



HAUTE AUTORITÉ DE SANTÉ

SERVICE ÉVALUATION ÉCONOMIQUE ET SANTÉ PUBLIQUE

Évaluation médico-économique des traitements de fond biologiques dans la prise en charge de la polyarthrite rhumatoïde

Rapport final

15 Avril 2019

Ce rapport est téléchargeable sur
www.has-sante.fr

Haute Autorité de Santé

Service Communication – Information

5, avenue du Stade de France – F 93218 Saint-Denis La Plaine Cedex

Tél. : +33 (0)1 55 93 70 00 – Fax : +33 (0)1 55 93 74 00

Sommaire

Évaluation médico-économique des traitements de fond biologiques dans la prise en charge de la polyarthrite rhumatoïde	1
Sommaire	3
Abréviations et acronymes	12
Introduction générale	15
Chapitre 1. Évaluation des traitements de fond biologiques dans la prise en charge de la PR : objectifs, contexte, méthodes	22
1. Objectifs et enjeux de l'évaluation.....	22
1.1 Saisine.....	22
1.2 Feuille de route.....	22
1.3 Objectif et enjeux de l'évaluation	23
1.4 Champs exclus de l'évaluation	24
2. Méthodologie de l'évaluation et processus d'assurance qualité	27
2.1. Méthodes d'analyse	27
2.1.1. Revue systématique de littérature clinique et réalisation de méta-analyses en réseau	27
2.1.2. Revue systématique de la littérature économique.....	28
2.1.3. Réalisation d'un modèle économique : modèle PR France	29
2.2. Mise en œuvre d'un processus d'assurance qualité	30
2.3. Validation avec un groupe de travail scientifique, des experts internationaux et les instances de la HAS.....	31
2.4. Calendrier d'élaboration du document.....	32
3. Contexte épidémiologique de l'évaluation	33
3.1. Prévalence de la PR	33
3.2. Mortalité, comorbidités et PR.....	34
4. Contexte clinique d'évaluation de la PR.....	40
4.1. Objectifs du traitement de la polyarthrite rhumatoïde.....	40
4.2. Synthèse des critères d'efficacité clinique utilisés dans le traitement de la PR.....	40
4.3. Synthèse des critères de qualité de vie utilisés dans le traitement de la PR.....	43
5. Prise en charge médicamenteuse et recommandations portant sur la PR.....	46
5.1. Les traitements de fond biologiques	46
5.2. Les traitements de fond synthétiques ciblés.....	47
5.3. Les recommandations de prise en charge médicamenteuse.....	47
6. Conclusion	54
Chapitre 2. Méthodologie de l'analyse de l'efficacité comparative des traitements de fond biologiques	55
1. Méthodes d'élaboration de la revue systématique de littérature.....	56
1.1. Identification des sources de données	57

1.2. Critères de sélection des essais cliniques dans la revue systématique.....	58
1.2.1. Critères d'inclusion.....	59
1.2.2. Les critères d'exclusion.....	61
1.2.3. Critères supplémentaires pour la réalisation des analyses de sensibilité.....	62
1.3. Processus de sélection, d'extraction et d'évaluation des études.....	63
2. Méthodologie des méta-analyses en réseau planifiées.....	64
2.1. Méthodes d'estimation.....	64
2.1.1. Choix d'une approche bayésienne.....	64
2.1.2. Choix d'une fonction de lien probit.....	64
2.1.3. Estimation des effets relatifs des traitements et des probabilités de réponse.....	65
2.2. Diagnostic de qualité des méta-analyses en réseau.....	67
2.2.1. Adéquation des modèles des méta-analyses en réseau.....	67
2.2.2. Évaluation de la cohérence des preuves.....	67
2.2.3. Analyse de l'hétérogénéité.....	68
2.2.3.1. <i>Les méta-régressions</i>	68
2.2.3.2. <i>Les analyses de sensibilité</i>	68
2.3. Présentation des résultats.....	68
3. Discussion.....	69
4. Conclusion.....	70
Chapitre 3. Efficacité comparative : Résultats de la revue systématique de la littérature et des méta-analyses en réseau.....	72
1. Études retenues.....	72
2. Évaluation de la qualité des essais cliniques retenus dans la mise à jour de la revue systématique réalisée par la HAS.....	80
3. Les critères d'efficacité retenus dans la revue systématique de littérature HAS & SchARR.....	86
3.1. Réponses ACR colligées dans la revue systématique de littérature HAS & SchARR.....	86
3.2. Réponses EULAR colligées dans la revue systématique HAS & SchARR.....	91
4. Résultats des méta-analyses en réseau HAS & SchARR.....	94
4.1. Les réseaux de preuves des méta-analyses en réseau.....	94
4.2. Estimation des effets de traitements et des probabilités de réponse.....	96
4.2.1. Évaluation de la qualité des ajustements des méta-analyses en réseau.....	96
4.2.2. Analyse principale – Patients ayant été traités par un csDMARD (MTX).....	98
4.3. Les analyses de sensibilité.....	107
4.3.1. Analyse de sensibilité n° 1.....	109
4.3.2. Analyse de sensibilité n° 2.....	114
4.3.3. Analyse de sensibilité n° 3.....	120
4.3.4. Analyse de sensibilité n° 4.....	125
4.3.5. Analyse de sensibilité n° 5.....	131
4.3.6. Analyse de sensibilité n° 6.....	136
4.4. Évaluation de la qualité statistique des méta-analyses HAS & SchARR.....	141

4.4.1. Analyse de l'hétérogénéité.....	141
4.4.2. Analyse de la cohérence des preuves.....	142
4.4.3. Cohérence des intervalles de confiance associés aux réponses ACR et EULAR.....	143
5. Synthèse et discussion	144
5.1. Synthèse	144
5.2. Discussion.....	146
6. Synthèse d'efficacité comparative sur les tsDMARDs et le sarilumab	149
6.1. Synthèse des évaluations cliniques de la HAS (avis CT).....	149
6.2. Synthèse des méta-analyses récentes.....	151
6.2.1. Revue systématique et méta-analyses de l'Agence canadienne des technologies de santé (CADTH).....	151
6.2.2. Revue systématique et méta-analyses en réseau de l' <i>Institute for Clinical and Economic Review</i> (ICER).....	151
6.2.3. Revues systématiques et méta-analyses en réseau de la Cochrane	152
6.3. Synthèse des critiques des évaluations du NICE.....	152
7. Conclusion	154
Chapitre 4. Revue systématique de la littérature économique portant sur les traitements de fond biologiques.....	156
1. Méthodes pour l'élaboration de la revue systématique de la littérature économique.....	159
1.1. Les questions spécifiques de la revue systématique de la littérature économique	159
1.2. Identification des sources de données	159
1.3. Critères de sélection des études éligibles à la revue systématique de la littérature.....	162
1.4. Processus de sélection, d'extraction et d'évaluation des résultats	164
2. Résultats de la revue systématique de la littérature économique.....	167
2.1. Sélection des références.....	167
2.2. Synthèse des résultats de la revue systématique de la littérature économique	169
3. Analyse critique de la revue systématique de littérature économique	177
3.1. Évaluation selon les recommandations CHEERS adaptées à la PR.....	177
3.2. Évaluation critique selon une liste de questions critiques provenant de plusieurs grilles	178
3.3. Les aspects présentant une amélioration par rapport aux évaluations économiques identifiées dans la littérature	180
3.4. Les limites actuelles de l'évaluation économique des séquences de traitement dans la littérature	182

4. Discussion de la revue systématique de la littérature économique.....	184
5. Conclusion	187
Chapitre 5. Modélisation pour l'évaluation économique de la prise en charge de la PR en France.....	189
1. Le modèle PR France.....	190
1.1. Les raisons du choix de modèle PR France	190
1.2. Structure du modèle PR France	190
1.2.1. Les événements.....	192
1.2.2. Les conditions	194
1.3. Exécution du modèle.....	197
1.4. Validation du modèle.....	198
1.4.1. Adéquation des résultats avec l'intuition : <i>face validity</i>	198
1.4.2. Validation interne.....	199
1.4.3. Validation externe	199
2. Choix structurants de l'évaluation économique.....	200
2.1. Type d'analyse	200
2.2. Stratégies à comparer, stratégies de l'analyse de référence, méthode d'analyse.....	200
2.3. Population d'entrée du modèle	202
2.4. Horizon temporel et prise en compte de la dimension temporelle	202
2.5. Perspective.....	203
2.6. Actualisation	203
2.7. Les coûts et leur valorisation	204
2.8. Résultats de santé et leur valorisation.....	204
2.9. Présentation des résultats.....	205
2.10. Analyses de sensibilité.....	206
3. Sources de données mobilisées dans le modèle PR France.....	208
3.1. Les sources de données de l'efficacité comparative extraites à partir d'essais cliniques.....	208
3.1.1. Données en 1 ^{re} ligne de traitements biologiques.....	208
3.1.2. Données en 2 ^e ligne et plus de traitements biologiques.....	208
3.2. Les données en vie réelle	210
3.2.1. Sources des données françaises en vie réelle portant sur les patients atteints de la PR.....	210
3.2.2. Registres et cohorte français identifiés.....	211
3.2.3. Pourquoi la cohorte ESPOIR a-t-elle été sélectionnée ?.....	213
3.2.4. La cohorte ESPOIR.....	214
3.3. Les données portant sur la tolérance	215
3.3.1. Types de données de tolérance utilisées dans le modèle PR France	215
3.3.2. Sources des données décrivant les événements indésirables.....	216
3.4. Les données portant sur la mortalité	216
3.5. Les estimations des désutilités associées aux événements indésirables.....	217
3.6. L'estimation des coûts.....	217
3.6.1. Les coûts d'acquisition des traitements de fond synthétiques et biologiques.....	218

3.6.2. Les coûts d'administration.....	218
3.6.3. Les coûts de suivi : consultations, examens biologiques	219
3.6.4. Coûts des événements indésirables.....	219
4. Les équations du modèle PR France.....	220
4.1. La survie.....	221
4.2. Temps pour développer une érosion	224
4.3. Les coûts relatifs aux consultations (rhumatologue, médecin généraliste), examens biologiques et imagerie	224
4.4. Calendrier des visites.....	225
4.5. Calendrier des doses	225
4.6. Bornes inférieure et supérieure de la variation du DAS 28	226
4.7. L'activité (DAS 28)	227
4.8. Changement de l'activité.....	227
4.9. L'incapacité fonctionnelle (le HAQ).....	227
4.10. Changer le traitement.....	229
4.11. Temps de changement de traitement.....	230
4.12. Les utilités	234
5. Les entrées du modèle PR France.....	236
6. Discussion.....	242
7. Conclusion	246
Chapitre 6. Résultats de l'évaluation économique des séquences de traitements de fond biologiques dans la prise en charge de la PR en France	247
1. Les résultats de l'analyse de référence.....	248
1.1. Les résultats de santé	249
1.2. Les coûts	250
1.3. L'analyse de l'efficience	251
2. Exploration de l'incertitude sur l'évaluation économique des séquences de traitements de fond biologiques.....	252
2.1. Les analyses de scénarios déterministes	252
2.1.1. Choix des analyses de scénarios déterministes.....	252
2.1.2. Les résultats des analyses de scénarios déterministes	254
2.2. Les analyses de sensibilité probabilistes.....	261
2.2.1. Choix des séquences dans l'analyse probabiliste	261
2.2.2. Spécifications des distributions des paramètres intégrés dans les analyses probabilistes	262
2.2.3. Traitement et interprétation des résultats des ASP	266
2.2.4. Résultats	267
3. Analyses complémentaires.....	270
3.1. Analyses complémentaires n° 1.....	270
3.1.1. Rationnel	270
3.1.2. Séquences sélectionnées	270
3.1.3. Données utilisées	270

3.1.4. Résultats	271
3.2. Analyses complémentaires n° 2.....	272
3.2.1. Rationnel	272
3.2.2. Séquences sélectionnées	272
3.2.3. Données utilisées	273
3.2.4. Résultats	276
4. Discussion.....	279
5. Conclusion	284
Conclusion générale / Synthèse	286
Références bibliographiques	294
Équipe projet & remerciements	303

Messages clefs

- Ce rapport répond à la saisine de la Direction de la Sécurité sociale portant sur « ***l'évaluation médico-économique des biothérapies dans la prise en charge de la polyarthrite rhumatoïde (PR)*** » et propose un modèle de décision économique pour comparer des stratégies de prise en charge de la PR en France (y compris des analyses de scénarios sur les traitements de fond synthétiques ciblés (tsDMARDs)).
- Pour les patients ayant une PR modérée à sévère **en échec de traitement de fond synthétique conventionnel csDMARD** (ex. : méthotrexate), les méta-analyses en réseau réalisées dans le cadre de ce projet ont mis en évidence que, si le comparateur commun était un csDMARD (ex. : méthotrexate), les associations du méthotrexate avec l'étanercept (ou son biosimilaire) et avec le tocilizumab étaient caractérisées par des réponses ACR et EULAR élevées. En revanche, les résultats des méta-analyses en réseau ne détectaient pas de différences statistiquement significatives entre les biothérapies (bDMARDs). **Une synthèse exploratoire de l'efficacité comparative des traitements de fond synthétiques ciblés (tsDMARDs) récemment évalués par la HAS complète les résultats des méta-analyses réalisées.** Il convient de noter que les associations des tsDMARDs avec le MTX étaient également plus efficaces que le MTX en monothérapie et qu'on ne peut pas différencier en termes de bénéfice clinique un tsDMARD d'un bDMARD.
- Un modèle *de novo* « **modèle à conditions et événements discrètement intégrés** », nommé PR France, a été élaboré afin de répondre aux limites identifiées dans la littérature économique. Il repose sur **un algorithme décisionnel de la prise en charge de la PR en pratique clinique et simule le parcours médicamenteux des patients atteints de PR modérée à sévère via 180 séquences de bDMARDs (et leurs biosimilaires) dans l'analyse de référence. Trente séquences supplémentaires ont inclus les produits récemment disponibles** (inhibiteurs de Janus kinase « JAK », un nouvel inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 et un nouveau biosimilaire) dans des analyses de scénarios.
- **Des données françaises en vie réelle longitudinales** issues de la **cohorte ESPOIR** « Étude et suivi des polyarthrites indifférenciées récentes », coordonnée par la SFR, ont été notamment mobilisées pour estimer les équations du modèle.
- L'analyse de référence a identifié deux séquences se situant sur la frontière d'efficacité : **biosimilaire de l'étanercept→abatacept→infliximab** et **biosimilaire de l'étanercept→tocilizumab→abatacept**. Les valeurs des ratios différentiels coût-résultat (RDCR) pour la deuxième séquence sur la frontière (et pour celles qui ne

sont pas dominées dans les scénarios) par rapport à la séquence (**biosimilaire de l'étanercept→abatacept→infliximab**), **étaient élevées et supérieures à 100 000 €/QALY**.

- Le différentiel moyen actualisé des années de vie pondérées par la qualité de vie liée à la santé (QALY) entre les séquences étudiées, résultant de l'efficacité clinique mesurée par la réponse EULAR, **était très faible** (égal à 1 %).
- Le différentiel de coûts entre les séquences **était plus important** que pour les QALY. Alors que la différence dans les QALY actualisés n'était que de 1 % en moyenne, les coûts actualisés différaient de plus de 35 %. Les séquences débutant par le biosimilaire de l'étanercept étaient moins coûteuses que les séquences correspondantes débutant par d'autres anti-TNFα. Les séquences les moins coûteuses étaient : **biosimilaire de l'étanercept→abatacept→infliximab** (analyse de référence) et **biosimilaire de l'adalimumab→abatacept→infliximab** (analyse complémentaire).
- Les coûts d'acquisition des traitements de fond biologiques constituent le poste qui contribue le plus au coût direct total d'une séquence de traitements sur l'horizon de vie du patient atteint de PR.
- **À défaut d'une diminution différenciée importante des prix des molécules mères de bDMARDs** (ou de leurs biosimilaires) ou des tsDMARDs :
 - le biosimilaire de l'étanercept (ou le biosimilaire de l'adalimumab) pourrait être utilisé en première **ligne de traitement biologique (après échec d'un csDMARD)** ;
 - il ne semble pas « efficient » d'utiliser successivement des anti-TNFα dans une perspective économique.
- **Les séquences incluant les tsDMARDs n'apparaissent pas comme stratégies « efficaces » ou « moins coûteuses » dans une analyse de scénario exploratoire.** La robustesse de ce résultat pose problème car il est entaché d'une très forte incertitude liée au niveau de preuve des données utilisées et notamment au faible recul en termes d'efficacité et de tolérance des tsDMARDs comparativement aux biothérapies disponibles.
- Des analyses de scénarios prenant en compte la triple association de csDMARDs (méthotrexate, sulfasalazine et hydroxychloroquine) en choix alternatif à une biothérapie au même stade de la stratégie **ont permis d'identifier la séquence la moins coûteuse se situant sur la frontière d'efficience**. Cependant, le bénéfice médico-économique de la trithérapie reste étroitement lié à la démonstration de la qualité de l'**adhésion thérapeutique**.
- La place des biosimilaires dans le parcours médicamenteux des patients atteints de PR doit être consolidée par des analyses d'impact budgétaire portant sur les

biosimilaires actuels et futurs. **Le renforcement de la mise en concurrence des traitements de fond biologiques et leurs biosimilaires est une opportunité pour générer des économies tout en garantissant un accès à des traitements de qualité à de nombreux patients.**

- **Une actualisation de cette évaluation économique s'avérera nécessaire après avoir collecté suffisamment de données d'efficacité et de tolérance en vie réelle sur les nouveaux traitements de fond ciblés tsDMARDs, la trithérapie (association de trois traitements de fond synthétiques), les nouveaux et futurs biosimilaires.**

Abréviations et acronymes

ACPA	Anticorps anti-protéines citrullinées
ACR	<i>American College of Rheumatology</i>
ACR20	Amélioration de 20 % de la réponse ACR
ACR50	Amélioration de 50 % de la réponse ACR
ACR70	Amélioration de 70 % de la réponse ACR
AIB	Analyse d'impact budgétaire
AIMS	<i>Arthritis Impact Measurement Scale</i>
ALD	Affection de longue durée
AMO	Assurance maladie obligatoire
ANRS	Agence nationale de recherche sur le sida et les hépatites virales
ANSM	Agence nationale de sécurité du médicament et des produits de santé
ASMR	Amélioration du service médical rendu
ATIH	Agence technique de l'information sur l'hospitalisation
bDMARD	<i>Biological Disease-Modifying Antirheumatic Drugs</i>
BSBR	<i>British Society for Rheumatology</i>
CADTH	<i>Canadian Agency for Drugs and Technologies in Health</i>
CDAI	<i>Clinical Disease Activity Index</i>
csDMARD	<i>Conventional Synthetic Disease-Modifying Antirheumatic Drugs</i>
CEESP	Commission évaluation économique et santé publique
CEPS	Comité économique des produits de santé
CHEC	<i>Consensus on Health Economic Criteria</i>
CODA	<i>Convergence Diagnosis and Output Analysis</i>
CNAM	Caisse nationale de l'assurance maladie
CCP	Peptides cycliques citrullinés
CRP	Protéine C réactive
CRL	Intervalle de crédibilité
CT	Commission de la transparence
CT-P13	Biosimilaire de l'infliximab
DAS 28	<i>Disease Activity Score 28</i>
DAS-CRP	<i>Disease Activity Score C-Reactive Protein</i>
DCIR	Datamart consommation inter-régimes
DMARD	<i>Disease-Modifying Antirheumatic Drugs</i>
DOM	Département d'outre-mer
DREES	Direction de la recherche, des études, de l'évaluation et des statistiques
DSS	Direction de la Sécurité sociale
ECOBIAS	<i>Bias in Economic Evaluation</i>

EI	Événement indésirable
ECR	Essai clinique randomisé
EGB	Échantillon généraliste de bénéficiaires
EMA	<i>European Medicines Agency</i>
EMIR	Échelle de mesure de l'impact de la polyarthrite rhumatoïde
ESPOIR	Étude et suivi des polyarthrites indifférenciées récentes
EULAR	<i>European League Against Rheumatism</i>
EVA	Échelle visuelle analogique
FR	Facteur rhumatoïde
ICER	<i>Institute for Clinical and Economic Review</i>
ISPOR	<i>International Society for Pharmacoeconomics and Outcomes Research</i>
ITT	Population en intention de traiter
i.v.	Administration intraveineuse
GRADE	<i>Grading of Recommendations Assessment, Development and Evaluation</i>
GERS	Groupement pour l'élaboration et la réalisation de statistiques
HAQ	<i>Stanford Health Assessment Questionnaire</i>
HAS	Haute Autorité de Santé
HRQoL	<i>Health-Related Quality of Life</i>
HTAi	<i>Health Technology Assessment International</i>
INSEE	Institut national de la statistique et des études économiques
Intensive cDMARDs	Association de deux ou trois csDMARDs
IMR	Taux de mortalité incidents
IC	Intervalle de confiance
ISPOR	<i>International Society for Pharmacoeconomics and Outcomes Research</i>
KCE	<i>Belgian Health Care Knowledge Centre</i>
LDA	<i>Low Disease Activity</i>
LEEM	Les entreprises du médicament
LFSS	Loi de financement de la Sécurité sociale
MAR	Méta-analyse en réseau
mITT	<i>Modified Intention to Treat</i>
MONO	Monothérapie
MP	Méthylprednisolone
MTX	Méthotrexate
NAD	Nombre d'articulations douloureuses
NHS	<i>National Health Service</i>
NICE	<i>National Institute for Health and Care Excellence</i>
NIRD	Maladies rhumatismales non inflammatoires

NSAIDs	<i>Nonsteroidal Anti-Inflammatory Drugs</i>
ONDAM	Objectif national des dépenses d'assurance maladie
OMS	Organisation mondiale de la santé
OMERACT	<i>Outcome Measures in Rheumatology</i>
PBO	Placebo
PFHT	Prix fabricant hors taxes
PHRC	Programme hospitalier de recherche clinique
PMSI	Programme de médicalisation des systèmes d'information
PPTTC	Prix public toutes taxes comprises
PR	Polyarthrite rhumatoïde
PrI	Intervalle prédictif
QALY	Année de vie ajustée sur la qualité de vie
RAQoL	<i>Rheumatoid Arthritis Quality of Life</i>
RDRC	Ratio différentiel coût-résultat
ROC	<i>Rotation or Change of Biotherapy</i>
s.c.	Administration sous-cutanée
SB2	Biosimilaire de l'infliximab
SB4	Biosimilaire de l'étanercept
ScHARR	<i>School of Health and Related Research</i>
SD	<i>Standard Deviation</i> (Déviation standard)
SDAI	<i>Simplified Disease Activity Index</i>
SE	Standard Error (Erreur type)
SEESP	Service évaluation économique et santé publique
SEM	Service évaluation des médicaments
SFR	Société française de rhumatologie
SMR	Taux de mortalité standardisé
SNIIRAM	Système national d'information inter-régimes de l'assurance maladie
SSZ	Sulfasalazine
STRASS	<i>Spacing of TNF-blocker Injections in Rheumatoid Arthritis Study</i>
TNF	<i>Tumor Necrosis Factor</i>
VS	Vitesse de sédimentation

Introduction générale

La polyarthrite rhumatoïde (PR) est une maladie rhumatismale auto-immune, inflammatoire, chronique, évoluant par poussées, provoquant des gonflements et des douleurs au niveau des articulations périphériques (principalement les poignets, les mains, et les pieds), déformant et détruisant progressivement les articulations. À l'échelle internationale, la prévalence de la PR est estimée entre 0,4 % et 1,3 % (1). En France, elle avait été estimée à 0,31 % en 2001 dans la population âgée de plus de 18 ans (2). Le nombre de personnes en affection de longue durée (ALD) pour PR « grave évolutive » était de 226 640 au 31 décembre 2015 (3). Selon la même source, ce nombre a augmenté de 3,4 % en 2013, de 4,4 % en 2014 et de 4,6 % en 2015. La proportion des femmes atteintes est deux à trois fois plus élevée que celle des hommes (2). La PR altère la qualité de vie du patient, conduit à une incapacité fonctionnelle et à la difficulté d'exercer une activité professionnelle. Dans certaines situations, elle nécessite le recours à la chirurgie (arthroplastie, arthrodèse, synovectomie).

Les recommandations de la Haute Autorité de Santé (HAS), relatives à la prise en charge des patients atteints de PR, datent de 2007. Elles ont été suspendues en 2011 en attendant de leur actualisation¹. En 2014, la Société française de rhumatologie (SFR) (4) a émis des recommandations fondées sur celles de l'*European League Against Rheumatism* (EULAR) de 2013 (Smolen *et al.*, 2014) (5). Les recommandations de l'EULAR ont été actualisées en 2017 (Smolen *et al.*, 2017) (6) et celles de la SFR en 2018 (Daïen *et al.*, 2019) (7). Selon ces recommandations, la prise en charge repose principalement sur le diagnostic précoce et la prescription d'un traitement de fond dit *Disease-Modifying Antirheumatic Drug* (DMARD), afin d'induire une rémission clinique et biologique permettant d'empêcher la progression structurale de la maladie ainsi que la dégradation fonctionnelle à long terme. En attendant l'efficacité du traitement de fond, une corticothérapie peut être proposée en respectant une dose cumulée faible, si possible sur une période maximale de 6 mois.

Ces recommandations soulignent l'importance de mettre en place un suivi régulier du patient et de proposer des adaptations thérapeutiques fréquentes (tous les 1 à 3 mois) tant que l'objectif n'est pas atteint. Ces ajustements s'inscrivent dans une stratégie dynamique de traitement fondée sur un objectif clairement défini « *treat to target* » et un suivi « serré ». En l'absence d'amélioration dans les 3 mois suivant le début du traitement ou si l'objectif thérapeutique n'a pas été atteint à 6 mois, le traitement doit être ajusté (4, 7).

1 En 2011, les dernières recommandations relatives à la prise en charge de la PR ont été suspendues par le Collège de la HAS à la suite d'une analyse qui a décelé des faiblesses de forme dans la gestion des liens d'intérêts.

En première ligne de traitement, le méthotrexate (MTX) est le médicament de fond conventionnel de référence (csDMARD)² de la PR. En cas de contre-indication ou d'intolérance au MTX, le léflunomide ou la sulfasalazine, qui ont fait la preuve de leur efficacité symptomatique et structurale, peuvent être prescrits.

