

Université de Sherbrooke

Utilisation de l'analyse des correspondances multiples et de la classification hiérarchique
pour modéliser la valeur pronostique des marqueurs présents chez les patients avec
polyarthrite d'installation récente

Par
Nathalie Carrier
Service de Rhumatologie
Département de médecine

Mémoire présenté à la Faculté de médecine
en vue de l'obtention du grade de
maître ès sciences (M.Sc.) en Sciences Cliniques (Biostatistiques)

Date : 22 avril 2008

Évaluateurs (¹ Directeur, ² Interne, ³ Externe):

Gilles Boire¹, M.D., M.Sc., Service de rhumatologie, Département de médecine
Bernard Colin¹, Ph.D., Département de mathématique
Jean-Marie Moutquin¹, M.D., M.Sc., Département d'obstétrique-gynécologie
Gina Bravo², Ph.D., Centre de recherche sur le vieillissement Université de Sherbrooke
Shengrui Wang³, Ph.D., Département d'informatique, Université de Sherbrooke

*Page 109 inexistante.
complet tel quel*



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-42940-2
Our file *Notre référence*
ISBN: 978-0-494-42940-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

■ ■ ■
Canada

Table des matières

Liste des tableaux.....	v
Liste des figures	vi
Table des abréviations.....	vii
1 – Introduction.....	1
1.1 Problématique	1
1.2 Objectif général et objectifs spécifiques	3
2 – Définitions et recension des écrits.....	5
2.1 Les définitions.....	5
2.1.1 Arthrite.....	5
2.1.2 La physiologie d’une articulation et problèmes associés à la polyarthrite rhumatoïde.....	5
Une articulation synoviale	6
Synovites.....	7
Érosions et pincements	8
2.1.3 Polyarthrite rhumatoïde	9
2.1.4 Maladie auto-immune	10
2.1.5 Facteurs déclencheurs	12
2.1.6 Diagnostic et pronostic	12
2.2 Recension des écrits.....	14
Dommages radiographiques.....	15
Capacité fonctionnelle	16
Activité de la maladie	17
Autoanticorps.....	18
Facteurs Génétiques	24
Vitesse de sédimentation et Protéine C-réactive.....	27
Globules blancs, polymorphonucléaires, hémoglobine	29
Âge et sexe.....	30
Tabagisme	32

Médication	34
3 - Matériel et Méthodes.....	37
3.1 Source des données.....	37
3.2 Devis de l'étude	37
3.3 Recrutement des sujets.....	38
3.3.1 Critères d'admissibilité.....	38
3.3.2 Échantillonnage.....	38
3.3.3 Taille de l'échantillon	38
3.3.4 Déroulement de l'étude.....	39
3.4 Variables et instruments de mesures.....	40
Variables socio-démographiques.....	41
Variables biologiques et sérologiques	41
Variables concernant l'état du patient.....	42
Médication	42
Capacité fonctionnelle	43
Indice d'activité de la maladie	43
Indice du degré des dommages articulaires	45
3.5 Stratégie d'analyse.....	46
3.5.1 Choix des types d'analyses statistiques	46
3.5.2 « <i>Data mining</i> »	47
Principe	48
3.5.2 Analyse des correspondances multiples.....	48
But et principe.....	48
Méthode et interprétation.....	49
3.5.2 La classification hiérarchique ascendante (agglomérative).....	51
But et principe (mesure de distance & indice d'agrégation).....	52
3.5.3 Application concrète aux données	56

3.4	Considérations éthiques	58
4	Résultats	59
4.1	Patients à l'étude et leurs caractéristiques	59
4.2	Résultats de l'analyse des correspondances multiples.....	64
	Interprétation graphique.....	66
4.3	Résultats de la classification hiérarchique	70
5	Discussion et Conclusion.....	87
5.1	Discussion des résultats	87
5.2	Forces et limites	93
5.2	Perspectives d'avenir	96
6	Remerciements.....	98
7	Bibliographie.....	100
8	Annexes	108
Annexe 1.	Formulaire du médecin (EUPA)	110
Annexe 2.	Formulaire de l'infirmière (EUPA)	112
Annexe 3.	Formule d'évaluation fonctionnelle.....	114

Liste des tableaux

Tableau 1.	Critères de classification de la polyarthrite rhumatoïde selon l'ACR	13
Tableau 2.	Résumé des variables et instruments de mesure	41
Tableau 3.	Patients à l'étude	60
Tableau 4.	Comparaison des caractéristiques à l'inclusion entre les patients inclus dans l'analyse et ceux perdus au suivi	61
Tableau 5.	Caractéristiques à l'inclusion (n=211)	62
Tableau 6.	Caractéristiques à l'inclusion et au suivi (n=211)	63
Tableau 7.	Décomposition de l'Inertie et du Khi-Carré	65
Tableau 8.	Coordonnées factorielles et contributions de chacune des variables à la création des axes	68
Tableau 9.	Comparaison des groupes résultant de deux classifications différentes: méthode utilisant les 4 premières coordonnées factorielles et méthode utilisant la métrique du Khi-carré sur les coordonnées originales	74
Tableau 10	Distribution des dommages radiographiques (score d'érosions) selon les 5 sous-groupes	78
Tableau 11.	Distribution de l'incapacité du patient dans le temps selon les 5 sous-groupes	82
Tableau 12.	Distribution de l'activité de la maladie dans le temps selon les 5 sous-groupes	83
Tableau 13.	Distribution des données socio-démographiques, inflammatoires et immunitaire selon les 5 sous-groupes; fréquence (%)	84
Tableau 14.	Distribution des traitements selon les 5 sous-groupes	85
Tableau 15.	Distribution des données sur la génétique (HLA-DRB1) selon les 5 sous-groupes	86

Liste des figures

Figure 1.	Articulation synoviale.....	6
Figure 2.	Articulation atteinte de PR.....	7
Figure 3.	Pincement articulaire	9
Figure 4.	Complexe tri-moléculaire : HLA / peptide dérivé de l'antigène / récepteur cellule T	11
Figure 5.	Organisation des gènes du HLA sur le chromosome 6.....	25
Figure 6.	Graphiques de l'ACM, dimensions 1 et 2, représentation des variables ..	66
Figure 7.	Graphiques de l'ACM, dimensions 1 et 3, représentation des variables ..	67
Figure 8.	Graphiques de l'ACM, dimensions 1 et 4, représentation des variables ..	67
Figure 9.	Dendrogramme issu de la classification hiérarchique	72
Figure 10.	Dendrogramme en utilisant la métrique du Khi-carré	72
Figure 11.	Dendrogramme en utilisant tous les facteurs de l'ACM.....	73
Figure 12.	Graphique de l'ACM, représentation des individus du groupe 1 et 5	75
Figure 13.	Graphique de l'ACM, représentation des individus des sous-groupes 2, 3 et 4.....	75
Figure 14.	Résumé des principales variables définissant chacun des niveaux du dendrogramme issu de la classification hiérarchique	77
Figure 15.	Évolution des dommages radiographiques de l'inclusion à 30 mois.....	78
Figure 16.1	Représentation individuelle de la différence d'érosions entre l'inclusion et 30 mois.....	80
Figure 16.2	Représentation individuelle des érosions à l'inclusion.....	81
Figure 16.3	Représentation individuelle des érosions à 30 mois	81
Figure 17.	Illustration des articulations évaluées par le médecin.....	111

Table des abréviations

Termes cliniques:

ACR : American College of Rheumatology

AINS: Anti-inflammatoires non stéroïdiens

Anti-CCP : Anti-cyclic citrullinated peptide antibodies, Autoanticorps

Anti-Sa : Anti-Sa/citrullinated vimentin, Autoanticorps

CHUS : Centre hospitalier universitaire de Sherbrooke

CRC : Centre de recherche clinique Étienne-Le Bel

CRP : Protéine C-Réactive

DAS : Disease activity score

DERAA : Regroupement d'allèles du HLA-DRB1 associé à une maladie moins sévère

DMARD : Disease Modifying Antirheumatic Drug, ou Agents anti-rhumatismaux à action lente

ELISA : Enzyme-Linked ImmunoSorbent Assay

FR : Facteur rhumatoïde, Autoanticorps

HLA : Human Leukocyte Antigen, ou antigène des leucocytes humains

M-HAQ : Modified Health Assessment Questionnaire

PMN: polymorphonucléaire

PR : Polyarthrite rhumatoïde

SE : Shared Epitope, regroupement d'allèles du HLA-DRB1 associé à la présence d'anticorps

VS : Vitesse de sédimentation

Termes mathématiques:

ACM : Analyse des correspondances multiples

ECD : Extraction de Connaissances à partir de données, plus connu sous nom de Data Mining

E.T. : Écart-type

EIQ : Étendue interquartile

Med. : Médiane

Moy. : Moyenne

Résumé

Utilisation de l'analyse des correspondances multiples et de la classification hiérarchique pour modéliser la valeur pronostique des marqueurs présents chez les patients avec polyarthrite d'installation récente

Par Nathalie Carrier, Service de Rhumatologie, Département de médecine
Mémoire présenté à la Faculté de médecine en vue de l'obtention du grade de
maître ès sciences (M.Sc.) en Sciences Cliniques (Biostatistiques)
Évalué par : Gilles Boire, Bernard Colin, Jean-Marie Moutquin,
Gina Bravo, Shengrui Wang

Ce projet de maîtrise a comme objectif d'évaluer, dans une cohorte de patients avec polyarthrite inflammatoire débutante, l'existence de différents sous-groupes de patients classifiables selon leurs caractéristiques cliniques lors de l'inclusion dans l'étude.

Les patients sont recrutés au Centre hospitalier universitaire de Sherbrooke (CHUS). Ils doivent être âgés de plus de 18 ans, se présenter avec au moins 3 articulations atteintes d'inflammation articulaire, et ce depuis plus d'un mois et moins de 12 mois. Les patients exclus de l'étude sont ceux atteints d'arthrite microcristalline (ex : Goutte), de collagénose (maladie du tissu conjonctif), ou d'arthrite septique, ainsi que les patients incapables de consentir. De plus, les patients sont suivis sur une période de 5 ans et réévalués tous les ans. Nous sommes donc en présence d'un dispositif d'observation avec cohorte prospective longitudinale. Chacun des patients est traité selon les traitements disponibles cliniquement et adaptés à sa situation clinique.

Une classification hiérarchique a été effectuée chez les 211 premiers patients consécutifs pour lesquels l'information clinique à l'inclusion et à 30 mois était complète. Les variables de l'inclusion qui ont été incorporées dans l'analyse statistique sont : les données démographiques (âge, sexe), certains autoanticorps, le décompte d'articulations gonflées, les paramètres d'inflammation systémique, les paramètres sanguins, si les patients remplissent au moins 4 des 7 critères de classification de la polyarthrite rhumatoïde (PR) selon l'American College of Rheumatology (ACR), la présence de traitements lors de l'inclusion et les dommages articulaires (érosions osseuses et pincements articulaires) sur les radiographies des mains et des pieds.

L'analyse statistique a permis l'identification de 5 sous-groupes de patients : 2 avec une meilleure issue clinique (Groupes 1 et 4), 2 avec une issue intermédiaire (Groupes 2 et 3) et 1 avec une très mauvaise issue radiologique et clinique (Groupe 5). Les patients des groupes 1 et 4 ont très peu d'anticorps ou d'érosions, peu d'inflammation systémique, et sont plus jeunes (63% et 47%, âge ≤ 50 ans

respectivement). Aucun des patients du groupe 4 ne remplit les critères de la classification de la PR, contrairement au groupe 1 (0% vs 100% respectivement). Les patients des groupes 2 et 3 ont généralement des dommages radiographiques à l'inclusion et leur nombre d'érosions osseuses augmente à 30 mois. Les patients du groupe 2 se distinguent par un très faible taux d'anticorps (entre 0% et 17%), tout en présentant fréquemment une inflammation très importante et de l'anémie. Le groupe 3 quant à lui a modérément d'anticorps (entre 40 à 50%), un taux élevé de globules blancs circulants et les patients sont généralement traités très tôt dans le développement de la maladie. Finalement, le groupe 5 est le groupe avec un nombre très important d'anticorps, beaucoup de dommages radiographiques à l'inclusion et à 30 mois, les critères de l'ACR sont généralement remplis et il y a présence d'inflammation systémique.

Comme les résultats le suggèrent, cette analyse semble compatible avec d'autres observations qui suggèrent que l'arthrite débutante représente plus d'une maladie.

Le groupe 3 ressemble au groupe 5 pour les caractéristiques de l'inclusion. Ceci pourrait indiquer que de traiter très tôt peut, dans certains cas, modifier l'expression des anticorps et l'évolution de la maladie. D'autres analyses plus approfondies seront nécessaires pour améliorer la classification des divers sous-groupes. Éventuellement, un modèle de régression multivarié pourra être utilisé pour prédire l'évolution de la maladie chez de nouveaux patients à partir de leurs caractéristiques initiales. Un tel modèle pourrait alors être très utile en clinique pour établir le bon traitement à administrer aux patients en considérant leurs caractéristiques cliniques lors de leur première visite avec le rhumatologue.

Mots clés : Polyarthrite inflammatoire débutante, classification hiérarchique, analyse des correspondances multiples, autoanticorps, érosions articulaires

1 – Introduction

1.1 Problématique

La polyarthrite rhumatoïde (PR) est un véritable problème de santé publique, tant par le nombre croissant, en raison de la population vieillissante, des personnes concernées que par ses répercussions sur la qualité de vie des patients.

Il existe des centaines de maladies pouvant donner des manifestations articulaires pathologiques, mais seulement certaines sont de type inflammatoire et appelées arthrites inflammatoires chroniques. Selon la Société d'arthrite du Canada, les arthrites inflammatoires chroniques touchent 2% de la population, dont la moitié est atteinte de PR, ce qui représente 300 000 Canadiens.

La polyarthrite rhumatoïde a de fortes répercussions sur la qualité de vie des patients pour de multiples raisons. Tout d'abord, elle est la forme la plus invalidante de l'ensemble des formes chroniques d'arthrite. En effet, cette maladie est une des causes majeures d'invalidité, de morbidité et de mortalité prématurée dont la prévention est possible si elle est diagnostiquée rapidement et traitée adéquatement. Sinon, environ 50% des malades arrêtent leurs activités professionnelles 5 ans après le début de la maladie et leur durée de vie est réduite en moyenne de 5 à 10 ans. De plus, la PR aura évidemment un impact important financièrement sur les familles des personnes atteintes (MacGregor, Silman 2003).

Les facteurs déclencheurs de la maladie sont encore inconnus des chercheurs et cliniciens, mais ils sont clairement multiples. Certaines composantes pathogéniques

pourraient être en cause, dont les facteurs génétiques, l'environnement (bactéries, virus...) ou un dérangement du système immunitaire. D'autres origines pourraient être en cause, dont les troubles hormonaux, le stress, et les facteurs alimentaires. D'autres facteurs potentiels restent à découvrir.

En plus de comprendre difficilement le développement de la PR, il n'existe encore aucun moyen de guérir cette maladie. Heureusement, si elle est diagnostiquée rapidement et bien traitée, on peut la maîtriser et éviter l'apparition de lésions articulaires graves, rétablir et conserver la capacité fonctionnelle et éviter l'invalidité. On pourra alors envisager une rémission qui se définit généralement par une absence d'inflammation articulaire et systémique, de douleur, de gonflements articulaires, de manifestations extra-articulaires et évidemment une absence de progression des dommages articulaires. Une fois diagnostiquée, certaines mesures peuvent être prises pour maîtriser le développement de ces complications, dont la prise de médication, l'exercice physique et un régime alimentaire équilibré. Lorsque la douleur devient intolérable, ou s'il y a perte de mouvement et limitation fonctionnelle, une procédure chirurgicale peut être considérée (La Société d'arthrite 2007).

En pratique, il est difficile de classifier ou de pronostiquer la PR par rapport aux autres formes d'arthrite inflammatoire durant la première année de son développement. En effet, il y a alors peu de signes caractéristiques visibles à l'examen clinique, les symptômes ne sont pas spécifiques à la PR et il y a de multiples issues cliniques possibles. Les maladies inflammatoires peuvent entrer en rémission, persister sans progression des dommages radiographiques ou se prolonger en étant accompagnées d'une destruction plus ou moins rapide des articulations. La maladie articulaire initiale peut se

différencier et développer les caractéristiques de la PR (ex. érosions osseuses symétriques et progressives avec ostéopénie périarticulaire dans une distribution typique de la PR) ou d'une autre maladie chronique à début inflammatoire comme l'arthrite réactive, la sarcoïdose, l'arthrite psoriasique, la spondylarthropathie et bien d'autres. À la section 2.1.6 Diagnostic et Pronostic, une description plus détaillée de ces différentes maladies est présentée. Souvent, seule l'évolution permettra de confirmer le diagnostic, mais il sera alors trop tard pour intervenir avec le maximum d'efficacité.

Suite à cette problématique, il est donc important de mieux comprendre les stades initiaux du développement de la maladie et d'identifier des facteurs qui pourraient aider à prédire l'issue clinique chez le patient.

1.2 Objectif général et objectifs spécifiques

L'objectif de ce projet est d'**explorer** les relations entre les caractéristiques cliniques des patients avec polyarthrite inflammatoire afin de **mettre en évidence** des facteurs de risque potentiels pouvant **expliquer** l'évolution de la maladie.

De façon plus spécifique, la classification hiérarchique ascendante (agglomérative) sera utilisée pour étudier s'il existe différents **sous-groupes** au sein de l'échantillon, selon les symptômes et caractéristiques cliniques du patient à l'**inclusion**. Par la suite, dans l'éventualité où des sous-groupes seront mis en évidence, l'**évolution** de la maladie dans le temps sera étudiée et nous tenterons de déterminer s'il existe des **facteurs de risque** expliquant cette évolution différente ou la formation de ces sous-groupes.

Dans la prochaine section, nous expliquerons de façon détaillée la physiologie de la maladie pour permettre au lecteur de bien comprendre les termes cliniques spécifiques utilisés dans le domaine. Par la suite, nous présenterons un certain nombre de publications et de recherches faites sur la PR ou la polyarthrite débutante afin de mettre en évidence ce qui a été trouvé jusqu'à maintenant sur le sujet. Enfin, nous décrirons la méthodologie qui a été utilisée pour répondre à ces objectifs en détail.

2 – Définitions et recension des écrits

2.1 Les définitions

2.1.1 Arthrite

Le terme « arthrite » provient du mot grec «arthros» pour articulation et du suffixe «ite» pour indiquer l'inflammation, ce qui signifie que c'est une maladie inflammatoire ou dégénérative touchant les articulations. Parmi les centaines de formes d'arthrite, l'arthrose domine puisqu'elle regroupe environ 50% de tous les cas. La polyarthrite rhumatoïde (PR) est la plus invalidante de toutes les formes, et elle représente environ 50% des arthrites inflammatoires chroniques. C'est l'ensemble des arthrites inflammatoires débutantes qui nous intéresse plus particulièrement, ensemble dont la PR fait partie (La Société d'arthrite 2007).

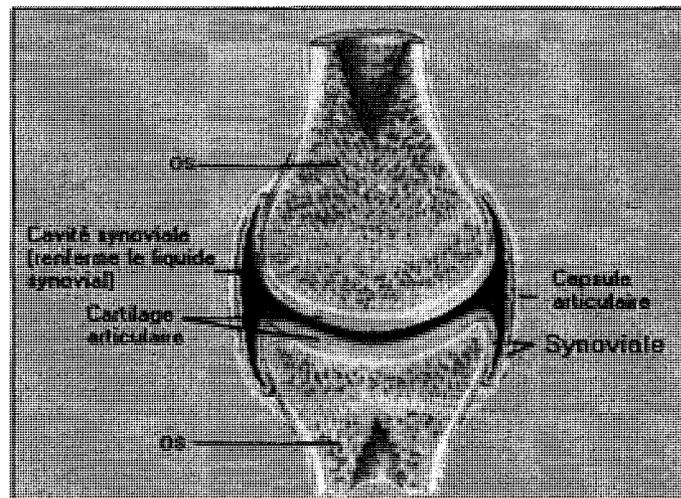
2.1.2 La physiologie d'une articulation et problèmes associés à la polyarthrite rhumatoïde

La PR est une maladie touchant les articulations, mais plus particulièrement les articulations synoviales. Avant d'approfondir la définition de cette maladie, nous allons définir quelques termes qui vont nous aider à mieux comprendre, par la suite, la physiopathologie.

Une articulation synoviale

L'image ci-dessous représente bien ce qu'est une articulation synoviale, donc le type d'articulations touchées par la PR. Elle se définit comme une articulation d'une grande mobilité dont les surfaces sont recouvertes de cartilage. L'articulation est composée d'une cavité articulaire remplie d'un liquide appelé synovie ou liquide synovial qui est formée par la membrane interne de l'articulation, appelée membrane synoviale ou plus simplement synoviale (Tortora, Grabowski & Parent 1994b). Un grand nombre d'articulations sont dites synoviales, en particulier les articulations des mains et des pieds, contrairement aux articulations discovertébrales qui ne contiennent pas de synoviales.

Figure 1. Articulation synoviale



Tiré de : www.genou.com/synoviale/synoviale.htm

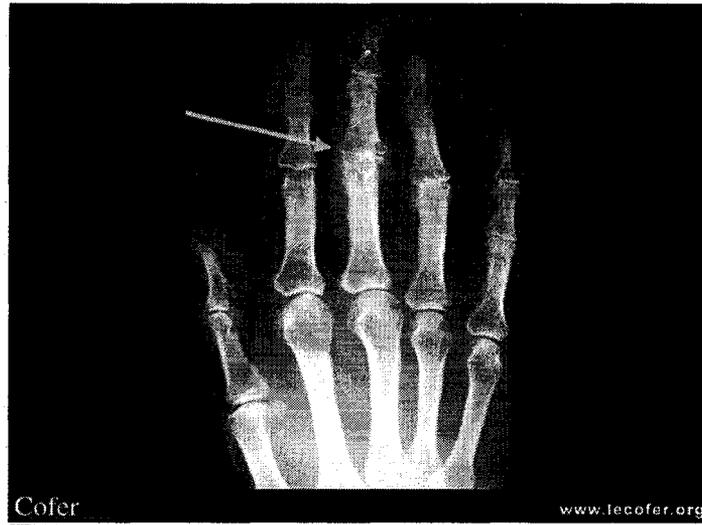
Érosions et pincements

Comme mentionné précédemment, l'inflammation synoviale persistante provoque, après un certain temps, une destruction articulaire qui entraîne érosions et pincements articulaires. Ces destructions sont visibles sur les radiographies des articulations.

Les **érosions** se définissent comme des lésions articulaires sous forme de destruction localisée du cartilage ou de l'os causées par le pannus synovial et les enzymes protéolytiques contenues dans le liquide synovial inflammatoire. Les tendons et les ligaments entourant l'articulation peuvent également être touchés, ce qui provoque une instabilité et une déformation articulaire. La figure 2 ci-dessous représente une articulation normale et une articulation atteinte de PR. Sur l'image de droite, les érosions du cartilage et de l'os sont indiquées à l'aide de flèches.

Le **pincement** résulte de l'amincissement graduel du cartilage. On l'observe donc sur les radiographies par un rétrécissement de l'espace entre les os. Le cartilage (qui est radiotransparent, i.e. invisible sur les radiographies) est le tissu recouvrant les surfaces de frottement des articulations et il amortit les chocs dus au mouvement. Sa surface est imbibée de liquide synovial qui permet un glissement facile et sans douleur entre les articulations, ce qui assure la mobilité de l'articulation. L'amincissement du cartilage provoque une plus grande friction entre les surfaces articulaires, souvent accompagnée de douleur lors de l'utilisation de l'articulation. Le pincement est représenté à l'aide d'une flèche sur une radiographie de la main à la figure 3 (Alparslan, Yu & Weissman 2005; Haraoui, Keystone 2001).

Figure 3. Pincement articulaire



2.1.3 Polyarthrite rhumatoïde

La polyarthrite rhumatoïde (PR) se définit comme une maladie inflammatoire chronique qui attaque principalement les articulations synoviales. Le début de la maladie est généralement insidieux avec une atteinte articulaire progressive qui cause une inflammation simultanée de plusieurs articulations. On peut également la définir comme une maladie auto-immune caractérisée par une inflammation systémique et synoviale (synovites), une atteinte symétrique, une tendance à provoquer des érosions de l'os ou du cartilage et dans certains cas le développement de manifestations extra-articulaires (La Société d'arthrite 2007).

La PR touche principalement les articulations des mains et des pieds, mais peut également toucher le reste des articulations synoviales (ex. genou, hanche, articulation apophysaire intervertébrale du cou) tandis que la colonne dorsolombaire est généralement épargnée. L'inflammation peut également s'attaquer aux membranes qui entourent le

cœur et les poumons, ainsi que toucher les glandes lacrymales et salivaires provoquant une sécheresse des yeux et de la bouche (La Société d'arthrite 2007). Chez près de 30% des personnes atteintes, la maladie s'accompagne de nodules ou bosses sous-cutanées dans les régions périarticulaires, tels les coudes et genoux.

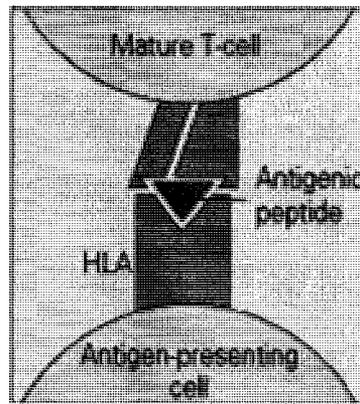
2.1.4 Maladie auto-immune

La PR étant une maladie auto-immune, il est donc intéressant de définir ce mécanisme pathologique. La maladie se développe lorsque le système immunitaire du patient s'attaque à certains constituants de son propre corps comme s'ils étaient des substances étrangères. Le corps produit alors des anticorps dirigés contre lui-même appelés des autoanticorps. Ceux-ci sont fabriqués par des cellules appelées lymphocytes B qui elles-mêmes sont stimulées par les lymphocytes T. Il y a donc un problème de régulation du système immunitaire (Tortora, Grabowski & Parent 1994a).

Plus spécifiquement, certaines molécules, dont le HLA, jouent un rôle important dans le mécanisme immunitaire. Le HLA provient de l'anglais «*Human leukocyte antigen*» qui se traduit par antigène des leucocytes humains. Ces molécules de HLA présentes dans la cellule vont entrer dans un compartiment intracellulaire où elles vont se charger d'un peptide d'une longueur de 12 à 20 acides aminés. Ce peptide est lui-même un dérivé d'un antigène. L'antigène est soit une composante de l'individu ou un agresseur comme un microbe. Le complexe constitué de molécules de HLA et de l'antigène va alors être transporté à la surface de la cellule. Les lymphocytes T porteurs d'un récepteur spécifique (protéine réceptive) vont chercher les cellules présentatrices d'antigènes qui portent le complexe HLA-peptide spécifique qu'elles ont été éduquées à reconnaître

(Figure 4 ci-dessous). Normalement, dans le cas d'un complexe HLA-peptide du soi, le lymphocyte T ne devrait pas réagir (i.e. qu'il devrait être tolérant par rapport à cet antigène), mais il le fait chez les patients atteints de maladies auto-immunes. Dans le cas d'un peptide dérivé d'un agresseur, le lymphocyte T sera prêt à réagir pour tenter de l'éliminer. Il va alors se transformer, se multiplier et appeler d'autres lymphocytes et d'autres cellules à l'aide. De plus, les lymphocytes T vont activer les lymphocytes B, qui vont fabriquer des anticorps qui aideront à détruire les antigènes ou les agresseurs qui les expriment. Dans le cas des maladies auto-immunes, le même processus se met en branle par suite d'un défaut dans la régulation de l'immunité, et l'activation du système immunitaire conduira à de l'inflammation chronique et à des dommages structuraux (Nepom, Nepom 2003).

Figure 4. Complexe tri-moléculaire : HLA / peptide dérivé de l'antigène / récepteur cellule T



Tiré de : www.rheumtext.com – Hochberg et al (eds)

2.1.5 Facteurs déclencheurs

Les facteurs déclencheurs sont encore inconnus, mais ils sont probablement multiples. Certaines composantes pathogéniques pourraient être en cause, dont les facteurs génétiques, l'environnement (bactéries, virus, tabac...) ou un dérangement du système immunitaire. D'autres origines pourraient expliquer son apparition comme des troubles hormonaux, le stress, des facteurs alimentaires, et potentiellement d'autres moins connues (La Société d'arthrite 2007).

