherche épidémiologique de traiter d'un exemple précis; il s'agit d'une vaste enquête sur l'épidémiologie de l'athérosclérose coronarienne menée depuis quelques années par divers organismes, aux buts multiples mais dont l'essentiel peut être ainsi résumé :

Déterminer les facteurs pronostics de la maladie; puis à partir de cette connaissance, sélectionner un groupe de sujets à haut risque qu'il sera peut-être possible de traiter préventivement.

On voit donc le principe de l'enquête : à partir d'un échantillon de sujets bien portants, mesurer sur eux un grand nombre de critères, les suivre régulièrement pendant un certain nombre d'années pour pouvoir comparer les sujets devenus malades à ceux restés sains, pour tous les facteurs étudiés.

On voit également ce qu'une telle étude implique : la maladie est assez rare, de sorte que si on veut avoir un nombre de malades suffisamment important pour permettre des

conclusions valables, il faut partir d'un échantillon très important (de l'ordre de plusieurs milliers de sujets); de plus on a intérêt à considérer un grand nombre de critères éventuels de risque (plusieurs dizaines).

Finalement, on se retrouve à la tête d'une quantité d'informations absolument considérable : même si on ne voulait effectuer que des calculs très simples sur cette population, l'ordinateur serait indispensable. En fait les calculs ne sont pas simples, car on se trouve non pas en situation expérimentale, mais en situation d'observation; il s'agit alors d'essayer de démêler l'écheveau très complexe de toutes les corrélations entre les divers critères que l'on risque de trouver. Ceci montre, la nécessité de l'ordinateur, et de l'écriture de programmes simples et puissants permettant de satisfaire facilement toutes les exigences du chercheur épidémiologiste.

J. Lellouch

LE CALCULATEUR ÉLECTRONIQUE DANS LA RECHERCHE EN BIOLOGIE

EXEMPLE D'APPLICATION : L'ÉTUDE DE LA CROISSANCE TUMORALE

A. LES MODÈLES

Avec cette application, le calculateur va introduire un changement réel dans les méthodes de raisonnement du chercheur en biologie.

Les modèles mathématiques ont été jusqu'ici peu employés en biologie et le début de leur emploi coïncide précisément avec l'arrivée du calculateur.

Qu'est-ce qu'un modèle?

> *C'est une construction mathématique substituée à l'objet réel étudié en vue d'un certain but poursuivi.*

Plutôt que de tenter une revue des différents modèles, je vais vous parler des techniques de *simulation* qui paraissent particulièrement intéressantes puisque :

— elles ont à voir avec les phénomènes soumis à variabilité, ce qui est le cas général en biologie;
— elles ne peuvent guère être appliquées qu'avec l'aide d'un calculateur.

L'exemple que je choisirai concerne des problèmes de croissance tumorale que nous étudions à l'I. G. R. (1) en collaboration avec le docteur Frindel et le professeur Tubiana.

Nous avons donc à étudier des populations de cellules sur lesquelles on peut faire un certain nombre d'hypothèses destinées à rendre compte des *observations* possibles grâce à des expériences qui sont :

— les croissances des tumeurs sous différentes conditions;
— les durées des phases du cycle cellulaire à différents moments de la croissance;
— les proportions de cellules dans les différentes phases du cycle cellulaire à différents instants.

1. Institut Gustave Roussy.

Les hypothèses qu'on peut faire concernent :

— la nature et la taille de la variabilité des durées de vie et des phases du cycle cellulaire;

— les probabilités qu'a une cellule selon son âge et selon l'âge de la tumeur à laquelle elle appartient de fuir la tumeur, de mourir, d'arrêter son processus de division.

A chaque ensemble d'hypothèses va correspondre une population théorique de cellules qu'on comparera à celle qu'on peut observer afin de rejeter ou valider les hypothèses faites.

Dans les *techniques de simulation*, on pourra simuler ces populations théoriques grâce à des calculs du type Monte-Carlo : ces calculs consistent à tirer dans des distributions de probabilité données des nombres au hasard qui représentent le temps que telle cellule reste en synthèse, la probabilité que telle autre a de se diviser, etc.