Les années 2000 ont été marquées par l'avènement des traitements de fond biologiques (bDMARD) qui ont modifié les modalités de prise en charge des patients atteints de PR (i.e. prescription nécessitant un bilan pré-thérapeutique poussé et une surveillance de l'activité de la maladie et de la survenue des événements indésirables spécifiques des bDMARDs tels que les infections). En juin et septembre 2017, la Commission de transparence (CT) de la HAS a évalué deux traitements (baricitinib et tofacitinib) représentant une nouvelle classe thérapeutique définie par l'EULAR en 2016 et nommée **traitements de fond synthétiques ciblés** (tsDMARDs). Ces traitements sont des inhibiteurs sélectifs et réversibles des Janus kinases (enzymes impliquées dans la physiopathologie de la PR à travers la transduction de signaux de nombreuses cytokines) et sont administrés par voie orale, alors que les traitements de fond biologiques sont administrés par voie parentérale. L'utilisation des bDMARDs (et plus récemment des tsDMARDs) représente actuellement une alternative thérapeutique en deuxième ligne et plus³ particulièrement chez les patients insuffisamment « répondeurs » au MTX.

Aujourd'hui, l'utilisation des bDMARDs et des tsDMARDs dans la prise en charge de la PR en France soulève les remarques suivantes :

1. les coûts de prise en charge de la PR restent élevés, et ce, même si des baisses de prix peuvent être négociées entre les industriels et le Comité économique des produits de santé (CEPS) lors de la fixation des remises dans le cadre d'accords prix-volume. Si les csDMARDs présentent un coût ne dépassant généralement pas une moyenne annuelle de 1 000 € par patient, le coût moyen annuel par patient des bDMARDs reste supérieur à 10 000 €. Une analyse des données de l'Échantillon généraliste des bénéficiaires (EGB) (8) a mis en évidence un coût moyen annuel de prise en charge d'un patient atteint de PR de 6 404 € en 2010. Ce coût s'explique notamment par les coûts des traitements et des hospitalisations. Stratifié sur le type de traitement, le coût annuel de prise en charge atteint 15 757 € pour les patients recevant un traitement de fond biologique ;

² Dans la littérature, en plus de l'appellation « médicament de fond conventionnel », d'autres termes comme « médicament de fond classique » ou « médicament de fond synthétique » (sDMARD) sont utilisés. Le terme « médicament de fond synthétique conventionnel » (csDMARD) est retenu dans ce document de travail. L'abréviation tsDMARD est retenue pour les nouveaux traitements de fond synthétiques dits « ciblés » comme les inhibiteurs de JAK (baricitinib et tofacitinib).

³ La Commission de transparence a estimé que, bien qu'ayant une AMM chez les patients non précédemment traités par MTX, la prescription d'un bDMARD (en association avec le MTX ou en monothérapie), quel qu'il soit, ne se justifie pas en 1^{re} ligne. La seule situation exceptionnelle, dans laquelle la prescription en 1^{re} ligne d'un bDMARD pourrait être envisagée, serait le cas d'une contre-indication à tous les csDMARDs.

2. avec l'arrivée des biosimilaires, il convient de tenir compte de leur impact sur les coûts de prise en charge et de l'efficacité des séquences de traitements de fond biologiques ;
3. aujourd'hui, la diminution de la mise en invalidité des patients atteints de PR ne peut pas s'expliquer seulement par l'utilisation des bDMARDs. L'amélioration de la prise en charge des patients, fondée sur une initiation rapide et efficace des traitements de fond synthétiques, contribue également à la diminution de la mise en invalidité (9). De même, les effets des bDMARDs sur la diminution du recours aux prothèses articulaires restent à démontrer (10). L'étude récente de Nystad *et al.* (2018) (11) a montré que des patients atteints de PR présentaient des risques faibles de recours à la chirurgie orthopédique après l'introduction des bDMARDs. Ce résultat a permis de différencier l'impact du diagnostic et de la prise en charge précoce de la PR sans que l'on puisse isoler clairement l'impact des bDMARDs de celui des csDMARDs dans le recours à la chirurgie orthopédique ;
4. dans leurs recommandations, l'EULAR (Smolen *et al.*, 2017) (6), la SFR (Gaujoux-Viala *et al.*, 2014) (4) et la mise à jour des recommandations de la SFR fin 2018 (Daïen *et al.*, 2019) (7) préconisent de tenir compte du coût de la PR ainsi que de ses traitements dans les stratégies de prise en charge thérapeutique, aussi bien à l'échelle individuelle que sociétale. Par exemple, selon Daïen *et al.* (2019) (7), si un patient est en rémission persistante sans corticoïdes, il est possible d'envisager de diminuer la fréquence d'administration ou de réduire la posologie des bDMARDs ou des tsDMARDs (traitements de fond synthétiques ciblés), notamment si ce traitement est associé à un csDMARD.

Dans ce contexte, la Direction de la Sécurité sociale (DSS) s'interroge sur le poids économique supporté par la collectivité pour ces traitements. Ainsi, une évaluation économique des bDMARDs⁴ dans la prise en charge médicamenteuse de la PR s'avère très utile pour :

- identifier la (les) stratégie(s) de prise en charge la (les) plus efficace(s) ;
- orienter les choix thérapeutiques sur des critères à la fois cliniques et économiques et contribuer à l'amélioration des pratiques, dans un contexte d'optimisation budgétaire.

Depuis les années 2000, une littérature assez abondante a porté sur l'analyse médico-économique des bDMARDs dans la prise en charge de la PR. Des évaluations médico-économiques ont également été réalisées par des agences d'évaluation des technologies de

⁴ Une actualisation de cette évaluation économique s'avérera nécessaire après avoir collecté des données d'efficacité et de tolérance en vie réelle sur les nouveaux traitements de fond ciblés.

santé, le NICE en particulier ; la plus récente a été publiée en avril 2016 (12) et concerne l'adalimumab, l'étanercept, l'infliximab, le certolizumab pegol, le golimumab, le tocilizumab et l'abatacept qui ont été comparés aux csDMARDs. Aux États-Unis, l'*Institute for Clinical and Economic Review* (ICER) (13) a publié son évaluation en avril 2017. Ces rapports ont principalement porté sur l'évaluation économique des bDMARDs en première ligne de traitement biologique, comparés également aux csDMARDs. En France, quelques travaux ont porté sur les aspects médico-économiques associés à la prise en charge de la PR. Certaines études ont abordé le poids économique des bDMARDs et leurs coûts (14-17). D'autres ont proposé des analyses coût-efficacité portant sur l'adalimumab, le rituximab et l'infliximab (Sany *et al.*, 2009 (18) ; Saraux *et al.*, 2010 (19)) ou des analyses évaluant le ratio coût-utilité d'une stratégie d'espacement des doses des anti-TNF administrés par voie sous-cutanée dans l'essai clinique français « *Spacing of TNF-blocker Injections in Rheumatoid Arthritis Study* » STRASS (20).

L'analyse économique proposée dans ce rapport ne répond pas uniquement à la rareté des études coût-efficacité/utilité portant sur la prise en charge de la PR en France, mais se distingue des évaluations des autres agences d'évaluation des technologies de santé (ex. : NICE) dans la mesure où la problématique centrale concerne l'analyse de l'efficacité des séquences de traitements de fond biologiques (stratégies de comparaison) et non de chaque médicament considéré à un instant donné du parcours médicamenteux du patient. Selon les recommandations cliniques actuelles et les données disponibles en vie réelle, l'évaluation économique proposée devrait contribuer à identifier des perspectives quant à la pertinence et l'efficacité du parcours de prise en charge médicamenteux dans la PR.

La PR est une maladie pour laquelle le recul sur l'évolution des pratiques de prise en charge cumulé à l'intégration rapide des nouveaux traitements sur le marché des soins exige une actualisation continue de la modélisation économique, tant au niveau conceptuel qu'à l'échelle technique (21). Par ailleurs, la réalisation d'études coût-efficacité/utilité portant sur la PR demeure confrontée aux difficultés suivantes (22-24) :

- le nombre d'essais cliniques comparant directement les bDMARDs entre eux est faible. Par conséquent, le recours aux méthodes de comparaisons indirectes s'impose ;
- l'objectif d'une prise en charge en vie réelle est la rémission ou l'atteinte d'une faible activité de la maladie. L'Agence européenne du médicament (*European Medicines Agency* (EMA)) recommande de considérer un critère principal défini en termes de rémission selon les critères de l'EULAR (DAS<2,6) ou *via* les critères stricts EULAR-

ACR (*American College of Rheumatology*) (25). Néanmoins, il existe peu d'essais cliniques utilisant la rémission comme critère principal ou secondaire ;

- l'analyse des résultats de santé de la PR, définis en termes d'années de vie ajustées sur la qualité de vie (QALY), ainsi que le choix d'un questionnaire de qualité de vie utilisé dans les essais cliniques pour mesurer le QALY ne font pas l'objet d'un consensus. Dans le contexte français, Saraux *et al.* (2010) (19) ont proposé des résultats de santé définis en fonction de l'activité de la PR (ex. : rémission ou temps de faible activité de la PR) dans le but de pallier les insuffisances du QALY en tant que résultat de santé de la PR ;
- les travaux de Goujoux-Viala *et al.* (2013) (26) sur les PR débutantes ont montré que si l'état de santé du patient était détérioré, le niveau d'utilité estimé était systématiquement plus faible avec l'EQ-5D qu'avec le SF-6D ;
- l'attention est attirée aussi sur la pertinence des méthodes utilisées pour estimer les utilités en explorant la relation entre l'EQ-5D et les instruments renseignant sur l'incapacité fonctionnelle et l'intensité de la douleur (12). Ces dernières se fondent sur des modèles mathématiques reposant notamment sur la corrélation entre l'EQ-5D et le critère d'incapacité fonctionnelle (*Health Assessment Questionnaire*, HAQ). Bien que ces techniques soient économétriquement pertinentes (Hernandez *et al.*, 2014) (27), le rôle et le poids très importants qu'elles attribuent au HAQ dans l'estimation de la qualité de vie ne sont pas transposables à tous les contextes de prise en charge de la PR ;
- la prédiction de la progression de la PR est multifactorielle et complexe. Il manque souvent des données de long terme décrivant la progression de l'activité de la PR, sa sévérité et sa mortalité. Par exemple, il n'est pas toujours facile de colliger des données de long terme renseignant à la fois sur l'activité de la PR et l'état structural articulaire ;
- aujourd'hui, il est difficile d'apparier des données issues des bases de données médico-administratives (riches en informations agrégées sur les coûts de prise en charge des patients) avec des données provenant de registres ou de cohortes (colligeant des informations cliniques décrivant le patient). À l'instar des rapports réalisés par les organismes étrangers d'évaluation des technologies de santé tels que le NICE et l'ICER, l'évaluation économique réalisée dans ce rapport n'utilise pas de données appariant les données cliniques et médico-administratives. L'appariement de ces deux sources de données est théoriquement possible, mais la mise en œuvre des protocoles et des procédures juridiques et techniques d'appariement (ex. : compatibilité, analyse de sensibilité des algorithmes d'identification des formes de PR)

exige la disponibilité de ressources techniques et cliniques et implique un investissement important dans les phases préalables à la production de ce type de données, dont le temps de mise en place dépasse les délais de réalisation de ce projet. À cet égard, un projet d'élaboration de ce type de données entre la SFR (promoteur de la cohorte française « Étude et suivi des polyarthrites indifférenciées récentes », ES-POIR) et le SNDS (Système national des données de santé) a récemment été mis en place. Les données ne sont pas disponibles au moment de la rédaction de ce rapport.

- L'objectif principal de ce projet consiste à évaluer l'efficacité des séquences de bDMARDs⁵ et des tsDMARDs des patients atteints de PR active modérée à sévère.
- Cette évaluation est fondée sur un nouveau modèle économique dans la prise en charge de la PR, appelé PR France.
- Le modèle PR France est fondé sur une méthodologie appropriée à la prise en charge de cette pathologie en France et est en accord avec les principales conclusions des avis de la Commission de transparence (CT) de la HAS portant sur les bDMARDs et des recommandations récentes des sociétés de rhumatologie française (SFR) et européenne (EULAR), à l'exception des stratégies « *treat to target* » en l'absence de données suffisantes sur la rémission.

La structure de ce rapport reflète, dans une certaine mesure, l'état actuel des connaissances cliniques et de modélisation économique sur la prise en charge médicamenteuse de la PR. Premièrement, les objectifs, les enjeux de l'évaluation et l'approche méthodologique de l'évaluation sont décrits. Deuxièmement, l'approche adoptée dans la revue systématique de la littérature clinique et la réalisation des méta-analyses en réseau est définie. Troisièmement, les résultats de cette revue systématique et de ces méta-analyses sont présentés. Quatrièmement, une revue systématique de la littérature économique portant sur les analyses coût-efficacité/utilité des traitements de fond biologiques (bDMARDs) est présentée. Le chapitre propose une synthèse de la littérature économique existante sur la population éligible au modèle économique en identifiant les aspects de l'évaluation économique communs à la plupart des études et ceux suscitant des interrogations

⁵ Les traitements de fond synthétiques ciblés (tofacitinib et baricitinib) et le sarilumab (inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6) n'ont pas pu être intégrés dans les méta-analyses réalisées puisque ces produits ont été évalués par la CT fin 2017 et début 2018. Cependant, ils ont fait l'objet d'une synthèse incluant des évaluations cliniques de la CT et du NICE, de revues systématiques et de méta-analyses en réseau récentes. La revue systématique de la littérature économique a intégré les études récentes portant sur les traitements de fond synthétiques ciblés. L'évaluation économique a également intégré ces produits dans une analyse de scénario.

méthodologiques. Cinquièmement, le modèle économique « PR France » et les données utilisées dans l'évaluation économique des traitements de fond biologiques dans la prise en charge de la PR en France sont décrits. Tout d'abord, le modèle PR France (sa structure, son implémentation technique et son processus de validation) est décrit. Ensuite, les choix structurants de l'évaluation économique et les sources de données mobilisées dans l'évaluation économique, incluant les données d'efficacité et de tolérance, sont présentés et discutés. Sixièmement, les résultats de l'évaluation économique (analyse de référence et analyses de sensibilité déterministes et probabilistes) sont présentés.

En conclusion générale, les principaux résultats d'efficacité comparative et de l'évaluation économique sont synthétisés. Leur prise en compte en tant qu'outil d'aide à la décision économique est discutée.

Chapitre 1. Évaluation des traitements de fond biologiques dans la prise en charge de la PR : objectifs, contexte, méthodes

L'analyse économique des traitements de fond biologiques (bDMARDs) dans la prise en charge de la PR, proposée dans ce rapport, représente la première évaluation réalisée par la HAS en France sur ces produits. Ce chapitre a pour objectif de présenter :

1. les principaux objectifs et enjeux de l'évaluation économique des traitements de fond biologiques dans la prise en charge de la PR en France ;
2. l'approche méthodologique adoptée pour sa réalisation. Elle concerne les travaux portant sur l'efficacité comparative au même titre que l'évaluation économique des bDMARDs ;
3. les dimensions épidémiologiques et cliniques contextualisant le projet.

1. Objectifs et enjeux de l'évaluation

1.1 Saisine

Le projet d'évaluation répond à une demande de la Direction de la Sécurité sociale (DSS), en concertation avec le Comité économique des produits de santé (CEPS) et s'inscrit dans le programme de travail 2016-2018 de la HAS. L'intitulé de la saisine est « *l'évaluation médico-économique des biothérapies dans la prise en charge de la polyarthrite rhumatoïde (PR)* » (Annexe A1-1). Une réunion de cadrage entre la DSS et la HAS a eu lieu le 10 juillet 2015. Elle a permis de préciser les attentes du demandeur :

1. disposer d'une évaluation économique des traitements de fond biologiques dans la prise en charge de la PR afin d'identifier la (les) stratégie(s) de prise en charge la (les) plus efficiente(s) ;
2. orienter les choix thérapeutiques sur des critères à la fois cliniques et économiques et contribuer à l'amélioration des pratiques, dans un contexte d'optimisation budgétaire.

1.2 Feuille de route

Une feuille de route a été proposée à la Commission d'évaluation économique et de santé publique (CEESP) afin d'évaluer la pertinence et la faisabilité de l'évaluation ainsi que la disponibilité de la littérature, de définir le périmètre de l'évaluation, la méthodologie et le calendrier prévisionnel de ce projet. Les recommandations de la CEESP ont mis l'accent sur la nécessité que l'évaluation envisagée réponde à deux besoins :

1. l'importance de réaliser une évaluation économique complète des traitements biologiques dans la prise en charge de la PR. Les évaluations économiques portant sur la PR en France ne sont pas nombreuses, et la plupart des études réalisées avaient principalement pour objectif de décrire les coûts des bDMARDs (ex. : Fautrel *et al.*, 2016) (8) ou de comparer les utilités correspondant aux instruments décrivant la qualité de vie (ex. : l'EQ-5D ou le SF-6D) (28) ;
2. l'intérêt d'une analyse des stratégies de prise en charge de la PR, définies en termes de séquences de traitements, qui contribuerait à décrire dans la mesure du possible le parcours efficient de traitement de la PR et qui nourrirait la mise à jour des recommandations de prise en charge de la PR de la HAS.

Cette feuille de route a été validée en novembre 2015 par le Collège de la HAS (29).

1.3 Objectif et enjeux de l'évaluation

En lien avec les questions d'évaluation définies dans la feuille de route, l'objectif principal de ce projet est de réaliser une évaluation médico-économique des stratégies de prise en charge, incluant les traitements de fond biologiques et synthétiques dans le traitement de la PR. Dans cette évaluation :

- les stratégies de prise en charge sont définies en termes de séquences de traitements selon leur classe ;
- l'intégration des biosimilaires et des traitements de fond synthétiques ciblés arrivés plus récemment sur le marché, est conditionnée par la disponibilité des données cliniques et économiques.

En parallèle des travaux menés par la CNAM dans le cadre de son plan de maîtrise médicalisée sur le bon usage des traitements de fond biologiques, l'évaluation médico-économique proposée constitue un outil d'aide à la décision utile au décideur public et, dans un deuxième temps, aux prescripteurs, tant le poids financier actuel et futur de ces thérapies est important pour l'assurance maladie.

En l'absence de recommandations récentes de la HAS sur la prise en charge de la PR, l'évaluation économique des stratégies de traitement fondées sur des séquences de traitements de fond repose sur différentes sources permettant de décrire la prise en charge de la PR en France :

- une revue systématique de littérature clinique et des méta-analyses en réseau portant sur les bDMARDs. Cette source de données a fait l'objet d'une collaboration entre la HAS et *SchARR* (*School of Health and Related Research*, Université de Sheffield, Royaume-Uni) ;
- les données françaises disponibles de cohortes/registres en vie réelle, notamment les données de la cohorte « Étude et suivi des polyarthrites indifférenciées récentes » (ESPOIR) qui ont fait l'objet d'une convention entre la HAS et la SFR, signée en septembre 2017 ;
- les avis de la CT sur les bDMARDs et (récemment) sur les tsDMARDs qui définissent la place des différents traitements dans la stratégie thérapeutique (annexe A1-2) ;
- les recommandations des sociétés savantes les plus compatibles avec les avis de la CT et le contexte français de la prise en charge de la PR, telles que la Ligue européenne luttant contre les maladies rhumatismales « *European League Against Rheumatism* » (EULAR (5, 6) et la Société française de rhumatologie (SFR) (7) ;
- les échanges avec un groupe de travail multidisciplinaire (annexe A1-3) qui sont cohérents avec les avis de la CT et permettant de spécifier le modèle économique retenu (modèle PR France) pour évaluer la prise en charge de la PR ;
- les données issues de bases médico-administratives à l'instar des données sur les prix des médicaments de la Caisse nationale d'assurance maladie (CNAM).

1.4 Champs exclus de l'évaluation

Les aspects suivants ont été exclus de l'évaluation économique des traitements de fond biologiques dans la prise en charge de la PR en France :

- des analyses d'impact budgétaire (AIB) n'ont pas été envisagées. En effet, la saisine n'avait pas pour objectif de quantifier l'impact financier rétrospectif et/ou prospectif de l'introduction d'un bDMARD dans la prise en charge de la PR ;
- afin de limiter le périmètre d'analyse et en accord avec le groupe de travail multidisciplinaire, l'analyse principale de l'évaluation économique concerne exclusivement les patients souffrant d'une PR active (modérée à sévère) et qui présentaient une réponse inadéquate ou une tolérance insuffisante aux traitements de fond synthétiques

conventionnels (csDMARDs). **Une évaluation concernant les patients présentant une PR sévère et naïfs du MTX était envisagée lors du démarrage du projet et a fait l'objet d'une méta-analyse en réseau. Cependant, la CT, dans ses avis récents depuis 2016, estime que, bien qu'ayant une AMM chez les patients non précédemment traités par MTX, la prescription d'un bDMARD (en association avec le MTX ou en monothérapie), quel qu'il soit, ne se justifie pas en 1^{re} ligne. Le seul cas, dans lequel la prescription en 1^{re} ligne d'un bDMARD pourrait être envisagée, serait celui d'une contre-indication à tous les DMARDs conventionnels. L'évaluation économique portant sur les patients présentant une PR sévère et naïfs du MTX n'a donc pas été réalisée ;**

- l'évaluation économique porte seulement sur les coûts directs de la prise en charge de la PR. Elle est conforme aux recommandations actuelles de la HAS dans son guide méthodologique sur l'évaluation économique des produits de santé (30). En dépit de leur importance, les coûts indirects ne seront pas intégrés dans une analyse complémentaire en raison de l'absence de recommandations spécifiques portant sur leur mesure dans le guide actuel de la HAS (30) et de l'absence d'un consensus sur les méthodes de mesure des coûts indirects auprès des chercheurs et des sociétés savantes en économie de la santé en France (Verstappen *et al.*, 2015) (31) ;
- conformément à l'objectif de la saisine, **l'évaluation économique n'intègre pas la prise en charge non médicamenteuse de la PR.** Celle-ci est considérée comme une option complémentaire aux traitements pharmacologiques. Les interventions non médicamenteuses ont pour objectif de restreindre les conséquences fonctionnelles, sociales et professionnelles liées à la douleur, à la fatigue et aux destructions articulaires progressives, afin de soulager le patient et de lui donner une qualité de vie optimale malgré sa maladie chronique. Tout comme celles de la SFR et l'EULAR, les recommandations de la HAS (10) soulignent que la stratégie thérapeutique non médicamenteuse doit être définie et adaptée au même titre que le traitement pharmacologique, et prendre en compte l'avis du patient. Une prise en charge non médicamenteuse nécessite une évaluation clinique préalable de la douleur, des déficiences et incapacités fonctionnelles, de l'état psychologique et de l'environnement social et professionnel du patient ;
- les doses et les durées de traitement utilisées dans le modèle économique (cf. chapitres 5 et 6) ont été collectées à partir des résumés des caractéristiques des produits (RCP). Ce choix est compatible avec la démarche adoptée dans le cadre du modèle PR France qui propose une modélisation selon les bonnes pratiques de prise en charge fondées sur les recommandations européennes et françaises. Les rédu-

tions des doses et des durées de traitement en vie réelle (ex. : données extraites à partir du Système national d'information inter-régimes de l'assurance maladie (SNIIRAM) ou de l'Échantillon généraliste des bénéficiaires (EGB)) ont été également envisagées pour tenir compte des stratégies ciblées (ex. : les stratégies « *treat to target* » T2T), **mais l'absence de données appariées à partir de ces bases avec des cohortes décrivant des critères cliniques de la PR, tels que les variables renseignant sur l'activité de la maladie (ex. : DAS 28) à 3 mois, n'a pas permis de modéliser ce type de stratégies. Lorsque le projet de la SFR portant sur l'appariement des données de la cohorte de données françaises (ESPOIR) avec celles du SNIIRAM aura été réalisé, ces stratégies pourront être incluses dans une version future du modèle PR France en mettant à jour les équations portant sur les doses, les visites et le changement de traitement.**

L'évaluation économique réalisée dans ce rapport est centrée sur l'analyse coût-efficacité/utilité des stratégies intégrant les séquences de traitements de fond synthétiques et biologiques et les biosimilaires (et les traitements de fond synthétiques ciblés) selon la disponibilité des données. La population d'entrée du modèle : elle concerne l'indication de la PR la plus fréquente et adaptée au contexte actuel de la prise en charge de la PR : patients atteints d'une PR active (modérée à sévère) et insuffisamment répondeurs (ou dont la tolérance est insuffisante) au méthotrexate (MTX).

2. Méthodologie de l'évaluation et processus d'assurance qualité

2.1. Méthodes d'analyse

L'approche méthodologique d'évaluation adoptée dans ce projet est fondée sur les méthodes décrites ci-après.

2.1.1. Revue systématique de littérature clinique et réalisation de méta-analyses en réseau

Afin d'évaluer l'efficacité comparative des traitements de fond biologiques, une revue systématique de la littérature et des méta-analyses en réseau portant sur les principales indications de la PR ont été menées. Ces travaux ont été réalisés dans le cadre d'une collaboration entre la HAS et l'équipe SchARR (*School of Health and Related Research*) de l'université de Sheffield. Ils ont porté sur l'indication principale relative aux traitements de fond biologiques (bDMARDs) prescrits dans la prise en charge de la PR en France. Cette indication concernait les patients atteints d'une PR active modérée à sévère, qui avaient été traités par des traitements de fond synthétiques (csDMARDs) tels que le méthotrexate (MTX) et qui présentaient une réponse inadéquate ou une intolérance insuffisante associée à ces traitements. La population incluant les patients atteints d'une PR sévère non traités par des csDMARDs (les patients naïfs de MTX) est présentée dans l'annexe A3.11.

La revue systématique de la littérature présentée dans ce rapport est une mise à jour d'une revue systématique réalisée par le NICE en 2013. La recherche de la littérature a interrogé plusieurs bases de données électroniques et a porté sur la période de mai 2013 à avril 2017 (période postérieure à la recherche documentaire de l'évaluation du NICE réalisée en mai 2013). La description du protocole de la revue systématique (processus d'interrogation des bases de données électroniques, critères d'éligibilité et grille qualitative d'appréciation des biais afférant aux études cliniques sélectionnées) et l'approche méthodologique retenue pour la réalisation des méta-analyses en réseau sont explicitées dans le chapitre n° 3.