2.1.6 Diagnostic et pronostic

Cliniquement, il est difficile de diagnostiquer la maladie et de prévoir son évolution, particulièrement au tout début de ses manifestations cliniques. Le diagnostic repose donc principalement sur les données cliniques. Les signes suggestifs de la polyarthrite rhumatoïde sont, surtout lorsqu'ils durent pendant plusieurs semaines (au moins 6 semaines pour remplir les critères de classification de la PR selon l'ACR, voir Tableau 1, p.13): les articulations enflées, chaudes, et parfois rouges; la présence de raideur matinale aux articulations de plus de 60 minutes; avoir des douleurs dans plus de 3 articulations en même temps; avoir des douleurs articulaires lorsque l'articulation est au repos; et finalement être atteint de façon symétrique. Le diagnostic repose aussi sur des caractéristiques biologiques. Dans le sang, on constate une augmentation des marqueurs de l'inflammation (vitesse de sédimentation et/ou protéine C-réactive) et on cherche également la présence du facteur rhumatoïde (FR) dans le sérum du patient. Pour ce qui est des dommages radiographiques, ils apparaissent généralement plus tardivement

lorsque l'inflammation de la synoviale a eu le temps d'induire des lésions cartilagineuses et/ou osseuses.

Jusqu'à maintenant, ce sont les critères de classification de l'ACR (American College of Rheumatology) créés en 1987 qui sont les plus utilisés pour classer la PR. Un patient répond aux critères de l'ACR s'il possède au moins 4 critères sur les 7 décrits ci-dessous (Tableau 1) et s'il a ces symptômes depuis plus de 6 semaines.

Tableau 1. Critères de classification de la polyarthrite rhumatoïde selon l'ACR

1	Raideur matinale de plus de 60 minutes
2	Avoir au moins 3 articulations touchées
3	Articulations des mains et des poignets atteintes
4	Atteinte symétrique des membres
5	Présence de nodules sous-cutanés
6	Présence du facteur rhumatoïde dans le sérum du patient
7	Changements radiographiques typiques

En pratique, les critères de l'ACR sont sensibles et spécifiques pour une maladie bien installée, mais ils sont d'une utilité limitée pour classer les polyarthrites d'installation récente. C'est donc un problème majeur car, il est important de confirmer le diagnostic le plus tôt possible pour intervenir avec le plus d'efficacité possible.

La principale raison pour expliquer qu'il est si difficile de diagnostiquer une PR au début de son évolution est que plusieurs types d'arthrites chroniques ont un début inflammatoire semblable à la PR. Par exemple, l'arthrite réactive se manifeste par de la douleur, de la raideur, de la rougeur ou de l'enflure tout comme la PR, mais ces manifestations résultent d'une réponse immunitaire à une infection bactérienne qui s'est produite ailleurs que dans l'articulation. Plusieurs autres maladies articulaires inflammatoires chroniques se présentent au début de façon très semblable à la PR mais

évoluent très différemment de la PR. Ces maladies comprennent entre autres, la spondylarthropathie qui touche principalement les articulations de la colonne vertébrale en plus des articulations périphériques, l'arthrite psoriasique qui se manifeste par des éruptions cutanées et de l'inflammation articulaire, et la sarcoïdose qui atteint généralement les poumons en plus des articulations. Par contre, ces maladies sont difficiles à identifier au début car elles présentent généralement de l'inflammation dans une ou plusieurs articulations périphériques, et ce pendant des mois ou des années avant que les autres signes et symptômes apparaissent. Suite à ces observations, on constate l'importance d'étudier si, dès le départ, il y a des sous-groupes de patients déjà identifiables en définissant des marqueurs pronostiques permettant de prévoir l'évolution de la maladie et d'administrer le traitement approprié.

2.2 Recension des écrits

Afin d'étudier l'évolution de la maladie, la plupart des chercheurs dans le domaine utilisent généralement un devis de type cohorte d'observations, le seul devis permettant d'étudier sur une longue période de temps un nombre considérable de patients. Ce type de cohorte permet de recueillir de l'information sur un certain nombre de marqueurs potentiels et d'étudier l'évolution de la maladie. La plupart des chercheurs tentent donc de prédire la sévérité de la maladie en cours d'évolution, généralement sur une période de 2 à 6 ans, en considérant un certain nombre de marqueurs potentiels présents à l'inclusion. Dans la plupart des cas, des modèles de régression sont utilisés. Nous ne nous attarderons donc pas à rediscuter du type de devis et du type d'analyse statistique dans la suite de la recension, sauf pour les cas inhabituels où nous élaborerons un peu plus.

Les différentes issues cliniques étudiées jusqu'à maintenant sont : la destruction ou la progression de la destruction des articulations telle que visualisée sur des radiographies simples ou par d'autres méthodes d'imagerie, une persistance de la maladie, l'incapacité fonctionnelle ou une rémission. L'atteinte radiologique est celle qui est la plus étudiée dans les articles puisque l'atteinte des articulations représente une partie considérable des dommages causés par la PR. Prévenir et diminuer les dommages articulaires sont donc des buts importants de la recherche. La persistance de la maladie et l'incapacité sont plus rarement observés dans la littérature. La rémission quant à elle est généralement utilisée dans les études évaluant l'effet de la médication.

Une revue de l'ensemble des marqueurs potentiels les plus couramment utilisés sera présentée dans cette section, et ce principalement par rapport aux dommages articulaires.

Dommages radiographiques

Dans la phase initiale de la PR, le développement inflammatoire de la synoviale, la synovite, est la cause des douleurs et des déformations. Très souvent, la synovite disparaît ou diminue grâce au traitement médical, généralement à base d'anti-inflammatoires ou d'infiltration locale de corticostéroïdes combinés à des antirhumatismaux à action lente. Au contraire, si la synovite résiste au traitement médical, elle peut entraîner des ruptures tendineuses, des distensions ligamentaires ou des destructions articulaires. Les déformations surviennent ainsi dans 75% des cas au niveau des mains et des poignets et touchent aussi les pieds. Des changements radiographiques ont lieu au cours des deux années qui suivent le début de la maladie chez 50 % à 70 %

des patients atteints de PR (van der Heijde 1995). Il est donc important d'établir un plan de surveillance des destructions par un bilan radiographique le plus tôt possible pour détecter et tenter de contrôler l'évolution. La progression radiographique de la maladie est considérée comme la valeur de référence pour mesurer l'efficacité des traitements et apparaît la meilleure issue pour étudier la maladie à court et à moyen terme (van der Heide et al. 1995; van Riel et al. 1995; van Zeben, Breedveld 1996). L'état initial des articulations est important à considérer pour étudier l'évolution de la maladie. La présence de dommages articulaires importants sur les radiographies articulaires tôt après l'installation de la maladie est généralement associée à une évolution sévère (Boire et al. 2005; Combe et al. 2001; Forslind et al. 2004; Kroot et al. 2000 Aug).

Capacité fonctionnelle

Beaucoup de patients finissent par avoir des difficultés à se livrer aux activités de la vie quotidienne. Le pronostic à long terme chez les patients atteints de la PR est défavorable. Un grand nombre d'entre eux sont condamnés à une invalidité croissante et à un décès prématuré. Généralement, les patients ont une maladie chronique mais fluctuante et l'incapacité fonctionnelle est variable non seulement chez le même individu mais d'un individu à l'autre. Il est possible d'avancer qu'après 10 à 15 ans d'évolution, très peu de patients, environ 1 patient sur 5 seulement, ne présentera aucun handicap ni déformation des articulations. Après 10 ans d'évolution, 1 patient sur 2 présentera une incapacité à son travail, et si plus de 20 articulations sont touchées, le patient risque de développer des anomalies articulaires importantes et un handicap sévère (Harris 2005).

Comme pour les dommages radiographiques, la capacité fonctionnelle initiale du patient aura aussi un impact sur l'évolution de sa qualité de vie ainsi que sur les dommages radiographiques. Mottonen et son équipe, en 1998, ont conclu que l'incapacité fonctionnelle est associée à une augmentation du risque d'avoir des dommages radiographiques à 1 an (OR=2.8, IC(95%)=1.4-5.6) et à une augmentation du risque de progression rapide (Mottonen et al. 1998). En effet, l'incapacité provient souvent d'une inflammation persistante qui finit à long terme par détruire les tendons et les articulations. Dans la même étude, on observe une forte corrélation entre les dommages radiographiques et l'incapacité fonctionnelle. Une autre étude plus récente a montré que les meilleurs marqueurs d'une incapacité sont la douleur, l'activité de la maladie, les dommages articulaires, et la durée de la maladie (Molenaar, Voskuyl & Dijkmans 2002). Par contre, d'autres auteurs n'observent aucune association entre l'incapacité et les dommages radiographiques, en particulier en début de maladie (Boire et al. 2005; Combe et al. 2001; Forslind et al. 2004).

Activité de la maladie

Le suivi de l'activité de la maladie peut se faire à l'aide de différents scores. Le plus utilisé en pratique clinique est le DAS28, calculé à partir de 4 paramètres : l'indice articulaire (nombre d'articulations douloureuses), l'indice synovial (nombre d'articulations gonflées), l'activité de la maladie évaluée par le patient sur une échelle de 0 à 100, et la vitesse de sédimentation (ou la protéine C-réactive, un autre marqueur sérique d'inflammation). Le décompte des articulations douloureuses ou enflées se fait en considérant les articulations métacarpo-phalangiennes et interphalangiennes proximales

des mains, les poignets, les coudes, les épaules et les genoux (voir section 3.4 Indice d'activité de la maladie pour plus de détails). Ce score est un indice pouvant expliquer l'amélioration de l'activité inflammatoire de la PR. De plus, le DAS28 semble être un bon marqueur pronostique des dommages articulaires et de l'incapacité fonctionnelle dans certaines études (Drossaers-Bakker et al. 2002; Kroot et al. 2000; Molenaar, Voskuyl & Dijkmans 2002; Mottonen et al. 1998). Par contre, certains auteurs n'obtiennent pas des résultats significatifs pour cette variable (Boire et al. 2005; Combe et al. 2001; Forslind et al. 2004), particulièrement lorsqu'il y a diminution de l'activité inflammatoire par les traitements. De plus, le DAS28 est un outil reconnu et utilisé dans les essais cliniques comme critère de base pour évaluer les réponses aux traitements. Il existe en effet un niveau critique de ce score à partir duquel on considère la maladie en rémission (Allaart et al. 2006; Quinn et al. 2005).

Autoanticorps

Depuis quelques années, certains autoanticorps ont été détectés dans le sérum des patients avec PR et leur présence pourrait être utile pour prédire les mauvaises issues cliniques chez les patients avec PR d'installation récente. Les principaux autoanticorps étudiés sont le facteur rhumatoïde (FR), l'«*Anti-cyclic citrullinated peptide*» (anti-CCP) et l'anti-Sa. D'autres autoanticorps pourraient être de bons marqueurs mais ne sont pas utilisés couramment pour des raisons techniques, puisque leur détection requiert des méthodes laborieuses pour lesquelles il n'y a pas présentement d'applications commerciales et pour lesquelles la standardisation entre les laboratoires est déficiente.

Généralement, le FR et l'anti-CCP sont les autoanticorps les plus utilisés en clinique, à cause de leur sensibilité et spécificité et de leur disponibilité commerciale.

Facteur rhumatoïde (FR)

Le FR a été mis en évidence en 1940 par le Norvégien Waaler dans le sang de sujets atteints de PR. Le FR est utilisé pour aider à diagnostiquer la polyarthrite rhumatoïde et pour la distinguer des autres formes d'arthrites avec des symptômes semblables comme la douleur, l'inflammation et la raideur articulaire. Un test négatif n'indique pas une absence de la maladie puisque environ 20% à 30% des patients atteints ne produisent pas de FR (i.e. qu'ils sont séronégatifs). Des résultats de FR positifs peuvent aussi être vus chez un petit pourcentage de personnes saines (en particulier chez les sujets âgés) et chez des patients présentant des conditions comme : une endocardite, un lupus érythémateux disséminé, la tuberculose, la syphilis, une sarcoïdose, un cancer, une infection virale, une maladie inflammatoire chronique du foie, du poumon ou du rein. Par ailleurs, le FR n'est pas utile pour diagnostiquer des conditions autres que la polyarthrite rhumatoïde.

Le FR présente certaines limitations pour diagnostiquer la PR. En effet, le FR n'est pas très sensible, soit autour de 50% à 60% pour les arthrites débutantes, mais assez spécifique, soit aux environs de 80% à 95% selon les études (moyenne de plus de 90%) pour les arthrites débutantes qui évolueront en polyarthrite rhumatoïde (Jansen et al. 2002; Niewold, Harrison & Paget 2007; Schellekens et al. 2000; Vittecoq et al. 2005). En combinaison avec d'autres signes et symptômes de la maladie, le FR peut être utilisé comme aide au diagnostic et au pronostic.

Selon plusieurs chercheurs, le FR est un bon prédicteur des dommages et de la progression radiographique, de sorte qu'il constitue un marqueur de référence depuis de nombreuses années. Par contre, les auteurs n'arrivent pas tous aux mêmes conclusions puisqu'il y a une grande controverse à savoir lequel des autoanticorps entre le FR et l'anti-CCP devrait être privilégié en pratique clinique. Nous allons d'ailleurs présenter certains éléments de cette controverse retrouvés dans plusieurs articles.

Combe (2001) a observé dans sa cohorte de patients que le FR était l'un des meilleurs prédicteurs des dommages et de la progression articulaire à 3 ans (**Dommages** : OR=3.5; $p<0.001$; **Progression** : OR=5.5; $p<0.001$) (Combe et al. 2001). Malheureusement, l'effet de l'anti-CCP n'a pas été analysé lors de cette étude. L'équipe de Kroot (2000) a aussi constaté que le FR est un des meilleurs prédicteurs des dommages et de la progression radiographique à 3 ans et à 6 ans (Kroot et al. 2000). Dans leur étude, l'anti-CCP avait une bonne valeur prédictive lors de l'analyse univariée, mais lorsque combinée avec d'autres facteurs pronostiques, la valeur prédictive de l'anti-CCP diminuait considérablement et ne contribuait pas significativement à la prédiction du pronostic. Mottonen a également observé que le FR est un des seuls marqueurs pouvant prédire les dommages radiographiques à un an (Mottonen et al. 1998). En plus d'être un bon marqueur des dommages articulaires, le FR est un des meilleurs pour prédire l'atteinte fonctionnelle à 3 et 6 ans (Kroot et al. 2000).

Anti-cyclic citrullinated peptide antibodies (Anti-CCP)

L'anti-CCP est un test plus récent que le FR qui semble aider à détecter les patients atteints de PR. Dans certains cas, le FR est négatif et l'anti-CCP positif, ce qui

pourrait alors aider à identifier certains patients à risque de développer cette maladie. Il existe plusieurs générations de tests ELISA (*Enzyme-Linked ImmunoSorbent Assay*) permettant de détecter la présence d'anti-CCP. La première génération a été celle la plus utilisée dans les études, mais d'autres chercheurs ont utilisé des ELISA de deuxième génération (celle actuellement commercialisée), ce qui rend difficile les comparaisons entre les études. Le test ELISA de première génération a une très forte spécificité variant entre 90% et 97% pour détecter les PR d'installation récente mais une sensibilité semblable au FR, soit entre 44% et 56% (Jansen et al. 2002; Niewold, Harrison & Paget 2007; Schellekens et al. 2000). Par contre, chez les patients diagnostiqués à l'aide des critères de classification de l'ACR, la sensibilité de l'anti-CCP est plus élevée, soit de 65%-70% (Schellekens et al. 2000).

Plusieurs études semblent suggérer que l'anti-CCP est associé au développement des dommages articulaires chez les patients atteints de PR et la présence de l'anti-CCP serait un important prédicteur des dommages radiographiques. En effet, dans une cohorte d'environ 400 patients avec polyarthrite d'installation récente suivis sur une période de 2 ans, l'anti-CCP présent dans le sérum initial des malades s'est avéré un très bon marqueur des dommages radiographiques importants (OR=4.7; p=0.0005) et de la progression des destructions articulaires (OR=3.0; p=0.0005) (Forslind et al. 2004). Visser et son équipe ont développé, chez une cohorte de 500 patients avec arthrite inflammatoire débutante suivis pendant 2 ans, un modèle prédictif d'une maladie persistante selon lequel l'anti-CCP serait associé à une arthrite érosive et ce avec une meilleure valeur prédictive que le FR (OR 4.58 vs 2.99, respectivement) (Visser et al. 2002). L'étude de Meyer a aussi constaté en 2003 que le risque d'une augmentation des

dommages radiographiques à 5 ans est significativement plus élevé chez les patients avec anti-CCP (OR=2.5, IC(95%)=1.2-5.0), mais que chez ceux avec FR il n'y a pas de différence significative (OR=0.7, IC(95%)=0.3-1.5) (Meyer et al. 2003). Dans l'article de Kroot publié en 2000 discuté précédemment dans la section du FR, les auteurs spécifiaient que l'anti-CCP perdait de l'importance lorsqu'il était inclus dans un modèle multivarié. Par contre, fait intéressant, leurs résultats démontrent que dans 70% des cas de polyarthrite débutante, l'anti-CCP est présent très tôt dans l'évolution de la maladie et ces patients développeraient plus de dommages radiographiques sévères. Finalement, une dernière étude constate également que l'anti-CCP avait une meilleure valeur prédictive que les autres anticorps sur une période de 2 ans; de plus, la prévision pour l'apparition de nouvelles érosions et pour une maladie progressive serait meilleure lorsque l'anti-CCP est combiné au FR (Vencovsky et al. 2003).

Anti-Sa

Pour ce qui est de l'anti-Sa, nous n'avons que très peu d'informations le concernant, car ce n'est que depuis quelques années que les chercheurs commencent à le considérer dans leurs études. Pour ce que nous en connaissons, cet anticorps serait associé à une maladie plus sévère avec une sensibilité de 30%-40% et une spécificité de 95%-100% (Boire et al. 2005; Lopez-Longo et al. 2006; Menard et al. 2000).

Une étude publiée en 2006 (Lopez-Longo et al. 2006) a observé que l'anti-CCP était plus sensible que l'anti-Sa pour la détection de patients avec diagnostic de polyarthrite rhumatoïde. La concordance entre anti-CCP et l'anti-Sa était seulement de

46% mais 12% des cas de PR présentaient un anti-Sa sans anti-CCP. Ces auteurs concluaient donc que l'anti-Sa pourrait être utile dans les cas où l'anti-CCP serait négatif.

Goldbach-Mansky et son équipe, en 2000, ont rapporté que l'anti-Sa était associé à de multiples érosions à 1 an, ainsi qu'à la prise de médicaments, ce qui semble indiquer qu'il serait un bon marqueur pour les maladies érosives, donc les maladies considérées comme plus sévères (Goldbach-Mansky et al. 2000). Hayem en 1999 a aussi démontré que l'anti-Sa était très bon pour prédire des dommages radiographiques sévères dans un modèle de régression logistique ajusté pour l'âge, le sexe et la durée de la maladie (OR=5.41, IC(95%)=1.76-16.6) (Hayem et al. 1999).

Notre équipe a déjà publié un article sur la contribution des autoanticorps, et d'autres informations cliniques (âge, sexe, synovites, raideurs, activité maladie, etc.), pour prédire le développement d'une maladie sévère à 30 mois chez une cohorte de 165 patients (maladie sévère définie par la présence de *dommages radiographiques* et/ou d'*incapacité fonctionnelle*) (Boire et al. 2005). Nous avons donc mis en évidence que l'anti-Sa, l'âge, et les dommages radiographiques à l'inclusion sont les meilleurs prédicteurs. Le FR et l'anti-CCP n'apportent aucune contribution dans le modèle multivarié, bien qu'il y ait une association significative lorsque ces anticorps sont étudiés individuellement.

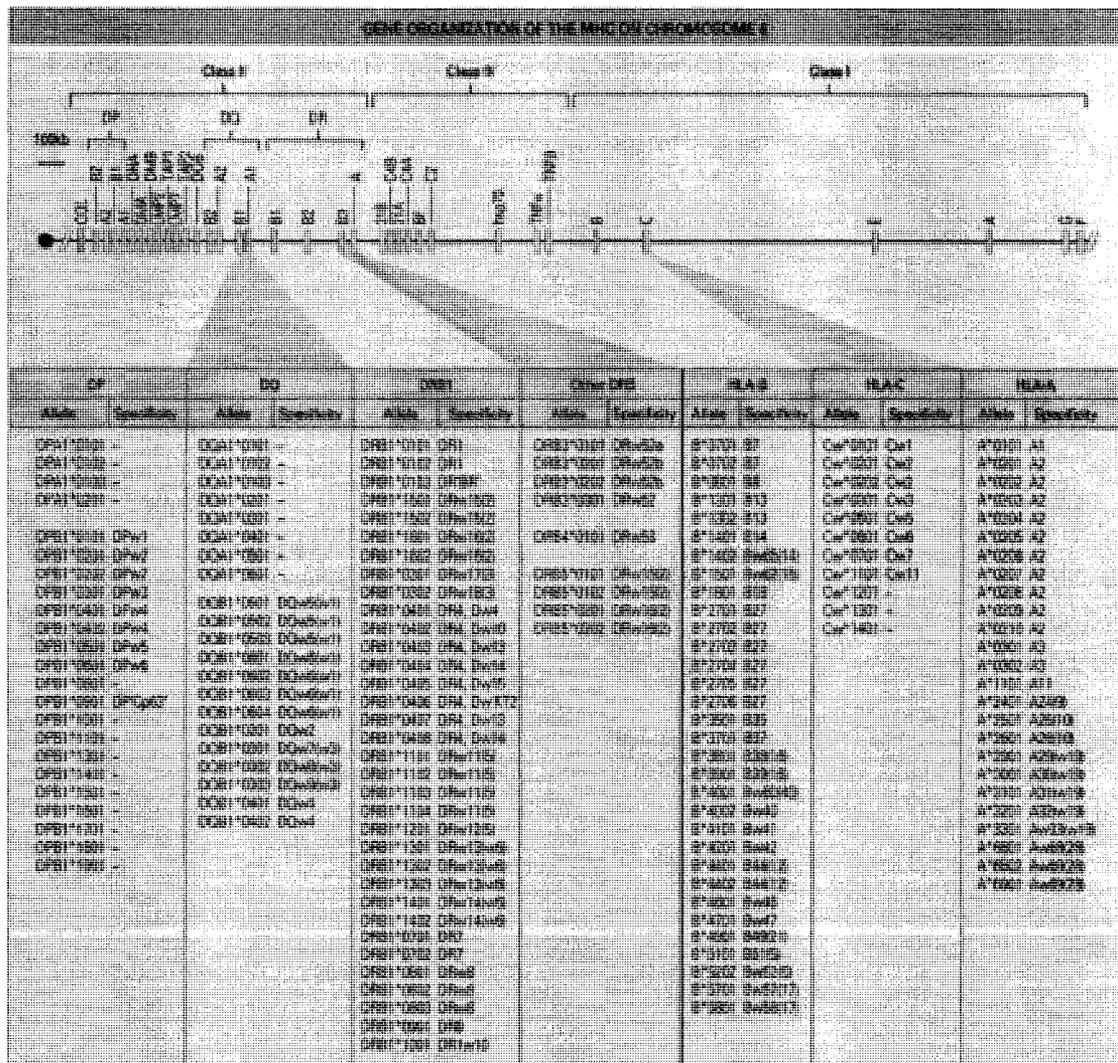
Suite à cette revue des autoanticorps pour prédire le développement des dommages radiographiques, on constate qu'il y a une très grande contradiction entre les auteurs. Ces articles semblent donc suggérer que l'utilisation des 3 anticorps pourrait être utile, éventuellement en combinaison.

Facteurs Génétiques

Plusieurs données probantes indiquent que les individus qui possèdent certains types d'antigènes HLA sont prédisposés à certaines maladies auto-immunes. Comme mentionné précédemment dans la section sur les maladies auto-immunes (section 2.1.4), le système immunitaire utilise entre autre les HLA pour reconnaître les cellules du corps et identifier les cellules et antigènes étrangers. La protéine présentatrice HLA qui se retrouve à la surface de la cellule est créée à partir de l'ADN. La séquence de cette protéine est encodée dans un gène à un endroit spécifique (locus) du chromosome 6 dans le noyau de la cellule. Le terme «allèle» réfère aux différentes formes très apparentées que peut prendre le gène au sein d'un même locus. En génétique, on utilise une combinaison de lettres et de chiffres pour désigner un allèle spécifique à un endroit précis (ou locus) d'un chromosome donné. Les molécules HLA sont parmi les protéines humaines qui présentent le plus d'allèles différents. On divise les molécules HLA en 3 grandes classes : I (HLA-A, -B et C), II (HLA-DR, -DP et -DQ), III (surtout des composantes du complément). Le HLA-DR représente le locus dont l'association avec la polyarthrite rhumatoïde est la plus fermement établie. Un complexe HLA-DR donné est constitué d'une chaîne protéique α qui est codée par le locus HLA-DRA, généralement constante, ainsi que d'une chaîne β qui est codée par un des locus HLA-DRB1, HLA-DRB2 ou HLA-DRB3. Les molécules HLA-DR d'intérêt clinique sont principalement les HLA-DRB1. La figure 5 représente l'organisation des gènes du locus spécifique du HLA sur le chromosome 6. On voit bien l'endroit où se situe le HLA-DR dans la classe II ainsi que la multitude d'allèles possibles pour ce locus que l'on retrouve dans la liste en dessous de l'image dans la section DRB1. Dans cette liste, on retrouve à gauche les

allèles définis par la séquence génomique (ex. DRB1*0101) et la spécificité telle que définie par de anticorps présents dans le sérum de mères multipares et qui reconnaissent des protéines apparentées qui peuvent être produites à partir d'un seul ou de plusieurs allèles. Comme les chromosomes sont présents par paires (combinaison des gènes hérités de la mère et du père), chaque individu exprime donc 2 molécules HLA (ou allèles) à chaque locus (Nepom, Nepom 2003).

Figure 5. Organisation des gènes du HLA sur le chromosome 6



Il semble maintenant bien établi que certains allèles spécifiques du locus HLA-DR facilitent la production d'autoanticorps comme le FR et l'anti-CCP, alors que d'autres allèles HLA-DR affectent négativement la production de ces autoanticorps.

Le regroupement le plus connu et étudié des allèles HLA-DR est le Shared Epitope (SE) qui est constitué des allèles suivants : HLA-DRB1*0101, *0102, *0105, *0401, *0404, *0405, *0408, *1001, *1402. Ces allèles partagent une séquence commune dans la région permettant la liaison entre la molécule HLA-DR et le peptide antigénique. Les allèles SE sont plus fréquents chez les personnes atteintes de PR que chez la population générale (soit respectivement 70% à 80% vs 30% à 40%) mais leur présence n'est pas essentielle au développement de la maladie. De plus, certains autoanticorps associés à la PR, comme le FR et l'anti-CCP, sont plus fréquents chez les proches parents des personnes atteintes de PR, ce qui suggère que les gènes du HLA favorisent le développement d'autoanticorps (Berglin et al. 2004; Huizinga et al. 2005; van der Helm-van Mil, A.H. et al. 2006). Par contre, comme la plupart des individus porteurs d'allèles SE ne développeront jamais la PR, on ne peut donc dire que cette maladie est héréditaire, ni que les allèles SE causent la maladie (MacGregor, Silman 2003; Tannenbaum, Russell 2001). Ce sous-groupe d'allèles a déjà été associé à une polyarthrite grave. Entre autres, la présence du HLA-DR4 chez les patients avec PR est associée à des manifestations d'une maladie plus sévère comme le facteur rhumatoïde (FR) positif, la présence de manifestations extra-articulaires, et des changements radiographiques plus sévères (Tannenbaum, Russell 2001). Par contre, d'autres auteurs rapportent que les patients avec SE présentent une maladie plus sévère seulement lorsqu'il y a présence d'anti-CCP (Berglin et al. 2004; Huizinga et al. 2005; van Gaalen

et al. 2004) et d'autres n'observent qu'une faible association voire pas d'association du tout (Harrison et al. 1999; Thomson et al. 1999; van der Helm-van Mil, A.H. et al. 2006). Une raison qui pourrait expliquer les discordances entre les études pourrait être le traitement administré en début de maladie, ce qui pourrait moduler l'influence du SE sur la sévérité de la maladie.