Sur cette « population » simulée, on peut *aussi effectuer une « expérience » qu'effectue le chercheur en laboratoire* (ainsi, la plus simple, consiste simplement à compter à des temps différents, le nombre de « cellules » présentes — courbes de croissance; une expérience plus compliquée permet d'estimer, à un moment donné, les durées des différentes phases — techniques autoradiographiques.)

Le chercheur pourra donc voir quels résultats il aurait obtenu s'il avait réalisé son expérience sur cette population *théorique.*

D'où deux conséquences :

— valider des hypothèses;

— donner les limites de ce que peut apercevoir une expérience :

Si deux corps d'hypothèses conduisent au même résultat sur le calculateur, on saura que les expériences actuelles ne permettent pas de trancher entre ces deux corps d'hypothèses : il faudra donc *trouver* une nouvelle expérience capable, elle, de trancher.

Le cas s'est produit chez nous où nous avons pu montrer grâce à des modèles de ce type que les expériences habituelles étaient incapables de renseigner sur l'existence des cellules à vie très longue : nous avons alors imaginé une nouvelle expérience que nous avons pu *tester* àpriori sur des populations simulées avant de la confronter avec la réalité.

Un mot pour dire l'utilité du calculateur dans ce dernier cas : notre simulation portait sur 1 000 lignées de cellules suivies sur 12 générations — soit à peu près 8×10^6 cellules : pour chacune d'elles, il fallait tirer au sort *6* nombres, si bien qu'au total une simulation représentait 50 millions de nombres tirés au sort...

Je ne discuterai pas ici les conditions dans lesquelles on doit préférer une solution mathématique analytique à l'utilisation du calculateur.

Je vous rappelle l'utilité de tels modèles : c'est

— juger de l'importance relative de plusieurs facteurs;

— trancher entre plusieurs hypothèses;

— suggérer des expériences;

— permettre l'extrapolation à des régions non accessibles directement à l'expérience (pour autant qu'on puisse juger que les hypothèses faites et confirmées par l'expérience restent valables dans ces régions);

— et aussi... permettre l'interprétation d'expériences aux résultats compliqués.

B. POUR MÉMOIRE : LE CALCULATEUR PERMET D'AUTOMATISER

— *Automatisation de certains dosages :* ainsi à l'Hôtel-Dieu le dosage de l'insuline sera réalisé automatiquement : la bande perforée sortant de l'appareil de dosage sera traitée

par un programme qui fournira au biologiste en retour les valeurs des différents dosages ainsi que la précision des estimations et les caractéristiques de la courbe standard qui change à chaque série de manipulation.

— *Machine à calculer perfectionnée* qui autorise le biologiste à employer certains plans d'expérience ou modes d'analyse tel l'analyse de covariance, qu'il se refusait auparavant non pas à cause de leur complication expérimentale, mais par suite des calculs qu'ils imposaient. Une autre application pour certaines expériences destinées à être souvent renouvelées : un plan de randomisation automatique.

C. OUTIL D'ANALYSE MATHÉMATIQUE DE CERTAINES DONNÉES

Signalé pour mémoire : ainsi la reconnaissance de formes (par exemple lecture automatique du cariotype), l'interprétation et l'analyse des données dimensionnelles inexploitables jusque-là ouvrent des champs d'application également intéressants pour le chercheur.

A.-J. VALLERON

EXPOSÉ SUR L'ENSEIGNEMENT PROGRAMMÉ

Je vais me borner à quelques idées très générales sur les possibilités d'utilisation de l'informatique dans l'enseignement médical en insistant bien d'emblée sur le fait qu'il s'agit beaucoup plus d'idées que de réalisations effectives, encore que nous ayons actuellement un petit programme opérationnel d'aide à l'enseignement de la statistique.

***

Bien entendu l'idée de programmer un enseignement — c'est-à-dire de codifier les *modes d'interaction* entre enseignants et enseignés — n'est pas exclusivement limitée au domaine de l'informatique (il existe aussi par exemple des « livres brouillés »). Néanmoins les possibilités offertes par les ordinateurs sont telles (mémoires, vitesse, etc.) que la tentation est grande d'identifier enseignement programmé et aide de l'ordinateur à l'enseignement. La confusion est faite quelquefois; c'est un premier point de réflexion.