Il est important de souligner que les récents inhibiteurs des Janus kinases (JAKs) tels que le baricitinib, le tofacitinib et le sarilumab (le nouvel inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6) n'ont pas été pris en compte dans les méta-analyses en réseau, car ils n'avaient pas été encore autorisés sur le marché au moment de la réalisation de la revue systématique et des méta-analyses en réseau (ou étaient en cours d'évaluation par la CT). Des données portant sur le tofacitinib existaient mais ne respectaient pas les critères d'inclusion (cf. justification des études exclues). Ces dernières ont seulement été incluses dans les analyses de sensibilité des méta-analyses en réseau. Afin de décrire le

positionnement de ces nouveaux traitements par rapport aux classes thérapeutiques des traitements de fond biologiques, une synthèse complémentaire sur les nouveaux tsDMARDs et le nouvel inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 est présentée dans le chapitre 3. Compte tenu de l'arrivée récente de ces produits et en attendant une actualisation de la revue systématique décrite plus haut, la synthèse proposée est fondée sur un résumé des avis de la CT relatifs aux trois produits et des résultats des méta-analyses les plus récentes intégrant ces produits, à l'instar des méta-analyses de l'Agence canadienne de l'évaluation des technologies de santé (CADTH) (32).

2.1.2. Revue systématique de la littérature économique

Une revue systématique de la littérature sur les analyses coût-efficacité des traitements de fond biologiques dans la prise en charge de la PR a été réalisée par la HAS. Ces travaux couvrent la période de septembre 2000 à juin 2018. La revue des études coût-efficacité/utilité a été limitée aux études publiées en anglais ou en français. Toutefois, le périmètre de la question de modélisation s'est étendu à 16 pays répartis sur trois continents (Amérique, Europe et Asie). D'une manière générale, la revue de la littérature économique s'articule autour de trois axes :

- une dimension **principale et spécifique** à la modélisation pharmaco-économique (coût-efficacité/utilité) portant sur les bDMARDs qui vise à répondre à la question suivante : quels sont les types de modèles, choix structurants, critères de santé, résultats et analyses de sensibilité présentés dans la littérature ? En particulier, la question suivante a été explorée : **les séquences de traitements et les événements indésirables spécifiques à la PR ont-ils été analysés dans la littérature ?**
- une dimension **complémentaire** et spécifique aux résultats de santé (ex. : qualité de vie) complétant la modélisation pharmaco-économique et répondant à la question : quels sont les aspects méthodologiques liés aux résultats de santé portant sur les traitements de fond biologiques ? Cette question **privilégie notamment les éléments méthodologiques relatifs aux études françaises portant sur les coûts et les questionnaires portant sur la qualité de vie ;**
- une dimension complémentaire portant sur les aspects contextuels épidémiologiques et cliniques (ex. : **les études portant sur la mortalité liée à la PR**).

Le protocole, le diagramme de sélection des études et l'appréciation de leur qualité méthodologique sont présentés dans le chapitre n° 4. La stratégie de recherche documentaire est présentée dans l'annexe A4-1.

2.1.3. Réalisation d'un modèle économique : modèle PR France

Un nouveau modèle pharmaco-économique consacré à la prise en charge médicamenteuse de la PR en France a été réalisé en partenariat avec le Pr Jaime Caro (Université McGill et *London School of Economics*). Ce modèle appelé « PR France » diffère des modèles actuellement disponibles dans la prise en charge de la PR (ex. : le modèle de l'agence britannique de l'évaluation des technologies de santé, *National Institute for Health and Care Excellence*, NICE (12) ; le modèle de l'organisme américain à but non lucratif de l'évaluation des technologies de santé, *Institute for Clinical and Economic Review*, ICER (13)). PR France est construit autour d'une structure appropriée à la prise en charge en France, en accord avec les principales conclusions des avis de la Commission de transparence (CT) de la HAS et des recommandations des sociétés savantes française (SFR) et européenne (EULAR). La structure du modèle PR France est présentée dans le chapitre n° 5. Pour construire ce nouveau modèle, les démarches suivantes ont été entreprises :

1. une spécification selon un modèle « à conditions et événements discrètement intégrés » (DICE, *Discretely Integrated Condition Event*) a été choisie. Cette approche de micro-simulation est fondée sur des travaux récents (Caro 2016 (33) ; Caro & Moeller, 2018 (34)). Le modèle PR France se fonde sur les réponses cliniques au sens des critères EULAR (cf. section 4.2). Les estimations ponctuelles des réponses EULAR ont été extraites à partir des méta-analyses en réseau réalisées par la HAS & SchARR en 2017 (cf. chapitres n° 3 et n° 4) ;
2. la réalisation d'une étude portant sur les données en vie réelle à partir de la cohorte française ESPOIR (Étude et suivi des polyarthrites indifférenciées récentes). Le choix de cette cohorte est fondé sur son *design* longitudinal et sur le fait qu'elle collecte des données sur l'activité de la maladie, l'incapacité fonctionnelle et la qualité de vie (Combe *et al.*, 2014) (35). Les analyses statistiques portant sur les données de cette cohorte selon les caractéristiques de la population d'analyse de l'évaluation ont permis de sélectionner un échantillon de patients de la population d'analyse et de créer les principales équations du modèle PR France (ex. : l'équation du changement de ligne de traitement à la suite d'une réponse inadéquate ou d'une tolérance insuffisante à un bDMARD, les équations de l'évolution de l'activité de la maladie et des scores d'utilité stratifiés sur l'activité de la PR et l'incapacité fonctionnelle liée à la PR). Le protocole de cette étude est décrit dans le chapitre n° 5 ;
3. un algorithme de changement des traitements (« *switch* ») et un calendrier décrivant l'utilisation des ressources selon les recommandations et les pratiques en vie réelle ont été réalisés au sein de la HAS et validés par les rhumatologues du groupe de travail

PR. Cet algorithme, décrit à l'aide du diagramme du modèle PR France, est présenté dans le chapitre n° 5 ;

4. une étude portant sur l'identification des profils des événements indésirables sévères et/ou graves spécifiques aux bDMARDs (selon leur classe thérapeutique) a été réalisée au sein de la HAS. Un protocole d'estimation de la fréquence des événements indésirables sévères et/ou graves entraînant des hospitalisations a été réalisé en réconciliant plusieurs sources d'informations : avis de la CT, les résumés des caractéristiques des produits (RCP), les rapports de l'Agence européenne du médicament (EMA) sur les médicaments évalués (*European public assessment report* (EPAR)), et la base de données électronique de l'EMA portant sur les événements indésirables des produits EudraVigilance. Cet algorithme est présenté dans le chapitre n° 5.

2.2. Mise en œuvre d'un processus d'assurance qualité

Le processus d'assurance qualité mis en œuvre dans le cadre de ce projet est fondé sur :

1. l'utilisation de l'approche « *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* » (PRISMA) (Moher *et al.*, 2009) (36) dans le cadre de la revue systématique et l'utilisation de méthodologies récentes et validées dans la réalisation des méta-analyses en réseau (Dias *et al.*, 2013) (37) ;
2. la validation des revues systématiques et des méta-analyses en réseau par des lecteurs indépendants ;
3. la création d'une grille de reporting spécifique à l'évaluation coût-efficacité de la PR à partir de la grille « *Consolidated Health Economic Evaluation Reporting Standards* » (CHEERS) (Husereau *et al.*, 2013) (38) et une grille d'analyse critique simplifiée combinant des items de la grille de Drummond (Drummond *et al.*, 2005) (39) et de la grille de « *Outcome Measures in Rheumatology* » (OMERACT) (Gabriel *et al.*, 2002) (40) ;
4. la double vérification indépendante des entrées du modèle sur les coûts, les événements indésirables et les autres résultats de santé (scores d'utilité calculés à partir de l'EQ-5D, le SF-6D, les survies estimées), ainsi que les entrées et les différentes équations du modèle PR extraites à partir de la cohorte française ESPOIR (cf. chapitre n° 5) ;
5. la méthodologie de validation du modèle PR France (ex. : validité intuitive, validation interne). Ces informations sont documentées dans le chapitre n° 6.

2.3. Validation avec un groupe de travail scientifique, des experts internationaux et les instances de la HAS

En plus des méthodes décrites dans les paragraphes précédents, le groupe de travail PR a participé à la validation de la méthodologie du rapport (méta-analyses, modèle PR France et analyses des coûts) et a contribué à l'élaboration du modèle PR France. Le travail présenté dans ce rapport a également été enrichi à partir d'échanges avec la CEESP et avec des experts spécialisés dans l'évaluation économique.

Les résultats ont été présentés à la CEESP pour discussion. Ce document a également fait l'objet d'une relecture par un groupe d'experts sur la période du 11 décembre 2018 au 15 janvier 2019. Ce groupe de relecture comprenait des experts cliniciens (rhumatologues, médecins généralistes), des économistes et des statisticiens. La liste complète des membres du groupe de lecture est donnée dans l'annexe A1-3.

Des experts internationaux ont également échangé avec le coordinateur du projet sur divers aspects méthodologiques :

1. le Pr Jaime Caro (Université McGill, *London School of Economics*) a collaboré :
 - à la réalisation, dans le cadre d'une convention de partenariat avec la HAS, du modèle PR France. Il a mis en place une modélisation fondée sur la simulation DICE,
 - à l'élaboration du protocole d'analyse de la cohorte ESPOIR et à la présentation des documents techniques destinés au GT PR.

Il a également examiné, d'un point de vue technique, l'interprétation des résultats du modèle PR France ;

2. le Pr Matthew D. Stevenson et son équipe (Université de Sheffield) ont participé à la validation de la méta-analyse en réseau (conception, méthodes et interprétation) ;
3. le Pr Pelham Barton (Université de Birmingham) a vérifié la première version du modèle PR France ;
4. le Dr Jörgen Möller (Université de Lund, Suède et Evidera, Londres) a réalisé la validation interne et indépendante du modèle.

2.4. Calendrier d'élaboration du document

Calendrier	Échéance
Présentation de la feuille de route du projet à la CEESP	20 septembre 2016
1 ^{re} réunion du groupe de travail PR	1 ^{er} juillet 2016
Réalisation des méta-analyses en réseau	15 juin 2017
2 ^e réunion du groupe de travail PR	21 juin 2017
Acquisition de la cohorte de patients ESPOIR	3 septembre 2017
Présentation de la 1 ^{re} version du modèle PR France	12 février 2018
3 ^e réunion du groupe de travail PR	17 juillet 2018
1 ^{re} version du rapport	26 octobre 2018
Relecture (Groupe de travail PR)	20 novembre 2018
Relecture externe	12 décembre 2018 – 15 janvier 2019
Présentation CEESP	12 février 2019
Soumission au Collège de la HAS	13 mars 2019
Diffusion du document	16 avril 2019

3. Contexte épidémiologique de l'évaluation

3.1. Prévalence de la PR

3.1.1. Estimation de la prévalence de la PR en France

En France, l'estimation de la prévalence de la PR repose sur une étude assez ancienne (Guillemin *et al.*, 2005) (2). En 2001, elle a été estimée à 0,31 % [IC 95 % = 0,18 - 0,48 %] dans la population âgée de plus de 18 ans. La proportion des femmes touchées est deux à trois fois plus élevée que celle des hommes. Après 60 ans, ce ratio diminue et la prévalence est comparable entre les deux sexes, les hommes étant atteints plus tardivement.

En se fondant sur cette estimation de la prévalence et en considérant la population française métropolitaine estimée au 1^{er} janvier 2017 par l'Insee (67 millions)⁶, le nombre d'individus présentant une PR serait d'environ 207 700 personnes.

Une estimation plus récente a pu être réalisée à l'aide des données de la CNAM portant sur le nombre de personnes en ALD (affection de longue durée) pour « polyarthrite rhumatoïde grave évolutive » en France. Le nombre de patients en ALD pour PR grave évolutive était, au 31 décembre 2015, de 226 640 (3). En 2013, 2014 et 2015, des augmentations respectivement de 3,4 %, 4,4 % et 4,6 % du nombre des patients pris en charge pour cette ALD ont été observées. En supposant que le nombre de personnes en ALD pour PR ait continué à augmenter à un rythme d'environ 5 % par an (41), le nombre de personnes en ALD pour PR serait d'environ 237 972 en 2016 et 249 871 en 2017. En considérant que les données de la CNAM couvrent 88 % de la population française, on pourrait estimer à 240 000 le nombre de patients atteints de PR évolutive grave en France en 2017.

3.1.2. Estimation de la prévalence à l'échelle internationale

À l'échelle internationale, la prévalence de la PR est estimée entre 0,3 et 1 % : selon une méta-analyse fondée sur neuf études d'incidence, dix-sept études de prévalence et deux études estimant à la fois l'incidence et la prévalence (42), la prévalence médiane était de 0,33 % (variant de 0,31 % à 0,5 %) pour les pays d'Europe du Sud, de 0,5 % (variant de 0,44 % à 0,8 %) pour les pays d'Europe du Nord.

En Angleterre et au Pays de Galles, selon Stevenson *et al.* (2016) (12), environ 400 000 personnes étaient atteintes de la PR avec environ 10 000 cas incidents par an. En 2002, la PR était plus fréquente chez les femmes (1,16 %) que chez les hommes (0,44 %). La majorité des PR est diagnostiquée chez des patients âgés de 40 à 80 ans.

⁶ Insee. Bilan démographique 2017. <https://www.insee.fr/fr/statistiques/3305173>.

Aux États-Unis (USA), selon l'évaluation de l'ICER de 2017 (13), entre 1,3 et 1,8 million de patients étaient atteints de PR, soit une prévalence de l'ordre de 0,6 % de la population américaine. Une étude réalisée par *Rochester Epidemiology Project* (43) montre qu'une diminution régulière de la prévalence de la PR aux USA a été observée sur la période 1955-1994. Néanmoins, en poursuivant la même analyse de l'évolution de l'incidence de la PR sur la période 1995-2007, l'étude de Myasoedova *et al.* (2010) (44) semble montrer une légère augmentation de l'incidence de la PR chez les femmes jusqu'en 2007, accompagnée d'une augmentation de la prévalence.

Au Canada, environ 1 % de la population est atteint de la PR (Rapport surveillance de la polyarthrite rhumatoïde, 2017) (45). Une étude fondée sur un suivi de 15 ans de 1996 à 2010, réalisée sur une population de la province de l'Ontario (46), a montré que la prévalence de la PR a presque doublé (passant de 0,49 % en 1996 à 0,9 % en 2010). Cette tendance a été confortée par les résultats d'une autre étude populationnelle rétrospective réalisée au sein de l'Institut de santé publique du Québec (Rapport surveillance de la polyarthrite rhumatoïde, 2017) ; elle a montré qu'entre 2000-2001 et 2014-2015, la prévalence ajustée de la PR a augmenté de 0,5 % à 0,8 %. Il est important de noter que cette augmentation de la prévalence est liée en partie à un changement de définition : nouveaux auto-Ac (ACPA) et nouveaux critères de classification permettant un diagnostic plus précoce de la maladie.

Les études internationales portant sur l'évolution de la prévalence/incidence de la PR reflètent une relative disparité selon les pays mais ne semblent pas infirmer une augmentation de la prévalence de la PR.

En France, l'estimation de la prévalence de la PR « grave évolutive » fondée sur les données de la CNAM semble concorder avec la tendance internationale. Toutefois, ce constat fondé sur les données médico-administratives devrait être confirmé par la mise en place d'études épidémiologiques (populationnelles) sur la prévalence de la PR permettant d'actualiser les résultats de la seule étude française réalisée en 2001.

3.2. Mortalité, comorbidités et PR

La revue de littérature a porté sur vingt-huit études sélectionnées, dont six méta-analyses, vingt-neuf analyses sur cohortes et trois études françaises portant sur les certificats de décès. Les méta-analyses ont été menées sur des études portant sur des cas prévalents ou des nouveaux cas (cas incidents). Le processus de sélection des articles est décrit dans la figure globale (figure 1, chapitre n° 4) qui présente la sélection des publications de la revue de littérature économique et relative aux aspects contextuels.

3.2.1. Principaux messages issus de la revue de littérature

► Études françaises portant sur la mortalité

Trois études françaises ont été recensées. Elles proposent des analyses en causes multiples des certificats de décès en France : sur la période 1970-1990 (Coste *et al.*, 1994) (47), la période 1970-2002 (Ziadé *et al.*, 2008) (48) et la période 2000-2011 (Avouac *et al.*, 2017) (49). Dans ces études fondées sur des méthodologies statistiques similaires, les informations rapportées dans les certificats de décès français mentionnent la PR comme cause « principale » ou « associée » au décès. L'étude de Ziadé *et al.* (2008) (48) réalisée sur la période 1970-2002 a montré que la PR avait un impact significatif sur la mortalité dans la population française, avec une tendance à l'augmentation des taux de mortalité associés aux groupes de population plus âgés à partir des années 1990. Le travail récent d'Avouac *et al.* (2017) (49) a mis en évidence :

1. un taux de décès entre 2000 et 2011 égal à 0,20 % (assez proche de Ziadé *et al.*, 2008) ;
2. un nombre de certificats de décès mentionnant la PR comme cause sous-jacente du décès diminuant de 13,9 % entre 2000 et 2011, alors que celui rapportant la PR comme cause associée a augmenté de 14 % au cours de la même période ;
3. des taux de mortalité cardiovasculaire, respiratoire et infectieuse fortement associés à la PR.

► Études internationales portant sur la mortalité

Six constats issus de l'analyse des études internationales sur la mortalité associée à la PR peuvent être synthétisés :

1. *la mortalité chez les patients atteints de PR (notamment les formes les plus graves) et les patients présentant un risque cardiovasculaire est plus élevée que celle de la population générale :*
 - dans l'étude de Minaur *et al.* (2004) (50), conduite sur 40 ans et à partir du diagnostic de la PR, la médiane de survie était réduite de 10 ans pour les hommes et 11 ans pour les femmes. Le ratio de mortalité standardisé était généralement compris entre 1,2 et 1,5 dans les études menées aussi bien sur des cas prévalents qu'incidentes. Une méta-analyse de Dadoun *et al.* (2013) (51) portant sur des publications couvrant la période 1955-1995, a rapporté une surmortalité de 47 % par rapport à la population générale. Par ailleurs, l'analyse d'une cohorte américaine de patients atteints de PR et de maladies rhumatismales non inflammatoires (NIRD) (England *et al.*, 2014) (52), entre 1998 et 2011, a montré qu'après ajustement sur l'âge et le sexe, le risque relatif (RR) de mortalité lié à la PR était plus élevé que celui lié aux NIRD et

demeurait constant au cours de la période d'analyse (RR = 1,29 ; IC_{95%} = [1,18 - 1,56]),

- la littérature montre un risque de mortalité cardiovasculaire élevé chez les patients atteints de PR (Agca *et al.*, 2017) (53). Par exemple, une analyse menée sur une cohorte danoise (Lindhardsen *et al.*, 2011) (54) a montré que le risque d'infarctus du myocarde était environ 70 % plus élevé que dans la population générale ;
2. *des travaux assez récents (méta-analyses et analyses de cohortes) rapportent une tendance à la diminution de la mortalité observée chez les patients atteints de PR tout en restant supérieure à celle de la population générale. Trois études récentes illustrent particulièrement ce constat :*
- la méta-analyse de Dadoun *et al.* (2013) (51), réalisée sur 11 études longitudinales initiées de 1955 à 1995, a porté sur un total de 51 819 patients. Dans cette étude, le taux de mortalité incident (IMR) a diminué de 2,3 % par an (IC_{95%} = [2,1 - 2,6]) mais est demeuré supérieur à celui de la population générale. Le taux de mortalité standardisé (SMR) était de 1,48 (IC_{95%} = [1,19 - 1,83]),
 - des résultats similaires ont été observés dans l'étude de Hoek *et al.* (2017) (55) portant sur des données de mortalité d'une cohorte de patients néerlandais. Ces dernières ont été collectées entre 1997 et 2012. Cette étude a montré qu'une fois ajustée sur l'âge, le sexe et l'année civile, la mortalité observée chez les patients atteints de PR était supérieure de 54 % à celle de la population générale. Elle tendait à diminuer au cours du temps (2 % annuellement, p-valeur=0,07),
 - dans une étude canadienne (Widdifield *et al.* 2015) (46) réalisée dans la province d'Ontario sur des patients résidents âgés de 15 ans et plus (avec PR vs. ceux sans PR) de 1996 à 2009, le taux de mortalité toutes causes diminuait de 21,4 %, comparativement à la population générale (21,4 % vs 13,4 %). Une légère tendance à la diminution du ratio de mortalité standardisé a été constatée mais elle demeure non significative dans la population étudiée ;
3. *d'autres travaux plus fins en termes d'analyses suggèrent que la tendance à la diminution de la surmortalité chez les patients atteints de PR diffère selon les périodes étudiées (ex. : années 1990 vs années 2000), en lien probablement avec une évolution de la prise en charge de la PR. Ce constat est notamment illustré par trois études récentes :*
- dans l'étude de Lacaille *et al.* (2017) (56) portant sur des données médico-administratives de la province de Colombie-Britannique, deux cohortes ont été comparées (1996-2000 vs 2001-2006). Il a été rapporté que la mortalité dans la population incidente des patients atteints de PR diminuait dans le temps. Toute-

- fois, la mortalité toutes causes était significativement plus élevée chez les patients atteints de PR sur la période 1996-2001 (Hazard Ratio [HR] = 1,55 ; IC_{95%} = [1,43-1,68]), mais elle ne l'était plus entre 2001 et 2006 (HR = 0,9, IC_{95%} = [0,90-1,60]),
- l'étude de Zhang *et al.* (2017) (57) portant sur des données médico-administratives anglaises comparant la mortalité toutes causes sur deux cohortes incidentes (PR diagnostiquée au cours de la période 1999-2005 et suivie jusqu'à la fin de 2005 ; PR diagnostiquée au cours de la période 2006-2012 et suivie jusqu'à 2012) par rapport à la population générale. Les résultats associés à la seconde cohorte montraient une diminution significative du HR (p-valeur=0,027) même si la mortalité au sein de cette cohorte demeurait élevée par rapport à celle de la population générale (HR = 1,21, IC_{95%} = [1,05-1,39]),
 - l'étude de Myasoedova *et al.* (2017) (58) portant sur une population incidente de patients américains atteints de PR (âge ≥ 18 ans, critère ACR vérifié au cours de la période 1980-2007) vs des patients non atteints de PR présentant les mêmes caractéristiques démographiques (âge, sexe, année civile). Les résultats montrent une diminution significative de la mortalité cardiovasculaire pour les patients dont la PR a été diagnostiquée en 2000-2007 vs 1990-1999 (HR= 0.43 : IC_{95%} = [0,19 - 0,94]). De plus, la mortalité globale à 10 ans pour les patients atteints de maladie cardiovasculaire est statistiquement comparable entre les patients atteints et non atteints de PR (cohorte incidente 2000-2007) ;
4. *la littérature est en faveur d'un meilleur contrôle de l'inflammation qui réduit le risque de mortalité cardio-vasculaire.* Des études ont montré que les csDMARDs (en particulier le MTX) et les bDMARDs (tels que les anti-TNF α) étaient associés à une réduction du risque de mortalité cardio-vasculaire chez les patients atteints de PR (Roubille *et al.*, 2015 ; Ajeganova *et al.*, 2013 ; Barnabe *et al.*, 2011 ; Rao *et al.*, 2015) (59-62). Dans certaines études, le risque de maladie cardio-vasculaire semble diminuer encore plus après une utilisation à long terme (Bili *et al.*, 2014 ; Innala *et al.*, 2011) (63, 64) ;
5. *une augmentation du risque de mortalité a été observée chez les patients atteints d'une PR très active et traités avec des doses élevées de glucocorticoïdes.* Comparés aux traitements de fond synthétiques (ex. : le MTX), Listing *et al.* (2015) (65) ont noté qu'il existait une association entre l'utilisation des bDMARDs (ex. : les anti-TNF α , le rituximab) et la réduction de ce type de mortalité.

3.2.2. Discussion

Les conclusions issues des résultats des analyses de mortalité et leur extrapolation au contexte français doivent être prudentes en raison des éléments suivants :

1. la littérature est consensuelle sur le fait que la mortalité chez les patients atteints de PR est plus élevée que celle de la population générale et qu'elle tend à baisser dans le temps, quelques travaux ne partagent pas ces conclusions et rapportent, par exemple, que la mortalité résultant des cas incidents de la PR est comparable à celle de la population générale (étude de Puolakka et al., 2010, réalisée en Finlande) (66) ;
2. certaines variables d'ajustement ne sont pas toujours rapportées dans les données analysées (ex. : consommation de tabac, précisions sur les caractéristiques des diagnostics de la PR dans les données médico-administratives). Aux États-Unis, une recherche récente (Sparks et al., 2016) (67) menée sur des données de la *Nurses' Health Study* auprès de 121 700 infirmières a ajusté la mortalité des femmes sur l'âge, la consommation de cigarettes avant le diagnostic et d'autres co-variables telles qu'un tabagisme supérieur à cinq paquets-années après le diagnostic de PR. Les auteurs ont indiqué que les femmes fumant après le diagnostic de PR présentaient un risque de surmortalité par rapport à celles n'ayant jamais fumé après le diagnostic de PR (HR = 2,69, IC_{95%} = [1,33 - 5,46]) ;
3. concernant la puissance et le *design* des études, les cohortes incidentes utilisées reposent sur des effectifs insuffisamment élevés (étude de Myasoedova et al., 2017) (58), un suivi de patients entaché de données manquantes, un recueil non systématique d'indicateurs sur la mortalité. Par exemple, dans la recherche de Dadoun et al. (2013) (51), 92 études parmi 141 ont été exclues en raison de l'absence d'études longitudinales, de polyarthrites précoces ou de données de mortalité ;
4. l'analyse en termes de causes multiples de certificats de décès, particulièrement utilisée dans les études françaises, repose exclusivement sur les informations rapportées dans les certificats de décès. L'absence de données portant sur l'âge au moment du diagnostic de la PR ainsi que d'autres variables socio-économiques décrivant les patients rendent difficile la comparaison de ce type d'études avec les analyses menées sur les cohortes ;
5. il n'est pas toujours facile de dissocier les raisons de la diminution de la mortalité liée à la PR elle-même, de celles induites par une diminution de la mortalité dans la population générale, même si une amélioration du parcours de la prise en charge de la PR pourrait contribuer à la diminution de la surmortalité liée à la PR ;
6. il est à noter que des travaux récents (Juge et al., 2018) (68) suggèrent l'importance de l'affection pulmonaire interstitielle comme facteur de mortalité potentiel dans la PR.

