

LES NON-RECOMMANDATIONS DANS LE TRAITEMENT DE L'ARTHROSE ET DE LA POLYARTHRITE RHUMATOÏDE, OU POURQUOI LA MÉDECINE BASÉE SUR LE NIVEAU DE PREUVE NE S'APPLIQUE PAS (ENCORE) À LA RHUMATOLOGIE

M. G. MALAISE (1)

RÉSUMÉ : La médecine basée sur le niveau de preuve ("Evidence-Based Medicine, EBM") est un concept dont la pénétration dans le langage médical est ahurissante. Toutes les disciplines se sentent concernées même si le concept s'applique mieux à certaines qui étudient des maladies et des paramètres d'évolutivité clairement définis (par exemple vie-décès), deux situations qui ne sont pas encore fréquemment rencontrées en rhumatologie. Nous passerons en revue les différentes raisons qui font que l'EBM ne peut valablement s'appliquer maintenant à la rhumatologie et analyserons les incertitudes qui pourraient transformer l'"Evidence-Based Medicine" en "evidence-biased medicine". Nous repréciserons aussi pourquoi l'EBM ne peut se substituer à l'expérience clinique, mais doit plutôt constituer un catalogue raisonné des pratiques médicales pour le plus grand bénéfice des patients mais aussi du médecin.

La médecine basée sur le niveau de preuve ("Evidence-Based Medicine, EBM") propose de nouveaux comportements pour les médecins limitant le rôle de l'intuition et de l'empirisme par une utilisation rationnelle des connaissances établies par une démarche scientifique rigoureuse (1). Alors qu'elle semble mieux s'appliquer à des disciplines comme la cardiologie (2) où les diagnostics sont clairs et les paramètres d'évolution tranchés (décès/survie, par exemple), elle suscite néanmoins l'inquiétude de l'expert et du clinicien par la transformation de l'approche traditionnelle, la médecine clinique, en médecine de cohortes (1). Elle en réduit fatalement le rôle autrefois central ("experience-based medicine") (3) en l'obligeant à se référer par rapport à une norme censée représenter la vérité. Elle en arriverait à inquiéter aussi l'économiste, car les preuves scientifiques sont indépendantes des coûts (1). Voulant combattre l'empirisme clinique (diagnostique et surtout thérapeutique), "l'essai thérapeutique contrôlé (avec tests statistiques ad hoc) n'est - pourtant - qu'un moyen historique d'accéder à cette connaissance empirique" (4). Cette approche ne peut en autant cas soustraire la discipline médicale à une approche physiopathologique la plus développée qui soit. Claude Bernard avait déjà radicalement pris position

(1) Chargé de Cours, l'Université de Liège, Département de Médecine Interne, Chef du Service de Rhumatologie, CHU Sart Tilman

WHY IS EVIDENCE-BASED MEDICINE SO DIFFICULT TO APPLY IN CLINICAL RHEUMATOLOGY TO-DAY ?

SUMMARY : The speed with which the phrase "Evidence-Based Medicine" has entered our language has been astonishing. Virtually all medical areas are concerned, but it probably applies best when clearcut diagnosis can be made or when mortality or well-defined morbidity are the outcomes, two conditions not really frequently encountered in rheumatology. We will discuss why rheumatology is probably not ready yet for the concept and will systematically review why "Evidence-Based Medicine" could in fact be "evidence-biased medicine". We will emphasize why evidence-based medicine can not replace the personal experience and knowledge of patients, two major corner-stones in the clinician's daily work, but should simply be an access to scientific data for the benefit of both patient and physician.

KEYWORDS : *Evidence-Based Medicine - Guidelines - Rheumatoid arthritis - Osteoarthritis - Rheumatology*

dans son "Introduction à l'étude de la médecine expérimentale" déclarant que "pour tous les phénomènes dont la cause est déterminée, la statistique n'a rien à faire; elle serait même absurde" (5). Comme le souligne si bien Ravaud (6), "c'est faute d'atteindre à la connaissance des causes que nous devons nous contenter d'une connaissance des relations". Ou encore : "ce n'est que faute de mieux que la statistique se substitue à une physiologie évanescence" (6). Or, cette démarche qui reste le but ultime de la connaissance scientifique (cf les mécanismes "moléculaires" identifiés dans de nombreuses maladies par les progrès foudroyants de la biologie) ne serait pas indispensable dans l'élaboration du nouveau paradigme (1).

Nous allons passer en revue les différents écueils qui freinent à ce jour la validité d'une mise en place d'un raisonnement clinique, diagnostique ou thérapeutique en Rhumatologie basé uniquement sur le niveau de preuve. Ils sont liés non seulement au développement actuel de la discipline rhumatologique et aux limites de la méthodologie. Dans tous les cas, ils constituent des freins significatifs qu'un esprit critique se doit de connaître. Nous les appellerons les 12 incertitudes. Dans la limite de nos possibilités, nous les illustrerons à travers deux grands domaines cliniques : celui de la polyarthrite rhumatoïde (PR) et celui de l'arthrose.

L'INCERTITUDE NOSOLOGIQUE

La validité d'un essai clinique, en double-insu et randomisé, dépend de la mise en place d'une méthodologie rigoureuse dont l'homogénéité du groupe de malades étudié n'est pas la moindre. Cette étape de classification n'est pas encore achevée en rhumatologie. En effet, à ce jour, le cadre nosologique rhumatologique reste abondamment descriptif. La capacité de réaction de la synoviale étant limitée, un même mode d'entrée clinique peut cacher des réalités physiopathologiques très différentes (7). Rappelons qu'il a fallu 22 siècles, d'Hippocrate à Landré-Beauvais, pour sortir la PR (appelée goutte asthénique primitive) de la goutte (8), laquelle nouvelle entité englobait dans les milieux scientifiques anglo-saxons jusqu'à la moitié des années 60 des maladies comme les spondylarthrites (9, 10). Les protocoles d'études cliniques les plus récents dans la PR comportent, par ailleurs, toujours une liste impressionnante d'autres affections à exclure (qui, pour la plupart d'entre elles, ne font pas l'objet de critères ni de classification, ni de diagnostic...).