***

Si l'on se limite à l'utilisation de l'ordinateur, on peut encore concevoir — au moins d'un point de vue théorique — deux grandes orientations :

1° se servir de l'ordinateur essentiellement comme d'un *outil* destiné à améliorer *globalement* l'enseignement (en homogénéisant la qualité des enseignants, en s'adaptant au rythme individuel des étudiants, en procédant à des interrogations permettant un certain renforcement des connaissances, etc.), sans remettre en cause les bases même de l'enseignement délivré;

2° au contraire profiter des possibilités offertes par l'ordinateur pour *repenser* complètement les *fondements* de l'enseignement soit en étudiant les mécanismes d'apprentissage (modèles, simulations, etc.) soit même en remettant en cause la nature du raisonnement

médical (problème du diagnostic : il existe évidemment des liens très étroits entre la recherche diagnostique telle qu'elle a été évoquée dans un autre exposé, et l'enseignement du diagnostic).

On retrouve dans cette dualité, aide à l'enseignement ou analyse des mécanismes, un clivage cher aux statisticiens entre problèmes pragmatiques et problèmes de recherche...

*
* *

Le troisième point dont je voudrais dire un mot est l'importance du *degré de liberté* laissé à l'étudiant. C'est en effet une dimension fondamentale pour fixer les niveaux d'enseignement où l'ordinateur peut apporter une aide (un enseignement simple peut a priori être très directif, un enseignement complexe beaucoup moins) et pour déterminer les éléments techniques qui doivent être mis en œuvre (terminaux plus ou moins élaborés; langage-machine : ce langage peut être relativement pauvre dans un enseignement directif — ce sera un langage « maître », très codifié — ou au contraire riche et diversifié — langage « élève » — si l'enseignement cherche à être peu directif).

Dans le domaine de l'enseignement supérieur en particulier on tend à laisser la plus grande liberté possible aux étudiants, au minimum sous la forme d'un enseignement « tutoriel » (enseignement à embranchements) et avec comme perspective et quelques réalisations la « résolution de problèmes » qui permet à l'étudiant de poser des questions, de classer les informations, de proposer des cheminements pour la solution, etc.

*
* *

Je cite, enfin en vrac d'autres thèmes de réflexions : choix des critères de comparaison des résultats de divers enseignements, choix des sujets (étudiants) sur lesquels portera l'étude, définition des stratégies pédagogiques — autant de mots qui évoquent des analogies avec le problème des choix thérapeutiques. C'est ce dernier thème de réflexion que je voulais suggérer : de l'enseignement programmé considéré comme une thérapeutique... Il reste à souhaiter que le critère d'efficacité choisi ne soit pas la survie de l'étudiant !...

P. Lazar

LA DOCUMENTATION AUTOMATIQUE A L'INSTITUT GUSTAVE-ROUSSY RÉALISATIONS ACTUELLES ET PROJETS EN COURS

Le Système automatique de bibliographie, d'information et de recherche en carcinologie (S. A. B. I. R.), créé à l'Institut Gustave-Roussy par le Service de documentation automatique, est un système spécialisé qui dépouille la littérature carcinologique de tous les pays et dont les services de Recherche rétrospective et de dissémination sélective d'information possèdent une diffusion internationale.

Ce système automatique est pour l'heure opérationnel en français, en anglais et en allemand ce qui permet de procéder actuellement à l'établissement d'un réseau international d'information en carcinologie.

Le but du système S. A. B. I. R. est de permettre une information détaillée et de

haute tenue scientifique tant par le choix des documents retenus que par la finesse des critères qui permettront de les appréhender.

La finesse même de la définition des concepts permet une grande liberté dans l'établissement de la stratégie de recherche. Dans le cadre d'un système spécialisé où les concepts univoques se répartissent en champs sémantiques donnés, selon des relations de hiérarchie et d'environnement connues, une recherche peut s'effectuer :

— soit *en une étape* par sélection directe des documents coordonnés à des concepts obéissant aux opérations logiques du principe de sélection que constitue la question;

— soit par *ajustements* successifs des paramètres de recherche après contrôle des réponses obtenues par le premier moyen.

Ce mode adaptatif de recherche peut être utilisé soit directement par le demandeur soit par les membres du service de documentation.