- La littérature **est consensuelle** sur le fait que la mortalité chez les patients atteints de PR (notamment les formes les plus sévères et les patients présentant un risque cardio-vasculaire) est plus élevée que celle de la population générale. **La surmortalité associée à la PR a été prise en compte dans l'estimation de la survie dans l'analyse de référence du modèle PR France (cf. chapitre 5).**
- La littérature est **globalement concordante** sur la diminution de la mortalité de la PR dans le temps. Des tendances en termes d'amélioration de la survie rapportées par des études de cohortes récentes montrent une association positive entre une prise en charge par traitements de fond synthétiques et biologiques et la réduction du risque cardio-vasculaire chez les patients atteints de PR.

4. Contexte clinique d'évaluation de la PR

Cette section propose une synthèse des objectifs de la prise en charge de la PR et des critères (cliniques, qualité de vie) permettant leur poursuite.

4.1. Objectifs du traitement de la polyarthrite rhumatoïde

Les objectifs du traitement de la PR sont d'induire, de maintenir une rémission clinique⁷ (ou une faible activité si la PR est ancienne) et de limiter la progression de la dégradation articulaire et, par conséquent, le handicap ultérieur. Afin d'atteindre ces objectifs, dans un premier temps, l'activité de la PR est mesurée à l'aide de critères composites validés incluant les indices articulaires. Ensuite, l'appréciation par le patient de son état de santé est mesurée à l'aide de critères de retentissement fonctionnel et de plusieurs questionnaires de qualité de vie.

4.2. Synthèse des critères d'efficacité clinique utilisés dans le traitement de la PR

Les principaux critères utilisés dans l'évaluation de l'efficacité clinique des traitements de la PR sont :

► **L'indice composite d'activité de la PR (DAS 28)**

Le DAS 28 est une simplification du DAS (*Disease Activity Score*) développé initialement pour 44 articulations (69). Le DAS 28 tient compte de l'évaluation de la douleur et du nombre de synovites sur les 28 sites articulaires proposés par l'EULAR, de la valeur de la vitesse de sédimentation globulaire et de l'appréciation globale de la maladie par le patient sur une échelle visuelle analogique de 100 millimètres (Annexe A1-5).

Les stades d'évolution de la PR sont définis en fonction des seuils du DAS 28 suivants :

- PR en rémission : $DAS\ 28 < 2,6$;
- PR faiblement active : $2,6 \leq DAS\ 28 \leq 3,2$;
- PR modérément active : $3,2 < DAS\ 28 \leq 5,1$;
- PR très active : **DAS 28 > 5,1**.

► **L'indice composite simplifié de l'activité de la PR (SDAI, *Simplified Disease Activity Index*) (70)**

Le SDAI est la somme des paramètres suivants :

- le nombre d'articulations gonflées (28 articulations analysées) ;

⁷ La rémission clinique est définie par l'absence de signes et symptômes d'activité inflammatoire significative.

- le nombre d'articulations douloureuses (28 articulations considérées) ;
- l'activité de la maladie jugée par le patient (échelle visuelle analogique de 0 à 10) ;
- l'activité de la maladie jugée par le praticien (échelle visuelle analogique de 0 à 10) ;
- la C-réactive protéine (mg/dl).

Les valeurs du SDAI sont comprises entre 0 et 86 :

- PR en rémission : $0 \leq \text{SDAI} \leq 3,3$;
- PR faiblement active : $3,4 \leq \text{SDAI} \leq 11$;
- PR modérément active : $11,1 \leq \text{SDAI} \leq 26$;
- PR très active : $26,1 \leq \text{SDAI} \leq 86$.

► **L'indice composite d'activité de la PR (CDAI, *Clinical Disease Activity Index*)**
(71)

Le CDAI est un critère d'activité fondé uniquement sur des paramètres cliniques :

- le nombre d'articulations gonflées (28 articulations analysées) ;
- le nombre d'articulations douloureuses (28 articulations considérées) ;
- l'activité de la maladie jugée par le patient (échelle visuelle analogique de 0 à 10) ;
- l'activité de la maladie jugée par le praticien (échelle visuelle analogique de 0 à 10).

Les valeurs du CDAI sont comprises entre 0 et 76 :

- PR en rémission : $0 \leq \text{CDAI} \leq 2,8$;
- PR faiblement active : $2,9 \leq \text{CDAI} \leq 10$;
- PR modérément active : $10,1 \leq \text{CDAI} \leq 22$;
- PR très active : $22,1 \leq \text{CDAI} \leq 76$.

► **La réponse au traitement au sens des recommandations de l'ACR ou de l'EULAR**

Ce critère est utilisé depuis 2011 dans les essais cliniques (72).

a) afin qu'un patient soit considéré comme répondeur selon l'ACR, il est nécessaire qu'il ait au minimum :

- ≥ 20 % de réduction du nombre d'articulations douloureuses (NAD),
- ≥ 20 % de réduction du nombre de synovites (NAG),
- ≥ 20 % d'amélioration de trois des cinq items suivants :
 - évaluation de la douleur par le patient,

- évaluation globale par le patient,
- évaluation globale par le praticien,
- auto-questionnaire évaluant le statut fonctionnel,
- marqueur biologique de l'inflammation (ex. : vitesse de sédimentation « VS » ou la protéine C-Réactive « CRP »).

Les critères de réponse ACR 20 % représentent le minimum de « niveau de réponse » mais la réponse est également évaluée à 50 % et à 70 % ;

- b) la réponse selon l'EULAR est fondée sur l'amélioration du DAS ou du DAS 28 en fonction de la variation de l'activité et du statut final des scores DAS ou DAS 28.

La classification de la PR en fonction des scores DAS 28 et DAS 44 est présentée dans le **tableau 1**.

Tableau1. Classification de l'activité de la PR selon le DAS 28 et le DAS 44. Source : Guide HAS ALD, Polyarthrite rhumatoïde évolutive grave, avril 2008 (73)

Activité de la PR	Score DAS 28	Score DAS 44
Rémission	DAS 28 < 2,6	DAS < 1,6
Faible	2,6 ≤ DAS 28 < 3,2	1,6 ≤ DAS < 2,4
Modérée	3,2 ≤ DAS 28 < 5,1	2,4 ≤ DAS < 3,7
Forte	DAS 28 ≥ 5,1	DAS ≥ 3,7

La réponse thérapeutique au sens de l'EULAR sur l'activité de la maladie entre deux mesures selon l'évolution du DAS 28 au cours d'une période est présentée dans le **tableau 2**.

Tableau 2. Réponse thérapeutique au sens de l'EULAR. Source : Guide HAS ALD, Polyarthrite rhumatoïde évolutive grave, avril 2008 (73)

Activité de la PR	Diminution du score DAS 28		
	> 1,2	> 0,6 et ≤ 1,2	≤ 0,6
Faible (2,6 ≤ DAS 28 < 3,2)	Bonne	Modérée	Nulle
Modérée (3,2 ≤ DAS 28 < 5,1)	Modérée	Modérée	Nulle
Forte (DAS 28 ≥ 5,1)	Modérée	Nulle	Nulle

- Depuis les années 1990, des critères d'efficacité mesurant l'activité de la PR ont été proposés pour évaluer la rémission des patients. Le critère le plus fréquemment étudié est le DAS 28 et les réponses cliniques les plus connues sont la réponse ACR (ACR20, ACR50 et ACR70) et la réponse EULAR.
- Dans un premier temps, la revue systématique de la littérature clinique et la réalisation de méta-analyses (chapitres 2 et 3) permettront d'identifier les critères cliniques fréquemment utilisés dans les essais cliniques portant sur la PR et les limites ou les contraintes sous-jacentes à leur utilisation.
- Dans un second temps, les critères cliniques fréquemment utilisés dans la prise en charge de la PR seront sélectionnés et utilisés dans le modèle PR France.

4.3. Synthèse des critères de qualité de vie utilisés dans le traitement de la PR

Dans cette section, une synthèse des critères décrivant la qualité de vie des patients (spécifiques ou génériques) qui peuvent être utilisés dans le traitement de la PR est proposée. Une description et une discussion approfondie des critères de qualité de vie retenus dans le modèle PR France sont proposées dans les chapitres n° 4 et 5.

4.3.1. Les critères d'incapacité fonctionnelle

► Le *Stanford Health Assessment Questionnaire* (HAQ)

Le HAQ est le questionnaire le plus utilisé pour mesurer l'incapacité fonctionnelle spécifique à la PR et l'aptitude du patient à la réalisation des actes de la vie quotidienne (74, 75). Le questionnaire HAQ est le plus fréquemment proposé dans les études cliniques (essais cliniques, registres et cohortes). Il est également utilisé dans la définition de la réponse ACR. Il explore huit domaines relatifs à l'activité physique : s'habiller et se préparer, se lever, manger, marcher, l'hygiène, atteindre et attraper un objet, la préhension (annexe A1-6). Quatre possibilités de réponse sont notées de 0 à 3 (sans aucune difficulté, avec quelques difficultés, avec beaucoup de difficultés, incapable de le faire). Le score HAQ est compris entre 0 et 3 : un score de « 0 » indique l'absence d'incapacité fonctionnelle, alors qu'un score égal à 3 exprime une incapacité maximale. Le score HAQ présente les normes suivantes : HAQ=0,22 (population générale) ; HAQ<0,5 : pas d'handicap, fonctionnement normal ; HAQ compris entre 0,5 et 1 : handicap minime ; HAQ compris entre 1 et 3 : handicap modéré ; HAQ supérieur à 3 : handicap sévère. **Les dimensions sociales et psychologiques de la PR ne sont pas prises en compte dans le HAQ.**

► **L'échelle visuelle analogique (EVA) spécifique à l'incapacité fonctionnelle du patient présentant une PR**

L'EVA-Fatigue a été proposée dans le but de compléter les critères décrivant la capacité fonctionnelle fondés sur des catégories de questions (ex. : le HAQ) (Pollard *et al.*, 2006) (76). Elle décrit le niveau de fatigue du patient à l'aide d'un score variant de 0 à 100, 0 représentant une absence de fatigue et 100 des problèmes majeurs.

4.3.2. Les critères génériques et spécifiques décrivant la qualité de vie du patient (Health-Related Quality of Life (HRQoL))

La qualité de vie pour un patient présentant une PR est mesurée soit par des instruments génériques, soit par des instruments spécifiques à la PR. Les instruments génériques les plus fréquemment utilisés sont l'EQ-5D, l'échelle visuelle analogique évaluant la fatigue (EVA-Fatigue), le SF-36 et le *Short Form-Six Dimensions* (SF-6D). Des instruments spécifiques comme l'*Arthritis Impact Measurement Scales* (AIMS) ou le *Rheumatoid Arthritis Quality of Life* (RAQoL) sont moins fréquemment proposés, comme le montrera la revue de littérature économique présentée dans le chapitre n° 4.

► **L'EQ-5D**

L'EQ-5D est un instrument générique intégrant les cinq domaines suivants : la mobilité, l'autonomie de la personne, les activités courantes, la douleur/gêne et l'anxiété/dépression. Dans sa version en trois niveaux (EQ-5D-3L), le patient choisit une réponse parmi trois possibilités dans chaque domaine : « aucun problème », « quelques problèmes ou problèmes modérés » et « problèmes extrêmes ». L'EQ-5D-3L couvre 243 états de santé différents. La version 5L (cinq niveaux) de l'EQ-5D, développée par l'EuroQol Group en 2009, est seulement disponible dans les études récentes mais contrairement à la version de l'EQ-5D en trois niveaux, elle n'est pas encore validée en population française. Dans la cohorte française ESPOIR (Étude et suivi des polyarthrites indifférenciées récentes) qui sera exploitée dans ce projet, l'EQ-5D à trois niveaux est utilisé (cf. chapitre n° 5).

► **Le SF-36**

Le SF-36 (<http://www.sf-36.org/tools/sf36.shtml>) est l'instrument générique qui inclut des items décrivant les aptitudes et les incapacités physiques des patients. Il peut être utilisé pour évaluer l'incapacité fonctionnelle des patients atteints de PR.

Le SF-36 contient 36 items répartis en huit domaines évaluant les limitations d'activité physique relatives aux problèmes de santé, les limitations d'activité sociale liées aux problèmes physiques ou psychologiques, les limitations d'activité usuelle à cause de problèmes de santé d'ordre physique, la douleur corporelle, la santé mentale générale, les limitations d'activité

usuelle à cause de problèmes psychologiques, la vitalité et la perception ressentie par le patient concernant sa santé générale. Le SF-36 intègre les dimensions de l'énergie et de la fatigue. Il permet d'obtenir deux scores : un score résumé physique et un score résumé psychique.

► **Le SF-6D**

Le SF-6D regroupe six domaines (limitations physiques, limitation de rôle, fonction sociale, douleur, santé mentale et vitalité) contenant quatre à six items chacun et permet de définir 9 000 états de santé. Il est dérivé à l'aide d'un algorithme du SF-36.

► **Échelle de mesure de l'impact de la polyarthrite rhumatoïde (AIMS)**

Cette échelle a été développée en 1980 (Meenan *et al.*, 1980) (77) afin de décrire l'état de santé de patients présentant une PR selon sept critères démographiques et 55 critères décrivant l'état de santé.

Une version plus courte de cet instrument est traduite en français : « l'Échelle de mesure de l'impact de la polyarthrite rhumatoïde » (EMIR) (Leplège *et al.*, 1998) (78). Elle comporte 57 items couvrant 12 dimensions (mobilité, marche, fonction de la main et des doigts, bras, soins personnels, tâches ménagères, douleur, travail, activités sociales, soutien de la famille et des amis, tension, humeur). Chaque critère est mesuré à l'aide d'une échelle en cinq points.

► **L'échelle *Rheumatoid Arthritis Quality of Life* (RAQoL)**

C'est un questionnaire spécifique de la PR (Maska *et al.*, 2011) (79) comportant 30 questions binaires décrivant des besoins non satisfaits et liés à des activités courantes. Le score final est compris entre 0 et 30. Plus le score RAQoL est élevé, moins la qualité de vie perçue par le patient est bonne.

- Depuis le début des années 2000, des instruments de mesure de l'incapacité fonctionnelle (ex. : HAQ), en plus des instruments génériques (EQ-5D-3L) et spécifiques, ont été proposés pour décrire la qualité de vie des patients présentant une PR.
- Dans un premier temps, la revue systématique de littérature économique (chapitre 5) permettra d'identifier les critères de qualité de vie fréquemment utilisés dans les modélisations économiques et les limites ou les contraintes sous-jacentes à leur utilisation.
- Dans un second temps, les critères de qualité de vie génériques ou spécifiques au contexte français seront sélectionnés et utilisés dans le modèle PR France.

5. Prise en charge médicamenteuse et recommandations portant sur la PR

Les traitements de fond dits « synthétiques » (« csDMARDs », *Conventional Synthetic Disease-Modifying Antirheumatic Drugs*) sont recommandés en première intention dans la prise en charge de la PR. Il s'agit principalement du méthotrexate (MTX), du léflunomide, de la sulfasalazine et de l'hydroxychloroquine. **Le traitement de fond de référence en première intention dans la PR active est à ce jour le méthotrexate.** Les autres csDMARDs (ciclosporine, sels d'or, D-pénicillamine et tiopronine) n'ont plus de place dans la stratégie thérapeutique de la PR en France (la CT a attribué le 23/07/2014 un SMR insuffisant à ces quatre produits). Quant au cyclophosphamide, il n'est plus utilisé dans les PR classiques, y compris dans les formes sévères. Toutefois, il reste un traitement de fond, de recours, dans les polyarthrites rhumatoïdes compliquées de vascularites.

La prise en charge de la PR a considérablement évolué depuis les années 2000 à la suite de l'apparition de nouveaux traitements de fond, notamment les traitements de fond biologiques (bDMARDs) utilisés en association (ou non) avec un csDMARD (ex. : MTX). Leur prescription nécessite la réalisation d'un bilan exhaustif afin de vérifier l'absence des principales contre-indications que sont la tuberculose, une néoplasie ou une hémopathie maligne. Une synthèse des principales caractéristiques des bDMARDs utilisées dans le modèle PR France (posologie, rythme d'administration) est proposée dans le chapitre 5.

5.1. Les traitements de fond biologiques

Six classes de traitements de fond biologiques ciblant le processus inflammatoire sont actuellement utilisées :

► **Les anti-TNF α** (les inhibiteurs du facteur de nécrose tumorale)

Ils représentent les bDMARDs les plus prescrits avec cinq molécules ayant une indication dans la PR : l'étanercept (ENBREL[®]) et ses biosimilaires (BENEPALI[®], ERELZI[®]) ; l'adalimumab (HUMIRA[®]) et ses biosimilaires (AMGEVITA[®], IMRALDI[®], HYRIMOZ[®], HULUO[®]) ; l'infliximab (REMICADE[®]) et ses biosimilaires (INFLECTRA[®], REMSIMA[®], FLIXABI[®], ZESSLY[®]) ; le certolizumab (CIMZIA[®]) et le golimumab (SIMPONI[®]). Ils sont proposés en cas de réponse inadéquate ou d'intolérance aux traitements de fond conventionnels comprenant le méthotrexate. Ils peuvent aussi être prescrits en première intention (patients avec une contre-indication au MTX) dans certaines formes actives et sévères de PR.

► Les autres classes utilisées concernant les bDMARDs ayant d'autres cibles que le TNF

Il s'agit des spécialités suivantes :

- un anti-CD20 (lymphocyte B) : le rituximab (MABTHERA®) et son biosimilaire (RIXATHON®, TRUXIMA®) ;
- un CTLA4 (lymphocyte T) : l'abatacept (ORENCIA®) ;
- un anti-interleukine 6 : le tocilizumab (ROACTEMRA®) ; sarilumab (KEVZARA®) ;
- un anti-interleukine 1 : l'anakinra (KINERET®).

L'abatacept et le tocilizumab ont une AMM chez les patients en échec des traitements de fond conventionnels dont le MTX, ou en échec d'au moins un anti-TNF. Le rituximab, quant à lui, a une AMM en cas d'échec d'au moins un anti-TNF ou en cas de contre-indication aux anti-TNF. L'anakinra a une place limitée dans la stratégie thérapeutique, compte tenu d'une efficacité qui semble inférieure à celle des autres bDMARDs. Une description détaillée des avis de la CT relatifs à ces produits est fournie dans l'annexe A1-2 et des informations liées à leurs coûts sont données dans le chapitre 5.

5.2. Les traitements de fond synthétiques ciblés

Une nouvelle classe thérapeutique a récemment été définie par la Ligue européenne contre les rhumatismes (EULAR) (Smolen *et al.*, 2014 (5) et Smolen *et al.*, 2017 (6)). Elle concerne les traitements de fond synthétiques ciblés, « *targeted synthetic DMARDs* ». Deux produits ont été récemment évalués par la CT : le baricitinib (OLUMIANT®) (41) qui est un nouvel agent chimique inhibiteur sélectif et réversible des Janus kinases (JAK1 et JAK2) et le tofacitinib (XELJANZ® 5 mg) qui inhibe les JAK1 et JAK3 (80). **Ces produits n'ont pas été intégrés dans les méta-analyses réalisées par la HAS et le SchARR puisque ces produits ont été évalués par la CT fin 2017. L'actualisation du projet intégrera cette nouvelle classe de produits.**

5.3. Les recommandations de prise en charge médicamenteuse

Des recommandations de la HAS, relatives à la prise en charge médicamenteuse de la PR, datent de 2007. Elles ont été suspendues par le Collège de la HAS en 2011 et sont en attente d'actualisation.

Les recommandations actuellement mises à disposition par la HAS sont les suivantes :

1. synthèse des recommandations professionnelles. Polyarthrite rhumatoïde : aspects thérapeutiques hors médicaments et chirurgie – aspects médico-sociaux et organisationnels, 2007 ([10](#)) ;
2. guide ALD, Polyarthrite rhumatoïde évolutive grave, avril 2008 ([73](#)) ;
3. guide ALD, Polyarthrite rhumatoïde évolutive grave, avril 2008 ([73](#)) ;
4. guide ALD, La prise en charge de votre polyarthrite rhumatoïde, vivre avec une polyarthrite rhumatoïde, décembre 2008 ([81](#)).

Dans l'attente de nouvelles recommandations de la HAS sur la prise en charge médicamenteuse de la PR, l'accent est particulièrement mis sur les recommandations de la Société française de rhumatologie (SFR) et de la Ligue européenne contre le rhumatisme (EULAR) qui sont en accord avec les conclusions des avis de la CT relatifs à la réévaluation des bDMARDs et l'évaluation des nouveaux de traitements de fond (biologiques et synthétiques ciblés).

Dans un premier temps, cette section propose une synthèse portant sur les recommandations de prise en charge médicamenteuse des sociétés savantes en rhumatologie (EULAR, ACR et SFR). Dans un second temps, une synthèse sur les recommandations de la SFR est présentée.

5.3.1. Les recommandations de l'EULAR

Les récentes mises à jour des recommandations de l'EULAR (Smolen et al., 2017) ([6](#)) ont été élaborées par un groupe international d'experts (*Task Force*) composé notamment de rhumatologues, de patients et des autres professionnels de santé. Les recommandations ciblent des rhumatologues, des patients, des médecins et autres professionnels de santé impliqués dans les soins aux patients atteints de PR, les sociétés de rhumatologie, les gestionnaires d'hôpitaux, les représentants de l'assurance maladie et les acteurs politiques. Les recommandations de l'EULAR sont fondées sur trois revues systématiques de littérature et des avis d'experts.

La mise à jour de 2016 des recommandations de l'EULAR comprend quatre principes généraux et 12 recommandations.

Les principaux messages des recommandations EULAR sont les suivants :

1. les patients atteints d'une PR active doivent être surveillés tous les 3 mois, et le traitement doit être ajusté s'il n'y a pas d'amélioration à 6 mois ;
2. les recommandations intègrent les concepts de **cibles de traitement** et de **traitement jusqu'à l'obtention de la cible** (« *treat to target* »), rémission ou activité faible de la maladie (« *Low Disease Activity* » (LDA)) ;
3. les recommandations proposent une hiérarchisation de l'approche thérapeutique à l'aide d'une stratification par niveau de risque défini, une fois que le traitement initial de csDMARDs et de glucocorticoïdes n'atteint pas la cible ;
4. tous les types de DMARDs (synthétiques, synthétiques ciblés, biologiques et biosimilaires) sont pris en compte dans les recommandations : le MTX est recommandé comme traitement de première intention ; la sulfasalazine (SSZ) ou le léflunomide peuvent être utilisés en substitution au MTX s'il existe des contre-indications au MTX. Les inhibiteurs du facteur de nécrose tumorale (anti-TNF α) ne sont plus les seuls produits biologiques recommandés pour les patients présentant une réponse insuffisante au MTX. Tous les produits biologiques sont considérés de manière similaire. **Les produits biologiques doivent être associés avec des csDMARDs** du fait d'une action synergique.

5.3.2. Les recommandations de l'ACR de 2016

Les récentes recommandations de l'ACR de 2016 (82) sont destinées aux patients, aux médecins et aux professionnels de santé. Elles ont été développées selon la méthodologie GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) qui est une méthode internationale qui spécifie clairement les groupes de patients, les interventions et les alternatives de traitements. Cette méthodologie gradue le niveau de preuve considéré tout en tenant compte des autres critères tels que les préférences des patients, l'expérience et l'expertise des rhumatologues et des professionnels de santé. Parmi ces recommandations, 23 % sont associées à un niveau de confiance élevé (ceci signifie que le panel des experts est confiant dans les effets attendus des recommandations) et 77 % sont conditionnelles (l'effet positif attendu des recommandations l'emporte sur leur effet négatif).

L'ACR souligne que la faible activité de la maladie ou la rémission est la cible pour tous les patients atteints d'une PR débutante ou active et recevant un csDMARD ou un bDMARD. Elle propose des recommandations spécifiques aux patients traités par un csDMARD et/ou un bDMARD et présentant des comorbidités (ex. : patients atteints d'hépatite, patients atteints d'un cancer).

5.3.3. Les recommandations de l'ACR 2015 vs. EULAR 2017

Même si les mises à jour récentes ont beaucoup contribué à rapprocher les dernières recommandations de l'EULAR et de l'ACR par rapport à leurs versions antérieures (2008/2010 pour l'ACR et 2012/2013 pour l'EULAR), elles présentent néanmoins quelques différences. Comparativement aux recommandations de l'ACR, celles de l'EULAR :

1. **préconisent d'utiliser les bDMARDs en association avec les csDMARDs (MTX) plutôt qu'en monothérapie ;**
2. ne distinguent pas les patients en fonction de la durée de la PR (débutante vs active) mais en fonction des phases de traitement (patients naïfs des csDMARDs, patients ayant utilisé des csDMARDs, patients ayant utilisé des bDMARDs) ;
3. intègrent les facteurs pronostiques dans la stratification de la hiérarchie thérapeutique.

5.3.4. Les recommandations de la SFR de 2018

En 2018, en se fondant sur les résultats d'une analyse systématique de la littérature et sur les avis des experts, la SFR a actualisé ses recommandations de 2014 ([83](#)) en proposant quatre principes généraux et 15 recommandations élaborées par un groupe d'experts composé de 12 experts rhumatologues (hospitaliers et libéraux), de deux patients atteints de PR et membres des associations de patients et d'une ergothérapeute. Les 15 recommandations de la SFR (Daien *et al.*, 2019) ([7](#)) soulignent l'importance d'une décision partagée entre le rhumatologue et le patient, d'une prise en charge globale (médicamenteuse et non médicamenteuse) de la PR, la nécessité de l'introduction la plus précoce possible d'un traitement de fond et la pertinence d'une évaluation régulière de l'activité de la maladie par des scores composites permettant de vérifier l'atteinte de la rémission.