Nous ignorons encore tout du cadre nosologique dans lequel il faut placer le rhumatisme palindromique et le RS3PE (Remitting Symetric Sero-negative Synovitis with Pitting Oedema) qui n'existent que sous une forme descriptive. Ces affections sont-elles des polyarthrites rhumatoïdes atypiques ? Les polyarthrites rhumatoïdes d'apparition tardive (late onset) sont-elles identiques aux formes précoces (early onset) ?

L'incertitude physiopathologique est encore plus grande dans l'arthrose, ou plutôt dans les maladies arthrosiques. Reléguées traditionnellement dans les processus dégénératifs, les arthroses constituent un domaine extrêmement hétérogène dont nous ne comprenons qu'à peine la diversité clinique et, pour ainsi dire, ignorons tout de l'hétérogénéité biochimique (11, 12). Pourquoi fait-on une arthrose de hanche ou de genou (déjà différentes dans leur mode évolutif) et ne fait-on pas spontanément une arthrose de cheville ?

Il faut encourager sans réserve les travaux permettant l'identification des affections par des paramètres immuno(bio)logiques, et créant ainsi les bases de nouvelles classifications.

L'INCERTITUDE DIAGNOSTIQUE

Sur cette incertitude nosologique se greffent d'autres incertitudes: comment appliquer rationnellement le choix des examens complémentaires biologiques ou techniques (7, 13) ? En

l'absence de critères de diagnostic (à ne pas confondre avec des critères de classification), comment raisonner le choix complémentaire de telle sorte que la valeur prédictive positive (la probabilité que la maladie existe lorsque le test est positif) se situe au moins au-delà de 10 % (c'est-à-dire ne générant que 10 résultats faussement positifs pour un vrai positif !). La démarche inverse, qui consiste à se servir de la valeur prédictive négative (la probabilité que la maladie n'existe pas quand le test est négatif), qui est souvent très élevée, allant jusqu'à 100 % est stérile sur le plan intellectuel et insoutenable sur un plan économique (13). Aujourd'hui encore, près de 12 ans après la publication des critères de classification de la PR (14), dont la sensibilité est de 96 % tout comme la spécificité d'ailleurs, leur valeur prédictive positive n'est que de 57 % (15) soulignant dès lors cruellement l'incertitude diagnostique qui subsiste.

L'INCERTITUDE DES BUTS : LE RÉPONDEUR

Pour démontrer l'efficacité thérapeutique, quel paramètre (clinique, biologique, radiographique, fonctionnel) choisir ? Est-il validé ? Si oui, quelle est sa signification réelle pour le patient ? Un exemple frappant vient des études récentes concernant l'utilisation thérapeutique des anticorps anti-TNF dans la PR (16, 17), approche thérapeutique par ailleurs révolutionnaire (18). L'amélioration clinique obtenue par un médicament (ou un placebo) chez un patient souffrant de PR peut s'évaluer par un score appelé ACR 20 (50 ou 70) (19). Le patient sera défini comme répondeur s'il présente une amélioration de 20 (50 ou 70) % du nombre d'articulations douloureuses, une amélioration de 20 (50 ou 70) % du nombre d'articulations gonflées et une amélioration de 20 (50 ou 70) % de 3 des 5 items suivants : l'évaluation de la douleur du patient, l'évaluation globale de l'activité de la maladie par le patient, l'évaluation globale de l'activité de la maladie par l'investigateur, l'évaluation du degré d'incapacité (HAQ) par le patient, les marqueurs d'inflammation (VS ou CRP). Après 6 mois de traitement par étanercept 25 mg 2 x semaine en injection sous-cutanée, 59 % des patients souffrant de PR satisfont le critère ACR 20 contre 11 % des patients recevant un placebo (16). Sachant que dans les deux groupes, les patients présentaient en moyenne 25 articulations gonflées et que la raideur matinale moyenne était d'environ 5 heures (16), 6 patients sur 10 recevant le traitement actif ont vu le nombre d'articulations gonflées se réduire seulement de 25 à 20 et 6 patients sur dix (les

mêmes ?) ont encore souffert d'une raideur matinale de 4 heures au lieu de 5. Un raisonnement similaire s'applique dans les études utilisant un autre anticorps anti-TNF, l'infliximab (ACR 20 à 6 mois : 50 % chez les patients infliximab + méthotrexate *versus* 20 % chez les patients placebo + méthotrexate) (17). Il s'agit là (16, 17) d'études rigoureuses avec des différences produit actif-placebo très significatives ($p \leq 0.01$), mais dont la pertinence clinique pourrait échapper si nous n'avons compris que nous étions à l'aube d'une révolution thérapeutique (18). Le niveau de preuve indiscutablement fourni dans ces publications satisfait-il le patient (et le médecin) ? Et au moment où l'information disparaît au profit de la communication (20), la diffusion immédiate par la presse du niveau de preuve (21) engendre le délicat problème de sa signification. Mis au courant en temps réel (et souvent avant le médecin clinicien!), le "consommateur-patient" ne peut bien évidemment qu'adhérer à cette connaissance du progrès, et bien sûr, la réclamer. Cette adhésion massive au concept présenté peut, de manière dangereuse, constituer une sorte de validation, qui, bien sûr, ne s'embarrasse pas de sa signification. Alors que la médecine par le niveau de preuve se voulait au départ indépendante du marché, elle se voit rattrapée par ce dernier (cf l'incertitude de marché). Il est hélas à redouter que les recommandations de prudence formulées par les milieux scientifiques ne portent plus guère car "l'information de hiérarchie est devenue consensuelle" (20).

L'INCERTITUDE DES BUTS : LE NON-RÉPONDEUR

Le paradoxe est encore un peu plus étonnant quand les patients considérés comme non-répondeurs répondent quand même. Ainsi, l'amélioration clinique et biologique induite par l'infliximab s'accompagne d'une stabilisation des paramètres radiologiques, ce qui est un progrès remarquable (22). Dans ces études menées de manière prospective et rigoureuse, il apparaît que les patients qui n'ont pas été considérés comme répondeurs sur le plan clinique présentaient pourtant une amélioration radiologique ! Ceci limite singulièrement la signification réelle des indices d'évolutivité développés. Et comment ne pas rappeler que l'insatisfaction des milieux scientifiques rhumatologiques dans le traitement de la PR vient de ce que presque tous les traitements dits de fond ont démontré une activité clinique convaincante contre placebo à court et à moyen terme, mais ne modifient pas le décours radiologique et fonctionnel de l'affec-

tion (23). Ces incertitudes ont alimenté une abondante littérature entre les partisans de l'approche pyramidale traditionnelle basée sur une escalade médicamenteuse lente et prudente (24) et ceux qui inversaient la pyramide et proposaient une polythérapie d'emblée (25, 26), par ailleurs tout aussi empirique. Ces controverses ont ouvert les yeux des rhumatologues sur l'absolue nécessité d'innover et de rester critiques devant les résultats des publications (cf l'incertitude du maintien du résultat).