C'est cette dernière option qui préside actuellement au fonctionnement du Service de documentation scientifique de l'Institut Gustave-Roussy.

L'ajustement des paramètres au niveau de spécificité désirable s'effectue à la lumière des tableaux sémantiques correspondants.

Le Service de documentation automatique s'oriente actuellement vers l'étude de l'établissement prévisionnel de l'énoncé optimum, cela sur la base de la fréquence de citation et de la cooccurrence probable des concepts dans un même champ ainsi qu'en fonction du degré de complexité du principe de sélection. L'abord de ces problèmes semble devoir être facilité par la spécialisation même des concepts retenus.

Un ensemble de tests portant sur le taux de récupération des informations ainsi que sur le niveau du bruit accompagnant les réponses est d'autre part en cours d'élaboration.

De plus, le Service de documentation automatique procède actuellement à une étude exploratoire de la contribution de la statistique à l'étude du langage naturel. Un premier travail actuellement en cours concerne une recherche de type lexical portant sur les fréquences de citation des mots du langage naturel et l'étude de leur écart éventuel à une distribution de Poisson dans divers ensembles de documents.

Par ailleurs l'évolution du système S. A. B. I. R. l'amène à se confronter au problème de la prise en compte et de l'exploitation automatique de la formule structurale des composés chimiques organiques, ceci dans la double optique de recherches documentaires et d'études de corrélation « structure-structure » et « structure-propriétés ». Depuis un an déjà le Service de documentation automatique a entrepris l'étude des divers systèmes existants et de leur aptitude tant à la recherche de structures complètes qu'à celles de motifs singuliers de celles-ci. Le système retenu devra permettre non seulement la collection de paramètres préétablis, structures complètes ou partielles, et cela selon les méthodes directes ou adaptatives déjà évoquées, mais encore la recherche de motifs communs non préétablis et, de façon générale, les études de corrélation.

Les options de départ résident essentiellement dans l'association de trois exigences : adaptation aux besoins des usagers, économie de gestion du système et capacité d'évolution de celui-ci vers la prise en compte et le traitement de composés à activité biologique présentant éventuellement de hauts poids moléculaires ainsi que vers la prise en considération des configurations spatiales : structure secondaire et éventuellement tertiaire.

Les systèmes existants se répartissent en trois groupes. Les systèmes à fragmentation, fondés sur une décomposition de la molécule en subunités, le mode de décomposition étant lui-même défini par un type donné de classification des composés organiques. Les fragments

ainsi reconnus sont alors codés et mis en mémoire avec, selon les systèmes, indication plus ou moins importante de leur mode de liaison. Ces systèmes sont adaptés à des recherches de type préétabli.

Les notations linéaires représentant les formules structurales par des suites linéaires ordonnées, relativement compactes, de symboles alphabétiques et numériques. L'ordre de ces suites émane de l'application d'un ensemble complexe de règles qui rendent la représentation univoque. Canoniques, concises, se répartissant naturellement dans un fichier en fonction de leur contenu structural, les notations linéaires conviennent à la constitution de fichier d'exploitation rapide et à la recherche portant sur l'ensemble de la structure des composés; par contre, malgré la possibilité de permutation des éléments de la notation, elles sont peu adaptées à la recherche de sous-structures.

Les systèmes topologiques rendant un compte explicite de l'existence et des positions relatives de chaque atome et de chaque liaison de la formule structurale. Ils se répartissent en deux groupes selon que, par programme, ils aboutissent ou non à une représentation canonique. Les opérations de codage et de traitement en sont peu économiques; cependant l'avantage des représentations topologiques réside dans le fait que leur nature explicite permet la recherche de motifs non prédéterminés; c'est là un apanage qui rend ces systèmes indispensables pour certains types de recherche.

Le Service de documentation automatique s'oriente vers l'élaboration d'un système intégré associant l'économie d'une présélection sur fichier à exploitation rapide (vraisemblablement du type fragmentation), aux possibilités de recherche offertes par les systèmes topologiques au sein des sous-fichiers issus de la présélection.

Tels sont dans les grandes lignes les réalisations actuelles et les projets en cours du Service de documentation scientifique de l'Institut Gustave-Roussy, toutes réalisations qui n'auraient pu se concevoir sans la possibilité d'utilisation d'un ordinateur puissant.