Les recommandations de la SFR s'articulent autour de quatre axes :

1. **Un processus de concertation** doit être mis en œuvre entre le rhumatologue et le patient dans le cadre d'une décision médicale partagée et fondée sur l'information et sur l'éducation du patient.
2. **Le rhumatologue** est le spécialiste qui doit prendre en charge les patients atteints de la PR. **Le médecin généraliste** intervient lors de la détection de la maladie et du suivi du patient en coordination avec le rhumatologue.
3. Tout patient atteint de PR devrait bénéficier d'une **prise en charge globale** comprenant les traitements médicamenteux et non médicamenteux (ex. : éducation thérapeutique, prise en charge des comorbidités, soutien psychologique, recours à la chirurgie, accompagnement socio-professionnel).
4. Les orientations thérapeutiques doivent prendre en compte le coût de la PR, de ses conséquences et de ses traitements à **l'échelle individuelle et sociétale**.

Les axes généraux des recommandations de 2018, tels que la décision médicale partagée entre le patient et le rhumatologue et la place du rhumatologue dans le processus de prise en charge de la PR, sont identiques à ceux de 2014. En revanche, les recommandations de 2018 diffèrent de celles de 2014, notamment au niveau de la prise en compte des nouveaux traitements de fond synthétiques ciblés (baricitinib, tofacitinib), l'introduction du choix préférentiel d'un traitement ayant un mode d'action différent en cas d'échec primaire à un bDMARD ou un tsDMARD et la place spécifique des mesures non médicamenteuses à l'instar de l'hygiène de vie et la gestion des comorbidités dans la prise en charge globale de la PR.

Concernant la prise en charge médicamenteuse de la PR, les recommandations 6 à 8 portent sur la première ligne de traitement (i.e. le MTX), alors que les recommandations 9 à 11 décrivent les stratégies correspondant à la deuxième ligne de traitement et plus.

5.3.5. Synthèse des recommandations retenues dans les avis de la CT et comparables à celles de la SFR

Cette section récapitule les principales recommandations retenues dans les avis de la CT et qui sont concordantes avec celles de la SFR et retenues dans le modèle PR France.

1. En première ligne de traitement, le MTX est le traitement de fond classique de référence de la PR. Une dose de 10 à 15 mg/semaine par voie orale est recommandée lors de l'initiation du traitement, avec augmentation rapide des doses, jusqu'à une dose optimisée autour de 0,3mg/kg/semaine (soit entre 15 et 25 mg/semaine en fonction du contexte clinique et de la tolérance au traitement).

En cas de contre-indication ou d'intolérance au MTX, le léflunomide ou la sulfasalazine peuvent être prescrits.

Bien que possédant une AMM chez les patients non précédemment traités par MTX, la prescription d'un bDMARD (en association avec le MTX ou en monothérapie) ne se justifie pas en première ligne. La seule situation, dans laquelle la prescription en première ligne d'un bDMARD pourrait être envisagée, serait le cas d'une contre-indication à tous les csDMARDs.

2. Traitement de deuxième ligne et plus, chez les patients insuffisamment répondeurs ou intolérants au MTX :
 - en l'absence de facteurs de mauvais pronostic, une association csDMARDs (méthotrexate et sulfasalazine ou hydroxychloroquine) ou le remplacement par un autre csDMARD (léflunomide, sulfasalazine) peut être proposé. En cas d'échec (ou de contre-indication), un bDMARD ou un tsDMARD doit être envisagé. Les bDMARDs envisageables dans ce cas sont les anti-TNF (adalimumab, certolizumab, étanercept, golimumab, infliximab), un antagoniste des récepteurs de l'interleukine 6 (tocilizumab, sarilumab), un modulateur de la costimulation des lymphocytes T (abatacept) ou le rituximab, uniquement dans certaines circonstances. Les tsDMARDs envisageables sont aujourd'hui le baricitinib et le tofacitinib ;
 - en présence de facteurs de mauvais pronostic (surtout atteinte structurale), l'addition d'un bDMARD ou d'un tsDMARD peut être proposée.
3. L'utilisation d'un bDMARD (ou d'un tsDMARD), quel qu'il soit, doit se faire préférentiellement en association avec le MTX. En cas de nécessité d'utilisation d'un bDMARD en monothérapie, il semble logique de privilégier le tocilizumab ou un tsDMARD, compte tenu de leur supériorité démontrée en monothérapie par rapport au méthotrexate seul. En cas d'échec d'un premier bDMARD, le remplacement par un autre bDMARD se justifie si l'activité de la maladie le nécessite. Les patients en

échec d'un premier anti-TNF peuvent recevoir un deuxième anti-TNF ou un bDMARD reposant sur un autre mode d'action.

4. Aujourd'hui, **les données ne permettent pas d'établir de hiérarchie au sein des bDMARDs au regard des données d'efficacité et/ou de tolérance, et notamment compte tenu de l'absence de facteurs prédictifs de réponses permettant d'orienter le choix du clinicien.** Il convient toutefois de préciser que la pratique actuelle est de débiter par un anti-TNF en raison de l'expérience clinique de 15 ans et de la bonne efficacité structurale de cette classe thérapeutique. L'arrivée des biosimilaires renforce cet usage. Des données d'efficacité ou de tolérance de long terme en vie réelle sur les tsDMARDs seraient très utiles pour observer si cette pratique va changer au profit des tsDMARDs.

6. Conclusion

Ce chapitre a permis d'expliciter les objectifs de l'évaluation économique des traitements de fond biologiques dans la prise en charge de la PR en tenant compte des contextes épidémiologiques et cliniques.

L'approche méthodologique adoptée dans ce projet met l'accent sur la rigueur du processus suivi dans la réalisation des différents travaux et repose sur les actions suivantes :

- la réalisation d'une revue systématique de données et de méta-analyses permettant d'évaluer l'efficacité comparative des bDMARDs (en 1^{re} ligne de traitement biologique) selon les standards méthodologiques scientifiques internationalement reconnus ;
- l'ajout d'une synthèse complémentaire sur l'efficacité comparative des traitements de fond ciblés évalués par la CT de la HAS en 2018 et des méta-analyses récentes (ex. : les méta-analyses de l'agence canadienne publiées en avril 2018). Cette synthèse renseigne sur le positionnement de ces nouveaux traitements par rapport aux classes thérapeutiques actuelles des traitements de fond biologiques ;
- la réalisation d'une revue de littérature complémentaire portant sur les questions principales de l'évaluation pharmaco-économique (ex. : modélisation du parcours du patient atteint d'une PR, prise en compte des séquences des traitements de fond), mais aussi les aspects secondaires liés aux principaux résultats de santé de l'évaluation économique (ex. : éléments contextuels nationaux et internationaux permettant d'appréhender la mortalité associée à la PR) ;
- la conception méthodologique d'un nouveau modèle (appelé « PR France ») adapté à la prise en charge médicamenteuse et la description des principales étapes nécessaires à sa construction.

Chapitre 2. Méthodologie de l'analyse de l'efficacité comparative des traitements de fond biologiques

Une revue systématique de la littérature et des méta-analyses en réseau (MAR) portant sur l'efficacité comparative des traitements de fond biologiques (bDMARDs) dans la prise en charge de la PR en France ont été réalisées dans le cadre d'une collaboration entre la HAS et l'équipe SCHARR (*School of Health and Related Research*) de l'université de Sheffield au cours de la période avril-décembre 2017 (84). Cette collaboration consistait à mettre à jour une revue systématique de la littérature et des méta-analyses réalisées par la même équipe en mai 2013 dans le cadre de l'évaluation du *National Institute for Health and Clinical Excellence* (NICE) portant sur l'adalimumab, l'étanercept, l'infliximab, le certolizumab pegol, le golimumab, le tocilizumab et l'abatacept, dans le traitement de la PR non précédemment traitée ou après échec des traitements de fond synthétiques (csDMARDs) (12).

La revue systématique et les méta-analyses en réseau ont porté sur les deux indications principales suivantes :

- les patients atteints d'une PR active modérée à sévère, précédemment traités par des csDMARDs tels que le méthotrexate (MTX) et présentant une réponse inadéquate/inappropriée ou une tolérance insuffisante associée à l'utilisation de ces csDMARDs ;
- les patients atteints d'une PR sévère n'ayant jamais été traités par des csDMARDs (les patients « naïfs » de MTX).

Il est important de noter que la Commission de la transparence (CT), dans ses avis les plus récents depuis 2016 (avis concernant le certolizumab pégol – 7 septembre 2016 ; avis étanercept – 20 juillet 2016 ; avis adalimumab – 20 juillet 2016 ; avis infliximab – 20 juillet 2016 ; et avis golimumab – 9 novembre 2016), a estimé que, bien qu'ayant une AMM chez les patients non précédemment traités par MTX, les traitements dans l'indication de la PR sévère, active et évolutive de l'adulte non précédemment traité par le méthotrexate, présentaient un service médical rendu insuffisant, au sens de l'article R. 163-16 du Code de la sécurité sociale pour le maintien de leur inscription sur la liste des spécialités remboursables aux assurés sociaux⁸. Ces recommandations de la CT ont été formalisées par une publication au Journal officiel le 10 novembre 2017 (85). **Par conséquent, l'analyse de l'efficacité comparative du MTX concernant les patients naïfs sera**

⁸ La seule option, pour laquelle la prescription en 1^{re} ligne d'un bDMARD pourrait être envisagée, serait le cas d'une contre-indication à tous les DMARDs conventionnels.

présentée dans les annexes du chapitre 3 (annexe A3.11) afin de conforter les futures mises à jour des recommandations de la HAS portant sur la prise en charge de la PR par des données dont le niveau de preuve contribuera à positionner les combinaisons de traitements de fond synthétiques (combinaison de deux ou trois traitements de fond synthétiques) dans le traitement des patients atteints d'une PR active et sévère. **Le groupe de travail a par ailleurs émis un avis favorable à la réalisation d'une évaluation économique sur l'indication principale des bDMARDs, à savoir les patients de la population précédemment traitée par des csDMARDs (ex. : MTX). Cette indication représente les caractéristiques de la population d'entrée dans le modèle PR France.**

Ce chapitre présente les aspects méthodologiques de l'analyse systématique de la littérature et des méta-analyses en réseau réalisées par la HAS et l'équipe SchARR de l'université de Sheffield. Ces choix ont fait l'objet d'un protocole de plan d'analyse validé en avril 2017, d'un premier rapport présenté lors de la deuxième réunion du groupe de travail en juin 2017 et d'un rapport final validé en 2019.

1. Méthodes d'élaboration de la revue systématique de littérature

La revue systématique des données de la littérature a été réalisée conformément aux recommandations *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) (www.prisma-statement.org) (Moher, 2009) (36). Les études portant sur les essais cliniques comparatifs randomisés (ECR) comparant des csDMARDs entre eux ou à un placebo ont été identifiées pour les deux indications d'analyse. Les critères de résultats portant sur les réponses « *American College Rheumatology* » (ACR) ou « *European League Against Rheumatism* » (EULAR) calculées à 24 semaines de suivi ont été retenus. La revue systématique de la littérature a porté sur les bDMARDs suivants : adalimumab « Humira® » (ADA), étanercept « Enbrel® » (ETN), infliximab « Remicade® » (IFX), certolizumab pegol « Cimzia® » (CTZ), golimumab « Simponi® » (GOL), tocilizumab « Roactemra® » (TCZ) et abatacept « Orencia® » (ABT), et les biosimilaires SB2, SB4 et CT-P13.

Il est à noter que l'anakinra « Kineret® » n'a pas été retenu dans les bDMARDs étudiés pour les raisons suivantes :

- dans son avis du 17 septembre 2014 portant sur l'anakinra (86), la Commission de la transparence (CT) a considéré que le service médical rendu était faible dans l'indication de la PR en association avec le MTX chez les adultes dont la réponse au méthotrexate seul n'était pas satisfaisante. Même s'il représente un traitement de

dernier recours dans la prise en charge de la PR, l'anakinra présentait donc un service médical (SMR) faible par rapport aux bDMARDs mentionnés ci-dessus ;

- la faible utilisation de l'anakinra dans la prise en charge de la PR en France. Selon le rapport de la CNAM, en 2011, le pourcentage de prescription de l'anakinra dans les dépenses en ville d'antirhumatisme spécifiques (traitant notamment la PR) était de l'ordre de 3 %.

Par conséquent, en accord avec le groupe de travail PR, la HAS et l'équipe SchARR réalisant les méta-analyses en réseau ne l'ont pas inclus dans leurs analyses statistiques.

1.1. Identification des sources de données

La recherche bibliographique a été réalisée selon une approche méthodologique similaire à celle utilisée par le NICE dans l'évaluation des traitements de fond biologiques en mai 2013 (12). Les termes utilisés dans les recherches Medline et Embase portant sur la PR et les termes d'intervention ont été combinés *via* des filtres pour les essais cliniques et les revues systématiques. Dans les bibliothèques interrogées (ex. : Cochrane), une recherche générique pour les termes et les types d'études a été effectuée. Toutes les recherches de la mise à jour proposées par la HAS et SchARR concernaient la période de mai 2013 à janvier 2017. Les bases de données électroniques interrogées sont présentées dans le tableau 1. Un exemple de la stratégie de recherche est donné dans l'annexe 2-1.

Tableau 1. Bases de données électroniques interrogées dans la revue systématique HAS & SchARR (84)

Bases de données interrogées	Recherche initiale	Mise à jour
Medline and Medline in Process (via Ovid SP)	1948 à mai 2013	Mai 2013 à janvier 2017
Embase (via Ovid SP)	1980 à mai 2013	
Cochrane Database of Systematic Reviews (via Wiley)	1996 à mai 2013	
Cochrane Central Register of Controlled Trials (via Wiley)	1898 à mai 2013	
Health Technology Assessment Database (via Wiley)	1995 à mai 2013	
Database of Abstracts of Reviews of Effect (via Wiley)	1995 à mai 2013	
Cumulative Index to Nursing and Allied Health Literature	1982 à avril 2013	
Toxicology Literature Online	Information en ligne jusqu'à juillet 2013	

En plus de la recherche menée dans les bases de données électroniques, les listes de références des revues systématiques et des essais cliniques randomisés inclus ont été également identifiées. Au besoin, des données complémentaires portant sur des essais cliniques ont été recherchées, soit en consultant les essais cliniques enregistrés sur le site public (*clinical trials.gov*, <https://clinicaltrials.gov>), soit en contactant les auteurs des publications portant sur les essais cliniques analysés.

1.2. Critères de sélection des essais cliniques dans la revue systématique

Dans cette section, les critères de sélection des essais cliniques dans la revue systématique seront présentés pour la population d'entrée du modèle PR France, c'est-à-dire « *les patients atteints d'une PR active modérée à sévère, précédemment traités par des csDMARDs tels que le méthotrexate (MTX) et présentant une réponse inadéquate/inappropriée ou une tolérance insuffisante associée à l'utilisation de ces csDMARDs* ». Les critères utilisés dans la sélection des essais de la population non utilisée dans l'évaluation économique, à savoir « *les patients atteints d'une PR sévère n'ayant jamais été traités par des csDMARDs* », sont décrits dans l'annexe A3.11.

1.2.1. Critères d'inclusion

➤ Population éligible

Les caractéristiques de la population éligible à la revue systématique HAS & SchARR et aux méta-analyses en réseau étaient les suivantes :

- les patients adultes atteints d'une PR active (définie par un critère d'activité de la maladie, DAS 28 \geq 5,1), ou modérée à sévère (définie par un score de DAS entre 3,2 et 5,1), traités par des csDMARDs (ex. : le MTX) ;
- et présentant une réponse inadéquate ou une tolérance insuffisante à la suite de l'administration de ce dernier.

L'appellation simplifiée « *patients ayant été traités par des csDMARDs (MTX)* » sera utilisée pour décrire cette population dans les paragraphes suivants du rapport.

➤ Les interventions

- Les interventions éligibles sont celles prescrites en monothérapie ou en association avec un csDMARD (MTX), conformément aux indications délivrées par l'Agence européenne du médicament (EMA). Seuls les bDMARDs prescrits en première ligne biologique ont été pris en compte dans les analyses. Il est important de souligner que les nouveaux traitements inhibiteurs des Janus kinases (JAK) tels que le baricitinib, le tofacitinib et le sarilumab (inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6) n'ont pas été pris en compte, car ils n'avaient pas été encore autorisés sur le marché au moment de la réalisation de la revue systématique et des méta-analyses en réseau (ou étaient en cours d'évaluation par la CT). Le baricitinib et le sarilumab n'ont pas été intégrés dans les analyses principales des méta-analyses en réseau car les données portant sur ces produits n'étaient pas disponibles au cours de la période d'analyse statistique. Des données portant sur le tofacitinib existaient mais ne respectaient pas les critères d'inclusion (cf. justification des études exclues). Ces dernières ont seulement été incluses dans les analyses de sensibilité des méta-analyses en réseau.
- Concernant les patients traités par des csDMARDs (MTX), les interventions retenues étaient les suivantes :
 - l'adalimumab ;
 - l'étanercept et son biosimilaire : SB4 (Benepali) ;
 - l'infliximab et ses biosimilaires : SB2 (Flixabi), CT-P13 (Inflectra, Remsima) ;
 - le certolizumab pegol ;
 - le golimumab ;
 - l'abatacept (formes IV et SC) ;

- le tocilizumab (formes IV et SC).

Les biosimilaires des molécules mères n'ont pas été intégrés dans les analyses de référence des méta-analyses en réseau. Ceci s'explique par le fait que les « bras » biosimilaires étaient équivalents à ceux de leurs molécules mères. Cependant, ils étaient inclus dans les analyses de sensibilité afin de fournir des estimations utiles pour leur prise en compte dans le modèle économique PR France. Les biosimilaires évalués par la CT après la réalisation de l'analyse des méta-analyses en réseau tels que les biosimilaires de l'étanercept (Erelzi) et du rituximab (Truxima) ou ceux évalués par la CT après juillet 2018 tels que les biosimilaires de l'adalimumab (Amgevita, Imraldi, Hyrimoz et Hulio) n'ont pas été inclus dans les analyses secondaires des méta-analyses en réseau.

➤ **Les comparateurs**

Les traitements comparés peuvent inclure (ou non) l'administration du placebo (selon le *design* de l'essai clinique analysé). Ils différaient selon les deux populations d'analyse citées ci-dessus.

- Les comparateurs considérés pour les patients naïfs de MTX étaient :
 - les associations de csDMARDs contenant le MTX et au moins un autre csDMARD (ex. : sulfasalazine, leflunomide, hydroxychloroquine). Ces dernières ont été nommées « csDMARDs intensifs » et incluaient notamment la trithérapie ;
 - les DMARDs en monothérapie avec escalade de doses ;
 - les bDMARDs comparés entre eux.
- Les comparateurs considérés pour les patients traités par des csDMARDs (ex. : MTX) étaient :
 - les stratégies de comparaison intégrant des csDMARDs (ex. : sulfasalazine, leflunomide), les AINS/les corticoïdes ;
 - les bDMARDs comparés entre eux.

➤ **Les critères d'efficacité comparative**

Les critères d'efficacité comparative retenus étaient ceux les plus fréquemment mentionnés et évalués dans les avis de la CT de la HAS :

- les réponses selon l'*American College of Rheumatology* (ACR) correspondant à une visite d'évaluation réalisée entre 22 et 30 semaines ;
- les réponses selon l'*European League Against Rheumatism* (EULAR) correspondant à une visite d'évaluation réalisée entre 22 et 30 semaines.

Le choix de la période d'évaluation clinique des réponses EULAR et ACR (entre 22 et 30 semaines) a été approuvé par le NICE dans ses évaluations les plus récentes des traitements de fond biologiques (Stevenson *et al.*, 2016) ([12](#)) et conjointement par l'équipe SchARR et le groupe de travail PR. Pour chaque essai, une seule visite de suivi

a été retenue⁹. **Le choix des critères cliniques définis en termes de réponses (notamment la réponse EULAR) s'explique essentiellement par :**

- **l'absence de données suffisantes colligées dans les essais cliniques et/ou l'absence d'un niveau de preuve robuste portant sur le critère rémission. Dans les méta-analyses récentes réalisées par l'Agence canadienne des technologies de santé (CADTH) (32), il y avait 29 études cliniques (25 à deux bras et quatre à trois bras) portant sur le critère rémission parmi un total de 91 études, alors que l'analyse de l'ACR 50 portait sur 57 essais cliniques (48 à deux bras, huit à trois bras et un à un seul bras) ;**
- **l'usage très fréquent des réponses ACR ou EULAR en tant que paramètres cliniques dans les modèles économiques (cf. chapitre 4). En effet, l'utilisation de la rémission n'est pas fréquente car elle risque d'ignorer les gains associés à la transition d'un patient d'un état « moins bon » à un état meilleur (même si le patient est en état de rémission) (Stevenson *et al.*, 2016) (12). En d'autres termes, la rémission clinique exprimée sous forme binaire ne correspond pas à un état sans douleur ; en fait, de nombreux patients en rémission éprouvent encore la douleur, qui est un aspect clé de la qualité de vie rapporté par le patient CADTH (2018) (32).**

➤ **Le *design* de l'essai clinique**

Toutes les études incluses dans les analyses principales des méta-analyses devaient être des essais cliniques comparatifs randomisés (ECR) réalisés sur la population d'analyse (ex. : population en intention de traiter, ITT) planifiée dans les protocoles d'études relatifs aux essais cliniques retenus.

1.2.2. Les critères d'exclusion

Les critères d'exclusion étaient les suivants :

- les études cliniques dont le *design* n'était pas comparatif randomisé (ECR) ;
- les essais cliniques portant sur des patients ayant un score de DAS 28 <3,2 ;
- les essais cliniques portant sur des patients ayant un score de DAS 28 <5,2, s'ils n'avaient pas été précédemment traités par le MTX ;
- les essais cliniques portant sur des patients ayant été traités avec un ou plusieurs bDMARDs.

Il est à noter que :

⁹ Les limites cliniques de ces critères sont abordées dans la discussion des résultats de la méta-analyse en réseau (chapitre 3).

- les essais cliniques dont les arrêts étaient prématurés pouvaient être inclus seulement si une imputation appropriée aux données manquantes s'avérait pertinente et faisable (ex. : imputation des non-répondeurs) ;
- afin de tester si l'hétérogénéité pouvait être en partie expliquée par l'ancienneté de la PR (en tant que modificateur de l'effet du traitement), l'ancienneté de la maladie a été considérée dans l'analyse de méta-régression de la réponse clinique (annexe A3.10).

1.2.3. Critères supplémentaires pour la réalisation des analyses de sensibilité

Même s'ils ne respectaient pas certains critères d'inclusion, des essais cliniques ont été retenus afin de réaliser des analyses de sensibilité permettant d'explorer l'hétérogénéité des critères cliniques retenus (réponses ACR et EULAR) induite par une modification de l'indication des produits ou de fournir des estimations des réponses cliniques relatives aux biosimilaires. Il s'agit des essais cliniques qui :

- rapportaient une proportion faible de patients ayant été traités par des bDMARDs ($\leq 20\%$ approximativement) ;
- incluaient le tofacitinib, permettant de créer un lien dans le réseau de la méta-analyse ;
- incluaient des patients précédemment traités par un csDMARD (ex. : le MTX) dont la durée de réponse était trop courte (ex. : arrêt prématuré) ;
- incluaient des patients précédemment traités par un csDMARD (ex. : le MTX) et d'autres qui étaient naïfs du MTX et représentaient une proportion faible (ex. : pourcentage $\leq 20\%$).
- incluaient les biosimilaires évalués par la CT et disponibles au moment de l'analyse statistique (avant octobre 2017).

1.3. Processus de sélection, d'extraction et d'évaluation des études

L'extraction des données a été réalisée par un premier évaluateur et vérifiée d'une manière indépendante par un second évaluateur.

La qualité méthodologique de chaque étude retenue a été examinée par un premier évaluateur et vérifiée d'une manière indépendante par un second évaluateur. La qualité des études sélectionnées a été analysée selon les critères de *Cochrane risk of bias tool* (87). Le tableau 2 synthétise une grille d'analyse simplifiée qui a été utilisée dans ce travail. Cette grille a été testée et validée dans l'évaluation des traitements de fond biologiques réalisée par le NICE en mai 2013 (12).

Tableau 2. Évaluation de la qualité méthodologique des études retenues dans la revue systématique HAS & SchARR (84)

Critères	Études (essais cliniques randomisés et comparatifs)			
	Étude 1	Étude 2	Étude N
La méthode utilisée pour générer la séquence d'attribution aux groupes de traitement était-elle adéquate ? (O/N/PC)				
L'attribution du traitement a-t-elle été assignée de manière adéquate ? (O/N/PC)				
Les groupes de traitement étaient-ils comparables à l'inclusion ? (O/N/PC)				
La procédure en double aveugle a-elle-été appliquée aux patients et aux investigateurs de l'étude ? (O/N/PC)				
Les évaluateurs des résultats (ACR ou EULAR) étaient-ils aveugles au traitement ? (O/N/PC)				
Les participants ont-ils été analysés dans leurs groupes de traitement assignés ? (O/N/PC)				
Tous les patients randomisés ont-ils été inclus dans les analyses d'efficacité ? (O/N/PC/mITT)				
Au moins 80 % des participants dans l'analyse finale faisaient-ils partie des patients randomisés ? (O/N/PC)				

O=oui ; N=non ; PC=pas clair ; mITT=population en intention de traiter modifiée « *modified intention to treat population* ». La mITT est une sous-population de la population en ITT où certains sujets randomisés sont exclus de manière justifiée (tels que les patients jugés inéligibles après la randomisation ou certains patients n'ayant jamais débuté le traitement).

2. Méthodologie des méta-analyses en réseau planifiées

Les méta-analyses en réseau ont été réalisées :

- en utilisant les données portant sur les **réponses ACR et EULAR** colligées dans la revue systématique dont la méthodologie a été décrite dans la section précédente ;
- en considérant une **approche bayésienne** (*Markov Chain Monte Carlo* (MCMC)) ;
- en retenant un modèle statistique à **effets aléatoires** tenant compte de :
 - o la spécificité des variables EULAR et ACR (variables catégorielles ordonnées),
 - o l'hétérogénéité des effets de traitements observés entre les études retenues.

2.1. Méthodes d'estimation

Ce paragraphe décrit le choix de l'approche d'estimation des méta-analyses en réseau, sa spécification statistique, les méthodes proposées pour analyser la cohérence des résultats, les analyses de sensibilité et la présentation des résultats des méta-analyses en réseau.