L'INCERTITUDE DES OBJECTIFS

Il est clair que l'avènement des traitements biologiques va modifier la stratégie thérapeutique de la PR. Des normes d'activité, des normes de rémission, des normes de réponse clinique et des normes de réponse radiologique sont en cours de validation ou sont déjà validées (27), et vont s'inscrire dans les desiderata requis pour l'élaboration des futurs protocoles d'évaluation (28). La rhumatologie n'a pas encore bénéficié des analyses performantes sur les avantages indirects que procurent les nouveautés thérapeutiques comme c'est le cas, depuis de nombreuses années, en hématologie-oncologie. Citons les scores FACIT (29) pour la quantification de la fatigue (Functional Assessment of Chronic Illness Therapy) et la détermination des QALYs (Quality Adjusted Life Expectancy) (30). Ces études à venir fourniront des renseignements nouveaux éclairant différemment le problème posé par les maladies chroniques rhumatologiques. Dès lors, on peut se poser la question de l'utilité de recommandations qui seraient construites sur la base des publications antérieures imparfaites sur le plan méthodologique et testant des molécules qui ne représentent aucun progrès tangible. C'est vrai pour la PR. C'est tout aussi vrai pour l'arthrose pour laquelle des critères thérapeutiques ont été proposés (31). Ceux-ci suggèrent de traiter la "crise" arthrosique par un antalgique plutôt que par un anti-inflammatoire, non pas tant sur la base d'un manque d'efficacité de ces derniers, mais plutôt en raison du risque digestif non prédictible qu'il fait courir. Mal acceptés par les milieux rhumatologiques européens surtout non-anglo-saxons, ces critères ne représentent pas une approche moderne du problème posé. Les AINS restent plus efficaces que les antalgiques sur de nombreux paramètres cliniques ou fonctionnels (WOMAC) (32) et le risque de toxicité digestive peut être très significativement modifié par l'association d'un AINS à du misoprostol, par exemple, ce qui réduit de moitié la

fréquence de pathologie ulcéreuse tant dans l'arthrose que dans la PR (33, 34). Sans oublier l'accessibilité prochaine à des AINS sélectifs Cox-2 (celecoxib, rofecoxib) pour lesquels les effets indésirables gastro-digestifs sont rigoureusement identiques à ceux du placebo alors que l'efficacité est comparable à celle des molécules classiques non sélectives.

L'INCERTITUDE DES PRIORITÉS

Il ne faut pas non plus tomber dans l'amalgame qui voudrait, à tort, assimiler l'absence de preuve d'efficacité à la preuve d'absence d'efficacité (35). Et l'absence de preuve vient peut-être simplement du fait que les études nécessaires pour l'identifier n'ont pas été réalisées. Pour de multiples raisons. Le manque de priorité du domaine à étudier, par exemple. Les investissements R&D de l'industrie pharmaceutique concernent surtout les secteurs cardio-vasculaires, l'oncologie, l'ostéoporose et globalement le secteur de la santé de la femme, garants de maladies à prévalence importante dans nos riches contrées occidentales autorisant l'espoir d'une rentabilité visible (2). Et clairement des millions d'individus vont en bénéficier. Mais les autres (36) ? Quid de l'intérêt pour des affections - rhumatologiques - dont l'incidence est moins de 1/10.000 et dont la prévalence ne dépasse pas 0,2/1.000 comme les maladies systémiques, la pseudopolyarthrite rhizomélique, la spondylarthrite ankylosante ? L'arthrose a une chance puisqu'elle se développe finalement chez tout le monde, sa prévalence étant liée à l'âge. De plus, la capacité qu'ont les instances scientifiques à organiser des essais cliniques indépendamment de toute intervention financière d'un partenaire industriel est limitée (2). Dans cette optique, le médecin clinicien aurait alors l'autorisation de se contenter d'un niveau de preuve moindre (de 2 à 5) (37) autorisant là l'empirisme et l'expertise clinique que pourtant cette méthode entend combattre. La médecine basée sur le niveau de preuve ne possède pas une indépendance décisionnelle. Elle ne parvient pas à s'affranchir des lois du marché qui, en économie libérale, ont toujours raison. L'incertitude des priorités est encore renforcée par la mise en route d'un système de recherche autocontrôlé par ceux qui doivent en bénéficier, que ce soit partenaires industriels, médecins et consommateurs-patients par l'intermédiaire de leurs associations (2). Ce système de production scientifique est résumé dans la figure 1. Il n'est pas très loin d'un autre système de production scientifique appelé l'effet "Gold" (voir plus loin).

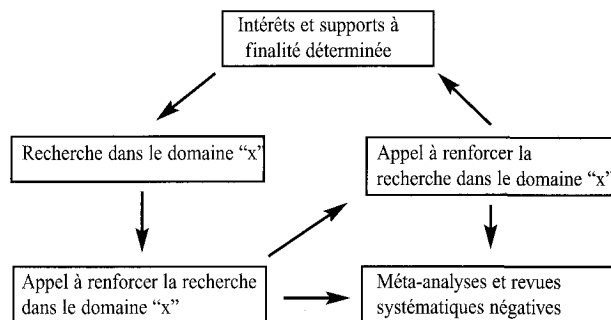


Fig. 1. Le "cycle de recherche à finalité auto-entretenu". Des groupes impliqués dans la finalité des recherches se rassemblent, partenaires industriels, médecins et patients définissant les hypothèses de recherches. Des travaux scientifiques sont générés mais en raison de différents biais (méthodologiques, de résultats ou de reviewing) engendrent des commentaires globaux, revues systématiques et autres méta-analyses dépourvues de signification positive, ce qui engendre la demande du groupe initiateur de renforcer les axes de recherche dans le domaine pré-défini. Ce type de mécanisme laisse peu de place au test d'hypothèses alternatives (2).