Quelques auteurs ont rapporté l'existence d'un groupe d'allèles appelé HLA-DR3, qui agit à l'inverse du SE et qui est associé à une diminution des anti-CCP (Irigoyen et al. 2005; Verpoort et al. 2005).

Plus récemment, certains chercheurs ont publié différents articles au sujet d'un nouveau sous-groupe d'allèles appelé DERA, qui diminue le risque de développer une maladie rhumatoïde sévère. Le DERA regroupe les allèles HLA DR*0103, *0402, *1102, *1103, *1301, *1302 (van der Helm-van Mil, A.H. et al. 2005; van Gaalen et al. 2004).

Vitesse de sédimentation et Protéine C-réactive

Deux tests sanguins sont utilisés couramment en clinique pour évaluer l'inflammation systémique : la vitesse de sédimentation (VS) et le niveau sérique de la protéine C-réactive (CRP).

La VS est une mesure non spécifique de l'inflammation qui a un intérêt essentiellement pronostique et qui fait partie de l'examen de routine dans la surveillance des maladies inflammatoires. C'est le temps nécessaire aux éléments cellulaires sanguins (globules blancs, globules rouges et plaquettes) pour sédimenter, c'est-à-dire tomber librement en bas d'une colonne de sang incapable de coaguler (grâce à un anticoagulant

utilisé lors du prélèvement). Elle est exprimée par la différence en hauteur (en mm) entre la colonne de sang initiale et la colonne de cellules sédimentées, mesurée au bout d'une heure et de 2 heures (des techniques plus rapides existent actuellement). En cas d'augmentation du taux de certaines protéines dans le sang, ce qui est le cas lors d'une inflammation, la VS accélère énormément. Les avantages à utiliser cette méthode résident dans le fait qu'elle est simple à mettre en œuvre, largement disponible, standardisée et peu onéreuse. Par contre, la VS ne peut être utilisée pour évaluer rapidement les changements cliniques de l'inflammation puisqu'elle dépend largement de la concentration de protéines dans le sang, et que cette concentration change relativement lentement. On sait par ailleurs que la VS augmente avec l'âge et serait aussi influencée par le sexe, ce qui complique l'interprétation de cette mesure. De plus en plus, la VS tend à être remplacée par le dosage de la protéine C-réactive (CRP) (Ballou, Kushner 2005).

La CRP est une protéine sérique spécifique de la réaction inflammatoire, synthétisée par le foie, et dont le niveau sérique augmente précocement au cours de la réponse inflammatoire. Les taux de CRP sont généralement les plus élevés dans des cas d'infection bactérienne. Il existe une bonne corrélation avec l'évolution de la PR, la persistance de l'inflammation indiquée par la CRP, et l'existence de lésions osseuses (Ballou, Kushner 2005). La CRP est la protéine de choix comme marqueur précoce de la réaction inflammatoire. De plus, cette mesure n'est pas affectée par l'âge ou le sexe, et elle peut être mesurée sur des échantillons entreposés, contrairement à la VS. Malheureusement, les techniques de détection varient selon les centres hospitaliers et ne sont pas uniformément disponibles. Certains cliniciens hésitent à utiliser la CRP car ils ont plutôt l'expérience de la VS.

Plusieurs chercheurs ont étudié la valeur pronostique de ces variables dans la PR débutante. Plusieurs études ont suggéré que les élévations soutenues de la VS et de la CRP sont associées à un développement érosif de la maladie. En 2004, Forslind a mis en évidence à l'aide d'une cohorte de 379 patients, que ces deux variables étaient associées à la destruction et à la progression radiographique avec un risque variant entre 2 et 2.7. Par contre, lors de l'analyse multivariée, seulement la VS restait significative dans le modèle pour les 2 issues différentes (Destruction : OR=2.0, p=0.025 et Progression : OR=1.8, p=0.045) (Forslind et al. 2004). Combe, dans sa cohorte de 191 patients suivis sur 3 ans, a aussi observé les mêmes résultats en univarié et en multivarié avec des risques semblables pour les mêmes issues (Combe et al. 2001).

D'autres études semblent suggérer que d'étudier l'évolution de la CRP et de la VS pourrait être utile dans l'évaluation du pronostic de la maladie. Une association significative entre une CRP élevée sur une longue période de temps et la sévérité de la progression des dommages radiographiques aurait été détectée chez des patients avec PR dans 2 études moins récentes, publiées en 1977 et 1988, respectivement (Amos et al. 1977; Larsen 1988). Par contre, dans les articles plus récents, il est rare que les paramètres de l'inflammation soient étudiés en continu. Généralement, ces paramètres sont mesurés et considérés seulement lors de l'inclusion du patient dans l'étude.

Globules blancs, polymorphonucléaires, hémoglobine

Le décompte des globules blancs est un indice de la présence ou l'absence d'une infection ou encore d'une réaction à un médicament. Comme les globules blancs combattent l'infection, le nombre de ces cellules tend à augmenter en présence d'une

infection, surtout bactérienne. L'analyse des globules blancs permet de détecter les effets secondaires causés par un médicament ou la présence d'une infection. On utilise aussi la formule leucocytaire pour déterminer le type de globules blancs (ex. polymorphonucléaires ou lymphocytes) et définir lequel est en croissance ou en décroissance. De plus, bien que les corticostéroïdes soient généralement utilisés pour diminuer rapidement l'inflammation, leur administration tend à augmenter la quantité de globules blancs dans le sang, en particulier les polymorphonucléaires (PMN). De plus, l'inflammation généralisée, lorsqu'elle n'est pas contrôlée par un traitement adéquat, cause notamment de la fatigue et de l'anémie (Harris 2005). Malgré le fait que ces variables soient simples à obtenir et à mesurer, ce ne sont pas des variables généralement considérées dans le cadre des études cliniques.

Âge et sexe

La PR est généralement diagnostiquée entre 40 et 50 ans chez les femmes et entre 50 et 60 ans chez les hommes. De plus, cette maladie invalidante touche principalement les femmes, soit trois fois plus souvent que les hommes. Il est difficile d'expliquer cette prédominance féminine, mais on a remarqué que la maladie apparaît généralement autour de la ménopause et lorsqu'elle se présente précocement, il y a régression de la maladie durant la grossesse. Il semble donc y avoir une influence hormonale importante associée à son développement (La Société d'arthrite 2007). De plus, il a été observé que les hommes ont généralement une maladie plus sévère que les femmes. En effet, selon une étude de Weyand en 1998, la survenue d'érosions est plus fréquente chez les hommes que chez les femmes (72% vs 55%, $p < 0.05$) et aurait tendance à apparaître plus rapidement (47% vs 31% pour une maladie érosive dans les 4 premières années) (Weyand et al.

1998). Par contre, malgré l'apparence d'un mauvais pronostic chez les hommes, une récente étude semble indiquer que les femmes s'en tirent moins bien à long terme, car elles seraient moins susceptibles d'entrer en rémission. Des chercheurs suédois ont mené une étude auprès de 700 patients atteints de PR dont la gravité apparaissait semblable au début de l'étude. Étant donné l'incidence importante chez les femmes, celles-ci constituaient les deux tiers de l'échantillon. Les chercheurs ont observé que les hommes étaient deux fois plus susceptibles d'être en rémission après 2 ans et 5 ans de suivi tandis que la maladie aurait progressé plus rapidement et s'était aggravée chez les femmes (Forslind et al. 2007). On s'explique mal comment la maladie peut se développer de façon si différente, mais cette variable semble être importante à considérer dans les études cliniques.

L'âge semble aussi être un facteur de risque associé à une maladie plus sévère mais aussi à une incapacité plus importante. En effet, notre équipe avait déjà remarqué que l'âge considérée en continu était significativement associée à une maladie sévère, soit être atteint de dommages radiographiques importants ou d'incapacité fonctionnelle, (OR=1.063 pour chaque année additionnelle, $p < 0.005$) (Boire et al. 2005). Le fait d'être âgé de plus de 55 ans avait aussi un impact sur les dommages et la progression radiographique à 2 ans dans la cohorte de Forslind en 2004 (OR=2.1, $p = 0.001$ dans les 2 cas) (Forslind et al. 2004). Quant à Kroot, il a trouvé une association significative pour l'âge, mais seulement pour l'incapacité à 3 et 6 ans (Kroot et al. 2000). Par contre, deux autres auteurs n'ont pas trouvé d'association significative entre l'âge et une maladie sévère (Combe et al. 2001; Mottonen et al. 1998). On a constaté également que lorsque la maladie apparaît plus tardivement, elle comporte quelques particularités. On l'appelle

souvent la polyarthrite du sujet âgé ou bien PR à début tardif. Comme mentionné précédemment, la PR touche plus fréquemment les femmes, mais à partir de 60 ans, la différence d'incidence selon le sexe diminue jusqu'à presque disparaître. En général, à l'âge avancé, la PR a un début plus aigu, qui peut parfois faire évoquer un épisode infectieux. Les grosses articulations sont plus fréquemment atteintes, il y a une installation rapide des manifestations articulaires et le FR est souvent séronégatif (Gabay 2007).

Tabagisme

Le tabagisme a été identifié dans plusieurs études comme un facteur de risque pour les maladies autoimmunes. En 2002, Harrison a souligné l'influence de la cigarette sur la PR en relatant un certain nombre d'articles sur le sujet (Harrison 2002). Tout d'abord, le tabac serait associé à un dysfonctionnement du système immunitaire (George, Levy & Shoenfeld 1997; Ginns et al. 1985; Hughes et al. 1985) et il semblerait que le tabagisme augmente l'apparition du FR, (Jonsson, Thorsteinsson & Valdimarsson 1998; Saag et al. 1997; Symmons et al. 1997) ce qui suggère par le fait même une augmentation du risque de PR puisque le FR est un marqueur important de cette maladie. L'unité de mesure du tabagisme est le paquet-année qui signifie fumer 1 paquet de cigarettes par jour pendant une année. Un paquet-année peut représenter fumer ½ paquet de cigarettes par jour pendant 2 ans, ou fumer 4 paquets par jours pendant ½ année. Un autre auteur (Saag et al. 1997) a donc démontré que le nombre de cigarettes par année était significativement associé au FR, aux érosions et à la présence de nodules chez les patients atteints de PR. Les fumeurs de plus de 25 paquets de cigarettes-années avaient un

risque 3.1 fois plus élevé d'être séropositif pour le FR, et 2.4 d'avoir des érosions. Plus récemment, il a été démontré que le tabagisme serait aussi associé à la présence de l'anti-CCP, mais plus spécifiquement lorsqu'il y a présence des allèles SE (Linn-Rasker et al. 2006; Pedersen et al. 2007). Forslind a également noté une association entre le tabagisme et la progression radiographique (Forslind et al. 2004). Une étude cas-témoins portant sur l'association entre le tabac et la PR faite chez 239 patients atteints de PR et 239 sujets sains, a montré que les patients étaient plus susceptibles d'être de gros fumeurs que les sujets du groupe contrôle. Ceux qui avaient fumé plus de 20 cigarettes par jour durant 40 à 50 ans avaient 13 fois plus de risques de présenter une PR. Comme un des facteurs connu pour le développement de la PR est les antécédents familiaux (i.e. la prédisposition génétique), ces individus ont donc été évalués pour vérifier si ce risque élevé n'était pas en fait expliqué par les antécédents. Plus de la moitié des patients n'avaient pas d'antécédents familiaux. De plus, les membres de la famille atteints de PR étaient plus nombreux à être de gros fumeurs au moment où la PR a été diagnostiquée (Hutchinson et al. 2001). Par contre, cette association du tabagisme avec une maladie sévère n'est pas générale. Harrison en 2001 a rapporté qu'effectivement il y avait une association entre la cigarette et la présence du FR et la présence de nodules, mais leurs patients n'avaient pas plus de dommages radiographiques et présentaient moins d'inflammation (Harrison et al. 2001). Par contre, Harrison a suggéré que ceci pourrait être dû à d'autres facteurs comme l'activité physique réduite chez les fumeurs.

Médication

Plusieurs médicaments sont utilisés dans la pharmacothérapie de la PR, dont les anti-inflammatoires non stéroïdiens (AINS), les antirhumatismaux (DMARD traditionnels ou biologiques) et les corticostéroïdes (American College of Rheumatology Subcommittee on Rheumatoid Arthritis, Guidelines 2002). Les AINS sont des médicaments anti-inflammatoires et analgésiques qui soulagent les signes et les symptômes articulaires tels que la douleur, les gonflements et la raideur. Cependant les AINS ne préviennent pas la progression des lésions articulaires. Les corticostéroïdes agissent rapidement et soulagent les inflammations douloureuses et destructives. Ils présentent toutefois des effets secondaires potentiellement sévères surtout en cas d'administration à doses élevées et à long terme. Finalement, les DMARD ont la capacité d'influencer favorablement la maladie par une réduction voire une prévention des lésions articulaires et par une conservation de la fonction des articulations (American College of Rheumatology Subcommittee on Rheumatoid Arthritis, Guidelines 2002). Traditionnellement, les DMARDs étaient généralement administrés au moins 3 mois après le diagnostic confirmé de PR, chez les patients qui ont toujours une inflammation persistante (douleur, raideur, gonflement) malgré le traitement à l'aide d'AINS. De plus, si un patient se présente avec des lésions articulaires dès le diagnostic ou s'il présente une activité importante de la maladie, les DMARD sont généralement considérés comme traitement de base. Les médicaments biologiques ont en commun de bloquer spécifiquement une cytokine, un type de protéines impliquées dans le mécanisme de l'inflammation. Le développement des médicaments biologiques a été une avancée importante dans le traitement de la PR. Actuellement, on utilise ces médicaments dans les

cas où les patients continuent d'avoir des symptômes ou ont une maladie évolutive malgré un traitement classique jugé optimal. Comme ces médicaments sont conçus pour supprimer diverses fonctions du système immunitaire, leur usage peut représenter un risque sérieux chez les patients sujets aux infections comme les diabétiques. Un autre inconvénient de ces médicaments est leur prix élevé, variant entre 17 000 \$ à 25 000 \$ et plus par année.

Plusieurs indices suggèrent qu'une fenêtre thérapeutique critique peut exister au cours des deux premières années qui suivent le début de la maladie, période au cours de laquelle il est encore possible d'induire une rémission (American College of Rheumatology Subcommittee on Rheumatoid Arthritis, Guidelines 2002; Landewe et al. 2002). De même, des changements radiographiques ont lieu au cours des deux années qui suivent le début de la maladie chez 50 % à 70 % des patients PR (van der Heijde 1995). Le Collège américain de rhumatologie suggère que le contrôle de la progression de la maladie devrait commencer à un stade précoce pour limiter les dommages articulaires causés par la PR (American College of Rheumatology Subcommittee on Rheumatoid Arthritis, Guidelines 2002). Plusieurs chercheurs ont constaté que de traiter tôt les patients a un effet bénéfique à long terme (Bukhari et al. 2003; Korpela et al. 2004; Quinn et al. 2005; van Aken et al. 2004). De plus, les études démontrent que la combinaison de plusieurs DMARDs (double ou triple thérapie) serait plus efficace que la monothérapie, tout comme l'association d'un ou de plusieurs DMARDs avec un corticostéroïde (Allaart et al. 2006; Korpela et al. 2004; Landewe et al. 2002; Quinn et al. 2005).

Il est particulièrement difficile de départager la relation de cause à effet pour ce qui est du traitement et de la sévérité de la maladie. Effectivement, certains chercheurs

ont conclu que l'intensité du traitement est associée à une maladie plus sévère. Par contre, si la maladie était déjà plus sévère au départ, c'est peut-être la raison pour laquelle les patients sont effectivement plus traités, et non parce que le traitement causerait une maladie plus sévère. Bukhari, en 2003, a donc utilisé une méthode appelée le «Propensity score» permettant d'ajuster les données selon la probabilité d'être traité à partir des données cliniques du départ (Bukhari et al. 2003). Il a conclu qu'après ajustement, traiter tôt les patients aurait un effet bénéfique à long terme.

Il est donc évident, en considérant la multitude de marqueurs potentiels déjà énumérés, qu'il est très complexe d'expliquer l'évolution des maladies inflammatoires débutantes. Il est important de continuer la recherche dans ce domaine afin de trouver éventuellement d'autres méthodes plus efficaces pour mieux diagnostiquer cette maladie et en définir des sous-groupes plus homogènes.

La section suivante portera sur la méthodologie concrète utilisée dans ce projet pour tenter de trouver de nouvelles méthodes pour étudier et comprendre cette maladie complexe.

3 - Matériel et Méthodes

3.1 Source des données

Les données sont tirées de la base de données EUPA fournie par le docteur Gilles Boire et ses collaborateurs du service de rhumatologie du centre de recherche clinique Étienne-Le Bel (CRC) du centre hospitalier universitaire de Sherbrooke (CHUS). Cette base de données comporte actuellement 376 patients et elle a pour but d'étudier la valeur pronostique de certains marqueurs présents à l'inclusion chez les patients avec polyarthrite d'installation récente afin de prédire les issues cliniques les plus importantes de la maladie (sévérité/rémission).

3.2 Devis de l'étude

Cette étude se définit comme un dispositif d'observation avec cohorte prospective longitudinale. Le dispositif d'observation est tout à fait approprié à ce type d'étude car il permet d'étudier à plus long terme le développement «naturel» de la maladie et permet également de toucher un plus large éventail de la population. Évidemment, il serait pour ainsi dire impensable et très onéreux de faire un essai randomisé pour étudier l'évolution de la maladie avec autant de patients sur une aussi longue période de temps compte tenu de la complexité de ce type de devis à long terme (suivi, décès, abandons).

3.3 Recrutement des sujets

3.3.1 Critères d'admissibilité

Les sujets de l'étude doivent être âgés de plus de 18 ans, se présenter au CHUS avec au moins 3 articulations présentant une synovite, et ce depuis plus d'un mois et moins de 12 mois. Les patients exclus de l'étude sont ceux atteints d'arthrite microcristalline (ex : Goutte), de collagénose (maladie du tissu conjonctif), ou d'arthrite septique, ainsi que les patients inaptes à consentir.

3.3.2 Échantillonnage

Un échantillonnage de convenance a été utilisé pour constituer la banque de données puisque le recrutement est fait chez l'ensemble des patients se présentant au CHUS et répondant aux critères d'admission.

La population cible est donc la population de l'Estrie et des régions environnantes atteinte de polyarthrite inflammatoire débutante. Plus spécifiquement, la population étudiée est l'ensemble des personnes atteintes d'arthrite inflammatoire débutante, âgées de plus de 18 ans, et qui se sont présentées au CHUS depuis le début de l'étude, soit à l'automne 1998.

3.3.3 Taille de l'échantillon

Étant donné le type de devis choisi, il n'y a pas de taille d'échantillon spécifique à atteindre. Toutefois, compte tenu des hypothèses initiales à tester, du taux d'attrition ou de perte au suivi estimé à 5% par an, un échantillon cible de 515 patients a été établi. Les patients répondant aux critères d'admissibilité et qui acceptent de participer à l'étude sont

donc encore inclus de façon consécutive et s'ajoutent à cette base de données. Jusqu'à maintenant, nous avons 376 patients. D'entre eux, 54 ont quitté l'étude pour diverses raisons et à différentes étapes du suivi. Les analyses ont été réalisées chez les 211 patients ayant complété leurs 3 premières visites de suivi. Étant donné que les analyses choisies sont de type exploratoire, il n'y a pas de calcul de taille d'échantillon à effectuer. Le nombre de sujets doit simplement être supérieur au nombre de variables incluses dans l'analyse.

3.3.4 Déroulement de l'étude

Le premier contact avec le patient se fait par l'entremise du rhumatologue à la clinique externe du CHUS. Le médecin effectue l'examen clinique de routine où il évalue le nombre d'articulations douloureuses ou gonflées, s'il y a présence de manifestations extra-articulaires, la durée de la raideur matinale ainsi que s'il y a atteinte symétrique des membres. À cette étape, les critères d'exclusion (ex. arthrite microcristalline, infection articulaire) sont aussi relevés. Lorsque le médecin constate que le patient remplit les critères d'inclusion et ne remplit pas les critères d'exclusion, donc qu'il est bien atteint d'arthrite inflammatoire débutante, le patient est invité à participer. Le rhumatologue explique brièvement le but de l'étude et son déroulement et, si le patient consent à participer, l'infirmière de recherche se déplace pour rencontrer le sujet. Celle-ci explique en détail le formulaire de consentement. Une fois ce formulaire signé, l'infirmière effectue une entrevue structurée face à face avec le patient afin de remplir certains questionnaires. Elle recueille alors les données socio-démographiques, ainsi que les réponses aux questions sur la capacité fonctionnelle et la perception du patient envers sa

maladie. Elle lui prescrit également un prélèvement sanguin pour la récupération de l'ADN génomique, pour la recherche de la présence d'autoanticorps et de paramètres inflammatoires ainsi que des radiographies des mains et des pieds pour évaluer la progression des dommages articulaires. Par la suite, le patient est contacté par téléphone pour fixer la rencontre subséquente (à 18, 30, 42 et 60 mois après le début de la maladie clinique), visite qui généralement coïncide avec une visite de routine du patient avec son rhumatologue. Le début de la maladie se réfère ici à l'apparition des signes et des symptômes comme la douleur, le gonflement et la raideur de l'articulation (mentionné plus précisément à la section 2.1.6 Diagnostique et pronostic). Dernièrement, le comité d'éthique a accepté que le projet se poursuive sur une période de 10 ans s'il y a persistance de l'arthrite. Le choix de réévaluer les patients à partir de la date du début des symptômes plutôt que selon la date du diagnostic, soit lors de leur première visite de consultation, permettra d'obtenir une cohorte plus homogène pour la durée de la maladie. Bien que la mesure ne soit pas exacte, car elle dépend de la mémoire du patient, elle est tout de même préférable à la date d'inclusion dans l'étude qui, elle, pourrait survenir entre 1 et 12 mois après le début des symptômes.

3.4 Variables et instruments de mesures

Cette section décrit en détail l'ensemble des variables mesurées et utilisées lors de l'analyse des données. Un tableau résumant l'ensemble des ces variables est joint ci-dessous.

Tableau 2. Résumé des variables et instruments de mesure

Variables	Instruments de mesures / Définitions
Socio-démographiques (âge, sexe, ethnie, scolarité, etc.)	Questionnaire socio-démographique
Biologiques/sérologiques (anticorps, ADN, paramètres inflammatoires)	Dans le sérum et le sang du patient (Mesurés en laboratoire)
Nb articulations avec synovites Nb articulations douloureuses Durée de la raideur matinale Symétrie Manifestations extra-articulaires	Évaluation du médecin
Médication / traitement	Dossier patient
Capacité fonctionnelle	M-HAQ (Modified Health Assessment Questionnaire) Questionnaire validé par Pincus, T. en 1983
Activité de la maladie	Utilisation du DAS (Disease Activity Score), Développé par DMFM van der Heijge dans les années 80 et la version DAS28-CRP a été validée par Prevoo en 1995
Dommages radiographiques : Érosions / pincements	Lecture des radiographies selon le score Sharp/van der Heijde (SvH) À l'insu du diagnostic et de l'évolution clinique

Variabes socio-démographiques

Les données socio-démographiques comme l'âge, le sexe, l'ethnie et le degré de scolarité proviennent du questionnaire rempli avec l'infirmière.

Variabes biologiques et sérologiques

Les variables biologiques et sérologiques comme la présence d'anticorps, l'hémoglobine, l'hématocrite, les globules blancs, la CRP, la vitesse de sédimentation et le typage HLA-DR à partir de l'ADN proviennent du sérum et du sang du patient. Les principaux autoanticorps étudiés sont le FR, l'anti-CCP et l'anti-Sa.

Variables concernant l'état du patient

Le médecin évalue le nombre d'articulations avec synovites ainsi que le nombre d'articulations douloureuses, la durée de la raideur matinale, l'atteinte symétrique des membres et la présence de manifestations extra-articulaires. Certaines de ces variables sont des critères de classification de la polyarthrite rhumatoïde et serviront aussi au calcul d'indices de sévérité de la maladie et d'autres seront plutôt des marqueurs pronostiques potentiels de la maladie.

La figure 17 à l'annexe 1 de la page 111 représente les articulations que le médecin évalue en clinique pour y identifier la présence ou non de synovites ou d'articulation douloureuses. Le médecin procède à deux décomptes articulaires. L'un est utilisé pour le calcul des articulations avec gonflements/synovites (articulations entourées) et l'autre pour le calcul des articulations douloureuses (articulations en noir).

Médication

La médication est une variable importante à considérer dans ce genre d'étude, car elle peut être une cause de biais important si elle n'est pas considérée lors de l'analyse. Cette information provient du dossier du patient qui a au préalable accepté, lors de la signature du formulaire de consentement, que son dossier soit consulté. Comme les corticostéroïdes sont généralement utilisés pour diminuer rapidement une inflammation importante et qu'ils ont un impact sur les paramètres inflammatoires, il était intéressant de les considérer séparément des autres DMARD (traditionnels ou biologiques).

Capacité fonctionnelle

La capacité fonctionnelle est mesurée à l'aide du questionnaire Modified Health Assessment Questionnaire (M-HAQ) et elle se définit comme le degré de difficulté à effectuer des tâches courantes comme le fait de s'habiller, de marcher ou de se lever du lit. Le questionnaire original, créé il y a plus de 25 ans, comportait 20 questions classées en 8 catégories de 2 ou 3 variables. Une version modifiée a été créée en 1983 pour ne conserver qu'une seule variable par catégorie. Ce nouveau questionnaire comprend maintenant 8 questions ayant un score de 0 à 3 où le score le plus élevé correspond à une incapacité fonctionnelle maximale. Ce score modifié a été validé par Pincus, en 1983 (Pincus et al. 1983). Le questionnaire a aussi été traduit en français et validé dans une cohorte de patients avec polyarthrite débutante (Guillemin, Briancon & Pourel 1992). Le score total obtenu est divisé par 8, d'où l'on obtient un score entre 0 à 3. Généralement, un score de 1 et plus est considéré comme une incapacité importante dans les études cliniques. Le questionnaire utilisé se retrouve à l'annexe 3 (section 1), page 114 à des fins de consultation.

Indice d'activité de la maladie

Le DAS (Disease activity score), développé par Désirée M.F.M van der Heijde et son équipe à l'Université de Nijmegen dans les années 80, est un indice permettant d'évaluer le degré d'activité de la maladie chez les patients avec PR (van der Heijde et al. 1990). Le DAS utilise un décompte de 44 articulations des mains, des poignets, des coudes, des genoux, des hanches, des chevilles et des pieds. Le DAS est calculé à partir de la présence ou l'absence de gonflement ou de douleur, d'une échelle visuelle analogue de 0 à 10 sur l'impact de la maladie chez le patient, et de la mesure de la vitesse de

sédimentation. Dans le cadre de cette étude, il a été impossible de calculer le DAS traditionnel pour 2 raisons. Tout d'abord, la technique utilisée au CHUS pour la VS n'était pas la technique utilisée (celle de Winthrobe) lors du développement du DAS (celle de Westergren), et il n'y a pas de conversion possible pour les résultats obtenus avec chacune de ces 2 méthodes. En conséquence, nous avons utilisé une variante du DAS utilisant la valeur de la CRP pour remplacer la VS. Comme mentionné dans la recension, la CRP est une mesure plus directe de l'inflammation que la VS et plus sensible aux changements à court terme. La production de CRP est associée à la progression radiographique chez les patients avec PR et est considérée aussi valide que la VS pour mesurer l'activité de la maladie (Inoue et al. 2007). De plus, nous avons utilisé une variante du DAS appelée DAS28 qui est calculé sur 28 articulations au lieu de 44. Ce choix était nécessaire, car à l'époque le DAS nécessitait une évaluation de la douleur articulaire appelée Indice de Ritchie qui impliquait 4 niveaux de réponse selon l'intensité de la douleur lors de la pression sur les articulations, une mesure qui n'est plus utilisée en clinique. Au contraire, le DAS28 utilise simplement la présence ou l'absence de douleur, permettant un décompte rapide plus compatible avec la pratique clinique. Le nouveau décompte articulaire a été réduit à 28 articulations en excluant les articulations des pieds, des chevilles et des hanches (voir articulations en noir en annexe 1, figure 17, p.111). Les auteurs ayant créé ce score justifient l'élimination de certaines articulations en indiquant, d'une part, qu'ils désiraient faciliter l'évaluation en pratique clinique, et d'autre part, parce qu'il est difficile de voir des gonflements au niveau des hanches. L'indice d'activité DAS28-CRP est alors calculé par l'expression suivante :

$$\text{DAS28 - CRP} = 0,56\sqrt{\text{TJC28}} + 0,28\sqrt{\text{SJC28}} + 0,36 * \ln(\text{CRP} + 1) + 0,014\text{VAS} + 0,96$$

Légende : TJC = Nombre articulations douloureuses; SJC = Nombre d'articulations présentant inflammation (Synovite); CRP = Protéine C-Réactive; VAS = Degré de douleur (échelle 0-10) ressentie par le patient

Le score DAS28 a été développé en utilisant une analyse canonique discriminante et il a été validé par Prevoo en 1995 (Prevoo et al. 1995). Le DAS28 nous donne alors une valeur d'activité de la maladie sur une échelle de 0 à 10. Dans la littérature, une valeur de DAS28 inférieure à 2.6 est généralement associée à une rémission; entre 2.6 et 3.2 à une activité de la maladie légère ou contrôlée; à partir de 3.2, la maladie est considérée active; et finalement une valeur supérieure à 5.1 est considérée comme sévère (Dept. of Rheumatology University Medical Centre Nijmegen - the Netherlands).