MM. Wolf-Terraine et D. Rimbert

DISCUSSION

Le Dr Norbert Marx demande d'abord de combien de lits dispose l'hôpital Gustave-Roussy et combien de malades pendant une année y sont : *a*) hospitalisés; *b*) vus dans la consultation externe.

En ce qui concerne le classement des maladies par leur symptomatologie, on peut dire qu'une maladie se « compose » d'un certain nombre de symptômes, autrement dit un malade présentant un certain nombre de symptômes doit être atteint d'une maladie portant le nom de... Théoriquement en donnant à un computer un certain nombre de symptômes, constatés chez un malade, la machine doit pouvoir donner un diagnostic correspondant à cet agglomérat de symptômes. Malheureusement un certain nombre de symptômes peut être identique pour plusieurs maladies et le computer devra laisser « au choix » plusieurs diagnostics possibles.

Le Dr Norbert Marx a étudié ce problème du diagnostic possible et probable en partant de tables de morbidité construites avec des diagnostics relevés aux contrôles médicaux de la Caisse primaire centrale de Sécurité sociale de la Région parisienne. Ce travail a paru dans les comptes rendus de la troisième Conférence internationale des actuaires et statisticiens de la Sécurité sociale qui s'est tenue 1962 à Madrid.

Ces tables de morbidité donnant une répartition de 101 489 diagnostics pour le sexe masculin et de 122 690 diagnostics pour le sexe féminin, les diagnostics furent répartis d'après les 17 groupes de maladies de la nomenclature internationale des maladies et par groupe d'âge.

Prenons comme unité un groupe de maladie, réparti par groupe d'âge.

Pour le groupe 1 : Maladies infectieuses et parasitaires nous constatons le pourcentage maximum pour les deux sexes pour le groupe d'âge de 20 à 24 ans.

Pour le Groupe 2 : Tumeurs, nous constatons le pourcentage maximum chez le sexe masculin pour le groupe d'âge de 60 à 64 ans; pour le sexe féminin nous trouvons le maximum pour le groupe d'âge de 45 à 49 ans.

Les mêmes tables nous permettent de prendre comme unité le groupe d'âge. Nous constatons ainsi pour les groupes d'âges de 20 à 24 ans précités que le pourcentage maximum se trouve chez le sexe masculin dans le groupe 1 des maladies : Maladies infectieuses et Parasitaires; mais chez le sexe féminin nous trouvons le maximum dans le groupe de maladies 11 : Accouchement et complications de la grossesse.

Dans le groupe d'âge de 60 à 64 ans précité, nous trouvons le maximum chez le sexe masculin dans le groupe de maladies 7 : Maladies de l'appareil circulatoire.

Pour le groupe d'âge 45 à 49 ans précité chez le sexe féminin nous trouvons le maximum dans le groupe 9 : Maladies de l'appareil digestif. Il s'agit dans ces cas de gastrites, hernies, etc.; mais pas de tumeurs.

Si nous pouvons ainsi localiser la répartition des maladies par sexe et par groupes d'âge, on pourrait envisager en prenant d'abord comme unité un groupe de maladies, en utilisant tous les numéros des maladies qui composent ce groupe et en calculant les pourcentages à arriver per exclusionem de détecter les diagnostics des maladies soupçonnées, mais rares.

En prenant ensuite comme unité le groupe d'âge et en procédant de la même manière, on constatera pour ce groupe d'âge les diagnostics rares et probables.

En connaissant pour des centaines de milliers de diagnostics la répartition des maladies les plus fréquentes et les plus rares, on aura des indications pour chercher per exclusionem les maladies qui étant en soi rares se trouvent à partir d'un nombre important de maladies ou disons sont à envisager à partir d'un très grand nombre de maladies.

Pour ces études il faut utiliser les diagnostics répartis par exemple pour la France : *a*) pour les grandes villes; *b*) pour la campagne et les villes au-dessous de 5 000 habitants; *c*) pour des régions à définir.

Le même principe doit être appliqué pour des statistiques comparatives pour les pays de la Communauté européenne qui ont une médecine presqu'identique et des systèmes de Sécurité sociale très proches.