L'INCERTITUDE DES RÉSULTATS

La publication d'un essai contrôlé et randomisé dans une revue prestigieuse est-elle le garant d'un niveau de preuve indiscutable pouvant dès lors servir de référence ? On aurait aimé pouvoir dire oui. Il faut hélas nuancer. Indépendamment du problème de la fraude scientifique, qui n'est pas aujourd'hui le centre de notre propos (38, 39), celui de l'indépendance de l'éditeur ne peut être écarté. La rocambolesque démission il y a 2 ans de l'éditeur en chef du New England Journal of Medicine, pour cette raison par ailleurs, ne nécessite me semble-t-il aucun commentaire supplémentaire. La publication dans le New England du 16 mars 2000 de l'article de Lovell (40) sur l'efficacité de l'éta nercept dans le traitement de l'arthrite chronique juvénile n'est pas exempte d'une critique méthodologique majeure. En effet, seuls les enfants considérés comme répondeurs au traitement [64 % ont satisfait à l'indice ACR50 (19)] ont participé secondairement à la phase double-insu, objet principal de la publication. Le biais est évident. Encore une fois, il est clair que le traitement par anti-TNF est une avancée inestimable en rhumatologie adulte et pédiatrique (18, 40), mais le problème posé est bien celui de la signification réelle du résultat. Une nuance indispensable aux yeux du scientifique critique que n'évoque ni l'expérimentateur (le paragraphe "discussion" de l'article scientifique est particulièrement court), ni l'éditorialiste (41), encore moins explicitée au grand public lors de la communication de l'information (42). Car, la réponse à un traitement contre un placebo dans la deuxième phase de l'essai n'est pas indépendante du fait qu'on y a déjà

répondu dans la première. Prenons l'exemple inverse : s'attendait-on à une différence d'efficacité du médicament par rapport au placebo si on avait sélectionné dans la seconde phase les enfants qui n'y avaient pas répondu ? Il est, de même, particulièrement inquiétant que l'issue des résultats d'études comparatives de l'efficacité et de la tolérance de différents AINS dans le traitement d'arthrites ne soit pas indépendante du commanditaire qui sponsorise la comparaison, quelle que soit cette dernière par ailleurs (43) ! L'"Evidence-Based Medicine" est-elle devenue l'"evidence-biased medicine" ? (44). Le problème de l'indépendance de l'éditeur pose le problème beaucoup plus vaste de l'indépendance de l'information (45, 46). Aux anciens flux de pouvoir se substituent ceux du pouvoir des flux (47). Il est très symptomatique, et inquiétant, que l'acquisition récente de Time Warner, maison contenant le "savoir", par AOL, société internet dorénavant chargée de le diffuser, se soit accompagnée d'une déclaration du premier soulignant que "l'objectivité" de son travail journalistique ne serait pas remise en cause !

L'INCERTITUDE DE LA TAILLE DE L'ICEBERG

Le médecin clinicien, qui reste encore dans notre système de santé l'interlocuteur privilégié du patient, serait trop occupé que pour pouvoir accéder seul à une information valable (47). La médecine par le niveau de preuve, aux mains d'épidémiologistes ou de médecins de recherche clinique, se veut donc d'être la nouvelle référence guidant le médecin clinicien dans le dédale des informations disponibles (48). Un nouveau phare d'Alexandrie, en quelque sorte. Ou un iceberg, qui peut aussi servir de guide sur une carte. Cette dernière comparaison me paraît assez intéressante. En effet, l'iceberg se déplace, tout comme les connaissances médicales évoluent, de telle sorte que sa signification dans le guidage n'est plus absolue, mais relative. Il dérive, comme d'ailleurs le raisonnement. C'est cette même confusion du risque relatif en épidémiologie avec celle du risque absolu qui engendre des idées fausses génératrices de comportements aberrants. La lecture du livre de Skrabanek et McCormick, "idées folles, idées fausses en médecine" illustre remarquablement ce point. S'il est clair, je cite, "que la plupart des pilotes d'avion, comparés au commun des mortels, ont un *risque relatif* d'être tués dans un accident d'avion de l'ordre de plusieurs milliers contre un, cependant ni eux, ni nous ne devrions en déduire qu'il faut refuser de prendre l'avion,

puisque le risque absolu est extrêmement faible" (49). Cette perversion du raisonnement est encore attisée par "les médias qui attirent constamment notre attention sur des dangers pour notre santé qui sont le plus souvent rares, car notre risque individuel est extrêmement faible et que, dans ces conditions, il devrait être ignoré" (49). Que n'a-t-on associé alimentation riche en fibres et prévention du cancer du côlon, induisant un comportement alimentaire neuf dans les populations ? Naïveté de la déduction épidémiologique ? Cynisme des producteurs de fibres ? Cette association forte née de la combinaison d'un sophisme de faisceau (la réunion de plusieurs arguments, dont chacun pris à part est suspect ou faible, constitue une preuve solide) et de l'argument d'autorité (50) (voir l'incertitude de l'argument d'autorité) est devenue évidente, puisque "tout le monde le dit", mais personne ne le prouve (49). L'édition du 20 avril 2000 du New England Journal of Medicine vient enfin de démontrer l'inexactitude de tels propos (51).

La deuxième caractéristique de l'iceberg est qu'on n'en voit que sa face émergée, qui représente souvent une toute petite partie. Il en va de même pour le concept de la médecine par le niveau de preuve. Elle interroge une partie de l'art médical et nous aide à répondre, partiellement, à certaines situations auxquelles le clinicien est confronté (6). Elle ne peut avoir d'autre ambition que "de nous fournir, à un moment donné de l'histoire, un catalogue raisonné des pratiques médicales distinguant : 1) ce qui est démontré comme utile; 2) ce qui est démontré comme inutile et 3) ce qui n'est pas démontré du tout. Pour toute pratique relevant de cette dernière catégorie, la médecine par le niveau de preuve ne peut prétendre fonder une quelconque recommandation" (4). En effet, l'absence de preuve n'est pas preuve d'absence, nous l'avons déjà souligné (2, 4, 35, 52). Outre le fait que l'absence de preuve peut être liée à un problème de priorités (cf l'incertitude des priorités), elle est aussi liée au fait que les résultats négatifs ne sont quasi jamais publiés (2). Le rhumatologue est quotidiennement confronté à cette situation. Enfin, il est clair qu'un biais culturel peut exister ainsi qu'en atteste Raffy (4) parlant de "l'ignorance des preuves". C'est un conflit idéologique lié à la culture McWorld (53, 54). Ainsi, le score européen d'activité d'une PR, mDAS (55), validé à l'échelon de l'individu (ce qui intéresse évidemment tant le médecin clinicien que le patient) est systématiquement écarté dans les études FDA au profit du score américain ACR (19) adapté à la cohorte et, dès lors, inexploitable au quotidien.