Indice du degré des dommages articulaires

L'indice du degré des dommages radiographiques est calculé par le score radiographique appelé Sharp/van der Heijde (SvH). Il est composé de la somme pondérée du nombre de pincements et d'érosions des articulations des mains, des poignets et des pieds. On utilise la somme pondérée car certaines articulations (ex. celles des pieds) ont plus de poids que les autres dans le calcul du score pour les érosions, tandis que toutes les articulations ont le même poids pour les pincements articulaires. Ainsi, pour les érosions, les articulations des pieds ont un score maximal de 10 tandis que celles des mains et des poignets ont un score maximal de 5. Pour les pincements, toutes les articulations ont un score maximal de 4. Ces scores différents ont été suggérés pour ne pas donner trop d'importance aux articulations des mains, car elles comptent plus d'articulations évaluables que les pieds (van der Heijde 2000). Comme la lecture des radiographies est

faite par une seule personne, à savoir le chercheur principal, le processus de lecture est standardisé et à l'insu de l'information clinique du patient dans le but de diminuer la possibilité de biais. Pour assurer l'insu de l'investigateur, seulement le numéro de patient apparaît sur l'enveloppe des radiographies. De plus, la fidélité test-retest a été évaluée chez un échantillon de 120 radiographies où chacune des lectures a été faite 2 fois à deux moments différents dans le temps. La plus petite différence détectable ou reproductible était de 5 unités sur une échelle de 0 à 448, ce qui est compatible avec l'information disponible dans la littérature (Bruynesteyn et al. 2002; Bruynesteyn et al. 2004). Différents scores résultent de ce décompte articulaire dont principalement le score total et le score d'érosions. Le premier consiste à la somme pondérée des érosions et des pincements présents dans les articulations des mains, poignets et pieds. Le score d'érosion quant à lui consiste seulement en la somme pondérée des articulations présentant des érosions.

3.5 Stratégie d'analyse

3.5.1 Choix des types d'analyses statistiques

Le but de cette étude est d'explorer les relations entre les variables cliniques à partir d'une approche peu utilisée jusqu'à maintenant et qui permet d'éviter d'imposer des liens «*a priori*» entre les variables. Comme la maladie peut évoluer de différentes façons et que la prévision en est difficile, nous avons décidé d'utiliser des méthodes statistiques exploratoires pour tenter de découvrir s'il existe des sous-groupes au sein de

notre cohorte et éventuellement comprendre les raisons biologiques de ces regroupements.

Les données fournies contiennent aussi bien des variables de nature quantitative (ex. l'âge, le nombre d'articulations avec érosions, la durée de la raideur matinale...) que qualitative (ex. le sexe, la présence ou non d'anti-Sa, la présence d'un allèle génomique au locus HLA-DR,...). Puisque nous nous intéressions à l'analyse simultanée d'un ensemble de variables de différentes natures, nous avons décidé de codifier les variables continues et d'appliquer ensuite un modèle d'analyse des correspondances multiples. Les choix des points critiques pour la codification ont été déterminés par le clinicien en considérant l'information clinique ou bien en observant la distribution des valeurs des patients par rapport aux valeurs normales dans la population générale. Finalement, c'est la classification hiérarchique qui a été utilisée pour révéler s'il existe des sous-groupes de patients dans la cohorte. L'analyse des correspondances multiples (ACM) et la classification hiérarchique sont toutes deux des méthodes d'analyse incluses dans ce qu'on appelle le «*Data Mining*» ou bien «l'extraction de connaissances à partir de données (ECD)».

3.5.2 «*Data mining*»

Ce regroupement de méthodes statistiques a vu le jour dans les années 80, quand les professionnels ont commencé à se soucier des grands volumes de données informatiques. Le data mining est de plus en plus utilisé de nos jours. On définit le data mining comme *"un processus non-trivial d'identification de structures inconnues,*

valides et potentiellement exploitables dans les bases de données (Fayyad, 1996)''. Il est aussi défini comme un processus de sélection, d'exploration, de modification et de modélisation de grandes bases de données afin de découvrir des relations entre les données jusqu'alors inconnues (Fayyad, Gregory Piatetsky-Shapiro & Smyth 1996; Han, Kamber 2001; Larose 2005; Tufféry 2005a).

Principe

Le data mining est un processus d'analyse dont l'approche est différente de celle utilisée en statistique classique. Cette dernière présuppose en général que l'on se fixe une hypothèse que les données vont nous permettre ou non de confirmer. Au contraire, le data mining adopte une démarche beaucoup plus empirique et essaye ainsi de faire émerger, à partir des données brutes, des hypothèses que l'expérimentateur peut ne pas soupçonner, et dont il aura à valider la pertinence.

3.5.2 Analyse des correspondances multiples

But et principe

L'analyse des correspondances multiples (ACM) permet de décrire les relations existant entre un groupe de variables qualitatives observées chez un nombre donné d'individus. Le but est d'obtenir un nombre réduit de composantes qui tiennent compte de l'ensemble des variables de l'analyse, donc de réduire la plus grande partie de l'information initiale en un nombre restreint de dimensions selon les correspondances entre les variables. Cette réduction est particulièrement intéressante lorsque le nombre

initial de dimensions est très élevé, donc s'il y a un nombre considérable de variables dans l'analyse. Cette méthode est l'une des seules permettant d'analyser les liens existant entre un ensemble de variables qualitatives. Les composantes factorielles ainsi obtenues sont donc alors continues, ce qui facilite l'utilisation de la méthode de classification choisie pour obtenir nos sous-groupes.

Un des objectifs importants de l'ACM consiste à comprendre la structure des variables étudiées à l'aide de la représentation graphique des observations. Lorsque le nombre de dimensions du problème est supérieur à trois, il devient impossible de représenter graphiquement toutes ses dimensions dans un seul graphique. Cette analyse représentera alors les individus et les variables dans un espace de dimension réduite en déformant le moins possible le nuage de points et elle permettra d'expliquer les liaisons éventuelles entre les variables originales. L'analyse va donc donner un sous-espace de dimension raisonnable, tel que la projection des variables ou des individus sur ce sous-espace retienne le plus d'information possible, donc tel que le nuage de points soit le plus dispersé possible (Baccini, Besse 2004).

Méthode et interprétation

La principale étape est de bien choisir les variables à analyser et de faire une codification des variables quantitatives de façon judicieuse. C'est généralement le clinicien qui fera le choix des catégories en se basant sur son jugement clinique.

Une fois l'ACM appliquée aux données, l'étape suivante consiste à analyser les résultats obtenus afin de réduire le nombre de dimensions. Cette réduction sera faite en considérant le pourcentage d'inertie expliqué (variabilité expliquée) par chacune des

composantes. L'endroit où il y a une chute importante d'inertie peut être un indice intéressant pour l'aide à la décision. Le choix final consiste en un choix éclairé entre une réduction logique du nombre de dimensions, de façon à simplifier le modèle et en un pourcentage d'inertie expliqué important.

Une fois le modèle réduit aux dimensions optimales, il est alors intéressant d'étudier chacun des axes principaux afin d'évaluer quelles variables contribuent essentiellement à sa création. On observe donc la qualité de la représentation de chacune des variables et les contributions relatives et absolues de chacune d'elles. La qualité de la représentation est mesurée par le cosinus carré de l'angle de la modalité par rapport à l'axe. Ce \cos^2 est le pourcentage pris par l'axe dans la dispersion de la modalité. Plus il est proche de 1, plus la variable est bien représentée. Notons aussi que la somme de l'ensemble des \cos^2 sur l'ensemble des axes vaut 1 (Tufféry 2005b).

Finalement, une représentation graphique de chacune des dimensions retenues sera réalisée de façon à comprendre les relations existant entre les variables de l'analyse. On notera que cette étape est généralement la partie la plus importante du processus. En pratique, lorsqu'un chercheur réalise une ACM, il ne publie que les plans engendrés par les 2 ou 3 premiers axes principaux et les tableaux des contributions seront parfois mentionnés dans le texte mais rarement publiés. Cette pratique est justifiable en partie par le fait que la représentation graphique est la plus riche en information, et que c'est l'avantage majeur de l'analyse, soit la représentation synthétique d'un grand tableau de chiffres.

Les variables les plus éloignées du centre sont généralement celles les mieux représentées, mais il est important de valider cette conclusion avec la qualité de

représentation de chacune d'elles. En effet, comme le nombre de dimensions est considérable mais que les variables sont représentées sur un seul plan, il est possible qu'une variable paraisse bien représentée alors qu'en réalité ce n'est pas le cas. Une fois bien définies quelles variables sont les mieux représentées, les oppositions et les proximités sont alors étudiées de façon à mettre en évidence la nature de ces liens. De là, un nom clinique est généralement attribué à chacun des axes selon l'identification des variables qui contribuent à leur création. On pourra alors reconnaître quelles variables se sont avérées les plus importantes pour résumer l'information initiale.

L'ACM est donc un outil pratique pour résumer de façon objective l'information contenue dans un ensemble de variables. Son utilisation est intéressante lorsque l'on a pour objectif de diminuer la dimensionnalité des données, de résumer l'information et de découvrir des relations entre les variables ou des associations entre les individus observés et les variables mesurées sur ces individus.

3.5.2 La classification hiérarchique ascendante (agglomérative)

La classification hiérarchique est fort utile dans de nombreux domaines. Elle est utilisée par exemple en biologie ou en zoologie afin de créer des taxonomies, c'est-à-dire des schémas de classification des animaux, des plantes, des insectes. Cette méthode permet également de décrire des populations par des modèles statistiques, et est fréquemment utilisée en marketing pour cibler le type de publicité selon la population visée. Finalement, dans notre domaine d'intérêt, la médecine, elle est utilisée de façon à affecter de nouveaux cas ou patients à une classe pour l'identification et pour établir un diagnostic (Anderberg 1973).

But et principe (mesure de distance & indice d'agrégation)

Soit un échantillon de données (sujets, personnes, cas), toutes décrites par un ensemble de variables (attributs, caractéristiques, mesures). Le but de la classification est de grouper ensemble les données dans des classes pour lesquelles le degré d'association entre les éléments d'une même classe est fort, et pour lesquelles le degré d'association est faible entre les éléments de classes distincts. Le principe est donc de regrouper les sujets les plus similaires ou les plus proches selon l'ensemble de leurs caractéristiques à partir d'une **mesure de distance** ou d'un **indice d'agrégation** (Anderberg 1973).

La **mesure de distance** ou de **similitude** entre les observations permettra d'établir le degré d'association entre les individus ou groupes d'individus. Ces distances seront présentées dans une matrice où les lignes et les colonnes seront les individus et celle-ci sera évidemment symétrique avec une diagonale nulle.

	Sujet 1	Sujet 2	...	Sujet k	...	Sujet n
Sujet 1	0					
Sujet 2	$d(2,1)$	0				
...			
Sujet k	$d(k,1)$	$d(k,2)$...	0		
...	
Sujet n	$d(n,1)$	$d(n,2)$...	$d(n,k)$...	0

De plus, il y a une multitude de façons de choisir ou de former la mesure de distance selon la nature des variables, qu'elles soient continues, nominales ou ordinales. Dans le cas où les variables seraient de nature continue, une métrique sera utilisée. Pour les variables catégoriques ou dichotomiques, ce sera plutôt une ultramétrique.

Une fois choisie la mesure de distance, il faut déterminer la méthode de regroupement des individus appelée : **Indice d'agrégation**. L'indice permet donc, à

partir de la distance (similarité) entre les sujets (ou groupes d'individus), de sélectionner les groupes les plus proches afin de les fusionner. Elle est appliquée sur la matrice des distances calculée à l'étape précédente. Les méthodes généralement rencontrées sont :

- Saut minimal ou plus petite distance (Single Linkage);
- Saut maximal ou plus grande distance (Complete Linkage);
- Distance moyenne (Average linkage);
- Critère d'agrégation selon l'Inertie appelé Critère de Ward;

Selon Tufféry, la méthode du saut maximal tend à produire des classes de diamètres égaux. Elle est très sensible aux données extrêmes et est donc peu utilisée. La distance moyenne entre deux observations tend à produire des classes de même variance. La méthode du saut minimal est moins adaptée pour détecter les classes de formes sphériques, mais elle a de bonnes propriétés théoriques et sait très bien détecter des classes allongées. La méthode de Ward est l'une de celles qui correspondent le mieux à l'objectif de la classification. Cette méthode tend à produire des classes de même effectif et de forme sphérique, mais elle est très sensible aux valeurs extrêmes. Malgré tout, c'est la méthode la plus utilisée, car elle s'applique bien aux problèmes réels. D'autres méthodes peuvent être utilisées et sont présentées dans certains livres sur le sujet (Anderberg 1973; Saporta 1990; Tufféry 2005b).

Voici une description de la méthode classique à appliquer pour effectuer une classification hiérarchique (Anderberg 1973).

- 1) Au départ, chaque sujet est considéré comme étant un groupe avec sujet unique;
- 2) Par la suite, les distances entre chacun des sujets/groupes selon l'ensemble des caractéristiques sont calculées;

- 3) Les deux groupes les plus proches sont fusionnés et remplacés par un seul groupe à l'aide de l'indice d'agrégation;
- 4) Les étapes 2 et 3 sont répétées jusqu'à l'obtention d'un seul et unique groupe;
- 5) Un **ARBRE DE CLASSIFICATION**, aussi appelé **dendrogramme** découlera de cette analyse;
- 6) L'utilisateur choisit alors le **nombre de classes / groupes** (Clusters) à partir de l'arbre ainsi obtenu tout en considérant les objectifs cliniques.

Dans notre cas, comme la classification est faite sur des coordonnées factorielles, nous avons décidé d'utiliser la distance Euclidienne au carré, mais il existe de nombreuses autres méthodes.

$$\text{Distance Euclidienne au carré : } d^2(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$

Comme nous utilisons les données transformées issues de l'ACM (coordonnées factorielles), il est possible d'utiliser des méthodes d'agrégation qui ne sont généralement applicables que sur des variables continues. Nous avons donc décidé d'utiliser la méthode de Ward. Par contre, peu importe la méthode, lorsqu'on utilise les données issues de l'ACM, il est préférable d'inclure tous les facteurs pour être en mesure de rendre compte de la totalité du phénomène. Dans notre cas, à des fins exploratoires, nous avons décidé d'effectuer la classification sur les 4 premiers facteurs, mais l'analyse a aussi été faite chez l'ensemble des facteurs à titre comparatif.

Supposons un échantillon de n observations partitionnées en g groupes différents donc le $i^{\text{ème}}$ groupe est composé de n_i observations avec une moyenne \bar{x}_i .

La somme des carrés des écarts intra-groupe défini pour chacun des groupes g est :

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

La méthode de classification de Ward, présentée ci-dessous, consiste donc à minimiser l'inertie intra-classe, donc à minimiser la variabilité à l'intérieur des groupes. Cette technique vise donc à fusionner les deux groupes qui résultent en la plus petite augmentation de W à chacun des stades de regroupement, donc le minimum de perte d'inertie (ou de variabilité) possible.

1. Au départ, chacun des groupes est constitué d'une observation unique
2. Les deux groupes pour lesquels l'augmentation de W est la plus petite seront regroupés pour n'en former qu'un seul
3. Un processus itératif sera alors enclenché jusqu'à l'obtention d'un seul et unique groupe final.

Au lieu d'appliquer la méthode de Ward sur les données transformées par l'ACM, il est aussi possible d'utiliser la même méthode mais sur les données d'origine, toujours en appliquant la métrique du Khi-carré. Pour ce faire, on utilise les données codées sous forme de tableau disjonctif complet. Un tableau disjonctif complet est l'ensemble des variables codées sous forme de 0 et de 1. Si la variable est constituée par exemple de trois catégories, 3 variables différentes seront créées pour représenter cette variable avec

seulement des 0 et des 1. Comme cette analyse nous semblait très intéressante, elle a aussi été faite à titre comparatif avec la méthode de Ward.

Évidemment, pour tout type d'analyse exploratoire, la principale étape à considérer lors de l'implantation de ce type d'analyse est de bien choisir les variables à inclure. C'est principalement le choix des variables qui aura la plus grande influence sur l'ultime résultat de l'analyse. L'algorithme détaillé que l'on retrouve dans le livre de Anderberg est présenté en annexe (Anderberg 1973).

3.5.3 Application concrète aux données

Concrètement, nous avons appliqué l'ACM et la classification sur les variables de l'inclusion dont les variables socio-démographiques (âge, sexe), les autoanticorps (anti-Sa, anti-CCP, FR), la présence ou non de traitements entre le début des symptômes et l'inclusion dans l'étude (DMARD, corticostéroïdes), les données biologiques (hémoglobine, PMN, VS, CRP) ainsi que les données cliniques et radiographiques (durée entre le début des symptômes et l'inclusion, durée de la raideur matinale dans les articulations, présence de synovites, érosions significatives). Nous avons donc un portrait des interactions entre ces variables et il est donc intéressant d'observer si des sous-groupes de la population se forment à partir de l'information déjà présente lors de l'inclusion des patients dans l'étude.

Malgré le fait que la banque de données était constituée de 376 patients, seulement 211 patients ont été considérés lors de l'analyse. Une première raison qui explique la diminution du nombre de sujets est le choix des analyses. En effet, dès qu'il y a des valeurs manquantes pour l'une des variables de l'analyse, le patient est

automatiquement exclu. Nous voulons aussi étudier, selon la formation des sous-groupes, comment la maladie se développe lors du suivi. Pour ce faire, nous devons sélectionner les patients ayant complété un certain nombre de visites, ce qui diminue encore une fois le nombre de sujets. Le choix de 30 mois depuis le début des symptômes a été fait de façon à obtenir un suivi suffisamment long tout en conservant un nombre assez important de patients pour l'analyse. Nous aurions pu choisir un intervalle plus court, i.e. 18 mois, mais nous considérons que ce temps de suivi n'est pas suffisamment long pour que le pronostic individuel des patients devienne cliniquement clair. En effet, le traitement commence généralement à être administré au début du suivi (qui peut aller de 1 à 12 mois après le début des manifestations cliniques) et un certain temps est nécessaire avant de trouver le type et le dosage de médication adéquats pour chacun des individus ou pour que cette médication atteigne son efficacité maximale. De plus, les patients qui évoluent plus lentement ont moins de temps pour se différencier les uns des autres. Inversement, nous aurions pu choisir un plus long temps d'observation (par exemple 5 ans). Toutefois, notre étude a débuté en l'an 2000, et bon nombre de nos patients n'ont pas encore atteint 5 ans d'observation. Plus la visite de suivi est éloignée (ex. 5 ans), moins il y a de patients ayant complété toutes les visites de suivi, ce qui diminue rapidement le nombre de patients à analyser. Le choix de 30 mois a donc été considéré comme un bon compromis. Les sous-groupes formés à partir des données à l'inclusion sont donc d'abord comparés entre eux selon leurs caractéristiques à l'inclusion, avant de les comparer selon les variables sur l'état du patient à 30 mois.

3.4 Considérations éthiques

Le projet de cohorte longitudinale prospective a été accepté par le comité d'éthique de la recherche du CHUS en 1998. Chacun des participants a reçu une description détaillée du déroulement de l'étude, ainsi que des avantages et inconvénients possibles de leur participation au projet de recherche. Personne n'a été obligé de participer et chacun était libre d'abandonner en cours d'étude sans pression de notre part. Chaque participant a donc donné un consentement libre et éclairé.

Une attention particulière a été portée afin d'assurer la confidentialité. La base de données a été dénominalisée et les sérums et le matériel d'ADN ont été codés pour assurer l'insu de l'investigateur et des laboratoires.

4 - Résultats

4.1 Patients à l'étude et leurs caractéristiques

Lors de l'analyse, la banque de données comportait 376 patients, dont 54 ont quitté l'étude pour diverses raisons et à différentes étapes du suivi. Le tableau 3 représente le nombre de patients à chacune des visites ainsi que le nombre de pertes au suivi. Il est important ici de noter qu'étant donné que nous sommes en présence d'une cohorte prospective, de nouveaux patients sont inclus régulièrement. Ainsi, si l'on considère le nombre de sujets à 18 mois, ils n'ont pas tous complété leur deuxième visite, donc la somme des pertes au suivi et du nombre de sujets ayant complété leur visite de 18 mois ne donnera pas le même total que le nombre de sujets à l'inclusion. De plus, le tableau 3 indique que 247 patients ont complété leur visite jusqu'à 30 mois, mais seulement 211 sujets sont inclus dans l'analyse. Cette différence s'explique par le fait que les analyses choisies nécessitent d'avoir l'information pour toutes les variables incluses dans l'analyse. Dès qu'il manque une seule valeur pour une variable, le patient est automatiquement exclu.

Il est intéressant de remarquer que la plupart des abandons ont lieu surtout au début de l'étude, en particulier chez des patients dont la maladie semble être spontanément rentrée dans l'ordre. Une fois l'étude en cours, il y a peu d'abandons et ceux-ci sont généralement le fait de patients qui décèdent entre deux visites consécutives.

Tableau 3. Patients à l'étude

Visite	Nombre de patients	Nombre de pertes au suivi		
		Total	Décès	Abandon
Inclusion	376	---	---	---
18 mois	305	18	2	16
30 mois	247	13	5	8
42 mois	209	7	5	2
60 mois	162	8	6	2
Après 60 mois	---	8	6	2

Les caractéristiques de base des patients inclus dans les analyses ont été comparées aux caractéristiques des patients perdus avant 30 mois (Tableau 4) afin de mettre en évidence un éventuel biais de sélection. Les autres patients perdus en cours d'étude ne sont pas analysés ici, car plusieurs ont abandonné ou sont décédés après 30 mois. Les patients ayant quitté après 30 mois se retrouvent dans les patients analysés, puisqu'ils ont complété leurs 3 premières visites. Aucune différence significative n'a été détectée entre les patients analysés et ceux perdus au suivi avant 30 mois.

Tableau 4. Comparaison des caractéristiques à l'inclusion entre les patients inclus dans l'analyse et ceux perdus au suivi

	Inclus dans les analyses (n=211)	Pertes au suivi (n=32)
Femmes (%)	124 (59)	16 (50)
Âge (années) - Moy. (ET)	57,5 (15,4)	59,3 (19,8)
Durée entre l'inclusion et le début de la maladie (mois) - Moy. (ET)	4,8 (3,0)	4,1 (3,3)
Critère de classification de la PR selon l'ACR \geq 4 (%)	172 (82)	25 (78)
FR \geq 40 IU/mL (%)	93 (44)	16 (50)
Anti-CCP > 20 U/mL; positif (%)	79 (37)	11 (38)
Anti-Sa (%)	51 (24)	5 (19)
CRP \geq 8.0 mg/L (%)	147 (70)	24 (75)
Capacité fonctionnelle		
M-HAQ - Méd. (EIQ)	0,75 (1,13)	0,75 (0,72)
M-HAQ score \geq 1.0 (%)	99 (47)	13 (41)
DAS28_CRP4_V1		
Dommages articulaires (Sharp)		
Score Total - Méd. (EIQ)	3,0 (7,0)	1,0 (7,0)
Score d'érosions - Méd. (EIQ)	1,0 (3,0)	1,0 (2,0)
Score d'érosion \geq 5 (%)	38 (18)	3 (13)

Aucune différence significative entre les patients inclus dans l'analyse et ceux qui ont quitté l'étude dans les 30 mois premier mois

Une analyse descriptive de l'ensemble des caractéristiques des patients a été effectuée afin de comparer nos résultats avec les autres études portant sur les polyarthrites débutantes (Tableau 5 et 6 ci-dessous). Comme mentionné dans la méthodologie (section 3.5.3, Application concrète aux données) les analyses ont été réalisées chez les 211 patients ayant complété leur 3 premières visites de suivi. Il est intéressant d'observer que les patients rencontrent les caractéristiques attendues de la population atteinte de polyarthrite, c'est-à-dire une population âgée d'environ 60 ans et majoritairement féminine. De plus, la durée entre le début des symptômes et l'inclusion dans l'étude est plutôt courte, soit d'une durée d'environ 4 mois, ce qui suggère que les

patients se présentent assez tôt dans le développement de la maladie. Dès le départ, près de 80% des patients remplissent les critères de classification de la PR selon l'ACR. De plus, environ 70% des patients ont une inflammation systémique importante au début de l'étude, ce qui n'est pas surprenant puisque l'un des critères d'inclusion consiste à avoir au moins 3 synovites, ce qui traduit un degré important d'inflammation locale des articulations. Finalement, le FR est présent chez près de 45% des patients, l'anti-CCP chez 37% et l'anti-Sa chez seulement 25%.

Tableau 5. Caractéristiques à l'inclusion (n=211)

Caractéristiques	Inclusion
Femmes (%)	127 (60.2)
Âge (années) – Moy. (ET)	57 (15)
Durée entre l'inclusion et le début de la maladie (mois) - Moy. (ET)	4.8 (3.0)
Critère de classification de la PR selon l'ACR ≥ 4 (%)	166 (78.7)
FR ≥ 40 IU/mL (%)	91 (43.3)
Anti-CCP > 20 U/mL; positif (%)	78 (37.1)
Anti-Sa (%)	52 (24.9)
CRP ≥ 8.0 mg/L (%)	146 (69.5)

Note : Moyenne (Moy.), Écart-type (ET)

En plus d'étudier les caractéristiques de l'inclusion, l'évolution à 30 mois a aussi été analysée. En considérant les données résumées au tableau 6, on observe que l'incapacité et l'activité de la maladie diminuent à 30 mois, mais que les dommages radiographiques augmentent. En effet, près de la moitié des patients avaient une incapacité importante au départ (M-HAQ score ≥ 1.0) comparativement à seulement 11% à 30 mois. Même chose pour l'activité de la maladie. Un DAS supérieur à 5.1 est considéré comme une activité sévère et une valeur inférieure à 2.6 est généralement associée à une rémission. On voit donc clairement que la maladie est très active au départ

et diminue pour devenir fréquemment en rémission. Comme mentionné précédemment, le score total pour les dommages articulaires fait référence à la somme des érosions et des pincements des mains, poignets et pieds tandis que le score d'érosions ne considère que la somme des érosions. Généralement, une fois l'articulation endommagée, il est rare que l'articulation se répare d'elle-même et dans le cas présent, on observe même une augmentation importante. On ne peut tout de même pas attribuer cette augmentation à l'ensemble des patients étant donné la variabilité observée, mais elle n'est pas négligeable. On peut donc constater que les dommages perdurent et même s'amplifient mais qu'il est tout de même possible de diminuer la douleur et l'incapacité à moyen terme malgré cette détérioration articulaire.