2.1.1. Choix d'une approche bayésienne

La réalisation des méta-analyses en réseau est fondée sur une méthode bayésienne « *Markov Chain Monte Carlo* » (MCMC). Un modèle à effets aléatoires a été retenu afin de tenir compte de l'hétérogénéité des effets de traitements intra-études. Le diagnostic Gelman Rubin (Brooks et Gelman, 1998) (88) a été prévu pour analyser la convergence de la MCMC vers des distributions cibles, qui a été évaluée en comparant les estimations entre les différentes chaînes de Markov et les variances intra-chaînes pour chaque paramètre du modèle. Les différences importantes entre ces variances indiquaient une non-convergence.

Toutes les analyses ont été effectuées à l'aide du logiciel R R2WinBUGS/R2OpenBUGS. Un échantillon de 2 000 probabilités aléatoires associées aux effets des traitements absolus a été généré afin qu'il puisse être utilisé dans les analyses de sensibilité probabilistes du modèle PR France élaboré dans ce projet (cf. chapitre n° 5). La taille de cet échantillon s'est avérée suffisante dans l'exécution du modèle économique anglais de la PR réalisé par l'équipe SchARR de l'université de Sheffield (12).

2.1.2. Choix d'une fonction de lien probit

Le modèle logistique est fréquemment utilisé dans les méta-analyses estimant les résultats d'efficacité se présentant sous forme de probabilités dérivées à partir de variables binaires (ex. : présence *versus* absence d'un événement) (Singh *et al.*, 2016) (89). Le modèle logistique est un modèle linéaire généralisé avec une fonction de lien logit. Le modèle linéaire

généralisé peut être appliqué à des variables continues ou à des catégories ordonnées en utilisant d'autres fonctions de lien, telles que l'identité, le logarithme et le probit (37).

Dans notre cas, le choix d'un lien probit s'explique par le fait que la réponse EULAR se présente sous forme de trois catégories ordonnées (absence de réponse, réponse modérée et bonne réponse). De manière similaire, la réponse ACR présente quatre catégories (absence de réponse, ACR20, ACR50 et ACR70). L'ACRXX signifie une amélioration de traitement d'au moins XX %. Le choix d'un lien probit permet d'analyser les catégories d'une façon mutuellement exclusive de telle sorte que les patients ne peuvent être inclus que dans une seule catégorie.

La fonction de lien probit appliquée au traitement des variables catégorielles ordonnées est décrite dans la littérature récente (37). Ses applications sur les réponses EULAR et ACR ont été testées et validées dans l'évaluation du NICE réalisée en 2013 (12).

2.1.3. Estimation des effets relatifs des traitements et des probabilités de réponse

Un lien probit suppose que l'effet du traitement est le même, quelle que soit la catégorie de réponse (aussi bien pour la réponse EULAR que pour la réponse ACR). Soit r_{ikj} le nombre de patients avec réponse dans le bras k de l'étude i dans la catégorie mutuellement exclusive $j = 1, 2, \dots, J$. Les réponses r_{ikj} sont supposées suivre une distribution multinomiale et se présentent comme des séries de distributions binomiales conditionnelles telles qu'une étude i reportant le nombre de patients, r_{ikj} , dans la catégorie $j = 1, \dots, J - 1$:

$$r_{ikj} \sim \text{Binomial}(q_{ikj}, N_{ikj}), j = 1, \dots, J - 1$$

Où :

$q_{ik1} = \text{Prob}(\text{réponse exprimée dans la catégorie 1 de l'étude } i) ;$

$q_{ik2} = \text{Prob}(\text{réponse exprimée dans la catégorie 2 de l'étude } i \mid \text{non dans la catdans } 1)$

...

q_{ikj}

$= \text{Prob}(\text{réponse exprimée dans la catiméede } j \text{ de l'étude } i \mid \text{non dans les categories } 1, 2, \dots, j - 1)$

Et

$$N_{ikj} = \begin{cases} n_{ik}, & \text{pour } j = 1 \\ n_{ik} - \sum_{u=1}^{j-1} r_{iku}, & \text{pour } = 2, \dots, J - 1 \end{cases}$$

Où N_{ikj} était le nombre total de patients dans la catégorie j du bras k de l'étude i , et n_{ikj} , le nombre de patients du bras k de l'étude i .

L'existence d'une variable latente continue est supposée en définissant un seuil z_{ij} , selon lequel un patient est susceptible de changer d'une catégorie à une autre dans une étude i (essai clinique l). Soit p_{ikj} la probabilité de réaliser à un seuil c , dans le bras k de l'étude i une réponse correspondant à la catégorie j .

Un lien probit a été utilisé pour transformer les probabilités p_{ikj} , de la manière suivante :

$$p_{ikj} = 1 - \Phi(\mu_{ij} + \delta_{i,bk}I_{k \neq 1} + z_{ij}).$$

L'effet du traitement était la variation du score probit du bras de contrôle ; $\delta_{i,bk}$ est l'écart type. Les z_{ij} sont supposées distribuées d'une manière aléatoire tout en restant constantes au sein de la même étude mais peuvent varier entre les études :

$$z_{ic} \sim N(v_c, \sigma_z^2).$$

Les effets de traitements spécifiques à l'étude $\delta_{i,bk}I_{k \neq 1}$, ont été supposés suivre une distribution de population commune telle que :

$$\delta_{i,bk} \sim N(d_{t_{ib}, t_{ik}}, \tau^2),$$

Où le traitement de référence d_1 était l'effet du traitement relatif aux csDMARDs dans la catégorie de non-réponse.

Si le nombre de sujets est suffisant, des distributions *a priori* sont considérées. Ceci devrait avoir une faible influence sur les résultats *a posteriori*.

Les distributions *a priori* de référence utilisées dans l'analyse sont :

- les valeurs à l'inclusion propres à chaque étude $\mu_i \sim N(0, 1\ 000)$;
- l'effet du traitement relatif au traitement de référence $d_{1t} \sim N(0, 1\ 000)$;
- l'écart-type inter-études des effets du traitement $\tau \sim U(0,1.1)$;
- le seuil relatif à la population en question $v_{c_j} = v_{c_{j-1}} + v_c', v_c' \sim U(0,5)$;
- l'écart-type inter-études des seuils $\sigma_z^2 \sim U(0,5)$.

Pour estimer les probabilités absolues de réponses associées à chaque catégorie et à chaque traitement, la vraisemblance d'une fonction binomiale est retenue pour caractériser

le nombre de patients r_{ik1} dans chaque étude classée « absence de réponse » lorsqu'il s'agissait des csDMARDs (MTX) de telle sorte que : $r_{ik1} \sim \text{Binomial}(n_{ik}, p_{ik1})$.

Une fonction de lien probit de type $p_{ik1} = \Phi(\mu'_i)$ est utilisée. Il est supposé que les valeurs à l'inclusion propres à chaque étude proviennent d'une population dont les effets sont tels que :

$$\mu'_i \sim N(\mu_b, \tau_b^2).$$

Le modèle est complété en utilisant les paramètres des distributions *a priori* :

- $\mu_b \sim N(0, 1000)$;
- $\tau_b \sim U(0, 1.1)$

Dans le cas où le nombre d'études est relativement faible, une distribution *a priori* informative est associée à l'écart-type inter-études ; une distribution log-normale $(-2.56, 1.74^2)$ peut être retenue pour estimer la variance inter-études si le résultat se présente sous forme d'odds ratio. L'*a priori* retenu est tronqué de la manière suivante : l'odds ratio « rapport des cotes » dans une étude ne doit pas être ≥ 50 fois à un autre et il a été transformé selon une échelle probit. L'*a priori* représentant une « croyance » d'une hétérogénéité, était fixé à 15 % (hétérogénéité faible), 66 % (hétérogénéité modérée) et à 19 % (hétérogénéité élevée).

2.2. Diagnostic de qualité des méta-analyses en réseau

2.2.1. Adéquation des modèles des méta-analyses en réseau

L'évaluation de l'adéquation des modèles est réalisée à l'aide de la déviance résiduelle (D_{res}).

2.2.2. Évaluation de la cohérence des preuves

L'évaluation des éventuelles inconsistances entre les preuves directes et indirectes *via* la méthode *node-splitting* (90) a été prévue dans le cas où des « preuves à réseau fermé » étaient disponibles. En effet, cette technique permet d'estimer l'effet du traitement, soit d_{XY} , selon deux composantes : les preuves directes fondées sur toutes les études comparant les interventions X et Y (d_{XY}^{Dir}), et les preuves indirectes relatives aux données restantes (d_{XY}^{Ind}). L'hypothèse de $d_{XY}^{\text{Dir}} = d_{XY}^{\text{Ind}}$ a été supposée pour examiner la distribution postérieure $\omega = d_{XY}^{\text{Dir}} - d_{XY}^{\text{Ind}}$, et pour calculer $p = (\omega > 0 \text{ or } \omega < 0)$.

2.2.3. Analyse de l'hétérogénéité

En plus de la spécification du modèle de la méta-analyse en réseau en termes d'effets aléatoires mentionnée dans le paragraphe 3.1.1, des méta-régressions et des analyses de sensibilité ont été prévues pour évaluer l'hétérogénéité inter-études.

2.2.3.1. Les méta-régressions

Des méta-régressions ont été prévues pour analyser l'hétérogénéité inter-études et plus particulièrement pour tester l'impact éventuel des principales caractéristiques cliniques et/ou thérapeutiques sur l'effet du traitement. Selon la disponibilité des variables considérées cliniquement pertinentes, il a été proposé de tester si l'ancienneté de la PR représentait un facteur de modification de l'effet du traitement.

2.2.3.2. Les analyses de sensibilité

Des analyses de sensibilité ont été réalisées pour explorer les résultats relatifs aux scénarios dans lesquels les critères d'inclusion n'ont pas été strictement vérifiés. Elles ont porté sur les scénarios suivants :

- inclusion des biosimilaires en supposant qu'ils étaient différents de leurs molécules mères ;
- population des patients traités par des csDMARDs (i.e. MTX) en incluant des études présentant un « mélange » de types de patients : une proportion faible de patients traités par des bDMARDs (pourcentage ≤ 20 %) et des patients naïfs de MTX ;
- population des patients traités par des csDMARDs (i.e. MTX) en incluant des études dont une proportion de patients (pourcentage ≤ 20 %) était naïve de MTX.

2.3. Présentation des résultats

Des méta-analyses en réseau ont été réalisées sur chacune des deux populations. Des analyses principales et de sensibilité (cf. paragraphe 3.2) ont été effectuées :

1. pour chaque intervention, l'estimation des probabilités de réponse (i.e. d'être dans chaque catégorie) était présentée en se fondant sur des distributions *a posteriori*. La médiane et les intervalles de crédibilité à 95 % étaient présentés ;
2. l'estimation de l'effet de chaque traitement relatif par rapport au comparateur de référence (csDMARDs) selon une échelle probit était présentée à l'aide de

graphiques en forêt (*forest plot*). La ligne présentait la valeur de 0. Les valeurs à gauche de la ligne indiquent un effet de traitement positif.

Les résultats (effet du traitement et probabilités de réponse) selon les intervalles de prédiction et les graphiques de classement des traitements en termes d'efficacité seront présentés dans les annexes du chapitre 3.

3. Discussion

La revue systématique de la littérature clinique était une mise à jour de la revue systématique de la littérature clinique de l'évaluation du NICE sur les bDMARDs en 2013 (Stevenson *et al.*, 2016). Elle intégrait en plus des essais cliniques randomisés retenus en avril 2013, ceux identifiés entre mai 2013 et mai 2017 et incluant (pour la première fois) le biosimilaire de l'étanercept (Benepali) et ceux de l'infliximab (Flixabi, Inflectra et Remisma) éligibles aux critères des deux populations d'analyse. Elle a porté sur deux critères cliniques (réponse EULAR et réponse ACR) et deux populations.

L'un des points forts de la méthodologie de notre revue systématique de la littérature était le choix strict et précis des critères cliniques et des populations conformément aux stratégies thérapeutiques de la PR (telles qu'elles sont définies par les critères EULAR et ACR), la sélection des indications existantes dans la prise en charge de la PR en France et la prise en compte des évaluations et des réévaluations des traitements de fond biologiques effectuées par la Commission de la transparence. Ce choix a été validé par les experts cliniques du groupe de travail PR.

L'approche méthodologique adoptée dans le rapport présentait quelques différences avec des revues systématiques de littérature portant sur l'efficacité comparative réalisées par d'autres institutions d'évaluation de technologies de santé telles que l'agence canadienne (CADTH, 2018) (32) et l'organisme américain (ICER, 2017) (13), la Cochrane (ex. : Singh *et al.*, 2016 (91) et Singh *et al.*, 2017 (92)), et d'autres publications de revues systématiques à l'instar de celle de Hazlewood *et al.* (2016) (93). Ces dernières publications incluaient de nombreux résultats (par exemple, la rémission, la progression radiographique) mais n'intégraient pas la réponse EULAR. De plus, la plupart du temps, les populations analysées dans l'évaluation proposée n'étaient pas complètement identiques à celle de l'agence canadienne (2018) ou de l'ICER (2017) (13). Ces deux dernières se sont focalisées sur l'analyse de l'efficacité comparative de la PR active, modérée à sévère chez les patients qui avaient une réponse inadéquate à un csDMARD (i.e. MTX). En ce qui concerne cette dernière population (qui est la population principale à l'entrée du modèle PR), les résultats

des méta-analyses en réseau (chapitre n° 3) étaient assez comparables à ceux de la réponse ACR (i.e. la réponse ACR50) obtenus dans les évaluations de la CADTH (2018) (32) et de l'ICER (2017) (13).

De même, l'approche des méta-analyses en réseau se distinguait des précédentes études mentionnées :

- le choix a été limité à la première ligne de traitement biologique de la PR (c'est-à-dire la deuxième ligne de traitement de la PR, après une réponse inappropriée ou une tolérance insuffisante au MTX) ;
- les essais cliniques des patients atteints de PR sévère et naïfs de MTX et ceux atteints de PR active modérée à sévère et ayant été traités par des csDMARDs (i.e. MTX) ont été analysés de manière séparée ;
- dans les analyses principales des méta-analyses en réseau, les essais présentant un suivi de 22 à 30 semaines ont été pris en compte ;
- la méthode d'estimation des effets de traitement par rapport au csDMARD de contrôle est une méthode originale quantifiant l'effet du traitement à l'aide d'un score de différence de probabilités (Dias *et al.*, 2013) (37) à la place d'un odds ratio. Les réponses ACR et EULAR ont été considérées comme des données catégorielles ordonnées, alors que les analyses de Hazlewood *et al.* (2016) (93), Singh *et al.* (2016) (91), Singh *et al.* (2017) (92) et de la CADTH (2018) (32) ont traité ces critères comme des variables binaires ne prenant pas en compte l'ordre et la corrélation entre les catégories. Dans les méta-analyses produites, le fait de traiter les réponses des patients comme des catégories mutuellement exclusives permet une analyse simultanée des données, y compris pour les études ne fournissant pas d'informations sur certaines catégories, ainsi qu'une estimation unique de l'effet du traitement.

4. Conclusion

Pour obtenir les données d'efficacité comparative à court terme (i.e. évaluations à 6 mois des interventions étudiées) nécessaires à la réalisation du modèle PR France, une revue systématique de la littérature et des méta-analyses en réseau ont fait l'objet d'une collaboration franco-anglaise (la HAS et l'équipe SchARR de l'université de Sheffield).

La méthodologie de la revue systématique se fondait sur un processus reposant sur :

- ✓ une description des critères d'éligibilité fondés principalement sur le critère comparatif et randomisé des essais cliniques ;

- ✓ une recherche documentaire interrogeant plusieurs bases de données électroniques ;
- ✓ une grille d'évaluation de la qualité des essais cliniques, fondée sur la méthodologie PRISMA, testée et validée lors de l'évaluation du NICE réalisée en 2013 ([12](#)).

La méthodologie des méta-analyses en réseau retenue est adaptée à la problématique étudiée sur l'efficacité :

- ✓ des analyses principales (respectant les critères d'éligibilité de la sélection des essais cliniques) et des analyses de sensibilité (intégrant les biosimilaires ou une proportion de patients ne respectant pas certains critères d'éligibilité des deux populations d'analyse) ont été proposées ;
- ✓ la spécification statistique des méta-analyses en réseau tenait compte :
 - des caractéristiques des critères cliniques fréquemment utilisées dans l'analyse de l'efficacité comparative dans la PR telles que les réponses EULAR et ACR. Ces dernières se présentent sous forme de variables catégorielles ordonnées,
 - des critères permettant de diagnostiquer la qualité de l'estimation des méta-analyses fondées sur une approche bayésienne (MCMC) : convergence et qualité d'ajustement des modèles retenus dans les analyses principales et des analyses de sensibilité,
 - des éventuels niveaux d'hétérogénéité inter-études (choix d'un modèle à effet aléatoire, méta-régression et analyses de sensibilité) et des incohérences des estimations.

Chapitre 3. Efficacité comparative : Résultats de la revue systématique de la littérature et des méta-analyses en réseau

Ce chapitre présente les résultats de la revue systématique de la littérature et des méta-analyses en réseau réalisées dans le cadre d'une collaboration entre la HAS et l'équipe ScHARR (*School of Health and Related Research*) de l'université de Sheffield. Ces travaux ont fait l'objet d'un rapport technique et d'un article publié dans le journal *International Journal of Technology Assessment in Health Care* en janvier 2019¹⁰ (Simpson *et al.*, 2019) (94). Les analyses ont porté sur les traitements de fond biologiques (bDMARDs) prescrits dans la prise en charge de la PR en France et concernant les deux indications suivantes :

- les patients atteints d'une PR active modérée à sévère, traités par des traitements de fond synthétiques (csDMARDs) tels que le méthotrexate (MTX) et présentant une réponse inadéquate/inappropriée ou une tolérance insuffisante associée à ces traitements ;
- les patients atteints d'une PR sévère non traités par des csDMARDs (les patients naïfs de MTX). Comme précisé dans l'introduction du chapitre 2, les résultats de la revue systématique de la littérature et des méta-analyses de cette indication non utilisée dans l'évaluation économique sont présentés dans les annexes de ce chapitre.

Les résultats de la revue systématique sont présentés dans la première section. Les résultats des méta-analyses (analyses principales et de sensibilité) sont présentés et interprétés dans la deuxième section. Une discussion des résultats est proposée dans une troisième section.

1. Études retenues

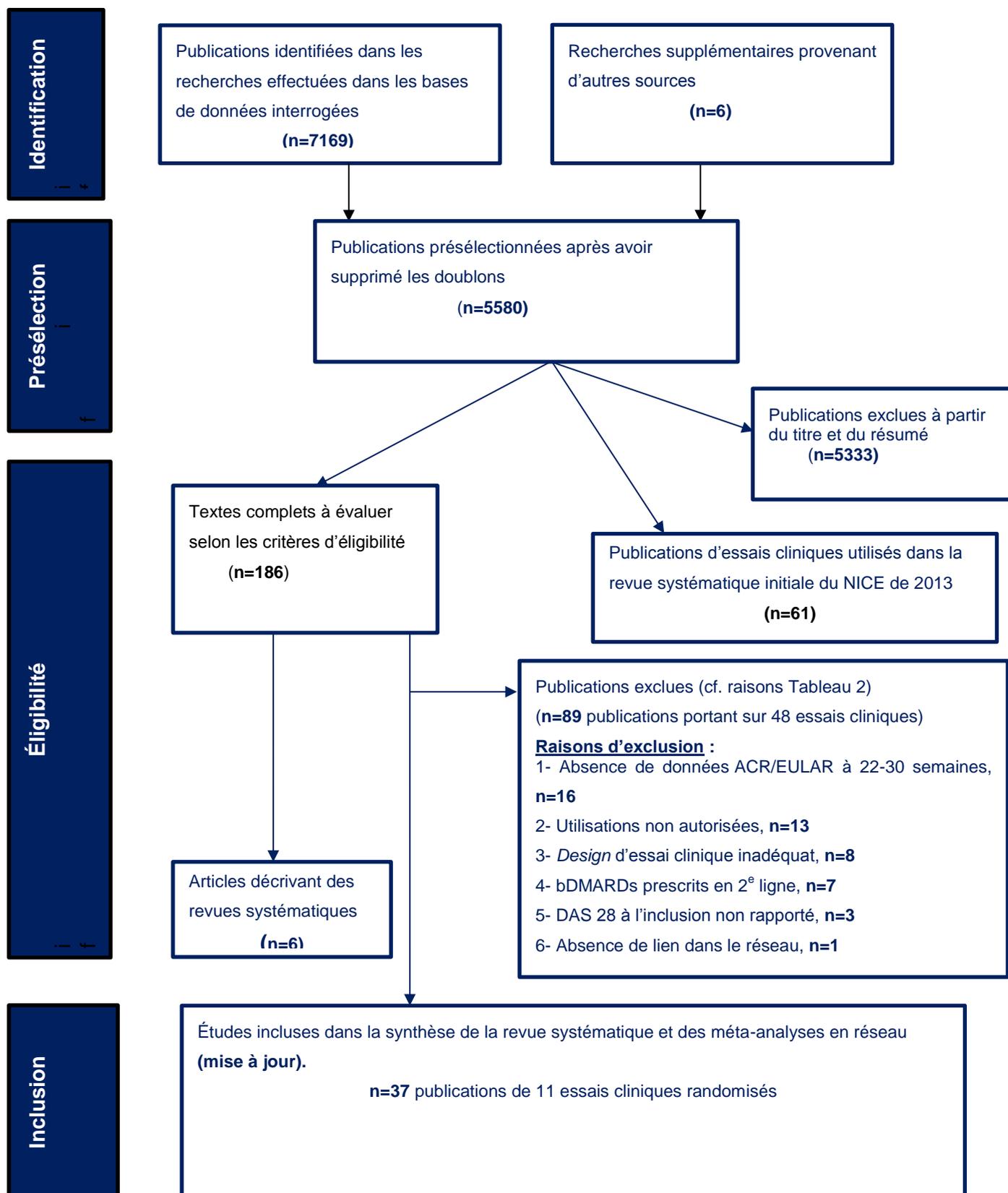
Le processus de sélection des études est présenté dans la figure 1. Sept mille cent soixante-neuf références ont été identifiées dans les bases de données électroniques interrogées. Après suppression des doublons, 5 574 références uniques restaient. Six recherches supplémentaires provenant d'autres sources ont été effectuées (ex. : revues systématiques et essais cliniques randomisés dont les auteurs ont été contactés pour fournir des informations complémentaires sur les caractéristiques des essais cliniques requises pour la revue systématique et la réalisation des méta-analyses en réseau HAS & ScHARR).

¹⁰ Simpson, E, Ren, S, Hock E-S., Stevens, J., Binard A., Pers Y-M., Archel R., Paisley S, Stevenson, M., Herpin C., Ghabri S. Rheumatoid Arthritis treated with 6-month of first-line biological or biosimilar therapy : an updated systematic review and network meta-analysis. *International Journal of Technology Assessment in Health Care*. 2019 ; 35(1). DOI : <https://doi.org/10.1017/S0266462318003628>.

Le filtrage par les titres a permis d'identifier 247 publications potentiellement pertinentes incluant 61 publications portant sur des essais cliniques comparatifs randomisés inclus dans l'évaluation du NICE réalisée en 2013 (12). Ces études sont synthétisées dans l'annexe A3.1 :

- six articles décrivaient des revues systématiques ;
- 48 études cliniques (publiées dans 88 articles) ont été exclues. Les raisons de l'exclusion de ces publications sont présentées dans la figure 1 et détaillées dans l'annexe A3.2 ;
- 37 publications portant sur 11 nouveaux essais cliniques randomisés (figure 1, tableau 1) respectant les critères d'inclusion décrits dans le deuxième chapitre ont été incluses. Les résultats de ces essais cliniques ont été publiés entre mai 2013 et avril 2017.

Figure 1. Diagramme de la sélection des études fondée sur l'organigramme **PRISMA** (36) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)



Le tableau 1 décrit les études portant sur les 11 nouveaux essais cliniques randomisés comparatifs (ECR) répondant aux critères d'éligibilité de la revue systématique. Neuf essais cliniques ont été identifiés pour les patients ayant été traités par des csDMARDs (MTX) :

- quatre essais cliniques portaient sur le TCZ : SURPRISE ([95](#)), BREVACTA ([96](#)), MUSHABI ([97](#)) et SUMMACTA ([98](#)) ;
- deux essais cliniques portaient sur l'ABT (NCT01001832 ([99](#))), et ETN ([100](#)) ;
- trois essais cliniques portaient sur des biosimilaires : SB4 ([101](#)), CT-P13 (PLANETRA ([102](#))) et SB2 (NCT01936181 ([103](#))).

Toutes ces études incluaient des réponses ACR. En revanche, les réponses EULAR étaient seulement colligées dans cinq essais cliniques (HOPEFUL1, SURPRISE, NCT01895309, PLANETRA et NCT01936181).