L'INCERTITUDE DU MAINTIEN DE LA PREUVE

Se peut-il que la preuve irréfutable constituant un niveau I de preuve lors d'études randomisées (phase III) ne se maintienne pas lorsqu'il s'agit d'appliquer le principe (souvent médicamenteux) en dehors du cadre strict de l'étude clinique, c'est-à-dire en phase IV. Oui, et c'est bien plus fréquemment que l'on ne croit. Pour pouvoir rencontrer les objectifs définis, les concepteurs de protocoles vont osciller entre deux écueils : d'une part, renforcer les critères d'inclusion pour sélectionner des patients dont l'affection présente un profil évolutif ou épidémiologique favorable permettant l'accomplissement de l'hypothèse de départ et, d'autre part, renforcer les critères d'exclusion de manière à ce qu'aucune interférence exogène ne compromette le résultat. Dans tous les cas, cela revient à sélectionner des patients "artificiels" totalement minoritaires dans leur pathologie, hors normes, alors qu'on les utilise justement pour en définir une. La présence d'une co-morbidité, et, par ailleurs, de co-médications, sont deux raisons incontournables en pratique clinique susceptibles de modifier la "beauté et la clarté" des résultats de phase III. Ceci concerne tant l'observation d'une efficacité moindre que celle d'une augmentation du nombre d'effets secondaires indésirables. Une utilisation incorrecte de la médication par le médecin est aussi une autre source de divergence de résultats. Prenons l'exemple de la toxicité gastro-intestinale des AINS. Il faut à la fois tenir compte des données obtenues lors des études prospectives et contrôlées contre un autre médicament ou contre un placebo et de celles obtenues lors d'études "post-marketing", surtout rétrospectives ou prospectives non contrôlées. Les divergences souvent observées s'expliquent aisément. En effet : (a) le rapport d'effets indésirables après la mise sur le marché d'un AINS n'est pas constant. Weber (56), étudiant 7 AINS, a, en effet, montré qu'il est maximal à la fin de la deuxième année de marketing, pour décliner de manière constante dans la suite, indépendamment de l'augmentation de prescription de la molécule; (b) les cas isolément rapportés à la Pharmacovigilance et l'étude des données rétrospectives de pharmaco-épidémiologie ne sont absolument pas le garant de l'utilisation correcte de la molécule tant au niveau du médecin prescripteur que du patient; (c) une discordance entre la fréquence d'effets indésirables dans les années post-marketing par rapport aux études contrôlées peut aussi venir d'une augmentation progressive de la posologie, engen-

drée par une certaine expérience clinique liée à la recherche d'une efficacité supérieure aboutissant à une toxicité qui l'est également.

Utilisant un outil statistique (!) particulièrement bien adapté à la réalité clinique, à savoir la méthode du taux de survie actuarielle selon Kaplan-Meier (57), Amor et coll. (58), suivis par Wolfe et coll. (59), ont montré que le taux de maintien d'un traitement de fond est particulièrement mauvais dans la PR (de 20 à 50 % à deux ans), en contradiction avec toutes les attentes que laissaient supposer les publications concernant ces médications. Il est évident que les études à court et moyen terme ne sont pas le garant du maintien des résultats à long terme. De plus, il est clair que si des études ont dû être menées sur une très large échelle pour mettre en évidence une différence, celle-ci a de fortes chances d'être négligeable dans la réalité. Seule la répétition de la différence par d'autres équipes indépendantes confèrera une crédibilité clinique. De toute façon, un médicament qui "marche" n'a pas besoin de toutes ces précautions méthodologiques. Le méthotrexate, redécouvert par les milieux scientifiques rhumatologiques mi-1980 (60), a efficacement franchi toutes les étapes sans faux pas: il n'y pas une étude divergente concernant son efficacité et son taux de maintenance thérapeutique est le meilleur observé parmi les traitements de fond (59), assurant une continuité de validité entre les phases III et IV. Positionné naturellement par les rhumatologues cliniciens comme "gold standard" dans le traitement de la PR, il n'a bénéficié ni de promotions, ni de "lobbying" et n'a fait l'objet d'aucun argument d'autorité, ni de consensus diplomatique (cf l'incertitude du consensus). Il n'en avait pas besoin. Le consensus s'était établi par la pratique clinique, ce qui représente probablement l'aboutissement suprême pour un médicament.

L'INCERTITUDE DE L'ARGUMENT D'AUTORITÉ

L'argument d'autorité consiste à définir quelque chose comme vrai parce la source d'information fait autorité. Nous l'avons déjà partiellement rencontré dans l'incertitude 7, celle des résultats publiés par une revue "qui fait autorité" et qui comporte cependant un biais. Ici, il s'agit de l'avis "autorisé" d'une personne qui "fait autorité". Les personnes qui "font autorité" se rencontrent à de multiples niveaux : le scientifique "reconnu" vis-à-vis du reste de la communauté, l'enseignant vis-à-vis de ses élèves (quel que soit le cycle), le médecin vis-à-vis de son malade. D'ailleurs le code pénal a reconnu

comme circonstance aggravante le fait de commettre un acte répréhensible lorsqu'on "a autorité" sur quelqu'un. L'argument d'autorité est détestable et profondément anti-scientifique. Il implique souvent un individu seul ou un groupe d'individus qui subissent progressivement une transformation paranoïde de leur personnalité, les dispensant de fournir les preuves de ce qu'ils avancent (49). Ce comportement touche tout autant les prosélytes de l'"Evidence-Based Medicine" que ceux de l'"experience-based medicine" (61). Plus le niveau d'intelligence des personnes qui font autorité est élevé, plus leur "profession de foi" risque d'être dénuée de sens (49). Le sens critique doit rester éveillé car "s'il peut être sage, au moins dans un premier temps, d'accepter ce qu'affirme quelqu'un qui fait autorité, il est déraisonnable et dangereux d'y croire" (49).