Tableau 6. Caractéristiques à l'inclusion et au suivi (n=211)

Caractéristiques	Inclusion	30 mois
Capacité fonctionnelle (M-HAQ) – Méd. (EIQ)	0.81 (1.13)	0.13 (0.63)
M-HAQ score \geq 1.0 (%)	98 (46.7)	23 (11.0)
Indice d'activité maladie (DAS28) – Moy. (ET)	5.3 (1.4)	2.7 (1.1)
Dommages articulaires (Sharp)		
Score Total – Méd. (IQR)	3.0 (7.0)	8.0 (12.8)
Score d'érosions – Méd. (IQR)	1.0 (3.0)	5.0 (9.0)
Score d'érosion \geq 5 (%)	41 (19.5)	114 (54.3)

Toutes les comparaisons entre l'inclusion et 30 mois sont statistiquement significative avec un $p < 0.0001$

Note : Moyenne (Moy.), Écart-type (ET), Médiane (Méd.), l'étendue interquartile (EIQ)

4.2 Résultats de l'analyse des correspondances multiples

Une analyse des correspondances multiples a été effectuée afin d'étudier les relations entre les variables et de faciliter l'analyse de classification qui nous intéresse plus particulièrement. Tout d'abord, le tableau 7 ci-dessous fournit les valeurs propres classées en ordre décroissant, ainsi que les pourcentages d'inertie expliquée pour chacun des axes principaux. La dernière colonne correspond au diagramme des valeurs propres rangées en ordre décroissant. Ce diagramme est un support visuel pour aider à choisir le nombre optimal de dimensions à conserver afin de réduire les dimensions du problème. En effet, dans le cas où il y aurait une grande chute d'inertie expliquée (grande différence entre les niveaux), seulement les dimensions les plus importantes sont conservées. Dans le cas présent, il n'y a pas une si grande différence entre les dimensions, mais on remarque une différence plus importante pour les axes 3 et 4. Les 4 premiers axes expliquent respectivement 9,70%, 8,90%, 8,07% et 6,31% de l'inertie totale pour un total de 32,98%. On a donc décidé de ne conserver que ces 4 premiers axes de façon à réduire les dimensions du problème tout en conservant un pourcentage d'inertie expliqué assez important pour ce type d'analyse.

Les valeurs de Khi-Carré présentées dans le même tableau donnent aussi un aperçu de l'existence d'une certaine dépendance entre les variables. En effet, si les axes étaient totalement indépendants, comme il y a 25 dimensions au total, les 4 premiers axes auraient représenté : $4 * \left(\frac{100}{25}\right) = 16\%$ et non 33 pourcent de l'inertie.

Tableau 7. Décomposition de l'Inertie et du Khi-Carré

Singular Value	Principal Inertia	Chi-Square	Percent	Cumulative Percent	2 4 6 8 10
					-----+-----+-----+-----+-----
0,39	0,153	538,76	9,70	9,7	*****
0,37	0,140	494,34	8,90	18,6	*****
0,36	0,127	448,03	8,07	26,67	*****
0,32	0,099	350,42	6,31	32,98	*****
0,30	0,088	308,88	5,56	38,54	*****
0,29	0,082	287,86	5,18	43,72	*****
0,28	0,078	276,16	4,97	48,7	*****
0,27	0,075	263,59	4,75	53,44	*****
0,27	0,071	251,06	4,52	57,96	*****
0,26	0,068	238,08	4,29	62,25	*****
0,25	0,061	215,31	3,88	66,13	*****
0,24	0,056	197,24	3,55	69,68	*****
0,23	0,054	190,78	3,44	73,12	*****
0,22	0,050	177,29	3,19	76,31	*****
0,22	0,049	172,86	3,11	79,42	*****
0,21	0,045	156,84	2,82	82,24	*****
0,20	0,041	143,55	2,58	84,83	*****
0,19	0,038	133,44	2,40	87,23	*****
0,19	0,036	126,41	2,28	89,51	*****
0,18	0,032	112,12	2,02	91,53	*****
0,17	0,028	99,13	1,78	93,31	****
0,15	0,024	83,97	1,51	94,82	****
0,15	0,022	76,75	1,38	96,21	***
0,14	0,019	68,36	1,23	97,44	***
0,13	0,016	55,71	1,00	98,44	***
0,12	0,014	49,38	0,89	99,33	**
0,10	0,011	37,24	0,67	100	**
Total	1,576	5553,54	100		
Degrees of Freedom = 8820					

Figure 7. Graphiques de l'ACM, dimensions 1 et 3, représentation des variables

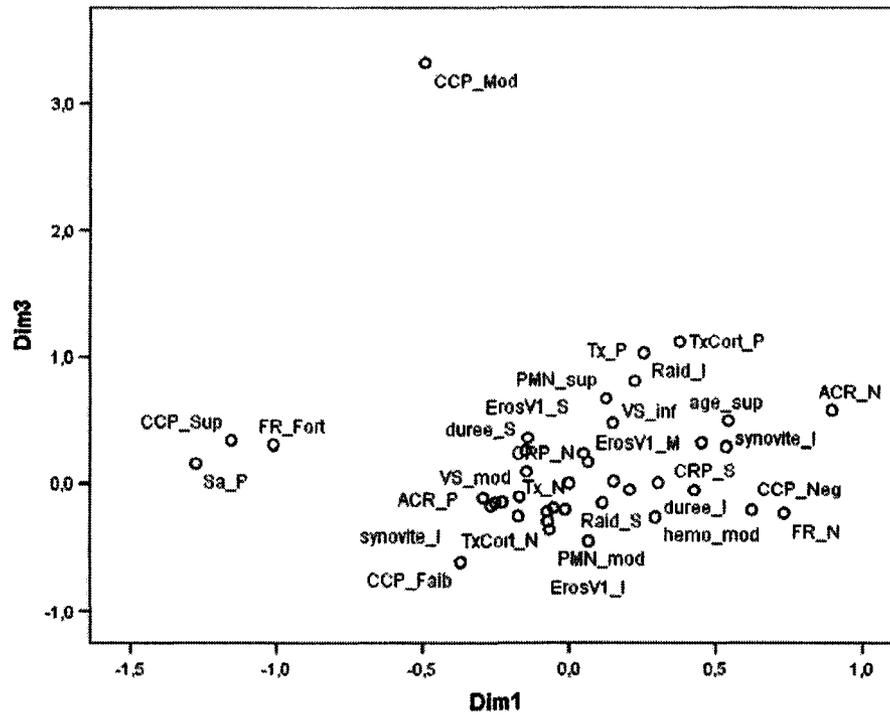


Figure 8. Graphiques de l'ACM, dimensions 1 et 4, représentation des variables

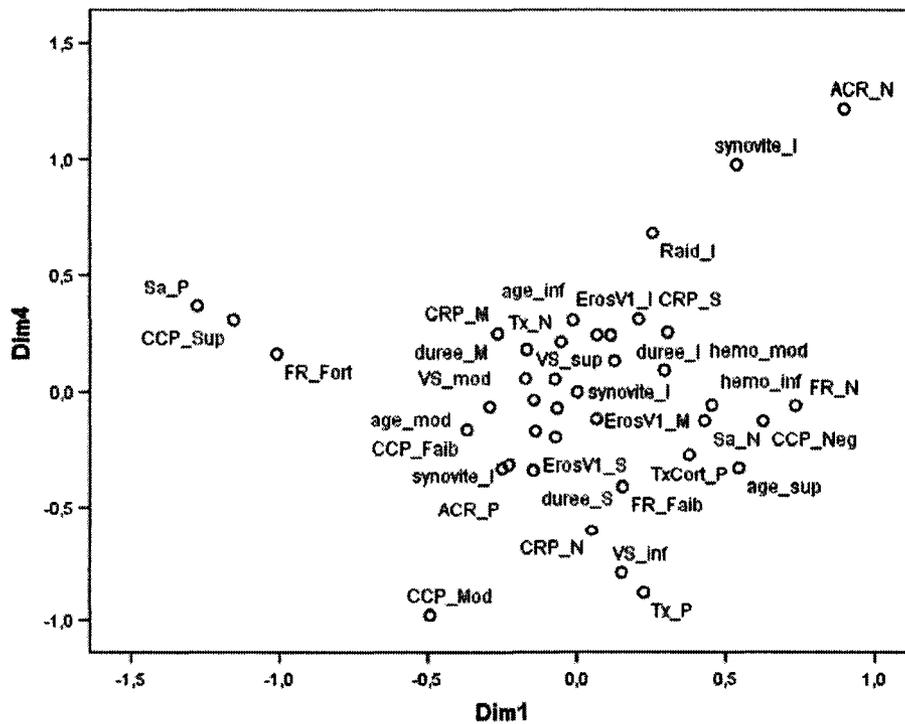


Tableau 8. Coordonnées factorielles et contributions de chacune des variables à la création des axes

Variables	Coordonnées factorielles				Contributions relatives				Qualité de la représentation			
	Dim1	Dim2	Dim3	Dim4	Contr1	Contr2	Contr3	Contr4	SqCos1	SqCos2	SqCos3	SqCos4
Sa P	-1,275	0,113	0,158	0,367	0,161	0,001	0,004	0,021	0,545	0,004	0,008	0,045
Sa N	0,430	-0,043	-0,051	-0,125	0,054	0,001	0,001	0,007	0,548	0,005	0,008	0,046
CCP Neg	0,624	-0,022	-0,205	-0,125	0,096	0,000	0,015	0,006	0,649	0,001	0,070	0,026
CCP Faib	-0,369	-0,080	-0,614	-0,163	0,002	0,000	0,009	0,001	0,006	0,000	0,017	0,001
CCP Mod	-0,492	-0,348	3,324	-0,980	0,001	0,001	0,090	0,009	0,003	0,002	0,155	0,014
CCP Sup	-1,154	0,059	0,342	0,307	0,166	0,000	0,021	0,019	0,621	0,002	0,054	0,044
FR N	0,734	-0,067	-0,231	-0,060	0,111	0,001	0,016	0,001	0,595	0,005	0,059	0,004
FR Faib	0,155	-0,124	0,018	-0,411	0,001	0,001	0,000	0,009	0,002	0,001	0,000	0,015
FR Fort	-1,010	0,106	0,309	0,161	0,158	0,002	0,021	0,006	0,662	0,007	0,062	0,017
hemo inf	0,454	1,393	0,324	-0,058	0,003	0,032	0,002	0,000	0,008	0,076	0,004	0,000
hemo mod	0,295	0,754	-0,262	0,091	0,010	0,069	0,011	0,001	0,034	0,221	0,027	0,003
hemo sup	-0,144	-0,392	0,091	-0,036	0,006	0,046	0,003	0,001	0,045	0,329	0,018	0,003
PMN mod	-0,066	-0,105	-0,363	-0,073	0,001	0,003	0,049	0,002	0,008	0,020	0,242	0,010
PMN sup	0,127	0,185	0,675	0,131	0,002	0,005	0,091	0,004	0,009	0,018	0,246	0,009
VS inf	0,151	-1,054	0,482	-0,795	0,002	0,101	0,028	0,083	0,006	0,291	0,061	0,166
VS mod	-0,168	-0,266	-0,103	0,180	0,005	0,013	0,003	0,009	0,021	0,053	0,008	0,024
VS sup	0,114	0,903	-0,152	0,240	0,002	0,130	0,005	0,013	0,007	0,470	0,013	0,033
CRP N	0,051	-0,758	0,241	-0,606	0,000	0,079	0,011	0,074	0,001	0,265	0,027	0,169
CRP M	-0,266	-0,222	-0,176	0,249	0,009	0,007	0,006	0,013	0,035	0,024	0,015	0,031
CRP S	0,210	0,886	-0,048	0,309	0,006	0,120	0,000	0,021	0,024	0,428	0,001	0,052
age inf	-0,013	-0,683	-0,205	0,306	0,000	0,069	0,008	0,020	0,000	0,240	0,022	0,048
age mod	-0,293	0,152	-0,112	-0,066	0,014	0,004	0,003	0,001	0,063	0,017	0,009	0,003
age sup	0,545	0,698	0,501	-0,330	0,028	0,050	0,034	0,016	0,092	0,151	0,078	0,034
ACR N	0,898	-0,510	0,584	1,218	0,065	0,023	0,040	0,190	0,205	0,066	0,087	0,377
ACR P	-0,228	0,126	-0,148	-0,313	0,016	0,006	0,010	0,049	0,203	0,062	0,085	0,384
duree I	0,306	0,372	0,007	0,253	0,013	0,021	0,000	0,014	0,050	0,073	0,000	0,034
duree M	-0,173	0,063	-0,253	0,056	0,004	0,001	0,012	0,001	0,015	0,002	0,032	0,002
duree S	-0,146	-0,483	0,265	-0,339	0,003	0,032	0,013	0,023	0,010	0,108	0,033	0,054
Raid I	0,256	-0,298	1,030	0,683	0,006	0,009	0,135	0,065	0,019	0,025	0,302	0,133
Raid S	-0,072	0,081	-0,294	-0,198	0,002	0,002	0,038	0,019	0,018	0,023	0,301	0,137
Tx N	-0,053	-0,050	-0,194	0,211	0,001	0,001	0,017	0,023	0,011	0,011	0,156	0,184
Tx P	0,226	0,191	0,812	-0,883	0,004	0,003	0,073	0,095	0,012	0,009	0,158	0,188
TxCort N	-0,074	-0,093	-0,221	0,052	0,002	0,003	0,023	0,001	0,027	0,043	0,242	0,014
TxCort P	0,379	0,444	1,117	-0,272	0,009	0,014	0,118	0,008	0,029	0,039	0,249	0,015
synovite I	0,538	-0,433	0,292	0,976	0,022	0,016	0,009	0,117	0,071	0,046	0,021	0,235
synovite M	0,001	-0,004	0,001	-0,001	0,000	0,000	0,000	0,000	0,003	0,015	0,002	0,003
synovite S	-0,250	0,272	-0,151	-0,330	0,013	0,016	0,007	0,035	0,067	0,079	0,024	0,116
ErosV1 I	0,068	-0,549	-0,451	0,242	0,001	0,048	0,043	0,014	0,003	0,177	0,119	0,034
ErosV1 M	0,066	-0,080	0,170	-0,119	0,001	0,001	0,005	0,003	0,002	0,003	0,013	0,006
ErosV1 S	-0,139	0,706	0,362	-0,169	0,002	0,069	0,024	0,006	0,009	0,232	0,061	0,013

La dimension 1 représente 9.70% de l'inertie totale et est principalement expliquée par les anticorps forts ou faibles et à un degré moindre par la présence d'au moins 4 critères de classification de la PR selon l'ACR. On observe graphiquement (Figure 6 ou Figure 7, en comparant les modalités de gauche à celles de droite) une opposition entre la présence et l'absence d'anticorps, ce qui est très logique biologiquement. De plus, la modalité «avoir moins de 4 critères de classification de la PR selon l'ACR» se retrouve avec les modalités représentant les anticorps absents. Cette association s'explique probablement par le fait que la présence de FR est un des 7 critères de l'ACR. Donc l'absence de FR pourrait être associée à des patients ne remplissant pas les critères diagnostiques d'une PR. Cet axe pourrait donc représenter l'opposition entre les polyarthrites séropositives et séronégatives.

L'axe 2 explique, quant à lui, 8.90% de l'inertie et est principalement créé par les modalités suivantes : une VS élevée, une CRP élevée, un taux important d'hémoglobine, et à degré moindre par une VS faible, une CRP faible, des patients plus jeunes, des érosions importantes et une hémoglobine diminuée modérément. La VS élevée, la CRP élevée, l'hémoglobine faible (i.e. l'anémie sévère) et modérée (i.e. l'anémie légère) ainsi que les érosions importantes se regroupent tous dans le haut du graphique (Figure 6). La CRP faible, la VS faible et les patients plus jeunes s'opposent au premier groupe. Nous avons vu dans la recension que lorsque l'inflammation est intense et dure assez longtemps, elle peut entraîner une anémie. Cet axe pourrait donc représenter les paramètres biologiques du patient, surtout l'aspect inflammatoire.

L'axe 3 explique 8,07% de l'inertie et est composé de la raideur matinale, des globules blancs et des corticostéroïdes. Une raideur de courte durée est associée au

traitement de corticostéroïde et aux PMN élevés. Évidemment ces variables sont opposées par la raideur plus prolongée, l'absence de traitement de corticostéroïdes et les PMN dans les limites de la normale. L'association entre les PMN et le traitement est logique puisque les corticostéroïdes à doses élevées tendent à faire augmenter les globules blancs, en particulier les PMN. Le fait qu'ils soient aussi associés à une raideur plus faible pourrait s'expliquer par l'efficacité anti-inflammatoire de ce traitement.

Finalement, le dernier axe explique 6.31% de l'inertie totale et se définit encore une fois par les critères diagnostiques mais aussi par le nombre peu élevé de synovites. On observe graphiquement que le fait de ne pas remplir les critères de l'ACR est associé au faible nombre d'articulations avec synovites; il faut noter que le nombre et la distribution des atteintes articulaires fait partie des critères de classification de la PR selon l'ACR.

Cette analyse ne nous apporte pas beaucoup de nouvelles connaissances cliniques mais elle est tout de même logique et confirme certain processus biologiques connus jusqu'à maintenant. Elle nous permet donc de passer à l'analyse suivante qui consiste à effectuer une classification hiérarchique sur les données réduites.

4.3 Résultats de la classification hiérarchique

Afin de mettre en évidence s'il existe des sous-groupes de patients dans l'échantillon, une classification hiérarchique a été effectuée sur les 4 premières dimensions obtenues à l'étape précédente. Une première analyse avait été faite en considérant les PMN et les GB dans les variables biologiques. Comme il existe déjà des

points critiques pour l'utilisation clinique de ces variables, leur classification s'est donc faite sur cette base. Suite à l'analyse des sous-groupes, il a été mis en évidence que l'une des classes de la variable GB n'était constituée que de 3 patients et que cette classe avait une qualité de représentation très médiocre. De plus, les PMN étant une catégorie de GB, le clinicien considérait que d'inclure les deux variables causait une sur-représentation de ces variables. Seulement les PMN ont donc été conservés dans le modèle final.

Le dendrogramme issu de cette analyse (Figure 9), représente tous les regroupements entre les individus à l'étude. Le niveau le plus bas correspond aux premiers regroupements faits et le dernier niveau est le rassemblement final de tous les individus. La hauteur des niveaux correspond à la distance entre les groupes ou bien à la perte d'inertie interclasse résultant de leur regroupement. En étudiant le dendrogramme, on observe que la population étudiée est très hétérogène étant donné l'existence de plusieurs niveaux dans le graphique et la formation d'un nombre important de sous-groupes. Le choix du niveau où l'arbre a été coupé, soit la droite pointillée sur l'image ci-dessous, a été fait de façon à obtenir des classes aussi homogènes que possible mais constituées d'un nombre assez important de sujets à des fins d'analyses ultérieures. Si l'arbre avait été coupé plus bas, le sous-groupe 5 aurait été divisé en deux et le groupe de gauche aurait été très petit. Au contraire, si l'arbre avait été coupé plus haut, les groupes 1 et 2 auraient été fusionnés malgré leur hétérogénéité illustrée par la hauteur des niveaux.

Figure 9. Dendrogramme issu de la classification hiérarchique

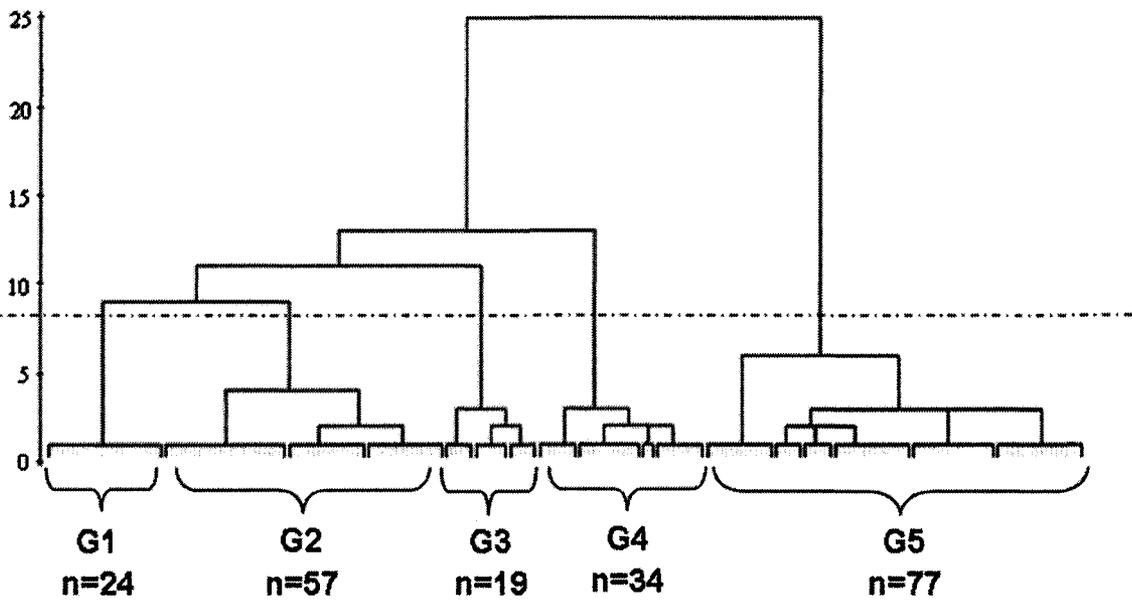


Figure 10. Dendrogramme en utilisant la métrique du Khi-carré

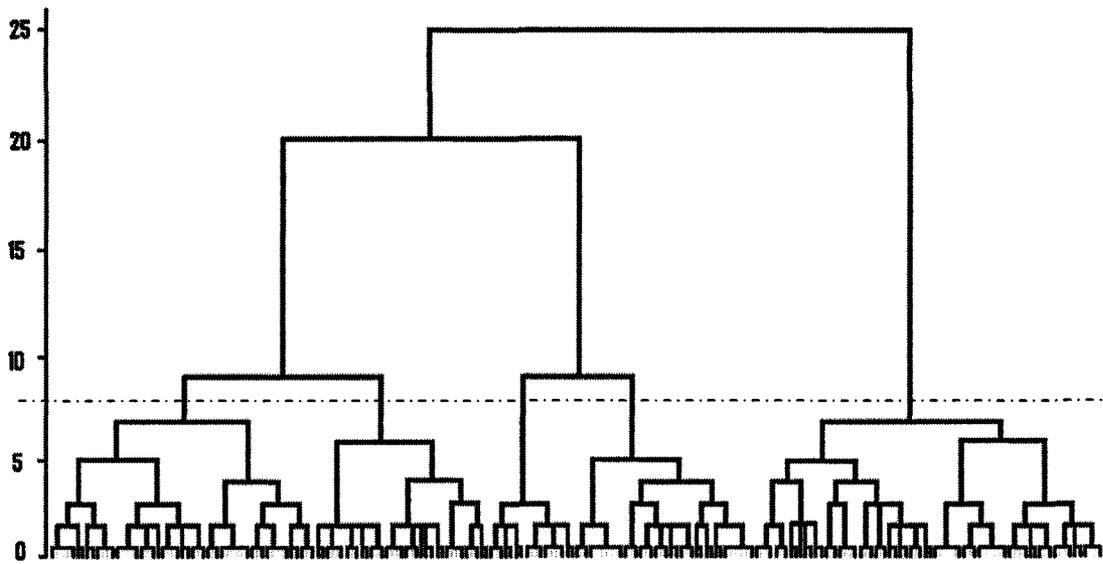
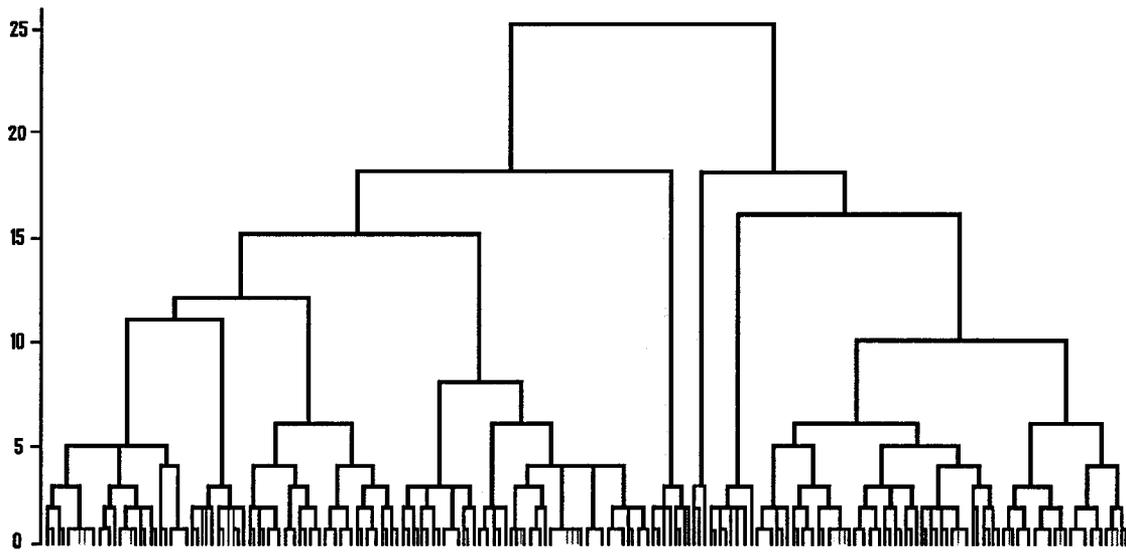


Figure 11. Dendrogramme en utilisant tous les facteurs de l'ACM



En comparant les différents modèles de classification, soit les dendrogrammes des figures 9, 10 et 11, on constate une certaine ressemblance dans les regroupements entre le modèle provenant de la classification des 4 premiers facteurs de l'ACM (Figure 9) et celui provenant de la métrique du Khi-carré sur les coordonnées non transformées (Figure 10).

Comparons alors ces deux méthodes. Le tableau 9 permet de comparer la distribution des patients selon les deux analyses. Par exemple, si l'on regarde le sous-groupe 1 de l'analyse selon les 4 premiers facteurs de l'ACM qui correspond à la première ligne du tableau, on observe que les patients se retrouvent principalement dans le groupe 1 selon la métrique du Khi-carré (20 patients sur les 24). On voit donc que les patients formant les sous-groupes 1 et 5 selon les 4 premiers facteurs de l'ACM se retrouvent principalement dans les mêmes sous-groupes (83% et 74% respectivement) en comparant la méthode utilisant la métrique du Khi-carré. On observe aussi que le sous-

groupe 1 de la deuxième analyse est beaucoup plus volumineux en patients et près de 60% des autres patients proviennent de l'ancien groupe 4 qui était un sous-groupe qui allait plutôt bien. Par contre, les individus des groupes 2, 3 et 4 ne se retrouvent plus dans les mêmes sous-groupes, ils sont donc plus difficiles à regrouper dans des sous-groupes vraiment distincts.

Tableau 9. Comparaison des groupes résultant de deux classifications différentes: méthode utilisant les 4 premières coordonnées factorielles et méthode utilisant la métrique du Khi-carré sur les coordonnées originales

		Groupe selon la métrique du Khi-deux					Total
		1	2	3	4	5	
Groupe selon 4^{er} facteurs de l'ACM	1	20	3	-	-	1	24
	2	9	17	2	28	1	57
	3	4	7	-	-	8	19
	4	20	5	-	6	3	34
	5	-	3	15	2	57	77
	Total	53	35	17	36	70	211

Les figures 12 et 13 confirment ces observations. Ces graphiques représentent les individus dans l'espace des deux premiers facteurs de l'ACM. La figure 12 représente bien le fait que les groupes 1 et 5 soient bien distincts l'un de l'autre. Si l'on compare ce graphique à la représentation des variables pour le même espace (Figure 6, p.64), on constate que les individus du groupe 5 se situent là où les anticorps de titre élevé étaient placés tandis que les individus du groupe 1 se regroupent à l'opposé où il y a absence d'anticorps. La figure 13 représente bien le fait que les 3 autres sous-groupes sont plus difficiles à distinguer les uns des autres selon ces deux dimensions.