Tableau 1. Les nouveaux essais cliniques retenus dans la revue systématique de littérature HAS & SchARR. Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Étude	Population	Intervention	Méta-analyse en réseau – Analyse principale	Méta-analyse en réseau – Analyse de sensibilité	Réponses ACR incluses dans l'essai clinique	Réponses EULAR incluses dans l'essai clinique
NCT01001832 Amano <i>et al.</i> , 2015 (99) Matsubara <i>et al.</i> , 2013 (104) Iwahashi <i>et al.</i> , 2014 (105)	Patients ayant été traités par des csDMARDs	ABT	Oui	Non	Oui	Non
NCT00445770 Takeuchi <i>et al.</i> , 2013 (100)	Patients ayant été traités par des csDMARDs	ETN	Oui	Non	Oui	Non
SURPRISE NCT01120366 Takeuchi <i>et al.</i> , 2014a (106) Takeuchi <i>et al.</i> , 2013b (107) Kaneko <i>et al.</i> , 2016 (95)	Patients ayant été traités par des csDMARDs	TCZ	Oui	Non	Oui	Oui

Étude	Population	Intervention	Méta-analyse en réseau – Analyse principale	Méta-analyse en réseau – Analyse de sensibilité	Réponses ACR incluses dans l'essai clinique	Réponses EULAR incluses dans l'essai clinique
BREVACTA Kivitz <i>et al.</i> , 2014 (96)	Patients ayant été traités par des csDMARDs	TCZ	Non	Oui (20 % ayant été traités par des anti-TNF)	Oui	Non
MUSHABI Ogata <i>et al.</i> , 2014 (97)	Patients ayant été traités par des csDMARDs	TCZ	Non	Oui (20 % ayant été traités par des anti-TNF)	Oui	Non
SUMMACTA NCT01194414 Burmester <i>et al.</i> , 2014 (98)	Patients ayant été traités par des csDMARDs	TCZ	Non	Oui (20 % ayant été traités par des anti-TNF)	Oui	Non

Étude	Population	Intervention	Méta-analyse en réseau – Analyse principale	Méta-analyse en réseau – Analyse de sensibilité	Réponses ACR incluses dans l'essai clinique	Réponses EULAR incluses dans l'essai clinique
NCT01895309 Emery <i>et al.</i> , 2017 (101) Emery <i>et al.</i> , 2016b (108) Emery <i>et al.</i> , 2016c (109) Vencovsky <i>et al.</i> , 2015a (110)	Patients ayant été traités par des csDMARDs	Biosimilaire ETN (SB4) Benepali	Biosimilaire	Biosimilaire	Oui	Oui

<p>PLANETRA NCT01217086 Hanyang <i>et al.</i>, 2013 (111) Lee <i>et al.</i>, 2014 (112) Yoo <i>et al.</i>, 2013 (113) Yoo <i>et al.</i>, 2013b (114) Yoo <i>et al.</i>, 2014 (115) Yoo <i>et al.</i>, 2016 (102)</p>	<p>Patients ayant été traités par des csDMARDs</p>	<p>Biosimilaire IFX (CT-P13) Inflectra/ Remsima</p>	<p>Biosimilaire</p>	<p>Biosimilaire</p>	<p>Oui</p>	<p>Oui</p>
<p>Choe <i>et al.</i>, 2015b (116) Choe <i>et al.</i>, 2016 (117) Durez <i>et al.</i>, 2016 (118) Smolen <i>et al.</i>, 2016 (119) Choe <i>et al.</i>, 2017 (103)</p>	<p>Patients ayant été traités par des csDMARDs</p>	<p>Biosimilaire IFX (Flixabi, SB2)</p>	<p>Biosimilaire</p>	<p>Biosimilaire</p>	<p>Oui</p>	<p>Oui</p>

2. Évaluation de la qualité des essais cliniques retenus dans la mise à jour de la revue systématique réalisée par la HAS

La qualité des essais cliniques retenus a été examinée en se fondant sur une grille d'analyse des risques de biais fondée sur la méthodologie des revues systématiques extraite à partir de la bibliothèque Cochrane (*Cochrane Risk of Bias*) (87). Toutes les interventions ont été commercialisées conformément aux autorisations de mise sur le marché (AMM) accordées par l'Agence européenne du médicament (EMA) et les autorités de santé françaises. Quatre essais cliniques présentaient des csDMARDs comme comparateurs (HOPEFUL1, IDEA, NCT00445770, BREVACTA). Trois essais cliniques (SUMMACTA, MUSASHI et NCT01001832) ont comparé différentes administrations de bDMARDs. Un essai clinique (SURPRISE) a comparé un bDMARD en monothérapie et en association avec le MTX. Quant aux essais cliniques portant sur les biosimilaires, ces derniers ont été comparés avec leur molécule mère (bDMARDs).

La qualité des essais cliniques était satisfaisante. Il existait généralement un faible risque de biais entachant ces derniers. Les tableaux 2 et 3 donnent les détails de la qualité des essais cliniques et la figure 2 quantifie le risque de biais afférent aux différentes études. Les items pour lesquels le risque de biais était le plus élevé étaient ceux portant sur la description de la procédure du double aveugle, la description de la procédure de randomisation et la comparabilité de certaines caractéristiques des populations entre les différentes études analysées.

Les deux essais cliniques portant sur les patients naïfs de MTX présentaient un faible risque de biais, bien que la procédure de randomisation de HOPEFUL1 (106) n'ait pas été explicitée (cf. tableau 2). De même, les essais cliniques portant sur les patients ayant été traités par des csDMARDs (MTX) présentaient un faible risque de biais (tableau 3), et ce, même si la procédure de randomisation n'a pas été clairement rapportée dans cinq des études : NCT01001832 (104), BREVACTA (96), NCT00445770 (100), SURPRISE (95), MUSASHI (97).

Un seul essai était ouvert (SURPRISE) (95). Les essais cliniques restants (8/11) ont été réalisés en double aveugle.

Parmi les essais cliniques, 90 % (10/11) ont porté sur des analyses en intention de traiter (ITT) ou en intention de traiter modifiée (mITT). Pour deux essais (SURPRISE (95), NCT01936181 (103)), les analyses d'efficacité ont été réalisées sur les populations ITT ou

MITT. Une seule étude (SASHI (97)) a été réalisée sur la population « per protocole » (PP) dans laquelle les patients randomisés inclus dans l'analyse d'efficacité étaient ceux qui ne présentaient pas de déviations majeures au protocole de l'étude (ex. : traitements concomitants, non-respect de la dose administrée, pas d'arrêt prématuré du traitement). Il est également à noter que des proportions très élevées de patients randomisés ont été incluses dans les analyses (SURPRISE 226/233 (96,9 %), NCT01936181 583/584 (99,8 %), MUSASHI 315/346 (91 %)).

La description de la randomisation présentait un risque élevé de biais ; la description de la procédure de randomisation a été peu explicitée. Parmi les essais analysés, 42 % d'entre eux ont rapporté une méthode adéquate de « *masquage des traitements attribués* », et de génération des séquences de randomisation.

Les caractéristiques des deux populations à l'inclusion sont présentées dans les tableaux A1 et A2 de l'annexe A3.3 ; les caractéristiques des patients à l'inclusion telles que l'âge, le sexe et le DAS 28 étaient globalement comparables. Il y avait quelques différences entre les essais cliniques portant sur les traitements antérieurs (pris par les patients avant l'inclusion à l'étude), par exemple, le nombre et les types de csDMARDs antérieurs. Cependant, lorsque les patients ont été regroupés par type de bDMARD prescrit, une « comparabilité » entre les essais cliniques en termes de caractéristiques de population et de qualité d'études a été globalement observée.

Les tableaux A3 et A4 de l'annexe A3.3 décrivent les principales caractéristiques des *designs* des études cliniques retenues : de nombreux essais cliniques incluaient des périodes où les patients ont été randomisés (période principale de l'essai clinique) suivies d'études d'extension de long terme. Seules les périodes randomisées ont été retenues. Les tableaux A4 et A5 de l'annexe A3.3 présentent les traitements antérieurs colligés lors des visites d'inclusion relatives aux différentes études retenues. Les informations sur les stratégies et les modalités d'administration des traitements sont présentées dans les tableaux A6 et A7 de l'annexe A3.3.

Par ailleurs, une probable hétérogénéité entre les essais pouvait être liée à l'ancienneté de la PR (annexe A3.4). L'ancienneté de la PR en tant que facteur de modification potentiel de l'effet du traitement a été explorée dans les méta-régressions concernant les réponses ACR. Les résultats de la section 2.5 et l'annexe A3.6 portent sur cette problématique.

Figure 2. Appréciation du risque de biais dans les études de la revue systématique HAS & ScHARR – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

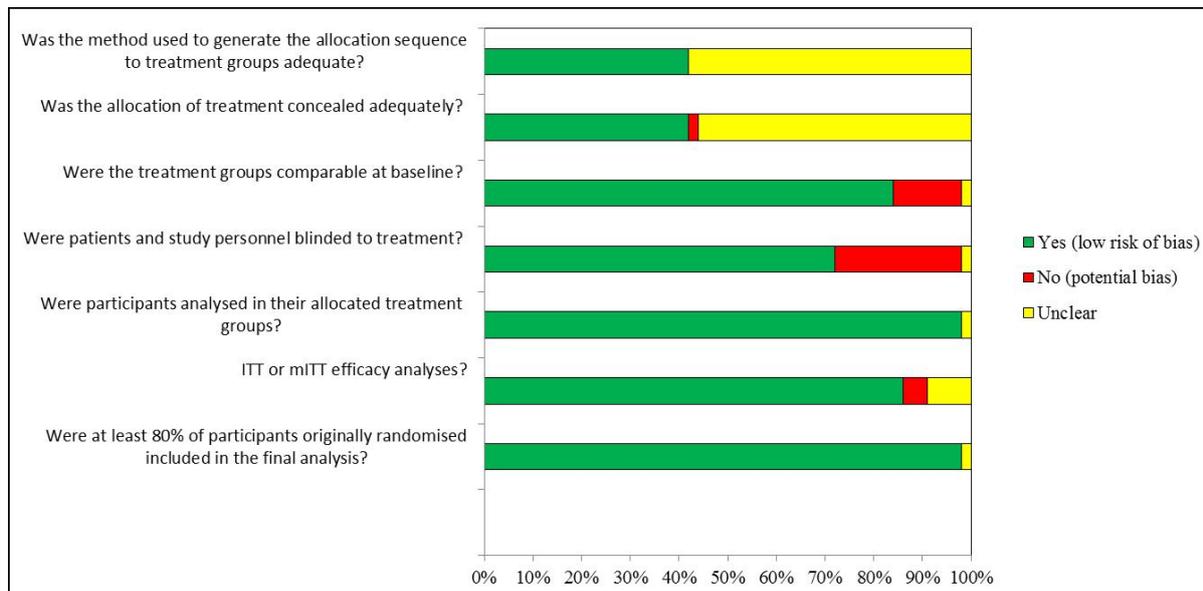


Tableau 2. Évaluation de la qualité des études retenues dans la revue systématique de littérature HAS & SchARR – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

<i>Item/Question</i>	NCT01001832 Matsubara <i>et al.</i> , 2013 (104) Iwahashi <i>et al.</i> , 2014 (105)	NCT004457 70 Takeuchi <i>et al.</i> , 2013 (100)	SURPRISE NCT01120366 Kaneko <i>et al.</i> , 2016 (95)	BREVACTA Kivitz <i>et al.</i> , 2014 (96)	MUSASHI Ogata <i>et al.</i> , 2014 (97)	SUMMACTA NCT011944 14 Burmester <i>et al.</i> , (98)	NCT018953 09 Emery <i>et al.</i> , 2017 (101)	PLANETRA NCT012170 86 Yoo <i>et al.</i> , 2016 (102)	NCT01936181 Choe <i>et al.</i> , 2017 (103)
La méthode utilisée pour générer la séquence d'attribution aux groupes de traitement était-elle adéquate ? (O/N/I)	I	O	I	O	I	O	O	O	O
L'attribution du traitement a-t-elle été assignée de manière adéquate ? (O/N/I)	I	I	O	I	I	O	O	O	O
Les groupes de traitement étaient-ils comparables à l'inclusion ? (O/N/I)	O	O	O	O	O	O	O	O	O

La procédure de double aveugle a-t-elle été appliquée aux patients et au personnel de l'étude ? (O/N/I)	O	O	N	O	O	O	O	O	O
Les évaluateurs des résultats (ACR ou EULAR) étaient-ils aveugles au traitement ? (O/N/I)	O	O	N*	O	O	O	O	O	O
Les patients ont-ils été analysés dans leurs groupes de traitement assignés ? (O/N/I)	O	O	O	O	O	O	O	O	O
Tous les patients randomisés ont-ils été inclus dans les analyses d'efficacité ? (O/N/I/mITT)	O	mITT	mITT	O	N	O	O	O	O
Au moins 80 % des patients dans l'analyse finale	O	O	O	O	O	O	O	O	mITT

faisaient-ils partie des patients randomisés ? (O/N/I)									
--	--	--	--	--	--	--	--	--	--

Légende : O : oui ; N : non ; I : inconnue ; mITT : Population en intention de traiter modifiée.

* Pour certains résultats de cette étude, les évaluateurs étaient aveugles, mais les évaluations des résultats de l'ACR et de l'EULAR n'ont pas été faites en aveugle.

3. Les critères d'efficacité retenus dans la revue systématique de littérature HAS & ScHARR

Cette section décrit les réponses portant sur les critères ACR et EULAR colligées dans les essais cliniques de la revue systématique de littérature HAS & ScHARR. Les données provenant des essais cliniques inclus dans la revue systématique initiale (12) sont présentées dans l'annexe A3.5.

3.1. Réponses ACR colligées dans la revue systématique de littérature HAS & ScHARR

Le tableau 3 décrit les réponses ACR portant sur les essais cliniques concernant les patients traités par des csDMARDs (MTX) dans les méta-analyses principales. Aucune différence significative entre les groupes de traitement dans l'essai NCT01001832 (ABT s.c. vs i.c.) ou SURPRISE (TCZ monothérapie vs association avec MTX) n'a été mise en évidence. Une différence clinique significative relative aux réponses ACR de l'intervention étudiée a été identifiée par rapport au comparateur retenu dans l'essai NCT00445770 (ETN vs. MTX). L'essai clinique NCT00202852 a montré une différence significative en faveur de l'IFX par rapport au MTX concernant les critères de réponse ACR20 (les réponses ACR50 et ACR70 n'étaient pas rapportées).

Les réponses ACR rapportées dans les essais inclus dans les analyses de sensibilité des principales méta-analyses en réseau sont présentées dans le tableau 4. L'essai BREVACTA a montré une différence significative pour le TCZ vs. placebo. Les réponses ACR étaient comparables dans les groupes de traitement de l'essai SUMMACTA (TCZ s.c. vs i.v.) ou MUSASHI (TCZ s.c. vs i.v.). Les biosimilaires ont été considérés équivalents aux molécules de référence dans les essais cliniques NCT01895309, PLANETRA et NCT01936181.

Tableau 3. Données ACR colligées dans les nouvelles études de la revue systématique de la littérature HAS & ScHARR – Analyse principale – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Étude / Auteur, année	Bras de traitement	Visite de l'évaluation	N ^{bre} de patients analysés	Nombre de patients avec une réponse ACR20	% de patients avec une réponse ACR20	Nombre de patients avec une réponse ACR50	% de patients avec une réponse ACR50	Nombre de patients avec une réponse ACR70	% de patients avec une réponse ACR70
NCT01001832 Matsubara <i>et al.</i> , 2013 (104)	ABT s.c. + MTX (n=59)	24 semaines	59	54	91,5	39	66,1	22	37,3
	ABT i.v. + MTX (n=59)	24 semaines	59	49	83,1	37	62,7	18	30,5
NCT00445770 Results from clinicaltrials.gov Takeuchi <i>et al.</i> , 2013 (100)	MTX + PBO (n=176)	24 semaines	176	99	56,3 ^b	57	32,4 ^b	21	11,9 ^b
	ETN mono 25mg twice weekly + PBO (n=182)	24 semaines	182	141	77,5	99	54,4	47	25,8
SURPRISE NCT01120366 Kaneko <i>et al.</i> , 2016 (95)	TCZ i.v. mono (n=115)	24 semaines	111	77	69,4	60	54,1	38	34,2
	TCZ i.v. + MTX (n=118)	24 semaines	115	86	74,8	63	54,8	38	33,0

^a = P<0.05, ^b = P<0.001 différence statistiquement significative entre l'intervention et le comparateur.

		s.c. TCZ plus PBO (i.v.) plus csDMARD (n=631)								
			24 semaines	631	429	68	290	46	151	24
NCT01895309 Emery <i>et al.</i> , 2017 (101)	Biosimilaire	ETN biosim (SB4) plus MTX (n=299)								
			24 semaines	298	220	73,8	128	43,0	69	23,2
		ETN 50 mg weekly plus MTX (n=297)								
			24 semaines	297	213	71,7	116	39,1	59	19,9
PLANETRA NCT01217086 Yoo <i>et al.</i> , 2013 (102)	Biosimilaire	IFX biosim (CT-P13) plus MTX (n=302)								
			30 semaines	302	184	60,9	106	35,1	50	16,6
		IFX + MTX (n=304)								
			30 semaines	304	178	58,6	104	34,2	47	15,5
NCT01936181 Choe <i>et al.</i> , 2017 (103)	Biosimilaire	IFX biosim (SB2) + MTX (n=291)								
			30 semaines	290	161	55,5	89	30,7	45	15,5
		IFX + MTX (n=293)								
			30 semaines	293	173	59,0	99	33,8	50	17,1

* $p < 0.001$, différence statistiquement significative entre l'intervention et le comparateur.

**Les données graphiques ont été estimées en utilisant le logiciel Engauge [version 4.1 ; Mark Mitchell, Los Angeles, CA, USA (2011)].

Tableau 4. Données ACR colligées dans les nouveaux essais cliniques dans la revue systématique de la HAS – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Étude / Auteur, année	Raison de l'analyse de sensibilité	Bras de traitement	Visite de l'évaluation	N ^{bre} de patients analysés	Nombre de patients avec une réponse ACR20	% de patients avec une réponse ACR20	Nombre de patients avec une réponse ACR50	% de patients avec une réponse ACR50	Nombre de patients avec une réponse ACR70	% de patients avec une réponse ACR70
BREVACTA Kivitz <i>et al.</i> , 2014 (96)	20 % des patients ayant utilisé des anti- TNF	TCZ (sc) + csDMARDs	24 semaines	437	266	61*	175	40	87	20
		PBO (sc) + csDMARDs	24 semaines	219	70	32	26	12	11	5
MUSASHI Ogata <i>et al.</i> , 2014 (97)	20 % des patients ayant utilisé des anti- TNF	i.v. TCZ plus PBO (s.c.) (n=173)	24 semaines	156	138	88,5	105**	67,1**	63**	40,4**
		s.c. TCZ plus PBO (i.v.) (n=173)	24 semaines	159	126	79,2	101**	63,5**	59**	37,3**
SUMMACTA NCT01194414 Burmester <i>et al.</i> , 2014 (98)	20 % des patients ayant utilisé des anti- TNF	i.v. TCZ plus PBO (s.c.) plus csDMARD (n=631)	24 semaines	631	442	70	297	47	170	27
		s.c. TCZ plus PBO (i.v.) plus csDMARD (n=631)	24 semaines	631	429	68	290	46	151	24

Étude / Auteur, année	Raison de l'analyse de sensibilité	Bras de traitement	Visite de l'évaluation	N ^{bre} de patients analysés	Nombre de patients avec une réponse ACR20	% de patients avec une réponse ACR20	Nombre de patients avec une réponse ACR50	% de patients avec une réponse ACR50	Nombre de patients avec une réponse ACR70	% de patients avec une réponse ACR70
NCT01895309 Emery <i>et al.</i> , 2017 (101)	Biosimilaire	ETN biosim (SB4) plus MTX (n=299)	24 semaines	298	220	73,8	128	43,0	69	23,2
		ETN 50 mg weekly plus MTX (n=297)	24 semaines	297	213	71,7	116	39,1	59	19,9
PLANETRA NCT01217086 Yoo <i>et al.</i> , 2013 (114)	Biosimilaire	IFX biosim (CT-P13) plus MTX (n=302)	30 semaines	302	184	60,9	106	35,1	50	16,6
		IFX + MTX (n=304)	30 semaines	304	178	58,6	104	34,2	47	15,5
NCT01936181 Choe <i>et al.</i> , 2017 (103)	Biosimilaire	IFX biosim (SB2) + MTX (n=291)	30 semaines	290	161	55,5	89	30,7	45	15,5
		IFX + MTX (n=293)	30 semaines	293	173	59,0	99	33,8	50	17,1

*p<0.001, différence statistiquement significative entre l'intervention et le comparateur.

**Les données graphiques ont été estimées en utilisant le logiciel Engauge [version 4.1 ; Mark Mitchell, Los Angeles, CA, USA (2011)].

3.2. Réponses EULAR colligées dans la revue systématique HAS & ScHARR

Un seul essai clinique, SURPRISE, rapportait une réponse EULAR dans l'analyse principale de la méta-analyse en réseau (tableau 5). La différence entre les bras de traitement n'était pas statistiquement significative au seuil $\alpha=5\%$ ($p=0,06$). Trois essais cliniques portant sur les biosimilaires rapportaient une réponse EULAR (tableau 6) démontrant l'équivalence entre le biosimilaire et le médicament de référence.

Tableau 5. Réponses EULAR colligées dans les nouvelles études de la revue systématique de littérature HAS & SchHARR – Analyse principale – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Étude / Auteur, année	Bras de traitement	Temps d'évaluation	N ^{bre} de patients analysés	N ^{bre} de patients avec absence de réponse EULAR	% de patients avec absence de réponse EULAR	N ^{bre} patients avec une réponse EULAR modérée/bonne	% de patients avec une réponse EULAR modérée/bonne
SURPRISE NCT01120366 Kaneko <i>et al.</i> , 2016 (95)	TCZ mono (n=115)	24 semaines	111	11	9,9	100	90,1
	TCZ + MTX (n=118)	24 semaines	115	4	3,5	111	96,5

Tableau 6. Données EULAR colligées dans les nouvelles études de la revue systématique de littérature HAS & SchHARR – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Nom de l'étude / Auteur, année	Raison de l'analyse de sensibilité	Bras de traitement	Temps d'évaluation	N ^{bre} de patients analysés	% de patients avec absence de réponse EULAR	N ^{bre} de patients avec absence de réponse EULAR	% de patients avec une réponse EULAR modérée	N ^{bre} de patients avec une réponse EULAR modérée	N ^{bre} de patients avec une bonne réponse EULAR	% de patients avec une bonne réponse EULAR
NCT01895309 Emery <i>et al.</i> , 2017 (101)		ETN biosim (SB4) plus MTX (n=299)	24 semaines	298	12,8	38	55,1	164	96	32,1
		ETN 50 mg								

Nom de l'étude / Auteur, année	Raison de l'analyse de sensibilité	Bras de traitement	Temps d'évaluation	N ^{bre} de patients analysés	% de patients avec absence de réponse EULAR	N ^{bre} de patients avec absence de réponse EULAR	% de patients avec une réponse EULAR modérée	N ^{bre} de patients avec une réponse EULAR modérée	N ^{bre} de patients avec une bonne réponse EULAR	% de patients avec une bonne réponse EULAR
	Biosimilaire	hebdomadaire plus MTX (n=297)	24 semaines	297	11,8	35	58,5	174	89	29,9
PLANETRA NCT01217086 Yoo <i>et al.</i> , 2016 (102)	Biosimilaire	IFX biosim (CT-P13) plus MTX (n=302)	30 semaines	245	15,5	38	58,4	143	64	26,1
		IFX + MTX (n=304)	30 semaines	249	17,7	44	58,2	145	60	24,1
NCT01936181 Choe <i>et al.</i> , 2017 (103)	Biosimilaire	IFX biosim (SB2) + MTX (n=291)	30 semaines	253	16,2	41	58,1	147	65	25,7
		IFX + MTX (n=293)	30 semaines	265	19,6	52	54,7	145	68	25,7

^a = P<0.05, ^b = P<0.001 différence statistiquement significative entre l'intervention et le comparateur.

4. Résultats des méta-analyses en réseau HAS & SchARR

Les méta-analyses en réseau principales ont porté sur :

- 38 essais cliniques issus de la revue initiale (tableau A1, annexe A3.6) ;
- et cinq essais cliniques issus de la mise à jour de la revue systématique de la littérature (tableau A2, annexe A3.6).

Les méta-analyses en réseau principales concernant les patients ayant été traités par des csDMARDs (indication de la population d'entrée du modèle PR France) ont porté sur 33 essais cliniques. En plus des essais retenus dans les méta-analyses en réseau principales, les analyses de sensibilité ont porté sur :

- 14 essais cliniques de la revue initiale (tableau A3, annexe A3.6) ;
- et six nouveaux essais cliniques recensés lors de la mise à jour de la revue systématique de la littérature (tableau A4, annexe A3.6).

4.1. Les réseaux de preuves des méta-analyses en réseau

Les réseaux de preuves associés aux réponses ACR et EULAR des analyses principales relatives aux deux populations étudiées sont présentés dans les figures 3 et 4. Quant aux réseaux de preuves relatifs aux analyses de sensibilité, ils sont présentés dans l'annexe A3.7.

Figure 3. Réseau de preuves : Réponse ACR – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

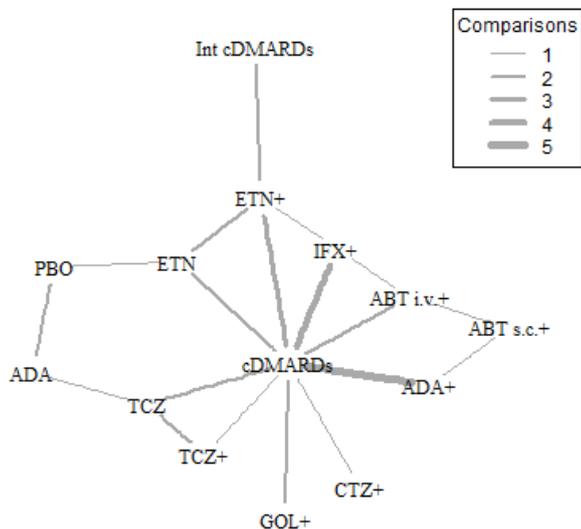
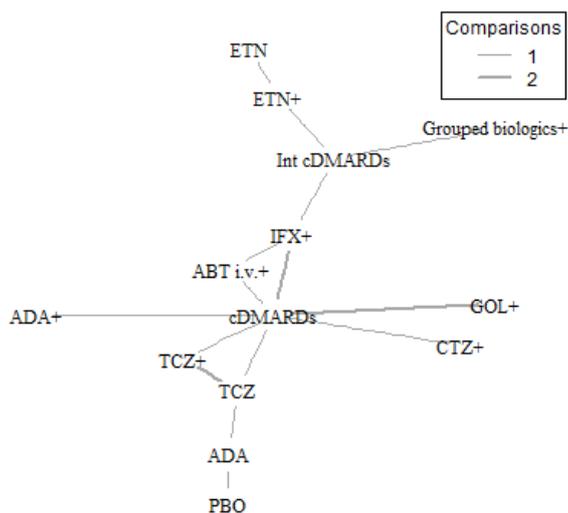


Figure 4. Réseau de preuves : Réponse EULAR – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)



Légende : (+) indique une association avec les csDMARDs (i.e. MTX) ; Int cDMARDs : association de 2 ou 3 csDMARDs ; Grouped biologics : choix du clinicien d'utiliser ADA ou ETN ou IFX (avec MTX) dans l'essai clinique TACIT.