L'INCERTITUDE DE LA VALEUR D'UN NOUVEAU CONCEPT, OU L'EFFET GOLD

Nous avons vu jusqu'ici quels étaient les écueils qui limitaient la signification des résultats scientifiques. Il s'agit d'exemples qui "illustrent des erreurs de raisonnement, l'utilisation d'arguments fallacieux, des fautes de logique, mais aussi diverses manières d'obscurcir la vérité, de la tordre ou de la déformer au point de la rendre méconnaissable sans la moindre intention avouée de l'altérer" (49). L'effet Gold (62) décrit comment construire un concept en y apportant les preuves scientifiques *a posteriori*. Le tableau I (d'après 50 et 62) résume les différentes étapes de la "construction". Son déroulement correct intègre à des moments divers les différents processus fallacieux précédemment (d)énoncés. L'effet Gold crée un besoin puis, après une campagne médiatique ciblée, recueille auprès de l'échantillon ainsi nouvellement informé la notion qu'il s'agissait bien d'une nécessité, d'une carence à combler au plus vite.

TABLEAU I. INSTALLATION D'UN NOUVEAU CONCEPT POURTANT NON PROUVÉ : "L'EFFET GOLD" (50, 62).

1. Emergence d'un concept nouveau par quelque personnes
2. Organisation d'une réunion de discussion de ses avantages et inconvénients à laquelle assiste un nombre plus élevé de personnes favorables au concept original que d'opposants
3. Désignation d'un comité représentatif pour préparer un recueil destiné à propager ce concept
4. Rédaction d'articles fondés sur le concept donnant l'impression d'un consensus croissant
5. Eventuelle création d'une publication spécialisée
6. Acceptation par le comité de rédaction et l'éditeur d'articles orthodoxes au concept
7. Accumulation de publications confirmant le concept de départ: passage du "il est généralement admis que" à "il est bien établi que", puis "il va de soi que"

Rassuré sur la justesse de leur "intuition", ses concepteurs consolident les fondations par l'édification d'une légitimité scientifique pouvant inclure la création de sa propre revue spécialisée, d'un comité de rédaction incluant les concepteurs initiaux et des leaders d'opinion (ce sont, en fait, généralement les mêmes) et par le rappel systématique du concept désormais "scientifiquement prouvé" par l'activation de multiples sociétés de consultance gérant une information dite de 3^{ème} cycle et dont la survie dépend quasi exclusivement de ce type de contrat. L'effet Gold se remarque particulièrement dans les rapports de consensus (49).

L'INCERTITUDE DU CONSENSUS

L'exploitation ultime de la médecine par le niveau de preuve devant servir de référence au monde clinique est l'élaboration d'un texte de consensus représentant l'opinion partagée par un groupe d'experts. Le Robert historique de la langue française nous apprend que la définition moderne (1970) du consensus, à la mode dans le vocabulaire politique, est celle d'un "accord social conforme aux vœux de la majorité" ou encore de "l'opinion d'une forte majorité". Une vérité scientifique ne s'établit pas d'après l'opinion de la majorité, mais pire, le consensus médical n'est finalement probablement qu'une sorte de PPCM (plus petit commun multiple), compromis satisfaisant toutes les parties. Hélas, il ne représente en fait rien du tout, car si toutes les parties partageaient une même idée physiopathologique ou conceptuelle sur la base de preuves irréfutables, la conférence de consensus serait inutile.

Il est ensuite tentant de passer du consensus à la norme. La norme est un terme nouveau apparu toujours selon Robert au 19^{ème} siècle, défini comme "représentant l'état habituel, régulier, conforme à la majorité des cas (1897), sans jugement de valeur", précise-t-il. Cette définition est terrifiante, car elle explicite clairement que la norme est indépendante de la qualification d'excellence et, dès lors, que les normes pourraient singulièrement s'écarter des évidences scientifiques que la médecine par le niveau de preuve entend prôner si une majorité s'en écarte. De plus, être hors norme pour quelque raison que se soit (de la diversité biologique au choix personnel) devient horriblement dangereux car on devient anormal, situation désagréable dans le monde économique et politique actuel nourri à l'aune de l'uniformité. En effet, le premier est confronté à la mondialisation, ce qui implique fatalement que normalisation soit synonyme de standardisation, c'est-à-

dire d'une uniformisation abaissant les prix de revient. Le second, obnubilé par la maîtrise des coûts de la santé, est tenté de saisir cette opportunité pour imposer une logique économique à la logique médicale ou à l'intérêt du patient, en témoigne la "rage normative" que les milieux médicaux ont connu dans un passé récent et connaissent par ailleurs toujours. La norme ouvre évidemment de suite la porte des recommandations médicales "opposables" (4).

EN CONCLUSIONS

Au terme de cette revue, il nous faut conclure. A la lueur du nombre d'incertitudes soulevées par la médecine selon le niveau de preuve, représentant autant d'écueils qui freinent son acceptation par l'esprit critique, faut-il d'emblée la condamner ? Par ailleurs, un esprit trop critique ne finit-il pas par tomber dans le septicisme, doutant de tout car puisque "si nous ne pouvons rien connaître de sûr, il faut nous abstenir de tout jugement et refuser tout dogmatisme" ? (63).

Il semble que ces craintes bien réelles venant du monde clinique illustrent une dérive que la médecine par le niveau de preuve n'envisageait pas au départ. Nous avons rappelé au début de l'article qu'elle arrivait à inquiéter aussi l'économiste car les preuves scientifiques sont indépendantes des coûts (1), et il n'est pas certain que les bénéfices de la standardisation en valent les coûts (64). Finalement, Sackett a peut-être été rattrapé par la dérive économique et par l'intérêt opportuniste de la tutelle politique puisqu'il éprouva le besoin de préciser (37) que "l'EBM se définit comme l'utilisation consciencieuse et judicieuse des meilleures données (preuves) actuelles de la recherche clinique pour la prise en charge personnalisée de chaque patient". Diplomatie courbe rentrante ? Probablement. Le monde clinique avait bien senti le souffle du boulet et jugé correctement la volonté de voir la médecine par le niveau de preuve se positionner comme une nouvelle discipline médicale (65), alors qu'elle ne doit prétendre qu'être une aide au service de la médecine clinique (4). La médecine clinique ne s'est pas laissée impressionner par la masse d'arguments "scientifiques" qui déferlaient et a fait reconnaître sa spécificité (66). Que de discussions pour confirmer qu'une situation médicale n'était finalement ni noire, ni blanche, mais souvent grise (67). Les deux tendances pourront se réconcilier. Que l'on ne se trompe cependant pas. Ces différences dans la pensée médicale ne sont pas indépendantes des influences idéologiques présentes aux USA, d'une part, en Europe, d'autre part. Il est cependant paradoxal

de remarquer qu'aux USA la promotion de la réussite individuelle (y compris dans le domaine de la santé), caricature du monde économique capitaliste, se base sur des normes établies par l'étude de cohortes alors que, dans nos sociétés européennes sociales libérales, la promotion du bien être collectif, priorité des états, se base d'abord sur celle de l'individu.