Figure 12. Graphique de l'ACM, représentation des individus du groupe 1 et 5

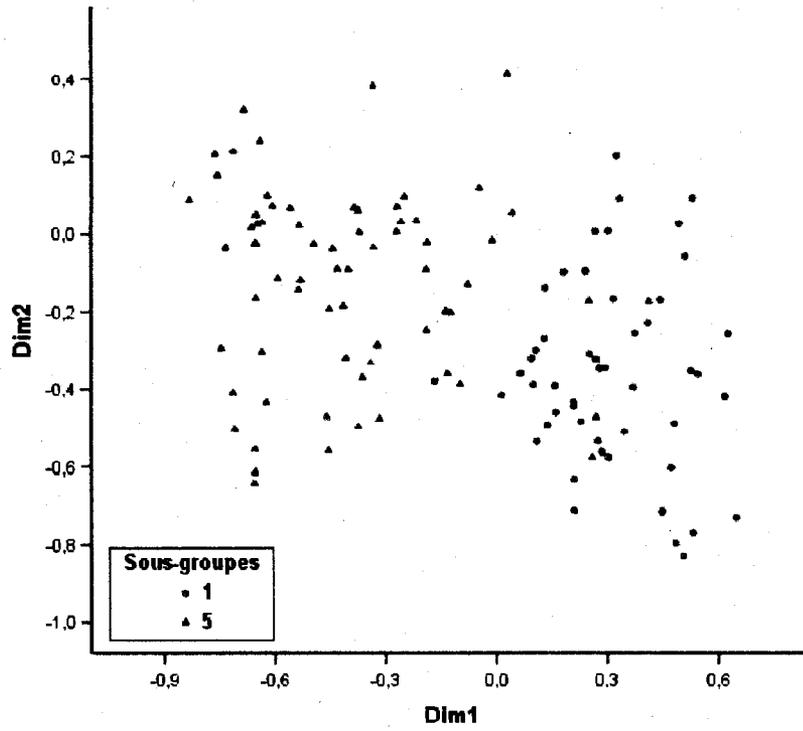
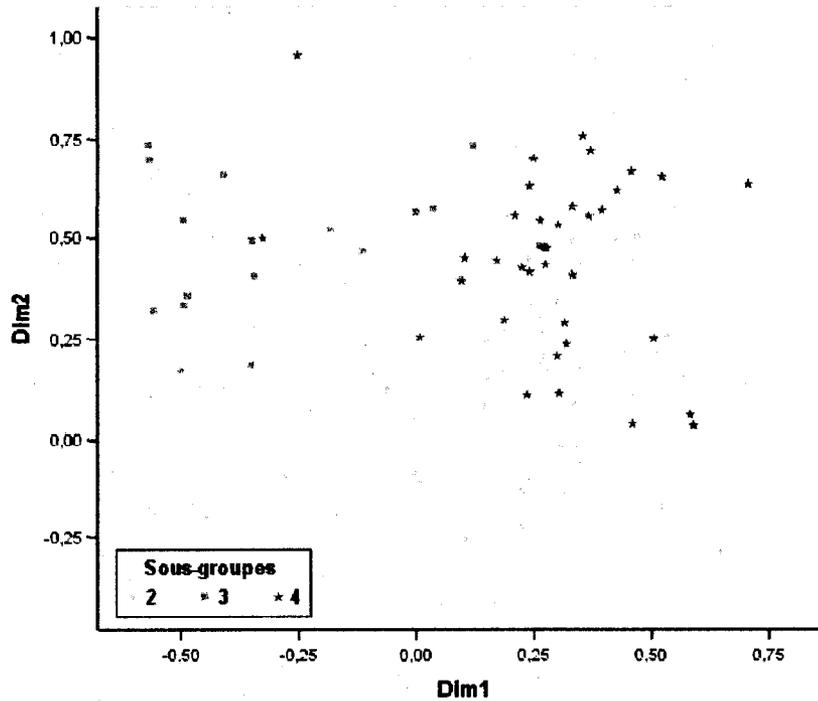


Figure 13. Graphique de l'ACM, représentation des individus des sous-groupes 2, 3 et 4

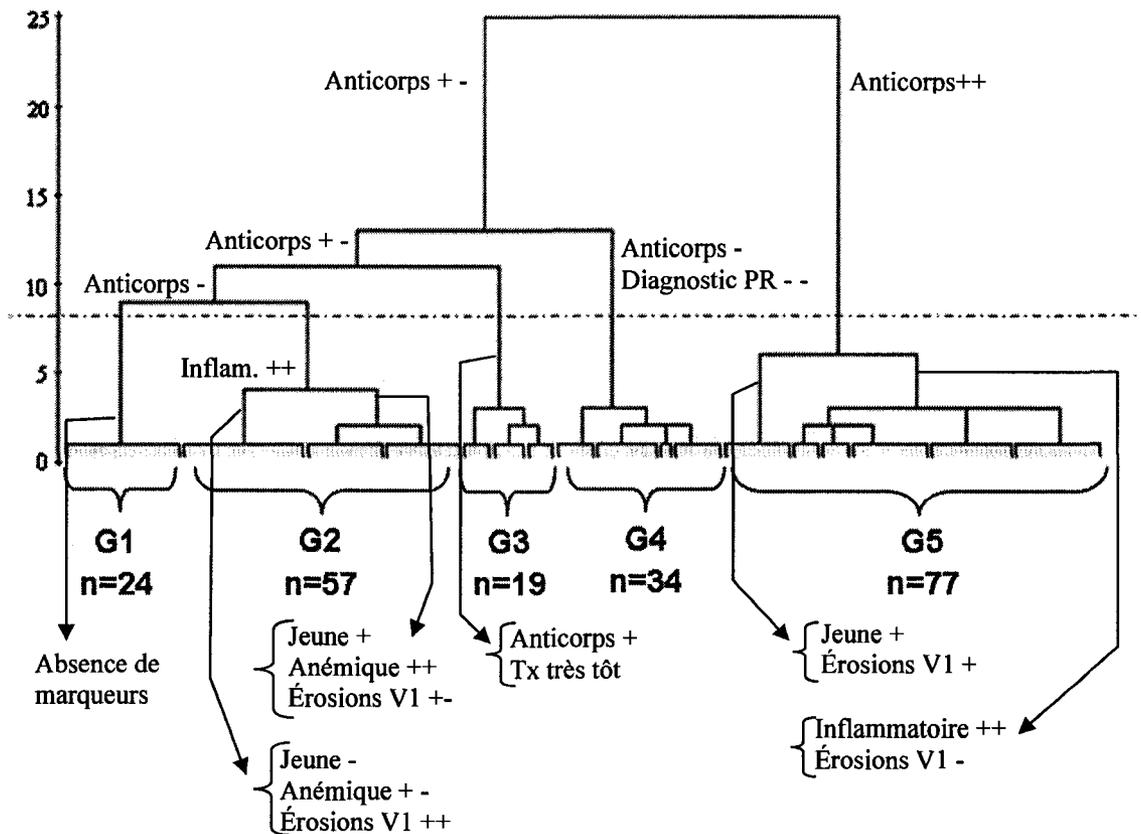


Si l'on considère le dendrogramme issu de la classification selon l'ensemble de toutes les coordonnées factorielles, il est beaucoup plus difficile d'identifier des sous-groupes principaux, probablement parce que la méthode utilise toute l'information disponible. Il y a donc beaucoup plus de distinctions entre les groupes, ce qui n'est pas nécessairement l'objectif de l'étude qui veut plutôt essayer de cibler un nombre restreint de sous-groupes et expliquer ce qui les définit. Nous allons donc analyser la classification qui résume un peu plus l'information clinique, soit celle qui utilise seulement les 4 premières composantes principales.

Étudions maintenant chacun des 5 sous-groupes pour évaluer ce qui les différencie, au début et en cours d'évolution. Les variables catégoriques ont été comparées par rapport aux sous-groupes en utilisant la méthode de Monte Carlo pour approximer le test du Khi-Carré. En effet, le test exact de Fisher était impossible à appliquer sur les données en considérant les dimensions importantes des tableaux. La plupart des variables sont significativement différentes en comparant selon les groupes.

La figure 14 représente les principales variables qui définissent chacun des niveaux. On voit donc que ce qui distingue le premier niveau est soit d'avoir beaucoup d'anticorps ou d'avoir peu ou pas du tout d'anticorps. Nous allons donc regarder séparément chacune des variables dans la prochaine section en commençant par les issues cliniques (érosions, capacité fonctionnelle et l'activité de la maladie). Une analyse de la distribution de l'ensemble des autres paramètres démographiques et cliniques terminera la section des résultats.

Figure 14. Résumé des principales variables définissant chacun des niveaux du dendrogramme issu de la classification hiérarchique



Légende : Inflammatoire (Inflam.), Traitements (Tx), Très présent (++), Moyennement présent (+), Moyennement ou pas présent (+ -), Peu présent (-), Absent (- -)

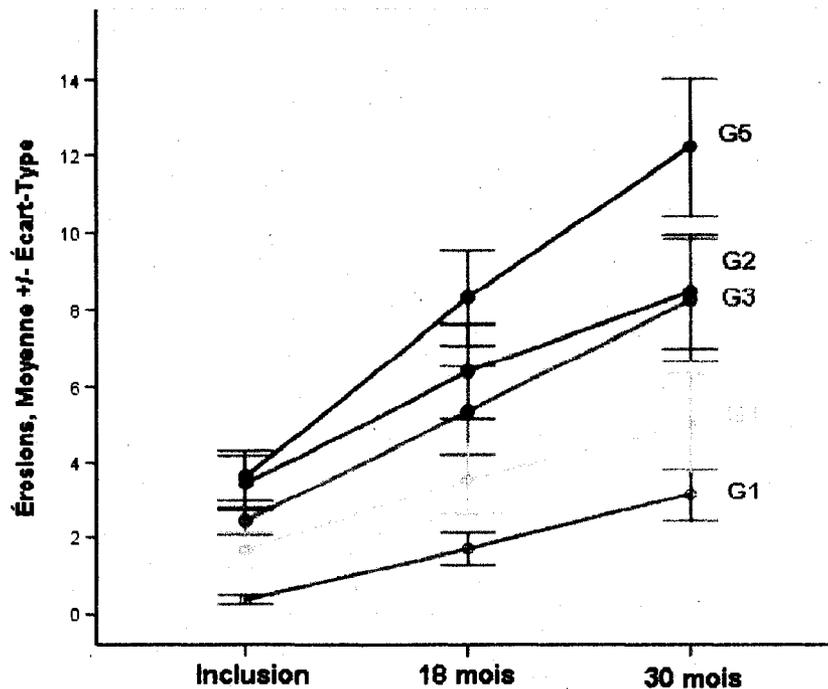
L'évolution des dommages radiographiques a été analysée à l'aide du test non paramétrique de Mann-Whitney, étant donné que les érosions ne sont pas distribuées normalement. La figure 15 ainsi que le tableau 10 ci-dessous montrent bien qu'il y a peu de dommages à l'inclusion mais que certains sous-groupes évoluent plus rapidement que d'autres. Les groupes 1 et 4 semblent rester plutôt stables tandis que le groupe 5 semble être très sévèrement atteint à 30 mois. Ces observations permettent d'obtenir une idée globale de l'évolution radiographique de chacun des sous-groupes.

Tableau 10 **Distribution des dommages radiographiques (score d'érosions) selon les 5 sous-groupes**

		G1	G2	G3	G4	G5	
		n=24	n=57	n=19	n=34	n=77	
Score d'érosions continu^a	Inclusion **	0 (1)	2 (4)	2 (2.3)	1 (2.3)	2 (5)	
	18 mois **	1 (4)	4 (7)	4 (4)	1 (4)	6 (8)	
	30 mois **	2 (5)	6 (7)	6 (10)	2 (7)	8 (12)	
Score d'érosions catégorisé^b	Inclusion	5-10	0 (0)	9 (16)	2 (11)	3 (9)	11 (14)
		≥10	0 (0)	4 (7)	0 (0)	1 (3)	8 (10)
	18 mois **	5-10	2 (8)	18 (32)	5 (26)	3 (9)	16 (21)
		≥10	0 (0)	8 (14)	2 (11)	5 (15)	24 (31)
	30 mois *	5-10	5 (22)	18 (32)	4 (21)	4 (12)	21 (28)
		≥10	1 (4)	15 (28)	7 (37)	7 (21)	30 (39)

Résultats présentés sous forme de : ^a médiane (étendue interquartile); ^b fréquences (%)
 * p = 0.003, **p<0.0001

Figure 15. **Évolution des dommages radiographiques de l'inclusion à 30 mois**



En observant plus attentivement les sous-groupes, nous avons constaté que certains patients du groupe le plus sévèrement atteint ont tout de même très peu d'érosions. Nous avons donc représenté sous forme de graphique la différence dans les érosions entre l'inclusion et 30 mois (Figure 16.1). Chacun des patients est donc représenté sur le graphique. Les différences ont été ordonnées en ordre croissant pour chacun des sous-groupes. Malgré le fait que le graphique 7 montre un nombre considérable d'érosions à 30 mois, le graphique 16.1 semble suggérer qu'une grande partie des patients du groupe 5 n'évoluent pas du tout ou très peu. Le groupe 5 reste tout de même le sous-groupe pour lequel on observe la plus grande progression, mais cette progression est surtout le fait d'une portion des patients qui progressent rapidement. Les mêmes constatations peuvent être faites chez les autres sous-groupes.

Nous voulions ensuite vérifier si les patients des groupes à mauvais pronostic qui démontraient peu d'augmentation de leur score d'érosions étaient en fait des patients qui avaient déjà un taux élevé d'érosions à l'inclusion. La figure 16.2 représente les érosions présentes à l'inclusion tout en conservant le même ordre des patients que la figure 16.1. Rappelons que la figure 16.1 ordonnait les patients selon le niveau de différence d'érosions entre l'inclusion et 30 mois. L'ordre des patients a été conservé de façon à pouvoir comparer les 2 graphiques. On constate que les patients des sous-groupes 1, 3 et 4 n'avaient pas ou peu d'érosions à l'inclusion sauf pour quelques rares exceptions. Par contre, les patients des groupes 2 et 5 sont beaucoup plus variables. Plusieurs ont plus de 5 érosions dès le départ, ce qui n'est pas le cas pour les groupes 1, 3 et 4. On retrouve néanmoins dans les groupes 2 et 5 des patients avec peu d'érosions au départ qui n'évoluent pas beaucoup plus à 30 mois. Des analyses plus poussées incluant d'autres

paramètres (e.g. traitements reçus, présence d'autres marqueurs non considérés dans le modèle) seront nécessaires pour expliquer pour quelle raison ces patients se retrouvent malgré tout dans les groupes où les patients sont généralement plus atteints à 30 mois. Pour sa part, la figure 16.3 permet d'observer le niveau de dommages radiographiques à 30 mois. L'allure de la courbe est beaucoup plus semblable à la distribution des différences d'érosions. Encore là, on peut observer que certains patients développent peu de dommages radiographiques à 30 mois malgré le fait qu'ils se situent dans le groupe le plus sévèrement atteint. Il faudra donc tenter de déterminer pourquoi ils se retrouvent malgré tout dans ce sous-groupe.

La variable «érosions présentes à l'inclusion» semble donc être un bon indicateur globalement pour différencier les sous-groupes, mais elle n'est néanmoins pas parfaite pour définir le pronostic de chacun des patients pris individuellement.

Figure 16.1 Représentation individuelle de la différence d'érosions entre l'inclusion et 30 mois

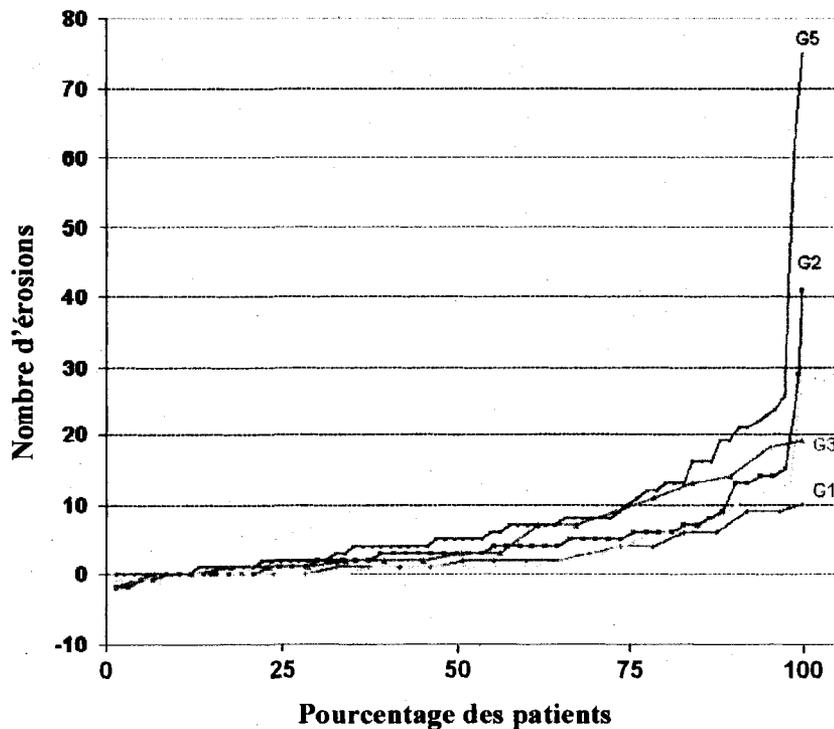


Figure 16.2 Représentation individuelle des érosions à l'inclusion

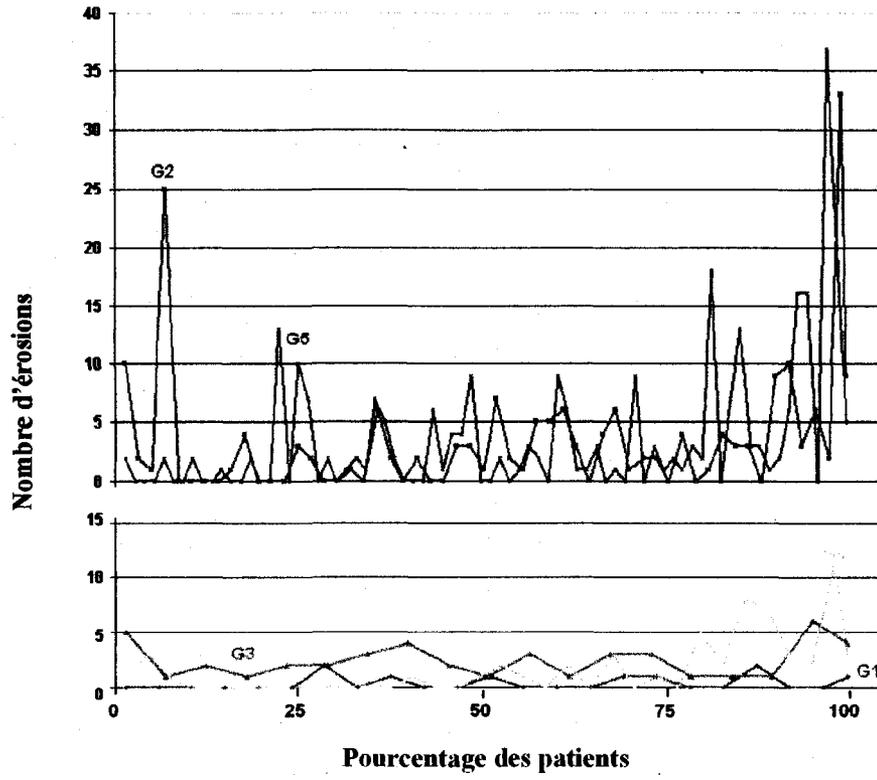
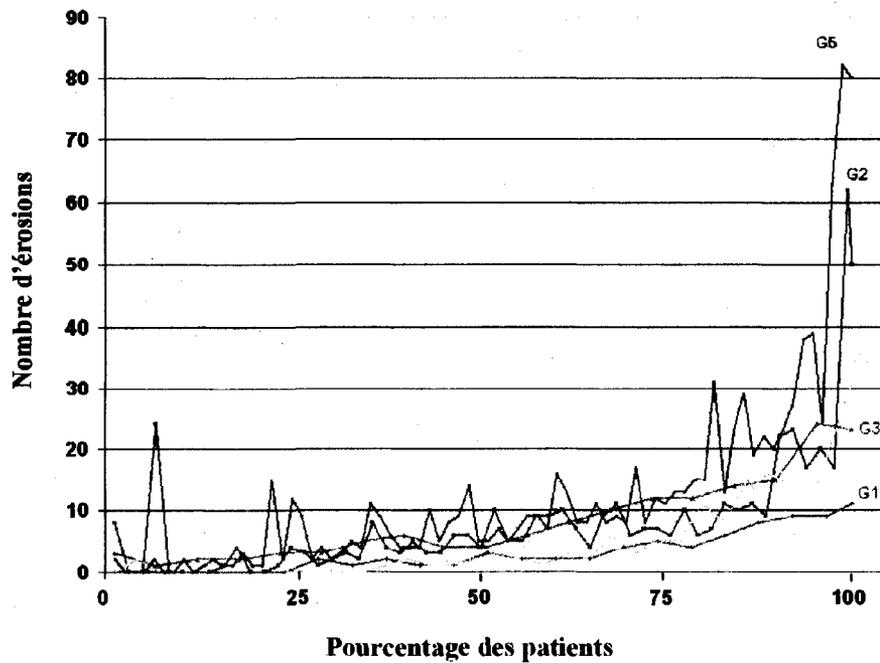


Figure 16.3 Représentation individuelle des érosions à 30 mois



Un autre paramètre intéressant à étudier dans le temps est l'évolution de l'incapacité fonctionnelle du patient, soit la difficulté à se livrer aux activités de la vie quotidienne (tableau 11). On observe que le groupe 2 a une incapacité plus importante lors de l'inclusion dans l'étude puisqu'un score de 1 ou plus est considéré comme une atteinte importante. Par contre, à 30 mois, il n'y a plus de différences par rapport aux autres groupes et l'incapacité semble plutôt stable ou diminuée, ce qui est positif pour le patient.

Tableau 11. Distribution de l'incapacité du patient dans le temps selon les 5 sous-groupes

M-HAQ	G1 n=24	G2 n=57	G3 n=19	G4 n=34	G5 n=77
^a Inclusion **	0,38 (0.9)	1.38 (1)	0,38 (0.6)	0,69 (0.9)	0,75 (1.1)
30 mois	0,60 (0.6)	0,13 (0.5)	0,38 (0.6)	0,25 (0.6)	0,25 (0.6)
^b Inclusion **	7 (29)	39 (68)	3 (16)	13 (38)	38 (49)
30 mois	4 (17)	5 (9)	4 (21)	3 (9)	9 (12)

Résultats présentés sous forme de : ^a médiane (étendue interquartile); ^b fréquence avec HAQ supérieur ou égal à 1.0 (%)

* p<0.05, **p<0.0001

Pour ce qui est de l'activité de la maladie (tableau 12), pratiquement tous les patients ont une maladie très active au départ (entre 85%-100%) et près de 50%-55% vont entrer en rémission après 30 mois, et ce peu importe le groupe dans lequel ils se situent. Il n'y a donc pas de différence statistiquement significative dans ce cas, ce qui suggère que cette variable n'est pas d'une très grande utilité dans ce modèle pour différencier les patients des sous-groupes. C'est tout de même intéressant cliniquement d'observer un taux aussi important de rémission à 30 mois.

Tableau 12. Distribution de l'activité de la maladie dans le temps selon les 5 sous-groupes

		G1	G2	G3	G4	G5
n=202		n=22	n=56	n=18	n=34	n=72
Inclusion	Rémission	1 (5)	0 (0)	0 (0)	1 (3)	2 (3)
	Contrôlée	1 (5)	0 (0)	1 (6)	4 (12)	3 (4)
	Active	20 (90)	56 (100)	17 (94)	29 (85)	67 (93)
30 mois	Rémission	12 (55)	31 (55)	8 (44)	19 (56)	41 (57)
	Sans DMARDs	3 (14)	9 (16)	2 (11)	12 (35)	5 (7)
	Contrôlée	5 (23)	7 (13)	4 (22)	7 (21)	12 (17)
	Active	4 (18)	18 (32)	6 (33)	8 (24)	19 (26)

Résultats présentés sous forme de fréquence (%). DMARD : Médicament anti-rhumatismaux.

Si l'on considère les variables du tableau 13, on observe que les anticorps se retrouvent en majorité dans le groupe 5, mais le groupe 3 possède un nombre non négligeable de patients avec FR et anti-CCP. Il y a un peu plus de femmes dans les groupes 1 et 3, tandis que les groupes 2 et 3 semblent être plus âgés puisque constitués d'un nombre important de sujets de plus de 70 ans. Les sujets des classes 2 et 4 se présentent un peu plus tard que les autres en clinique. Pour ce qui est du tabac, il y a un peu plus de patients fumeurs dans le groupe 5, mais il n'y a pas de différence significative entre les groupes. On observe un nombre élevé de PMN chez les patients du groupe 3. De façon intéressante, un seul groupe (groupe 4) ne semble pas rencontrer les critères de l'ACR et ceux-ci ont également moins de synovites et de raideur matinale. De plus, les patients du groupe 2 sont ceux ayant de plus fortes concentrations de paramètres inflammatoires, et ces patients sont aussi plus souvent atteints d'anémie.

Tableau 13. Distribution des données socio-démographiques, inflammatoires et immunitaire selon les 5 sous-groupes; fréquence (%)

		G1 n=24	G2 n=57	G3 n=19	G4 n=34	G5 n=77
Anticorps :						
Anti-Sa **		0 (0)	2 (4)	1 (5)	1 (3)	49 (64)
Anti-CCP **	Faible	3 (13)	1 (2)	0 (0)	1 (3)	4 (5)
	Modéré	0 (0)	0 (0)	3 (16)	0 (0)	0 (0)
	Fort	0 (0)	0 (0)	4 (21)	2 (6)	61 (79)
FR **	Modéré	3 (13)	5 (9)	3 (16)	2 (6)	4 (5)
	Fort	2 (18)	4 (7)	7 (37)	1 (3)	69 (90)
Socio-démographiques :						
Femmes *		20 (83)	30 (53)	14 (74)	22 (65)	38 (49)
Age **	+ 70 ans	0 (0)	26 (46)	10 (53)	8 (24)	6 (8)
	- 50 ans	15 (63)	10 (18)	3 (16)	16 (47)	28 (36)
Présente moins de 3 mois après début symptômes *		7 (29)	27 (47)	5 (26)	15 (44)	19 (25)
Fumeur ou ancien fumeur (n=195)		12 (50)	30 (53)	9 (47)	13 (38)	51 (66)
Inflammation :						
Critère de classification de la PR selon l'ACR **		24 (100)	54 (95)	17 (89)	0 (0)	73 (95)
Raideur matinale prolongée **		23 (96)	56 (98)	12 (63)	12 (35)	61 (79)
Synovites ≥ 5 **		23 (96)	47 (82)	16 (94)	15 (44)	67 (87)
PMN élevés **		0 (0)	19 (33)	13 (68)	15 (44)	27 (35)
Anémie/Anémie sévère **		1 (4)	35 (61)	3 (16)	10 (29)	17 (24)
VS élevée **		0 (0)	36 (63)	3 (16)	10 (29)	28 (36)
CRP ≥ 8 **		8 (33)	47 (83)	10 (52)	21 (62)	58 (75)

Résultats présentés sous forme de fréquence (%). Anti-CCP : anti-cyclic citrullinated peptide antibodies; FR: facteur rhumatoïde; ACR : American College of Rheumatology; PMN: polymorphonucléaire; VS: vitesse de sédimentation; CRP : Protéine C-Réactive
* p<0.05, **p<0.0001

La distribution des traitements a été plutôt surprenante (Tableau 14). En effet, un seul sous-groupe, soit le groupe 3, semble être traité beaucoup plus tôt que les autres dans le développement de la maladie, souvent même avant l'inclusion dans l'étude. Le groupe 3 est traité autant avec des DMARDs que des corticostéroïdes. Par la suite, le traitement

semble rejoindre celui dont bénéficient les patients des autres sous-groupes. On a déjà observé précédemment que les PMN sont très élevés dans ce sous-groupe. Une explication possible pour cette élévation des PMN est la présence d'un traitement avec des corticostéroïdes lors de l'inclusion, des médicaments bien connus pour induire une leucocytose.

Tableau 14. Distribution des traitements selon les 5 sous-groupes

Visite	Traitements	G1 n=24	G2 n=57	G3 n=19	G4 n=34	G5 n=77
Avant l'inclusion	DMARD	1 (4)	12 (21)	13 (68)	6 (18)	9 (12)
	Corticostéroïdes	1 (4)	6 (11)	13 (68)	7 (21)	8 (10)
Inclusion à 18 mois	DMARD	20 (83)	49 (86)	17 (89)	24 (71)	70 (91)
	Corticostéroïdes	3 (13)	22 (39)	11 (58)	10 (29)	24 (31)
18 à 30 mois	DMARD	19 (79)	42 (74)	14 (74)	19 (56)	69 (90)
	Corticostéroïdes	2 (8)	7 (12)	2 (11)	5 (15)	14 (18)

Résultats présentés sous forme de fréquence (%). DMARD : Médicament anti-rhumatismaux.