Ainsi au lieu de nous baser sur un nombre de symptômes qui sont souvent identiques pour plusieurs maladies, on pourrait arriver dans des cas douteux à des diagnostics de maladies rares.

Les services financiers de la Sécurité sociale pourraient utiliser les méthodes décrites en n'utilisant qu'une répartition des dépenses par groupes d'âge pour répérer immédiatement les groupes d'âge qui sont les plus coûteux pour la Sécurité sociale.

En procédant ainsi on pourrait approfondir ce problème vital pour la Sécurité sociale en étudiant avec les services médicaux les groupes les plus coûteux pour la Sécurité sociale.

M. Philippe Manchon. — Vous nous avez parlé d'informatique et de l'ordinateur de l'institut Gustave-Roussy. Sans vouloir en inférer tout ce qui se passe dans cet hôpital,

que je ne connais pas, je voudrais signaler une tendance à l'automatisation qui semble avoir maintenant une place importante dans nombre d'hôpitaux. Il s'agit de l'automatisation des dosages chimiques ou biochimiques réalisés habituellement dans les laboratoires centraux de chimie biologique, auxquels sont plus particulièrement attachés les pharmaciens. On a alors affaire à de l'information sans ordinateur; ce qui n'exclut nullement que cette information soit traitée ultérieurement par ordinateur. Mais cette information, plus facile et plus rapide de nos jours, apporte *elle aussi* son aide au diagnostic... plus souvent d'ailleurs à une révision ou à une précision du diagnostic. En effet l'automatisation des dosages et le développement de l'analyse multiple séquentielle permet souvent de fournir au médecin de *l'information qu'il n'a pas demandée.* Ayant son idée sur tel malade qu'il a examiné, il va faire une demande de deux ou trois de ces dosages biologiques. On va lui donner le résultat de douze ou quinze. Cela va l'aider à se faire un tableau plus complet du malade et parfois à réformer son diagnostic.

Une fois l'investissement réalisé, le coût de ces dosages multiples est vraiment très peu augmenté par rapport au coût de deux ou trois seulement d'entre eux. Néanmoins, il ne faut pas oublier que ces dosages ne sont pas les seuls examens possibles, et j'ai été très heureux d'entendre Mme Alperowitch nous dire qu'une technique était au point pour optimiser le coût des divers examens à entreprendre : coût financier et « coût-santé » pour le malade, ce qui est encore plus important!

A M. Lazar je dirai mon plein accord sur ce qu'il a dit à propos du « degré de liberté » à laisser aux étudiants de tel ou tel niveau. Cette notion — reconnue comme fondamentale en statistique — est peut-être un peu méconnue en pédagogie! Mais, dans le domaine médical et para-médical de la recherche, un tel enseignement pose un problème. Bien des efforts ont été faits (je ne pourrai en donner la liste exhaustive) pour que les biologistes et les médecins suivent des cours de statistique et d'informatique. Mais les plus âgés d'entre eux, dont beaucoup répugnent à l'abstraction des mathématiques, sont désarmés devant les progrès rapides de ces sciences.

Il semble qu'il faille doter les laboratoires d'instruments qui leur permettent de faire eux-mêmes leurs statistiques sans avoir recours à des statisticiens de métier servant un ordinateur puissant. Les calculateurs électroniques de table peuvent répondre à ce besoin. Mais il faut alors que ces utilisateurs soient conseillés par des gens ayant une formation statistique un peu plus poussée que la leur pour leur permettre d'éviter les écueils.

Le Dr Jean Robert Debray, de l'Institut, après avoir félicité les collaborateurs du Pr Schwartz pour le vif intérêt de leurs diverses communications, expose, à la demande de M. Brichler, les grandes lignes d'une récente expérience de médecine préventive, utilisant les techniques d'automatisation et le traitement de l'information par ordinateur.

Cette réalisation a été conçue par la commission de médecine préventive du Haut Comité médical des Assurances générales de France, que préside le Pr Turpin.