Quant à la rhumatologie, elle n'a finalement que peu intéressé la Cochrane Library. En plein essor, elle va, dès lors, représenter un très intéressant laboratoire parce que son développement ne pourra se faire que par une intégration des deux composantes fondamentales car, "en médecine clinique, l'excellence relève toujours d'un savant partage entre l'art de l'incertitude et la science de la probabilité" (62).

BIBLIOGRAPHIE

1. Fieschi M, Giorgi R, Gouvernet J, Degoulet P.— De la connaissance à la pratique médicale : l'introduction de la médecine basée sur le niveau de preuve. *Rev Med Intern*, 2000, **21**, 105-109.
2. Dieppe P, Szebenyi B.— Evidence based rheumatology. *J Rheumatol*, 2000, **27**, 4-7.
3. Guiraud-Chaumeil B, Even Ph.— Confrontation avec le système de santé américain, in *Les hôpitaux universitaires de l'an 2000*. Editions Prévat, Toulouse, 1996, 239.
4. Raffy O.— Les dangers de l'évidence based medicine. *Rev Prat*, 1999, **49**, 801-802.
5. Bernard C.— *Introduction à l'étude de la médecine expérimentale*. Flammarion, Paris, 1984, 194 pages.
6. Ravaut P.— Evidence-based medicine: du concept à la pratique. *Rev Med Intern*, 2000, **21**, 21-23.
7. Kavanaugh A.— The role of the laboratory in the evaluation of rheumatic diseases. *Clin Cornerstone*, 1999, **2**, 11-21.
8. Leca AP.— *Histoire illustrée de la rhumatologie. Goutte, rhumatismes et rhumatisants*. Editions Roger Dacosta, Paris, 1984.
9. Robecchi A, Daneco V.— I criteri diagnostici dell American Rheumatism Association e loro valore per la diagnosi di reumatismo brucellare e di reumatismo reiteriano. *Reumatismo*, 1961, **13**, 257-260.
10. Leca AP.— Les critères de la polyarthrite rhumatoïde (Etude critique). *Rhumatologie*, 1966, **18**, 151-162.
11. Dieppe PA.— Osteoarthritis : time to shift the paradigm. *Br Med J*, 1999, **318**, 1299-1300.
12. Brandt KD.— Osteoarthritis. *Rheum Dis Clin N Am*, 1999, **25**, 2.
13. Suarez-Almazor ME, Gonzalze-Lopez L, Ivan Gamez-Nava J, et al.— Utilization and predictive value of laboratory tests in patients referred to rheumatologists by primary care physicians. *J Rheumatol*, 1998, **25**, 1980-1985.
14. Arnett FC, Edworthy SM, Bloch DA et al.— The American Rheumatism Association 1987 revised criteria for