Pour ce qui est des paramètres génétiques (Tableau 15), il n'y a pas beaucoup de résultats significatifs. Dans le groupe 5, il y a un peu plus de patients qui ont le double SE, ce qui est parfois associé à une maladie plus sévère. Il y a également plus de DR3 chez les patients avec anti-CCP négatif, comme dans la littérature, mais ce taux plus élevé se retrouve seulement dans le groupe 4.

Tableau 15. Distribution des données sur la génétique (HLA-DRB1) selon les 5 sous-groupes

n=203		G1 n=22	G2 n=56	G3 n=19	G4 n=34	G5 n=72
SE *	+/+	0 (0)	3 (5)	2 (11)	2 (6)	18 (24)
	+/-	8 (36)	25 (45)	9 (47)	13 (38)	28 (37)
DERAA	+/- ou +/+	8 (36)	17 (30)	5 (26)	12 (35)	15 (20)
DR3		6 (27)	16 (29)	3 (16)	13 (38)	18 (25)

Résultats présentés sous forme de fréquence (%).

*p<0.05

Maintenant que les résultats ont été étudiés, il est important de critiquer leur intérêt clinique et de considérer toutes les forces et les limites de cette étude. La prochaine section permettra de répondre à ces questionnements et par le fait même, une discussion de la composition des sous-groupes sera faite plus en profondeur en établissant un lien avec la littérature.

5 - Discussion et Conclusion

5.1 Discussion des résultats

Les statistiques exploratoires utilisées ont permis, dans un premier temps, d'atteindre le premier objectif fixé. Ces méthodes avaient pour but d'explorer les relations entre les caractéristiques cliniques des patients avec polyarthrite inflammatoire débutante. Elles ont donc permis la découverte de 5 sous-groupes de patients formés à partir des caractéristiques cliniques mesurées à l'inclusion. Ces regroupements suggèrent donc que l'arthrite inflammatoire est effectivement composée de plusieurs maladies différentes qui pourraient s'expliquer par différents symptômes ou paramètres cliniques.

En étudiant chacun des sous-groupes obtenus, nous constatons que la maladie n'évolue pas de façon similaire si l'on considère l'atteinte radiographique des articulations. En effet, 2 sous-groupes évoluent très peu, 2 autres augmentent modérément et le dernier semble être beaucoup plus sévèrement atteint après 30 mois. Pour ce qui est de l'évolution de l'incapacité, il y a beaucoup plus de patients avec incapacité à l'inclusion qu'à 30 mois, mais on constate que le groupe 2 est constitué de patients avec des problèmes importants, soit 68% des patients avec une incapacité considérable. De plus, un très grand nombre de patients ont une maladie active au départ et 50%-55% de l'ensemble des sujets entre en rémission à 30 mois, et ce peu importe le groupe dans lequel ils se situent. Cette importante réduction d'activité de la maladie porte à croire qu'elle pourrait s'expliquer par l'efficacité du traitement conventionnel administré tout au long de l'étude. Effectivement, plusieurs études pharmacologiques

utilisent le DAS28-CRP comme indice d'activité de la maladie et ont y observe de forts taux de rémission suite aux traitements (Allaart et al. 2006; Quinn et al. 2005). La maladie semble donc moins active, avec une réduction de l'inflammation, des articulations avec synovite et de l'incapacité fonctionnelle, mais les articulations continuent de se détériorer au niveau radiographique dans certains sous-groupes. Étudions maintenant chacun des sous-groupes individuellement pour définir ce qui les caractérise.

Groupe 1 :

Les particularités du groupe 1 sont que très peu de patients ont des anticorps, en particulier aucun n'ont d'anti-Sa, et la majorité des FR et anti-CCP présents le sont avec un titre faible. Ce groupe est constitué d'une majorité de femmes, principalement âgées de moins de 50 ans; comme aucun patient de plus de 70 ans n'est présent, nous avons donc affaire à un groupe plutôt jeune. Les patients ont très peu d'érosions au départ et en développeront très peu en 3 ans, mais ils remplissent tous les critères de classification de la PR selon l'ACR. Ils ont pratiquement tous de la raideur matinale prolongée et des gonflements dans au moins 5 articulations. Par contre, ces patients ne sont pas très inflammatoires du côté systémique (CRP et VS peu élevées), ils ont des globules blancs normaux et ont été très peu traités avant la visite d'inclusion. Finalement, ces patients se présentent en clinique généralement tard, avec 70% du groupe qui se présente au moins 3 mois après le début des symptômes. En considérant l'ensemble de ces caractéristiques, on constate que ces patients ont donc une très bonne issue clinique. Comme il y a absence d'anticorps, que la maladie évolue lentement, mais que les sujets remplissent les critères

diagnostiques de l'ACR, cette maladie ressemble à la polyarthrite séronégative du sujet jeune. Certains patients de ce groupe pourraient aussi présenter une arthrite réactive d'évolution plus prolongée qu'à l'habitude ou une spondylarthropathie de type arthrite psoriasique (La Société d'arthrite 2007).

Groupe 2 :

Le groupe 2 est aussi défini par des patients avec très peu d'anticorps, comme le groupe 1. Les patients ont aussi des globules blancs normaux et sont aussi rarement traités avant la visite d'inclusion. Les patients remplissent aussi les critères de classification de la PR, il y a présence de raideur matinale prolongée et de synovites nombreuses. Ce qui distingue principalement le groupe 2 du groupe 1 est le nombre plus élevé d'érosions au départ et après 30 mois. Toutefois, le groupe 2 n'est pas le groupe le plus sévèrement atteint au niveau radiologique. Les patients sont autant des hommes que des femmes et sont beaucoup plus âgés que le groupe 1, avec près de 50% du groupe âgé de plus de 70 ans. Les patients se présentent plus tôt, les paramètres de l'inflammation systémique sont très élevés et ils sont aussi atteints d'anémie. C'est également le groupe où il y a le plus de patients avec une incapacité importante, probablement un facteur expliquant qu'ils se présentent en clinique de rhumatologie en moyenne moins de 3 mois après le début des symptômes. Le fait que les paramètres de l'inflammation soient très élevés pourrait expliquer le fort degré d'incapacité ainsi que la présence d'érosions précoces. En effet, la littérature a parfois constaté une association entre ces facteurs (Balédent 2007; Combe et al. 2001; Forslind et al. 2004). En se basant sur ces multiples caractéristiques, on pourrait suggérer que ce groupe est associé à la polyarthrite

séronégative du sujet âgé. En effet, nous avons vu dans la recension que cette maladie touche autant les hommes que les femmes. Le début peut faire penser à un épisode infectieux, d'où les paramètres inflammatoires présents ici, et les anticorps sont absents (Gabay 2007).

Groupe 3 :

Le groupe 3 ressemble beaucoup au groupe 2 si l'on compare l'évolution des érosions entre l'inclusion et 30 mois. Peu de patients de ce groupe ont des anti-Sa, quelques-uns ont des anti-CCP et un peu plus produisent du FR. Les patients de ce groupe remplissent aussi les critères de l'ACR, ont de la raideur matinale prolongée et des synovites multiples. Ils sont très peu inflammatoires au niveau systémique, mais ils présentent un décompte de globules blancs très élevé comparativement à tous les autres groupes. C'est aussi le seul groupe qui soit traité de façon importante par des corticostéroïdes et des DMARDs aussi tôt dans le développement de la maladie. Comme mentionné précédemment, on a vu que l'administration de corticostéroïdes provoque une augmentation du nombre de globules blancs circulants, ce qui expliquerait la présence de ces deux variables en nombre important dans ce groupe. Il y a aussi un peu plus de femmes que d'hommes, les patients sont aussi plutôt âgés et se présentent plus tard après le début des symptômes. Le fait que les patients soient traités avec des corticostéroïdes très tôt pourrait peut-être expliquer le fait qu'il y a peu d'anticorps présents et un peu moins d'érosions initialement que le groupe le plus sévère (groupe 5). En effet, plusieurs chercheurs ont constaté l'efficacité de traiter le plus tôt possible dès l'apparition des symptômes (Bukhari et al. 2003; Korpela et al. 2004; Quinn et al. 2005; van Aken et al.

2004). Ces observations suggèrent que ces patients pourraient être atteints de polyarthrite rhumatoïde « classique » dont l'évolution a été partiellement modifiée par le traitement.

Groupe 4 :

Le groupe 4 est très semblable au groupe 1. Les patients de ce groupe sont pratiquement tous séronégatifs pour l'ensemble des anticorps et l'évolution des érosions est aussi peu importante. Les patients sont majoritairement plus jeunes, soit moins de 50 ans, comme le groupe 1, mais le nombre de femmes rencontre les proportions normales de la population générale atteinte de polyarthrite, soit les deux tiers de femmes. Près de la moitié des patients se présentent tôt en clinique. Un des points majeurs de discordance entre le groupe 1 et le groupe 4 est le fait qu'aucun patient du groupe 4 ne remplisse les critères de classification de la PR selon l'ACR. Tous les patients des autres groupes remplissent ces critères dans une proportion de 90%-100%. De plus, très peu de patients du groupe 4 ont de la raideur matinale prolongée et des synovites importantes au départ. Fait intéressant, c'est le seul groupe, parmi les groupes ayant peu d'anti-CCP, qui a un taux plus élevé de HLA-DR3, mais seulement chez 40% des patients. Il a été rapporté que la présence de HLA-DR3 est associée à une absence d'anti-CCP, mais l'association n'est pas très forte à l'intérieur de nos sous-groupes (Irigoyen et al. 2005; Verpoort et al. 2005). Il ne semble donc pas que la détection du HLA-DR3 serait très utile pour définir les groupes. En considérant ces caractéristiques et principalement parce que les critères de classification de la PR ne sont pas présents chez les patients du groupe 4, ce groupe ressemble plus à un groupe hétérogène de patients atteints de spondylarthropathie, d'arthrite psoriasique ou de sarcoïdose. Ces maladies apparaissent souvent chez des

patients plus jeunes, soit entre 15-50 ans, et touchent principalement les femmes, sauf dans le cas de la spondylarthropathie. De plus, le FR est généralement absent chez ces patients, de sorte que lorsque les signes ou symptômes caractéristiques de ces maladies (ex. absence d'atteinte axiale dans le cas d'une spondylarthropathie) sont absents, ces patients présentent une maladie articulaire inflammatoire non spécifique (La Société d'arthrite 2007).

Groupe 5 :

Le groupe 5 est définitivement le groupe le plus sévèrement atteint. Les anticorps y sont très présents, il y a plus d'érosions au départ et c'est le groupe qui a le plus de dommages radiographiques à 30 mois. Les patients sont un peu plus jeunes que dans les autres groupes et il y a un peu plus d'hommes. Les patients remplissent les critères de classification de la PR, ont une raideur matinale prolongée, et les synovites sont présentes en grand nombre. La CRP est plutôt élevée chez ces patients, mais pas la VS. Les globules blancs sont normaux, les patients sont peu traités avec des corticostéroïdes lors de leur inclusion, et ils se présentent tard en clinique. Le fait qu'ils se présentent tard pourrait avoir un impact sur le développement de la maladie (anticorps et érosions), en comparaison avec le groupe 3. S'ils avaient été traités plus tôt, peut-être que les dommages auraient été moins importants. Par contre, le seul moyen d'y remédier est l'éducation de la population et des professionnels de la santé de première ligne, ce qui n'est pas facile et constitue un problème important limitant l'accès aux soins. Le médecin de famille doit aussi être informé et prendre conscience que plus le patient est traité tôt, plus il est facile d'agir avec le maximum d'efficacité. Un autre facteur qui pourrait être en

cause est que 66% des patients sont des fumeurs ou des anciens fumeurs. Nous avons vu que le tabac a un impact sur le développement des anticorps, surtout l'anti-CCP, et sur les dommages articulaires. La proportion importante de fumeurs n'est cependant pas statistiquement différente des autres groupes. Étant donné ces résultats, ce sous-groupe est celui qui ressemble le plus à la polyarthrite rhumatoïde séropositive classique.

Nous avons donc obtenu cinq sous-groupes plutôt différents par leurs caractéristiques à l'inclusion et leur évolution, ce qui répond grandement à nos objectifs du départ.

5.2 Forces et limites

L'un des plus grands avantages des méthodes utilisées est qu'elles permettent de traiter un nombre important de variables simultanément. De plus, comme ce sont des méthodes exploratoires, elles sont utilisées pour révéler la structure des données. Ces méthodes sont donc considérées comme des outils de découverte. Dans le cas présent, elles nous ont permis de mettre en évidence 5 sous-groupes importants au sein de notre cohorte et par le fait même de comprendre un peu mieux les raisons biologiques de la formation de ces sous-groupes. À partir de maintenant, nous pourrions élaborer de nouvelles hypothèses et utiliser un modèle prédictif comme la régression pour obtenir éventuellement un outil d'aide au diagnostic en clinique (Anderberg 1973).

Évidemment, toute étude comporte un certain nombre de biais et de limites qu'il est nécessaire de considérer. Un point positif est qu'il ne semble pas y avoir de biais de perte au suivi. En effet, aucune différence n'a été détectée pour les caractéristiques des patients à l'étude et ceux qui ont abandonné. Par contre, une limite importante dans notre

cas est évidemment la validité externe. Ce type d'analyse n'est pas fait pour être généralisable à d'autres populations, mais plutôt comme un outil de découverte des structures présentes dans la banque de données. Il n'est donc aucunement envisageable dans le cas présent de généraliser les résultats obtenus à l'ensemble des patients avec arthrite inflammatoire débutante. De plus, une autre raison supplémentaire qui rend difficile la généralisation des résultats est que nous sommes dans un centre unique de recrutement, en présence d'une population majoritairement francophone. Il est donc possible que la maladie se développe différemment dans une autre région, chez des individus avec une prédisposition génétique différente. Une autre limite importante à prendre en considération est le fait que nous avons un nombre considérable de variables et que nous n'avons qu'un nombre réduit d'individus. Cette faible taille d'échantillon constitue en effet une limite. Comme nous sommes en présence d'une étude observationnelle, il est évident que la médication risque d'être modifiée en cours d'étude en fonction de l'évolution clinique, et peut avoir un impact important sur les données à 30 mois. Nous avons considéré dans l'analyse exploratoire la présence de traitement dès l'inclusion du patient dans l'étude puisque nous voulions avoir un portrait de l'état clinique au début de la maladie. Par la suite, pour étudier l'évolution de la maladie selon les sous-groupes formés, nous avons étudié les traitements administrés entre les visites. Nous avons étudié le type de médication prise, mais il était plus difficile de contrôler la dose, la date du début ainsi que la durée du traitement, des variables probablement tout aussi importantes. L'évaluation de la médication représente à elle seule un autre projet de recherche. Finalement, le fait que les variables ont dû être catégorisées peut nuire à la reproductibilité de l'étude et en constitue une limite.

Il est important de noter ici que nous sommes en présence d'un modèle exploratoire et non d'une analyse prédictive. L'objectif de mon travail était d'explorer les données afin de vérifier si différents sous-groupes de patients pouvaient être définis à partir des caractéristiques de l'inclusion. Nos résultats confirment que c'est bien le cas. Par contre, en observant la distribution des patients, on constate que les sous-groupes définis ne sont pas parfaitement homogènes. Comme nous l'avons présenté plus haut, malgré le taux généralement plus élevé du score d'érosions dans les 3 groupes plus sévèrement atteints (groupes 2, 3, 5), certains patients de ces groupes n'ont tout de même pas beaucoup d'érosions, ni à l'inclusion ni à 30 mois, et ceci même dans le groupe 5. Ces patients se sont retrouvés dans le groupe 5 à cause de la présence d'anticorps et de la présence d'une arthrite très inflammatoire et affectant plusieurs articulations. Il est clair que ces marqueurs sont simplement des prédicteurs et non des agents de la progression des dommages articulaires. Il faudrait approfondir l'analyse et possiblement étudier chacun des paramètres des individus qui n'évoluent pas selon le modèle prédit et regarder s'ils ont des caractéristiques semblables autres que les érosions. La même chose devra être faite pour expliquer pourquoi un patient présentant déjà des érosions importantes à l'inclusion se retrouve malgré tout classifié dans un groupe moins sévèrement atteint. Ce modèle nous suggère donc une première façon de regrouper les patients selon leurs caractéristiques du départ, mais il devra être amélioré si l'on cherche éventuellement à obtenir un modèle prédictif.

Finalement, en lien avec le commentaire précédent, ce type d'analyse est moins souvent apprécié ou utilisé par les cliniciens puisqu'il ne permet pas de prédire de façon mathématique de quelle façon un patient va évoluer. Dans leur état actuel, ces résultats ne

sont donc pas utiles en clinique. Il faudra beaucoup plus approfondir et élargir notre population étudiée, ainsi que faire d'autres analyses pour envisager l'utilisation d'un outil prédictif en clinique.

5.2 Perspectives d'avenir

Il pourrait être intéressant de tester la stabilité du modèle en l'appliquant sur une cohorte semblable ou sur les 200 prochains patients. Évidemment, pour ce faire, il est nécessaire d'avoir les mêmes informations cliniques, car s'il manque une seule variable, les modèles ne seront pas comparables. Dans un cas semblable où une ou plusieurs variables seraient manquantes, il faudrait refaire un nouveau modèle avec notre cohorte en excluant ces variables, mais le modèle de base ne pourra alors être validé.

Il aurait aussi été intéressant d'étudier sur une plus longue période de temps l'évolution des sous-groupes. Par contre en augmentant la période d'observation minimale requise, le nombre de patients ayant atteint ce stade diminue considérablement. L'analyse pourra éventuellement être refaite dans quelques années pour évaluer la stabilité du modèle.

L'analyse pourrait aussi être refaite en réduisant les variables à celles prépondérantes observées au cours de l'ACM. Il pourrait être intéressant de voir si les patients se regroupent alors de la même façon.

Finalement, pour remédier au fait que les résultats ne sont pas encore accessibles présentement en clinique, il serait intéressant de créer un modèle prédictif à l'aide d'un modèle de régression logistique multivarié. Ce nouveau modèle pourrait alors, probablement après plusieurs essais et modifications du modèle de base, devenir un outil d'aide au diagnostic.

On peut aussi construire des modèles prédictifs en effectuant une classification de type métrique ou k-plus proches voisins ou encore en construisant une règle d'affectation bayésienne afin d'exploiter des connaissances et informations a priori. Il y a donc encore beaucoup de choses à faire dans ce domaine.

6 - Remerciements

Je tiens à remercier tout particulièrement mes directeurs Dr Gilles Boire et M. Bernard Colin qui m'ont grandement soutenu tout au long de ma maîtrise. Ils ont toujours été très présents et compréhensifs, particulièrement en ce qui concerne mes connaissances en biologie et en génétique!!! Cette recherche m'a beaucoup appris et m'a donné la possibilité de m'intégrer dans le domaine de la recherche clinique.

Merci également à mon troisième directeur Dr Jean-Marie Moutquin qui a permis la réalisation de cette recherche par ses connaissances et son implication en recherche clinique.

Merci aussi aux nombreux cliniciens de l'équipe (Dr Artur J. de Brum Fernandes, Dr Patrick Liang, Dr Sophie Roux, Dr Ariel Masetto), collaborateurs (Dr Pierre Cossette, Dr Henri A. Ménard, M. Claude Daniel et M. Maximilien Lora), les infirmières de recherche (Mme Céline Boulet et Mme Isabelle Gagnon), le personnel du laboratoire d'immunorhumatologie (Mme Carol Richard et Mme Céline St-Pierre). J'ai souvent dû entrer dans le bureau de Céline et Isabelle pour leur poser de nombreuses questions concernant le déroulement de l'étude du côté clinique et pour modifier la banque de données.

Évidemment, merci à l'équipe d'évaluateurs qui, je l'espère, auront trouvé le sujet et la lecture intéressants.

Je tiens à remercier ma famille, mes amies et collègues (Marie-Pierre Garant, Jennifer St-Laurent, Naomé Roy-Matton, Jolianne Renaud, Geneviève Leduc, Mathieu

Beaudoin et H  l  ne Ladouceur) qui m'ont soutenue et encourag  e. Ils ont aussi assist   plusieurs fois    mes pr  sentations de s  minaires avant la repr  sentation finale!

Les bourses que j'ai obtenues du *Canadian Arthritis Network* (CAN) et de l'Universit   de Sherbrooke (bourse institutionnelle) ont aussi aid      la r  alisation de cette ma  trise.

7 - Bibliographie

- Allaart, C. F., Y. P. Goekoop-Ruiterman, J. K. de Vries-Bouwstra, F. C. Breedveld, B. A. Dijkmans, and g. FARR study. 2006. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: The BeSt study. *Clinical & Experimental Rheumatology* 24(6 Suppl 43):S-77-82.
- Alparslan, L., J. S. Yu, and B. N. Weissman. Imaging. In: *Kelley's Textbook of Rheumatology*, E. D. Harris editor, 2005, p.744.
- American College of Rheumatology Subcommittee on Rheumatoid Arthritis, Guidelines. 2002. Guidelines for the management of rheumatoid arthritis: 2002 update. see comment. *Arthritis & Rheumatism* 46(2):328-346.
- Amos, R. S., T. J. Constable, R. A. Crockson, A. P. Crockson, and B. McConkey. 1977. Rheumatoid arthritis: Relation of serum C-reactive protein and erythrocyte sedimentation rates to radiographic changes. *British Medical Journal* 1(6055):195-197.
- Anderberg, M. R. Cluster analysis for application. London: Birnbaum, Z.W.; Lukacs, E., Academic Press Inc edition, 1973. (Volume 19 in Probability and mathematical statistics)
- Baccini, A., and P. Besse. Data mining 1. Exploration Statistique. Université Paul Sabatier Toulouse III, Laboratoire de Statistique et Probabilités, 2004.
- Balédent, F. 05/07/2007. Vitesse de sédimentation et CRP. *développement et santé, n° 146, avril 2000*, Available: http://devsante.org/article.php3?id_article=10875.
- Ballou, S. P. and I. Kushner. Laboratory evaluation of inflammation. In: *Kelley's Textbook of Rheumatology*, E. D. Harris editor, 2005, p.720.
- Berglin, E., L. Padyukov, U. Sundin, G. Hallmans, H. Stenlund, W. J. Van Venrooij, L. Klareskog, and S. R. Dahlqvist. 2004. A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis. *Arthritis Research & Therapy* 6(4):R303-8.
- Boire, G., P. Cossette, A. J. de Brum-Fernandes, P. Liang, T. Niyonsenga, Z. J. Zhou, N. Carrier, C. Daniel, and H. A. Menard. 2005. Anti-Sa antibodies and antibodies against cyclic citrullinated peptide are not equivalent as predictors of severe outcomes in patients with recent-onset polyarthritis. *Arthritis Research & Therapy* 7(3):R592-603.

- Bruynesteyn, K., D. Van Der Heijde, M. Boers, A. Saudan, P. Peloso, H. Paulus, H. Houben, B. Griffiths, J. Edmonds, B. Bresnihan, A. Boonen, and S. Van Der Linden. 2002. Detecting radiological changes in rheumatoid arthritis that are considered important by clinical experts: Influence of reading with or without known sequence. *Journal of Rheumatology* 29(11):2306-2312.
- Bruynesteyn, K., S. Van Der Linden, R. Landewe, F. Gubler, R. Weijers, and D. Van Der Heijde. 2004. Progression of rheumatoid arthritis on plain radiographs judged differently by expert radiologists and rheumatologists. *Journal of Rheumatology* 31(6):1088-1094.
- Bukhari, M. A., N. J. Wiles, M. Lunt, B. J. Harrison, D. G. Scott, D. P. Symmons, and A. J. Silman. 2003. Influence of disease-modifying therapy on radiographic outcome in inflammatory polyarthritis at five years: Results from a large observational inception study. *Arthritis & Rheumatism* 48(1):46-53.
- Combe, B., M. Dougados, P. Goupille, A. Cantagrel, J. F. Eliaou, J. Sibilia, O. Meyer, J. Sany, J. P. Daires, and A. Dubois. 2001. Prognostic factors for radiographic damage in early rheumatoid arthritis: A multiparameter prospective study. *Arthritis & Rheumatism* 44(8):1736-1743.
- Dept. of Rheumatology University Medical Centre Nijmegen - the Netherlands. Score DAS, disease activity score, Available: <http://www.das-score.nl/www.das-score.nl/>.
- Drossaers-Bakker, K. W., A. H. Zwinderman, T. P. Vlieland, D. Van Zeben, K. Vos, F. C. Breedveld, and J. M. Hazes. 2002. Long-term outcome in rheumatoid arthritis: A simple algorithm of baseline parameters can predict radiographic damage, disability, and disease course at 12-year followup. *Arthritis & Rheumatism* 47(4):383-390.
- Fayyad, U., G. Gregory Piatetsky-Shapiro, and P. Smyth. 1996. From data mining to knowledge discovery in databases. *AI Magazine* 17(3):37-54.
- Forslind, K., M. Ahlmen, K. Eberhardt, I. Hafstrom, B. Svensson, and G. BARFOT Study. 2004. Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: Role of antibodies to citrullinated peptides (anti-CCP). *Annals of the Rheumatic Diseases* 63(9):1090-1095.
- Forslind, K., I. Hafstrom, M. Ahlmen, B. Svensson, and G. the BARFOT Study. 2007. Sex: A major predictor of remission in early rheumatoid arthritis? *Annals of the Rheumatic Diseases* 66(1):46-52.
- Gabay, C. 2007. *Revue médicale suisse: Polyarthrites et connectivites à début tardif*, Available: <http://titan.medhyg.ch/mh/formation/print.php3?sid=21223>.
- George, J., Y. Levy, and Y. Shoenfeld. 1997. Smoking and immunity: An additional player in the mosaic of autoimmunity. *Scandinavian Journal of Immunology* 45(1):1-6.

- Ginns, L. C., J. H. Ryu, P. R. Rogol, N. L. Sprince, L. C. Oliver, and C. J. Larsson. 1985. Natural killer cell activity in cigarette smokers and asbestos workers. *American Review of Respiratory Disease* 131(6):831-834.
- Goldbach-Mansky, R., J. Lee, A. McCoy, J. Hoxworth, C. Yarboro, J. S. Smolen, G. Steiner, A. Rosen, C. Zhang, H. A. Menard, Z. J. Zhou, T. Palosuo, W. J. Van Venrooij, R. L. Wilder, J. H. Klippel, H. R. Schumacher Jr, and H. S. El-Gabalawy. 2000. Rheumatoid arthritis associated autoantibodies in patients with synovitis of recent onset. *Arthritis Research* 2(3):236-243.
- Guillemin, F., S. Briancon, and J. Pourel. 1992. Validity and discriminant ability of the HAQ functional index in early rheumatoid arthritis. *Disability & Rehabilitation* 14(2):71-77.
- Han, J., and M. Kamber. *Data mining, concepts and techniques*. Morgan Kaufmann Publishers edition, 2001.
- Haraoui, B., and E. Keystone. Chapter 5 - Rheumatoid arthritis. In: *Mechanisms in rheumatology*. H. Tannenbaum, and A. S. Russell, editors. 2001, p.41
- Harris, E. D. Clinical features of rheumatoid arthritis - hematologic abnormalities. In: *Kelley's Textbook of Rheumatology*, E. D. Harris editor, 2005, p.1057.
- Harrison, B., W. Thomson, D. Symmons, B. Ollier, N. Wiles, T. Payton, E. Barrett, and A. Silman. 1999. The influence of HLA-DRB1 alleles and rheumatoid factor on disease outcome in an inception cohort of patients with early inflammatory arthritis. *Arthritis & Rheumatism* 42(10):2174-2183.
- Harrison, B. J. 2002. Influence of cigarette smoking on disease outcome in rheumatoid arthritis. *Current Opinion in Rheumatology* 14(2):93-97.
- Harrison, B. J., A. J. Silman, N. J. Wiles, D. G. Scott, and D. P. Symmons. 2001. The association of cigarette smoking with disease outcome in patients with early inflammatory polyarthritis. see comment. *Arthritis & Rheumatism* 44(2):323-330.
- Hayem, G., P. Chazerain, B. Combe, A. Elias, T. Haim, P. Nicaise, K. Benali, J. F. Eliaou, M. F. Kahn, J. Sany, and O. Meyer. 1999. Anti-Sa antibody is an accurate diagnostic and prognostic marker in adult rheumatoid arthritis. erratum appears in *J rheumatol* 1999 sep;26(9):2069. *Journal of Rheumatology* 26(1):7-13.
- Hughes, D. A., P. L. Haslam, P. J. Townsend, and M. Turner-Warwick. 1985. Numerical and functional alterations in circulatory lymphocytes in cigarette smokers. *Clinical & Experimental Immunology* 61(2):459-466.
- Huizinga, T. W., C. I. Amos, van der Helm-van Mil, A.H., W. Chen, F. A. van Gaalen, D. Jawaheer, G. M. Schreuder, M. Wener, F. C. Breedveld, N. Ahmad, R. F. Lum, R. R. de Vries, P. K. Gregersen, R. E. Toes, and L. A. Criswell. 2005. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1

shared epitope for antibodies to citrullinated proteins. *Arthritis & Rheumatism* 52(11):3433-3438.