Dans une communication faite le 16 mai 1969, à la Société médicale des Hôpitaux de Paris, les auteurs ont, en huit textes, décrit le processus d'investigation développé dans cinq secteurs : savoir : pathologie cardio-vasculaire (Pr agrégé J. Himbert), appareil respiratoire (Pr agrégé J. Chrétien), appareil digestif (Pr agrégé J.-P. Hardouin); appareil urinaire (Pr Richet), dépistage du diabète sucré (Pr agrégé M. Gueniot).

Un questionnaire d'une centaine de questions « à choix forcé », a été élaboré par le Pr P. Pichot. Une quinzaine d'analyses biologiques — exécutées par le laboratoire du Dr Guigan — complètent l'ensemble des données. Données volontairement restreintes,

en fonction de deux limites. Limite de temps — une heure d'investigations (y compris les 10 à 15 mn nécessaires pour les réponses aux questionnaires). Limite de coût; le prix de revient étant fixé à environ 100 F.

Dans la communication du 16 mai, le Dr Claude Roussel a donné le schéma de la réalisation pratique : les données, le programme, les résultats (la programmation a été effectuée par le service des systèmes spéciaux de la Cie I. B. M.-France).

La phase expérimentale de cette entreprise qui a débuté fin mai 1969, se poursuit actuellement. Les patients examinés, tous volontaires, appartiennent au personnel des A. G. de France.

La conception même de cette recherche de points de repère physiologiques (R. P. R. P.) repose, indique M. Debray, sur les notions suivantes :

— division de l'examen du sujet en deux temps. Le deuxième temps étant l'indispensable examen clinique du médecin, choisi par l'intéressé.

— choix des investigations précliniques du premier temps, effectué en tenant compte des facteurs suivants : recueil d'éléments avant tout mesurables (biologiques, spirométriques), recours à des données graphiques significatives (enregistrement électrique, radiophotographie), enregistrement de données anamnétiques personnelles ou héréditaires.

Les buts que nous nous sommes assignés, conclu M. Debray, pourraient schématiquement être résumé ainsi : rendre service aux patients volontaires, donner un nouvel essor à la médecine préventive, en y associant les médecins traitants, permettre sur les plans scientifique et technique (grâce à la « mise en mémoire » des données recueillies et la répétition annuelle des investigations) d'étudier l'évolution des anomalies perçues parfois au stade infraclinique — et à concourir ainsi à une meilleure préhension et à une meilleure prévention de ce que l'on nomme « risques aggravés » en matière d'assurance.

M. Corcos. — 1° Question à M. Lellouch : Quelles ont été la valeur initiale et la variation des variables retenues dans la recherche *à priori* des facteurs explicatifs de la proportion des personnes ayant contracté telle maladie?

2° question à M. Valleron : Jusqu'à quel point pouvait-on accorder sa confiance aux résultats des simulations du développement cellulaire, surtout si cela revenait à remettre en question des expériences tenues jusque-là pour significatives?

Réponse de M. Lellouch : Notre étude n'ayant pas assez de recul, nous n'avons pas assez de malades pour pouvoir faire l'analyse des résultats. Cependant ceux de l'enquête de Framingham, la plus ancienne du genre, ont montré que la valeur initiale d'une variable suffisait à résumer le rôle prédicteur de cette variable dans l'apparition de la maladie.

Réponse de M. Valleron : Je répondrai d'abord en utilisant la formulation inverse de la vôtre : Jusqu'à quel point pourrait-on accorder sa confiance à l'emploi de simulations (et de façon générale de modèle) si cela ne revenait pas à pouvoir remettre en question des expériences « tenues jusque-là pour significatives »! Vous savez, en effet, que tel est bien un des grands emplois des modèles : rejeter une hypothèse et amener la nécessité d'une nouvelle expérience pour tester une nouvelle hypothèse (en revanche, on peut dire qu'au moins dans les premières étapes d'une recherche, un modèle trop en agrément avec les données est bien souvent suspect).

Enfin, dans le cas particulier de l'exemple présenté, je veux rappeler que j'emploie les modèles un peu dans l'esprit suivant « si par hasard la population de cellules sur laquelle j'expérimente avait telles propriétés (hypothèses du modèle), les apercevrai-je grâce à telle expérience »? Si oui, je garde l'expérience; sinon j'essaie d'imaginer une expérience dont les résultats soient sensibles à ces hypothèses...