- the classification of rheumatoid arthritis. *Arthritis Rheum*, 1988, **31**, 315-324.
15. Gabriel SE.— The sensitivity and specificity of computerized databases for the diagnosis of rheumatoid arthritis. *Arthritis Rheum*, 1994, **37**, 821-823.
 16. Moreland L, Schiff MH, Baugartner SW et al.— Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. *Ann Intern Med*, 1999, **130**, 478-486.
 17. Maini R, St Clair EW, Brredveld F, et al.— Infliximab (chimeric anti-tumor necrosis factor a monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomized phase III trial. *Lancet*, 1999, **354**, 1932-1940.
 18. Advances in targetted therapies : TNF- α blockade in clinical practice. *Ann Rheum Dis*, 1999, **58** (suppl 1): 1-130.
 19. Felson DT, Anderson JJ, Boers M et al.— American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*, 1995, **38**, 727-735.
 20. Ramonet I.— *La tyrannie de la communication*. Editions Galilée, 1999, 198 pages.
 21. Poncin J.— *Le 7^{ème} Soir*, 14 Août 1999.
 22. Lipsky P, St. Clair W, Furst D, et al.— 54-week clinical and radiographic results from the ATTRACT trial: a phase III study of infliximab (Remicade) in patients with active RA despite methotrexate. *Arthritis Rheum*, 1999, **42** (suppl), S1980. Abstract.
 23. Wilske KR, Healey LA.— Challenging the therapeutic pyramid : a new look at treatment strategies for rheumatoid arthritis. *J Rheumatol*, 1990 (suppl 25), **17**, 4-7.
 24. Schumacher HR.— *Primer on rheumatic diseases*. ed 9. Arthritis Foundation, Atlanta, 1988, 93.
 25. Fries JF.— Reevaluating the therapeutic approach to rheumatoid arthritis: the "sawtooth" strategy. *J Rheumatol*, 1990 (suppl 25), **17**, 12-15.
 26. O'Dell JR, Haire CL, Erikson N, et al.— Treatment of rheumatoid arthritis with methotrexate alone, sulfasalazine and hydroxychloroquine, or a combination of all three medications. *N Engl J Med*, 1996, **334**, 1287-1291.
 27. Molenaar E, van der Heijde D, Boers M.— Uptake on outcome assessment in rheumatic disorders. *Cur Op Rheumatol*, 2000, **12**, 91-98.
 28. Kreutz G.— European regulatory aspects on new medicines targetted at treatment of rheumatoid arthritis. *Ann Rheum Dis*, 1999, **58** (Suppl 1), 192-195.
 29. Cella DC. F.A.C.I.T MANUAL.— *Manual of the functional assessment of chronic illness therapy, (Facit) scales*. Version 4, Novembre 1997.
 30. Anand P.— QALYs and the integration of claims in health-care rationing. *Health Care Anal*, 1999, **7**, 239-253.
 31. Hoshberg MC, Altman RD, Brandt KD, et al.— Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. *Arthritis Rheum*, 1995, **38**, 1541-1546.
 32. Geusens P, Manicourt D, Masure J, Kaufman L.— Quality of life in patients with osteoarthritis treated with non-steroidal anti-inflammatory drugs. A case-controlled survey. *Arthritis Rheum*, 1999, **42** (Suppl) 1340, S294 (abstract).
 33. Agrawal N, Van Kerckhove HEJM, Erhardt LJ, Geis GS.— Misoprostol coadministered with diclofenac for prevention of gastroduodenal ulcers. A one-year study. *Dig Dis Sci*, 1995, **40**, 1125-1131.
 34. Silverstein FE, Graham DY, Senior JR, et al.— Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis on nonsteroidal anti-inflammatory drugs. *Ann Intern Med*, 1995, **123**, 241-249.
 35. Van Weel C, Knottnerus JA.— Evidence-based interventions and comprehensive treatment. *Lancet*, 1999, **353**, 916-918.
 36. Garattini S.— Financial interests constrain drug development. *Science*, 1997, **275**, 287.
 37. Sackett DL, Rosenberg WMC, Gray JAM, et al.— Evidence-based medicine: what it is and what it isn't. *Br Med J*, 1996, **312**, 71-72.
 38. Djerassi C.— *Cantor's dilemma*. Penguin Books, New-York, 1991, 230 pages
 39. Broad W, Wade N.— *La souris truquée*. Seuil, Paris, 1987, 294 pages.
 40. Lovell DJ, Giannini EH, Reiff A, et al.— Etanercept in children with polyarticular juvenile rheumatoid arthritis. *N Engl J Med*, 2000, **342**, 763-769.
 41. Pisetky DS.— Tumor necrosis factor blockers in rheumatoid arthritis. *N Engl J Med*, 2000, **342**, 810-811.
 42. Poncin J.— Les enfants, aussi, peuvent être perclus de rhumatisme. *Le Soir*, 21 mars 2000, p 13.
 43. Rochon PA, Gurwitz JH, Sims RW, et al.— A study of manufacturer supported trials of non steroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med*, 1994, **154**, 157-163.
 44. Grimley Evans J.— Evidence-based and evidence-biased medicine. *Age Ageing*, 1995, **24**, 461-463.
 45. Halimi S.— Un journalisme de racolage. *Le Monde Diplomatique*, Août 1998, 12-13.
 46. Chomsky N.— Machines à endocriner. *Le Monde Diplomatique*, Août 1998, 13.
 47. Castells M.— *La société en réseaux. L'Ere de l'information*. Fayard, Paris, 1998, 616 pages.
 48. Davidoff F, Haynes RB, Sackett DL, Smith R.— Evidence-based medicine: a new journal to help doctors identify the information they need. *Br Med J*, 1995, **310**, 1085-86.
 49. Skrabanek et McCormick— "*Idées folles, idées fausses en médecine*". O Jacob, Paris, 1992, 206 pages.
 50. Masson C.— Pièges & conséquences "pratico-pratiques" de l'argument d'autorité en rhumatologie. Le syndrome de Non-Maladie, la situation du diagnostic hypostasié. *Synoviale*, 2000, **88**, 38-43.
 51. Schatzkin A, Lanza E, Corle D, et al.— Lack of effect of a low-fat, high fiber diet on the recurrence of colorectal adenomas. *N Engl J Med*, 2000, **342**, 1149-1155.
 52. Croft P.— Admissible evidence. *Ann Rheum Dis*, 1998, **57**, 387-389.
 53. Barner B.— Culture McWorld contre démocratie. *Le Monde diplomatique*, Août 1998, 14-15.
 54. Brzezinski Z.— *Le grand échiquier. L'Amérique et le reste du monde*. Hachette Littératures 2000, 273 pages.

55. Prevoe MLL, van't Hof MA, Kuper HH, et al.— Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum*, 1995, **38**, 44-48.
56. Weber JCP.— Epidemiology of adverse reactions to non-steroidal anti-inflammatory drugs, in Rainsford and Velo (Eds), *Advances in Inflammation Research*, vol 6, Raven Press, New York, 1984.
57. Kaplan C, Meier P.— Non parametric estimation from incomplete observations *J Am Stat Ass*, 1958, **53**, 547-581.
58. Amor B, Herson D, Cherot A, et al.— Polyarthrites rhumatoïdes évoluant depuis plus de 10 ans (1966-1978). Analyse de l'évolution et des traitements de 100 cas. *Ann Med Int*, 1981, **132**, 168-173.
59. Wolfe F, Hawley DJ, Cathey MA.— Termination of slow acting-rheumatic therapy in rheumatoid arthritis: a 14-years prospective evaluation of 1017 consecutive starts. *J Rheumatol*, 1990, **17**, 994-1002.
60. Wilkens FR.— Resolve: methotrexate is the drug of choice after NSAIDs in rheumatoid arthritis. *Sem Arthritis Rheum*, 1990, **20**, 76-80.
61. Chabot JM.— Evidence-based medicine : des plus et des moins (II). *Revue Prat*, 1999, **49**, 1549-1550.
62. Lyttleton RA.— The Gold Effect, in *Lying truths, a critical scrutiny of current beliefs and conventions*, Duncan R, Weston-Smith M, eds. Pergamp, Press, Oxford, 1979, 182-198.
63. Minois G.— *L'Eglise et la Science. Histoire d'un mal-entendu. I. De Saint Augustin à Galilée*. Fayard, Paris, 1990, 477 pages.
64. Kassirer JP.— The quality of care and the quality of measuring it. *N Engl J Med*, 1993, **329**, 11263-1265.
65. Anonymous.— Evidence-based medicine, in its place (Editorial). *Lancet*, 1995, **346**, 785.
66. Mandt D.— Can randomised trials inform clinical decisions about individuals. *Lancet*, 1999, **353**, 743-746.
67. Naylor CD.— Grey zones of clinical practice: some limits to evidence-based-medicine. *Lancet*, 1995, **345**, 840-842.

Les demandes de tirés à part sont à adresser au Pr. M. Malaise, Service de Rhumatologie, CHU Sart Tilman, 4000 Liège.