Hutchinson, D., L. Shepstone, R. Moots, J. T. Lear, and M. P. Lynch. 2001. Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA.see comment. *Annals of the Rheumatic Diseases* 60(3):223-227.

Inoue, E., H. Yamanaka, M. Hara, T. Tomatsu, and N. Kamatani. 2007. Comparison of disease activity score (DAS)28- erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values. *Annals of the Rheumatic Diseases* 66(3):407-409.

Irigoyen, P., A. T. Lee, M. H. Wener, W. Li, M. Kern, F. Batliwalla, R. F. Lum, E. Massarotti, M. Weisman, C. Bombardier, E. F. Remmers, D. L. Kastner, M. F. Seldin, L. A. Criswell, and P. K. Gregersen. 2005. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: Contrasting effects of HLA-DR3 and the shared epitope alleles. *Arthritis & Rheumatism* 52(12):3813-3818.

Jansen, A. L., I. van der Horst-Bruinsma, D. van Schaardenburg, R. J. van de Stadt, M. H. de Koning, and B. A. Dijkmans. 2002. Rheumatoid factor and antibodies to cyclic citrullinated peptide differentiate rheumatoid arthritis from undifferentiated polyarthritis in patients with early arthritis.see comment. *Journal of Rheumatology* 29(10):2074-2076.

Jonsson, T., J. Thorsteinsson, and H. Valdimarsson. 1998. Does smoking stimulate rheumatoid factor production in non-rheumatic individuals? *APMIS* 106(10):970-974.

Korpela, M., L. Laasonen, P. Hannonen, H. Kautiainen, M. Leirisalo-Repo, M. Hakala, L. Paimela, H. Blafield, K. Puolakka, T. Mottonen, and G. FIN-RACo Trial. 2004. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment with disease-modifying antirheumatic drugs: Five-year experience from the FIN-RACo study.see comment. *Arthritis & Rheumatism* 50(7):2072-2081.

Kroot, E. J., B. A. de Jong, M. A. van Leeuwen, H. Swinkels, F. H. van den Hoogen, M. van't Hof, L. B. van de Putte, M. H. van Rijswijk, W. J. van Venrooij, and P. L. van Riel. 2000. The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. *Arthritis & Rheumatism* 43(8):1831-1835.

La Société d'arthrite. 2007. Site de la société d'arthrite du canada, Available: <http://www.arthrite.ca/types%20of%20arthritis/default.asp?s=1>.

Landewe, R. B., M. Boers, A. C. Verhoeven, R. Westhovens, M. A. van de Laar, H. M. Markusse, J. C. van Denderen, M. L. Westedt, A. J. Peeters, B. A. Dijkmans, P. Jacobs, A. Boonen, D. M. van der Heijde, and S. van der Linden. 2002. COBRA

combination therapy in patients with early rheumatoid arthritis: Long-term structural benefits of a brief intervention.see comment. *Arthritis & Rheumatism* 46(2):347-356.

Larose, D. T. *Discovering knowledge in data, an introduction to data mining.* John Wiley, 2005.

Larsen, A. 1988. The relation of radiographic changes to serum acute-phase proteins and rheumatoid factor in 200 patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology* 17(2):123-129.

Linn-Rasker, S. P., van der Helm-van Mil, A.H., F. A. van Gaalen, M. Kloppenburg, R. R. de Vries, S. le Cessie, F. C. Breedveld, R. E. Toes, and T. W. Huizinga. 2006. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. *Annals of the Rheumatic Diseases* 65(3):366-371.

Lopez-Longo, F. J., M. Rodriguez-Mahou, S. Sanchez-Ramon, A. Estecha, M. Balsera, R. Plaza, E. Fernandez-Cruz, and L. C. Perez. 2006. Anti-cyclic citrullinated peptide versus anti-Sa antibodies in diagnosis of rheumatoid arthritis in an outpatient clinic for connective tissue disease and spondyloarthritis. *Journal of Rheumatology* 33(8):1476-1481.

MacGregor, A. J., and A. J. Silman. Classification and epidemiology. In: *Rheumatology.* Mosby, Toronto: M. C. Hochberg, A. J. Silman, J. S. Smolen, M. E. Weinblatt, and M. H. Weisman, editors., 2003, p.757.

Menard, H. A., E. Lapointe, M. D. Rochdi, and Z. J. Zhou. 2000. Insights into rheumatoid arthritis derived from the Sa immune system. *Arthritis Research* 2(6):429-432.

Meyer, O., C. Labarre, M. Dougados, P. Goupille, A. Cantagrel, A. Dubois, P. Nicaise-Roland, J. Sibilia, and B. Combe. 2003. Anticitrullinated protein/peptide antibody assays in early rheumatoid arthritis for predicting five year radiographic damage. *Annals of the Rheumatic Diseases* 62(2):120-126.

Molenaar, E. T., A. E. Voskuyl, and B. A. Dijkmans. 2002. Functional disability in relation to radiological damage and disease activity in patients with rheumatoid arthritis in remission. *Journal of Rheumatology* 29(2):267-270.

Mottonen, T., L. Paimela, M. Leirisalo-Repo, H. Kautiainen, J. Ilonen, and P. Hannonen. 1998. Only high disease activity and positive rheumatoid factor indicate poor prognosis in patients with early rheumatoid arthritis treated with "sawtooth" strategy.see comment. *Annals of the Rheumatic Diseases* 57(9):533-539.

Nepom, G. T., and B. Nepom. Genetics of the major histocompatibility complex in rheumatoid arthritis. In: *Rheumatology,* M. C. Hochberg, editor, 2003, p.811.

- Niewold, T. B., M. J. Harrison, and S. A. Paget. 2007. Anti-CCP antibody testing as a diagnostic and prognostic tool in rheumatoid arthritis. *Qjm* 100(4):193-201.
- Pedersen, M., S. Jacobsen, P. Garred, H. O. Madsen, M. Klarlund, A. Svejgaard, B. V. Pedersen, J. Wohlfahrt, and M. Frisch. 2007. Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: A nationwide case-control study in denmark. *Arthritis & Rheumatism* 56(5):1446-1453.
- Pincus, T., J. A. Summey, S. A. Soraci Jr, K. A. Wallston, and N. P. Hummon. 1983 Nov. Assessment of patient satisfaction in activities of daily living using a modified stanford health assessment questionnaire. *Arthritis & Rheumatism* 26(11):1346-1353.
- Prevoo, M. L., M. A. van 't Hof, H. H. Kuper, M. A. van Leeuwen, L. B. van de Putte, and P. L. van Riel. 1995. Modified disease activity scores that include twenty-eight-joint counts. development and validation in a prospective longitudinal study of patients with rheumatoid arthritis.see comment. *Arthritis & Rheumatism* 38(1):44-48.
- Quinn, M. A., P. G. Conaghan, P. J. O'Connor, Z. Karim, A. Greenstein, A. Brown, C. Brown, A. Fraser, S. Jarret, and P. Emery. 2005. Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal: Results from a twelve-month randomized, double-blind, placebo-controlled trial. *Arthritis & Rheumatism* 52(1):27-35.
- Saag, K. G., J. R. Cerhan, S. Kolluri, K. Ohashi, G. W. Hunninghake, and D. A. Schwartz. 1997. Cigarette smoking and rheumatoid arthritis severity.see comment. *Annals of the Rheumatic Diseases* 56(8):463-469.
- Saporta, G. Probabilités, analyse des données et statistique. Paris: Technip edition, 1990.
- Schellekens, G. A., H. Visser, B. A. de Jong, F. H. van den Hoogen, J. M. Hazes, F. C. Breedveld, and W. J. van Venrooij. 2000. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide.see comment. *Arthritis & Rheumatism* 43(1):155-163.
- Symmons, D. P., C. R. Bankhead, B. J. Harrison, P. Brennan, E. M. Barrett, D. G. Scott, and A. J. Silman. 1997. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis: Results from a primary care-based incident case-control study in Norfolk, England. *Arthritis & Rheumatism* 40(11):1955-1961.
- Tannenbaum, H., and A. S. Russell. Mechanisms in rheumatology. Core Health Services Inc edition. Ontario, Canada. 2001
- Thomson, W., B. Harrison, B. Ollier, N. Wiles, T. Payton, J. Barrett, D. Symmons, and A. Silman. 1999. Quantifying the exact role of HLA-DRB1 alleles in susceptibility

- to inflammatory polyarthritis: Results from a large, population-based study. *Arthritis & Rheumatism* 42(4):757-762.
- Tortora, G. J., S. R. Grabowski, and J. Parent. Chapitre 22 - Le système lymphatique, les mécanismes de défense non-spécifiques et l'immunité. In: *Principes d'anatomie et de physiologie*, Centre Éducatif et Culturel Inc., 1994a, p.761
- Tortora, G. J., S. R. Grabowski, and J. Parent. Chapitre 9 - Les articulations. In: *Principes d'anatomie et de physiologie*, Centre Éducatif et Culturel Inc., 1994b, p.226.
- Tufféry, S. *Data mining et statistique décisionnelle, l'intelligence dans les bases de données.*, Technip edition, 2005a.
- Tufféry, S. *L'analyse des correspondances multiples.* In: *Data mining et statistique décisionnelle, l'intelligence dans les bases de données*, Technip edition, 2005b, p.104.
- van Aken, J., L. R. Lard, S. le Cessie, J. M. Hazes, F. C. Breedveld, and T. W. Huizinga. 2004. Radiological outcome after four years of early versus delayed treatment strategy in patients with recent onset rheumatoid arthritis. *Annals of the Rheumatic Diseases* 63(3):274-279.
- van der Heide, A., C. A. Remme, D. M. Hofman, J. W. Jacobs, and J. W. Bijlsma. 1995. Prediction of progression of radiologic damage in newly diagnosed rheumatoid arthritis. *Arthritis & Rheumatism* 38(10):1466-1474.
- van der Heijde, D. 2000. How to read radiographs according to the Sharp/van der Heijde method. *Journal of Rheumatology* 27(1):261-263.
- van der Heijde, D. M. 1995. Joint erosions and patients with early rheumatoid arthritis. *British Journal of Rheumatology* 34(Suppl 2):74-78.
- van der Heijde, D. M., M. A. van 't Hof, P. L. van Riel, L. A. Theunisse, E. W. Lubberts, M. A. van Leeuwen, M. H. van Rijswijk, and L. B. van de Putte. 1990. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. *Annals of the Rheumatic Diseases* 49(11):916-920.
- van der Helm-van Mil, A.H., T. W. Huizinga, G. M. Schreuder, F. C. Breedveld, R. R. de Vries, and R. E. Toes. 2005. An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility. *Arthritis & Rheumatism* 52(9):2637-2644.
- van der Helm-van Mil, A.H., K. N. Verpoort, F. C. Breedveld, T. W. Huizinga, R. E. Toes, and R. R. de Vries. 2006. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. *Arthritis & Rheumatism* 54(4):1117-1121.

- van Gaalen, F. A., J. van Aken, T. W. Huizinga, G. M. Schreuder, F. C. Breedveld, E. Zanelli, W. J. van Venrooij, C. L. Verweij, R. E. Toes, and R. R. de Vries. 2004. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. *Arthritis & Rheumatism* 50(7):2113-2121.
- van Riel, P. L., D. M. van der Heijde, I. H. Nuvér-Zwart, and L. B. van de Putte. 1995. Radiographic progression in rheumatoid arthritis: Results of 3 comparative trials. *Journal of Rheumatology* 22(9):1797-1799.
- van Zeben, D., and F. C. Breedveld. 1996. Prognostic factors in rheumatoid arthritis. *Journal of Rheumatology - Supplement* 44:31-33.
- Vencovsky, J., S. Machacek, L. Sedova, J. Kafkova, J. Gatterova, V. Pesakova, and S. Rzickova. 2003. Autoantibodies can be prognostic markers of an erosive disease in early rheumatoid arthritis. *Annals of the Rheumatic Diseases* 62(5):427-430.
- Verpoort, K. N., F. A. van Gaalen, van der Helm-van Mil, A.H., G. M. Schreuder, F. C. Breedveld, T. W. Huizinga, R. R. de Vries, and R. E. Toes. 2005. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. *Arthritis & Rheumatism* 52(10):3058-3062.
- Visser, H., S. le Cessie, K. Vos, F. C. Breedveld, and J. M. Hazes. 2002. How to diagnose rheumatoid arthritis early: A prediction model for persistent (erosive) arthritis [see comment]. *Arthritis & Rheumatism* 46(2):357-365.
- Vittecq, O., V. Goeb, X. Le Loet, and F. Tron. 2005. The preclinical phase of rheumatoid arthritis and lupus. *Joint, Bone, Spine: Revue Du Rhumatisme* 72(6):443-446.
- Weyand, C. M., D. Schmidt, U. Wagner, and J. J. Goronzy. 1998. The influence of sex on the phenotype of rheumatoid arthritis [see comment]. *Arthritis & Rheumatism* 41(5):817-822.

8 – Annexes

Annexe 1. Formulaire du médecin (EUPA)

Date d'inclusion dans EUPA : _____

Date de la dernière visite EUPA : _____

Date estimée du début de l'arthrite inflammatoire : _____

Durée de la raideur matinale : _____

Niveau fonctionnel: 1 2 3 4

Manifestations extra articulaires : Aucune

- | | | |
|---|--|--------------------------------------|
| <input type="checkbox"/> Nodosité | <input type="checkbox"/> Psoriasis | <input type="checkbox"/> Uvéite |
| <input type="checkbox"/> Fibrose pulmonaire | <input type="checkbox"/> pleurite ou péricardite | <input type="checkbox"/> Sjögren |
| <input type="checkbox"/> Felty | <input type="checkbox"/> Palindrome | <input type="checkbox"/> Fièvre (fu) |
| <input type="checkbox"/> RS3PE | <input type="checkbox"/> PMR-Like | <input type="checkbox"/> Autre _____ |
| | | <input type="checkbox"/> Vasculite |
| | | <input type="checkbox"/> Raynaud |

Diagnostic:

- | | | |
|---|---|--|
| <input type="checkbox"/> Polyarthrite rhumatoïde séropositive | <input type="checkbox"/> Polyarthrite rhumatoïde séronégative | <input type="checkbox"/> polyarthrite séronégative |
| <input type="checkbox"/> Arthrite psoriasique | <input type="checkbox"/> Spondyloarthropathie | <input type="checkbox"/> lupus |
| <input type="checkbox"/> Goutte | <input type="checkbox"/> autre _____ | |

Suivi de la médication :

MÉDICAMENTS	DOSE ACTUELLE	CHANGEMENT CE JOUR (DOSE)
DMARDS		
Méthotrexate		
Plaquenil		
Sulfazalazine		
Biologique		
Prednisone		
Injection I/A		
AINS		
Biphosphonate		
Statine		

Sévérité de la polyarthrite selon le médecin :



Est-ce que vous croyez que ce patient va développer :

- une polyarthrite chronique (> 3 ans) ? oui non ne sais pas
 Une rémission complète? oui non ne sais pas

Signature

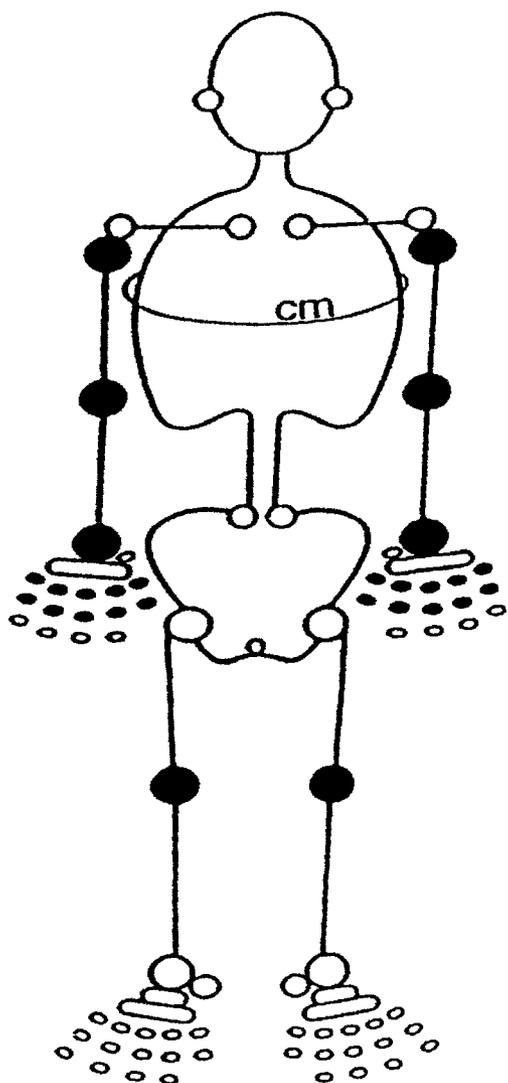
Date

Figure 17.

Illustration des articulations évaluées par le médecin

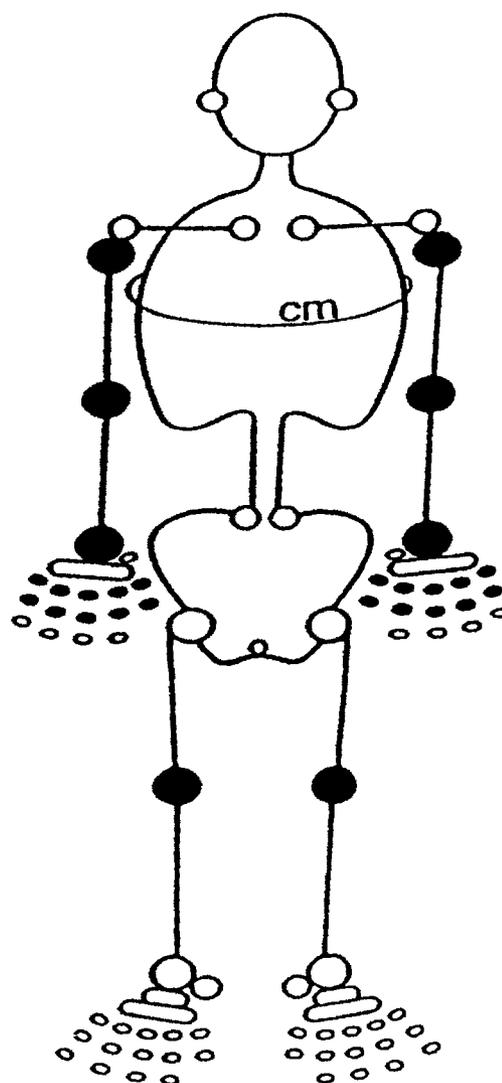
SYNOVITES

()



DOULEURS

X



Annexe 2. Formulaire de l'infirmière (EUPA)

Sexe : femme homme

Âge : _____

Scolarité (études complétées) :

école primaire : 1 2 3 4 5 6 7

école secondaire : sec. I II III IV V

CÉGEP 1 2 3

Université baccalauréat maîtrise Ph. D.

Autre _____

Emploi :

temps complet pas de travail rémunéré

temps partiel étudiant

retraité sans emploi (chômeur)

congé de maladie dû à l'arthrite oui non

Travail du patient (profession) : _____

Ethnicité :

blanc canadien français

anglais

africain canadien

amérindien

hispanique

Asie de l'Est / oriental / chinois

caucasien (origine européenne)

Asie du Sud / indien / philippin

autre _____

Statut marital :

célibataire

marié

séparé

veuf (ve)

divorcé

conjoint de fait

Avec qui vivez-vous actuellement?

seul

avec autre membre de la famille

avec époux ou conjoint

avec ami ou colocataire

avec enfant(s)

résidence

avec parent(s)

avec une personne rémunérée pour cela

Chirurgie articulaire : _____

Taille : _____

Poids : _____

Overall Status measure for rheumatoid arthritis : OSRA (grille)

Activity Score*	0	1	2
Well being [†]	Same/better	Worse	Much worse
Joint pain [†]	None/mild	Moderate	Severe
EMS [†]	< 30 min.	30-60 min.	> 60 min.
Active joint count	< 3	3-10	> 10
Extraarticular disease	None	Nodules	Other
Damage score*			
Function (problems) [†]	Mild	Eat/dress/stairs	Housebound
Social life (impairment) [†]	None	Moderate	Severe
Destroyed/replaced large jts	0	1-2	> 2
Small joint damage	Mild	Needs aids	Surgery
Organ Impairment	Mild	Skin/muscle	Major

Critères de PR (ACR)

Raideur péri-articulaire (1 h) _____

Gonflement (3 articulations) _____

Gonflement poignet, MCP, IPP _____

Gonflement symétrique _____

Nodosité rhumatoïde _____

Facteur rhumatoïde _____

Érosion ou porose aux mains/poignets
(4 critères ou plus pendant 6 sem.) _____

TOTAL : _____

Co-morbidité:

Effets secondaires :

Signature: _____

Date: _____

Annexe 3.
FORMULE D'ÉVALUATION FONCTIONNELLE
Index des activités de la vie courante

Les questions qui suivent concernent vos activités quotidiennes.
 Répondez à chaque question, même si vous pensez qu'elle n'a aucun rapport avec votre problème de santé.

IL N'Y A PAS DE BONNES OU DE MAUVAISES RÉPONSES.
 Répondez exactement comme vous le pensez ou comme vous le ressentez.

Visite : _____

S.V.P. cochez (✓) UNE SEULE réponse, celle qui correspond le mieux à votre situation au cours des 8 derniers jours.

1. Êtes-vous capable de :	Sans aucune difficulté	Avec un peu de difficulté	Avec beaucoup de difficulté	Incapable de le faire
a. Vous habillez, y compris nouer vos lacets de soulier et boutonner vos vêtements?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
b. Vous étendre au lit et vous lever du lit?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
c. Porter à votre bouche une tasse ou un verre de plein?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
d. Marcher en terrain plat à l'extérieur?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
e. Vous laver et vous sécher entièrement?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
f. Vous baisser pour ramasser un vêtement par terre?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
g. Ouvrir et fermer les robinets?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
h. Monter et descendre d'une auto?	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3

2. Comment vous sentez-vous AUJOURD'HUI en comparaison à il y a UN MOIS?

S.V.P. cocher UNE SEULE réponse.

- beaucoup mieux
- mieux
- le même
- pire
- beaucoup pire

4. Quel est votre degré de SATISFACTION par rapport à votre capacité de faire vos activités habituelles?

S.V.P. cocher UNE SEULE réponse.

- très satisfait
- assez satisfait
- assez peu satisfait
- très peu satisfait

3. Quelle phase vous décrit le mieux AUJOURD'HUI?

S.V.P. cocher UNE SEULE réponse.

- Je peux faire tout ce que je veux.
- Je peux faire la plupart des choses que je veux faire, mais avec un peu de limitations.
- Je peux faire certaines choses que je veux faire, mais pas toutes, et j'ai plusieurs limitations.
- Je ne peux presque rien faire de ce que je veux faire.

5. Quand vous vous levez le matin, vous sentez-vous raide ou enraidit?

- oui non

6. Si vous avez répondu OUI, combien de temps vous faut-il pour devenir aussi souple que vous pouvez le devenir dans la journée?

_____ minutes ou _____ heures

7. Combien de douleur avez-vous eue à cause de votre maladie pendant LA DERNIÈRE SEMAINE?

Faites un trait vertical « | » sur la ligne ci-dessous :

Aucune douleur |-----| Douleur la pire que je n'ai jamais eue

8. Combien de fatigue avez-vous ressentie pendant LA DERNIÈRE SEMAINE?

Faites un trait vertical « | » sur la ligne ci-dessous :

Pas de fatigue |-----| Fatigue extrême

Visit : _____

ACTIVITIES AND LIFESTYLE INDEX

The questions below concern your daily activities.
Please try to answer each question, even if you do not think it is Related to you or any condition you may have.

THERE ARE NO RIGHT OR WRONG ANSWERS.
Please answer exactly as you think of feel.

Please check (✓) the ONE best answer for your abilities:

1. Are you able to :

- a. Dress yourself, including trying shoelaces and doing buttons?
- b. Get in and out of bed?
- c. Lift a full cup or glass to your mouth?
- d. Walk outdoors on flat ground?
- e. Wash and dry you entire body?
- f. Bend down to pick up clothing from the floor?
- g. Turn regular faucets on and off?
- h. Get in and out of a car?

Without any difficulty	With some difficulty	With much difficulty	Unable to do
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3

2. How do you feel TODAY compared to ONE MONTH AGO?

Please check (✓) only one.

- much better today than one month ago
- better today than one month ago
- the same today as one month ago
- worse today than one month ago
- much worse today than on month ago

3. Which of following best describes you TODAY?
Please check (✓) only one.

- I can do everything I want to do.
- I can do most of the things I want to do, but have some limitations.
- I can do some, but not all, of the things I want to do, and I have many limitations.
- I can hardly do any of the things I want to do.

4. How SATISFIED are you with your ability to do your usual activities?

Please check (✓) only one.

- very satisfied
- somewhat satisfied
- somewhat dissatisfied
- very dissatisfied

5. When you get up in the morning, do you feel stiff?

- yes
- no

6. If you answer YES, how long is it until you are limber as you will be for that day?

_____ minutes or _____ hours

7. How much pain have you had because of you condition IN THE PAST WEEK?

Place a vertical line « | » below :

No pain

Pain as bad as it could be

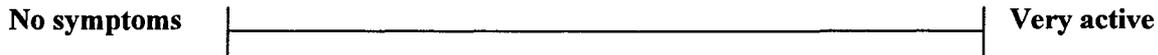
8. How much of a problem has UNUSUAL fatigue or tiredness been for you OVER THE PAST WEEK?

Place a vertical line « | » below :

Fatigue is no problem

Fatigue is a major problem

9. Disease activity, considering all the ways in which illness and health conditions may affect you at this time.
Place a vertical line « | » below :



How much of a problem has sleep (resting at night) been for you IN THE PAST WEEK?

Please mark ONE BOX BELOW (X) to indicate the severity of the problem.

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

Sleep is not a problem

Sleep is a major problem

TABACCO AND ALCOHOL USE

1. Tobacco and alcohol use (check all that apply)

<input type="checkbox"/> Never used	User of :	Ex-user of :
	<input type="checkbox"/> Cigarettes : # years _____ # packs/day _____ <input type="checkbox"/> Pipe <input type="checkbox"/> Cigars <input type="checkbox"/> Chewing Tobacco	<input type="checkbox"/> Cigarettes : # years _____ # packs/day _____ <input type="checkbox"/> Pipe <input type="checkbox"/> Cigars <input type="checkbox"/> Chewing Tobacco I stopped : _____

2. Alcohol Use (Check only one)

<input type="checkbox"/> Non-Drinker	Drinker :	Ex-drinker :
	_____ drink per _____ (indicate the number and check the frequency) <input type="checkbox"/> week <input type="checkbox"/> month	<input type="checkbox"/> Light (less than 2 drinks/day) <input type="checkbox"/> Moderate (2-4 drinks/day) <input type="checkbox"/> Heavy (more than 4 drinks/day)

Please list below all drugs or medicines taken over the last week (include birth control pills, aspirin, and any kind of drug or medicine bought without prescription).

	Name of the drug	Dose (if known)	How many per day?	How helpful is it?			Any side effects?		If side effect, is it		
				A lot	Some	None	Yes	No	GI	Skin	Other
1.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
2.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
3.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
4.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
5.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
6.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
7.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
8.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
9.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
10.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
11.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				
12.	_____	_____	_____	<input type="checkbox"/>	_____	_____	_____				