4.2. Estimation des effets de traitements et des probabilités de réponse

La lecture et l'interprétation des résultats se fondent sur les éléments méthodologiques suivants :

- dans le paragraphe 4.2.1, un premier aperçu est donné sur la qualité des ajustements des méta-analyses en réseau ;
- les résultats des analyses principales (paragraphe 4.2.2 et 4.2.3) et les analyses de sensibilité (section 4.3) :
 - o les effets de chaque intervention par rapport aux comparateurs (ex. : csDMARDs) dérivés des distributions *a posteriori* ont été exprimés selon une échelle probit et représentés graphiquement à l'aide de *forest plot* (graphiques en forêt),
 - o **le comparateur commun le plus fréquent est un csDMARD (ex. : MTX). Une différence négative relative à une intervention par rapport au comparateur (ex. : MTX) exprimait un effet positif de l'intervention en question par rapport au comparateur, à savoir une plus faible proportion de patients dans les catégories inférieures (ex. : non-réponse) ;**
- les probabilités absolues de réponse (ACR ou EULAR) ont également été fondées sur des distributions *a posteriori*. Elles sont présentées à l'aide de la médiane et de l'intervalle de crédibilité à 95 % (CrI) ;
- la section 4.4 donnera des détails complémentaires sur la cohérence des preuves et l'analyse de l'hétérogénéité inter-études ;
- les résultats des distributions prédictives des effets du traitement sont présentés dans l'annexe A3.8. Les probabilités de « classement » des traitements sont présentées dans l'annexe A3.9.

4.2.1. Évaluation de la qualité des ajustements des méta-analyses en réseau

Le tableau 7 présente les indicateurs de la qualité des ajustements. Pour toutes les analyses principales, les modèles retenus ajustaient les données d'une manière adéquate : la déviance résiduelle totale était très proche du nombre total des observations.

Tableau 7. Qualité des ajustements des méta-analyses en réseau HAS & SchARR – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Résultat clinique / indicateur statistique	Déviance résiduelle totale	Nombre total des observations
Réponse ACR	213	194
Réponse EULAR	63	54

4.2.2. Analyse principale – Patients ayant été traités par un csDMARD (MTX)

➤ Réponses ACR à la semaine 24

Des « graphiques en forêt » décrivant l'effet de chaque intervention par rapport aux csDMARDs selon une échelle probit sont présentés dans les figures 5a-5d. Les probabilités de réalisation d'au moins une réponse (ACR20 ou une réponse ACR50 ou ACR70) sont présentées dans le tableau 8.

Au niveau $\alpha=5\%$, les effets du traitement étaient statistiquement significatifs pour toutes les interventions (sauf pour l'ADA et les csDMARDs intensifs qui étaient situés à la limite de la significativité statistique du niveau $\alpha=5\%$). Les TCZ (avec et sans MTX) apparaissaient comme les interventions présentant les réponses ACR les plus élevées par rapport aux csDMARDs (médiane de rang 2, probabilité d'être la meilleure : respectivement 0,25 pour le TCZ et 0,29 pour le TCZ+, cf. annexe A3.9).

Figure 5a. Réponses ACR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

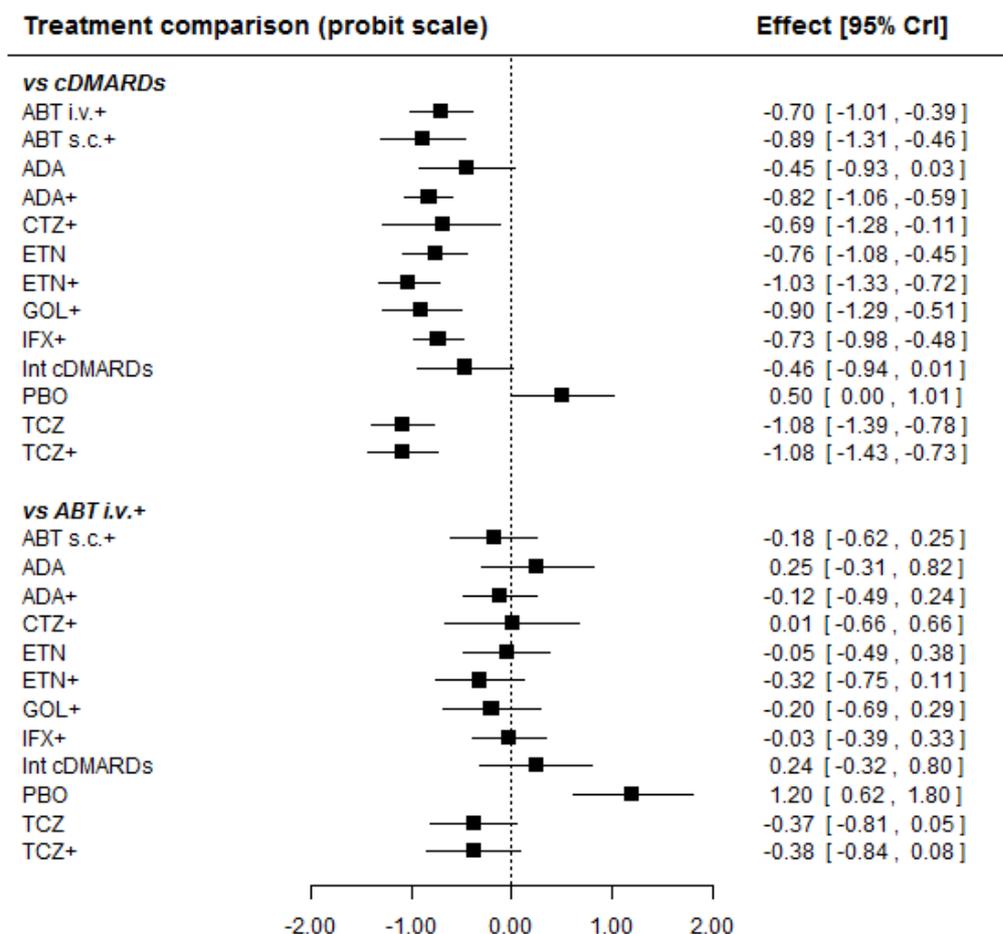


Figure 5b. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

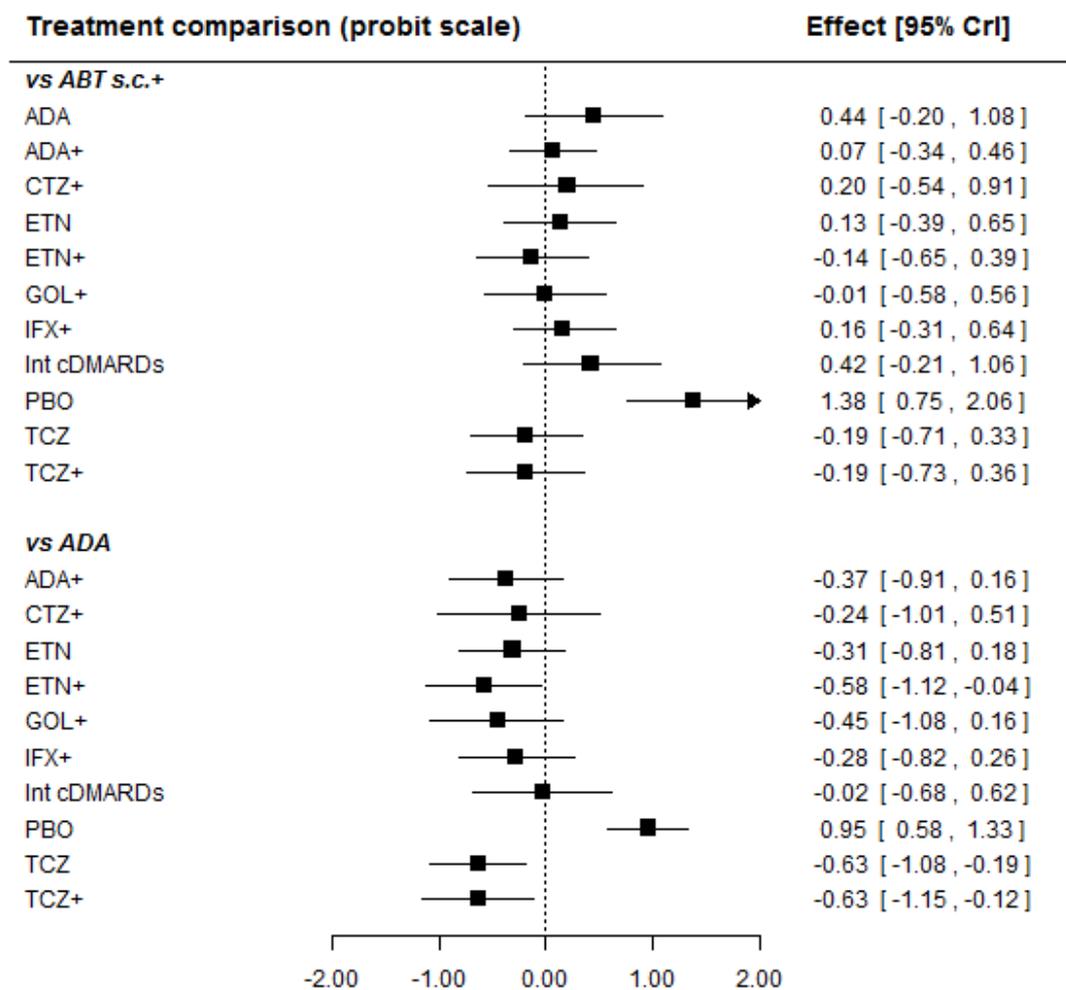


Figure 5c. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

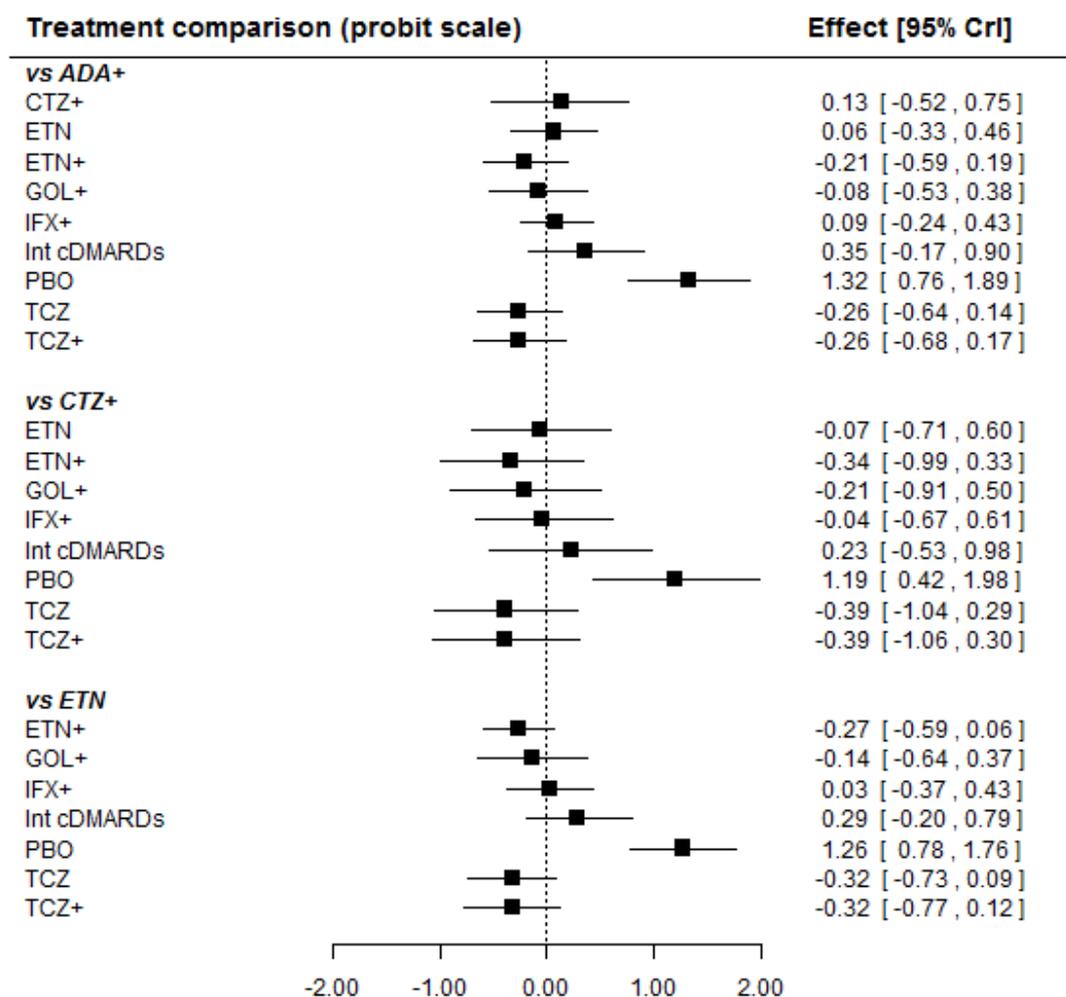
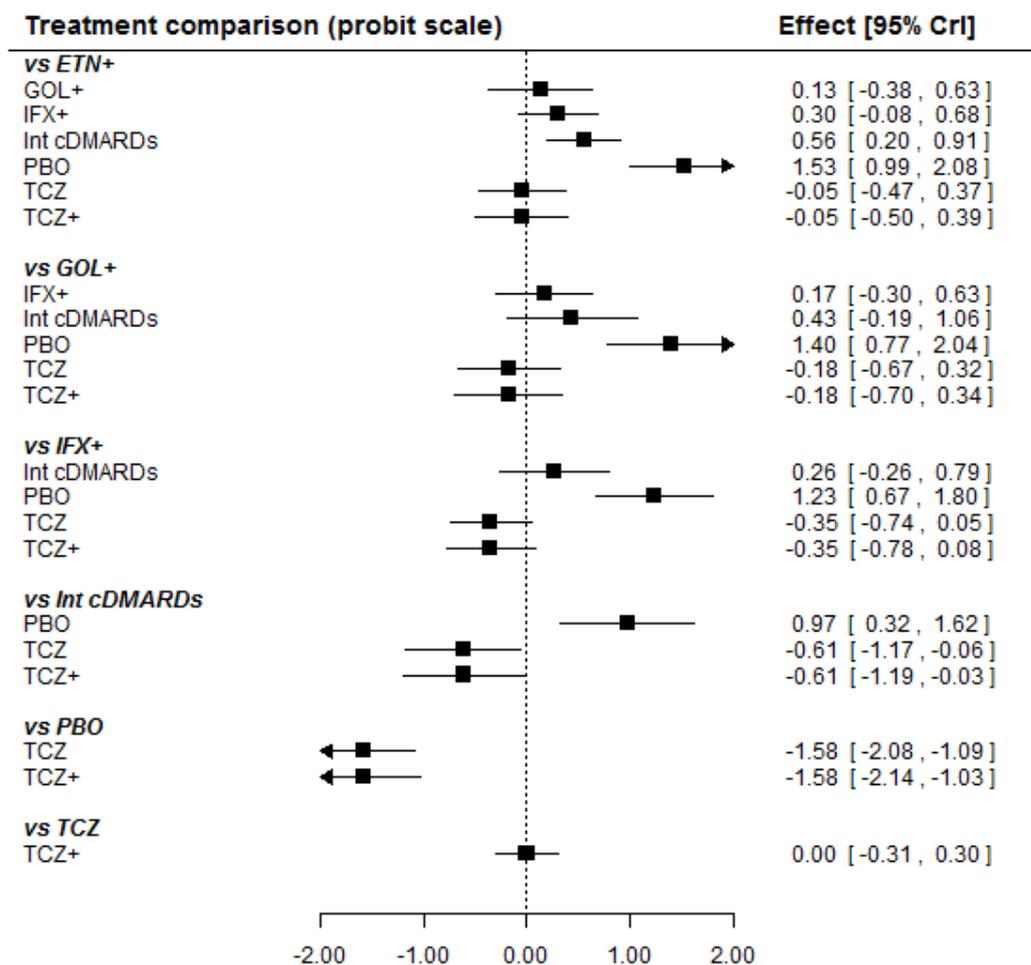


Figure 5d. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)



Légende :

- **ADA** : adalimumab ; **ETN** : étanercept ; **IFX** : infliximab ; **GOL** : golimumab ; **MP** : méthylprednisolone ; **Int cDMARDs** : intensifs cDMARDs (au moins 2 csDMARDs) ; **SU Int cDMARDs** : intensifs csDMARDs administrés par étape ;

- L'usage du signe **(+)** dans certains résultats des méta-analyses en réseau indique une association d'un bDMARD avec un csDMARD (exemple : **TCZ+** signifie tocilizumab en association avec le méthotrexate (MTX)).

Tableau 8. Réponses ACR : Probabilités – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	Réponse ACR20 [95 % CrI]	Réponse ACR50 [95 % CrI]	Réponse ACR70 [95 % CrI]
csDMARDs	0,31 [0,26 - 0,36]	0,13 [0,10 - 0,17]	0,04 [0,03 - 0,06]
ABT i.v.+	0,58 [0,45 - 0,71]	0,34 [0,23 - 0,47]	0,16 [0,09 - 0,26]
ABT s.c.+	0,65 [0,48 - 0,80]	0,41 [0,25 - 0,58]	0,21 [0,10 - 0,36]
ADA	0,48 [0,29 - 0,68]	0,25 [0,12 - 0,43]	0,11 [0,04 - 0,23]
ADA+	0,63 [0,52 - 0,73]	0,38 [0,28 - 0,49]	0,19 [0,12 - 0,27]
CTZ+	0,58 [0,35 - 0,79]	0,33 [0,15 - 0,57]	0,16 [0,05 - 0,35]
ETN	0,61 [0,47 - 0,73]	0,36 [0,24 - 0,50]	0,17 [0,10 - 0,28]
ETN+	0,70 [0,58 - 0,81]	0,46 [0,33 - 0,60]	0,25 [0,16 - 0,37]
GOL+	0,66 [0,50 - 0,79]	0,41 [0,26 - 0,58]	0,21 [0,11 - 0,35]
IFX+	0,59 [0,48 - 0,70]	0,35 [0,25 - 0,46]	0,17 [0,10 - 0,25]
Int cDMARDs	0,49 [0,30 - 0,68]	0,26 [0,12 - 0,44]	0,11 [0,04 - 0,23]
PBO	0,16 [0,06 - 0,32]	0,05 [0,02 - 0,14]	0,01 [0,00 - 0,05]
TCZ	0,72 [0,60 - 0,82]	0,48 [0,35 - 0,62]	0,27 [0,17 - 0,39]
TCZ+	0,72 [0,58 - 0,83]	0,48 [0,34 - 0,63]	0,27 [0,16 - 0,40]

➤ Réponses EULAR à la semaine 24

Des « graphiques en forêt » décrivant l'effet de chaque intervention par rapport aux csDMARDs selon une échelle probit ont été présentés dans les figures 6a-6d. Les probabilités de réalisation d'au moins une réponse EULAR (modérée ou bonne) sont présentées dans le tableau 9.

Comparativement aux csDMARDs (ex. : MTX), les effets du traitement étaient statistiquement significatifs pour CTZ+, ETN+, GOL+, IFX+, TCZ (en association ou non avec MTX) à un seuil de 5 %. Les TCZ (en association ou non avec MTX) étaient les interventions qui présentaient les réponses EULAR les plus élevées (médiane de rang 2, probabilité d'être meilleure 0,20 pour le TCZ et 0,38 pour le TCZ+, respectivement) (cf. annexe A3.9).

Figure 6a. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

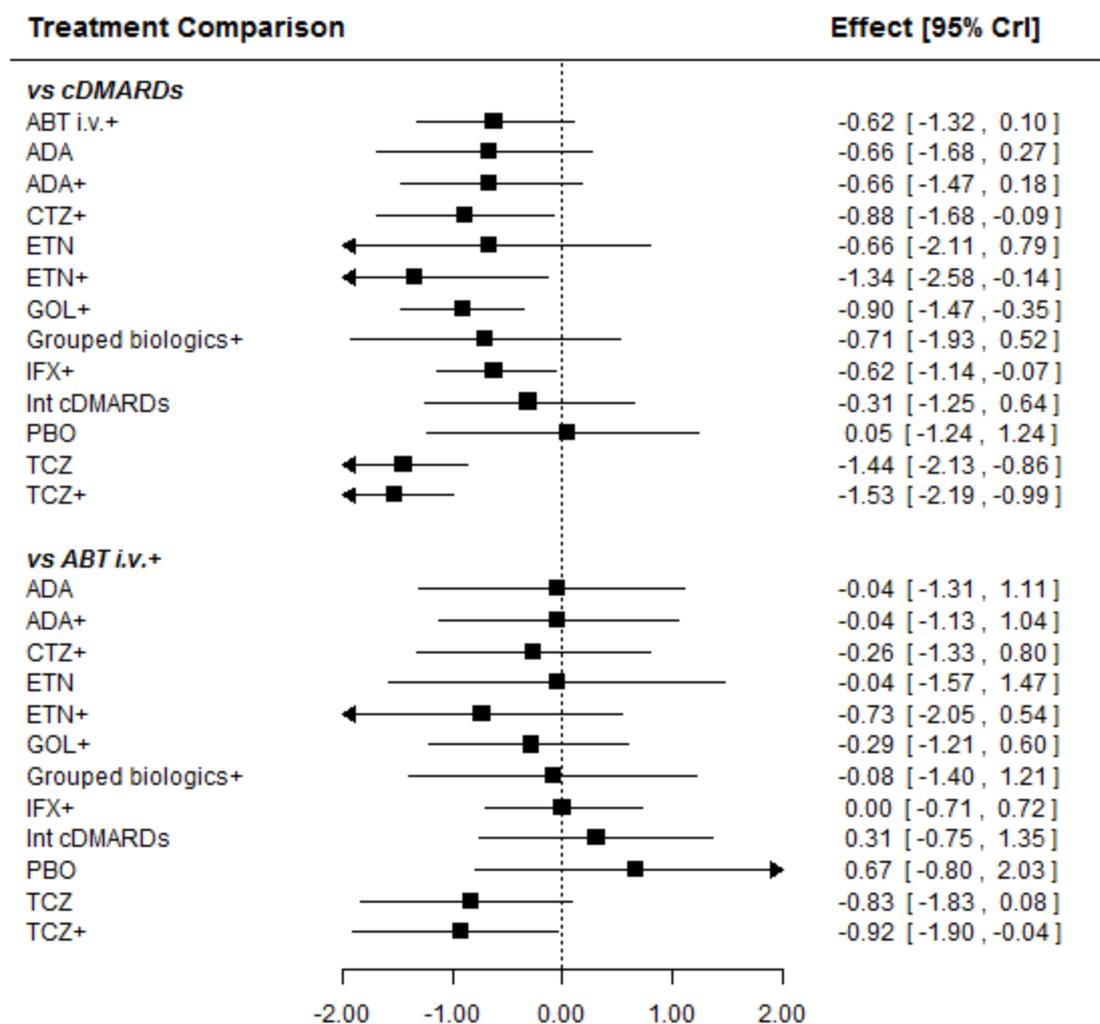


Figure 6b. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

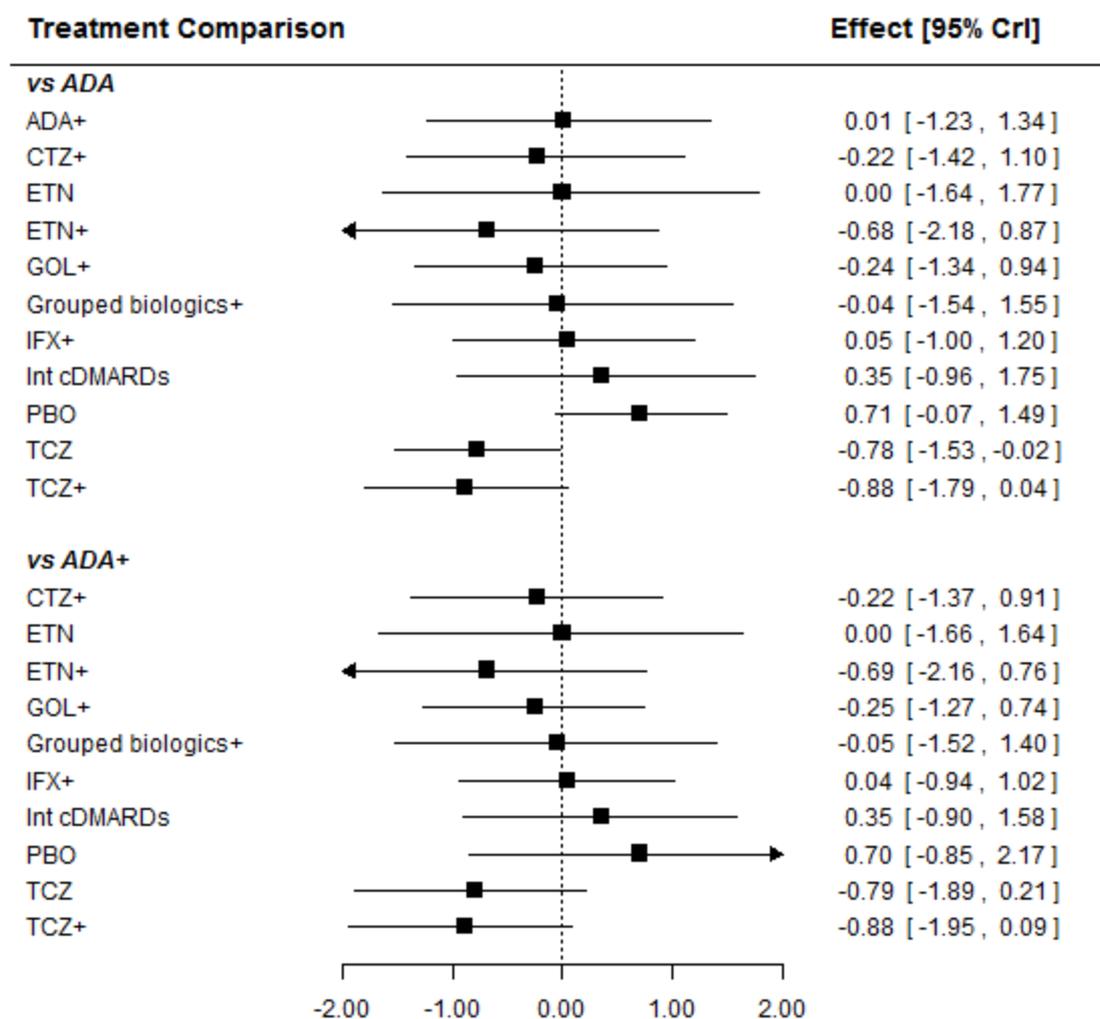


Figure 6c. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

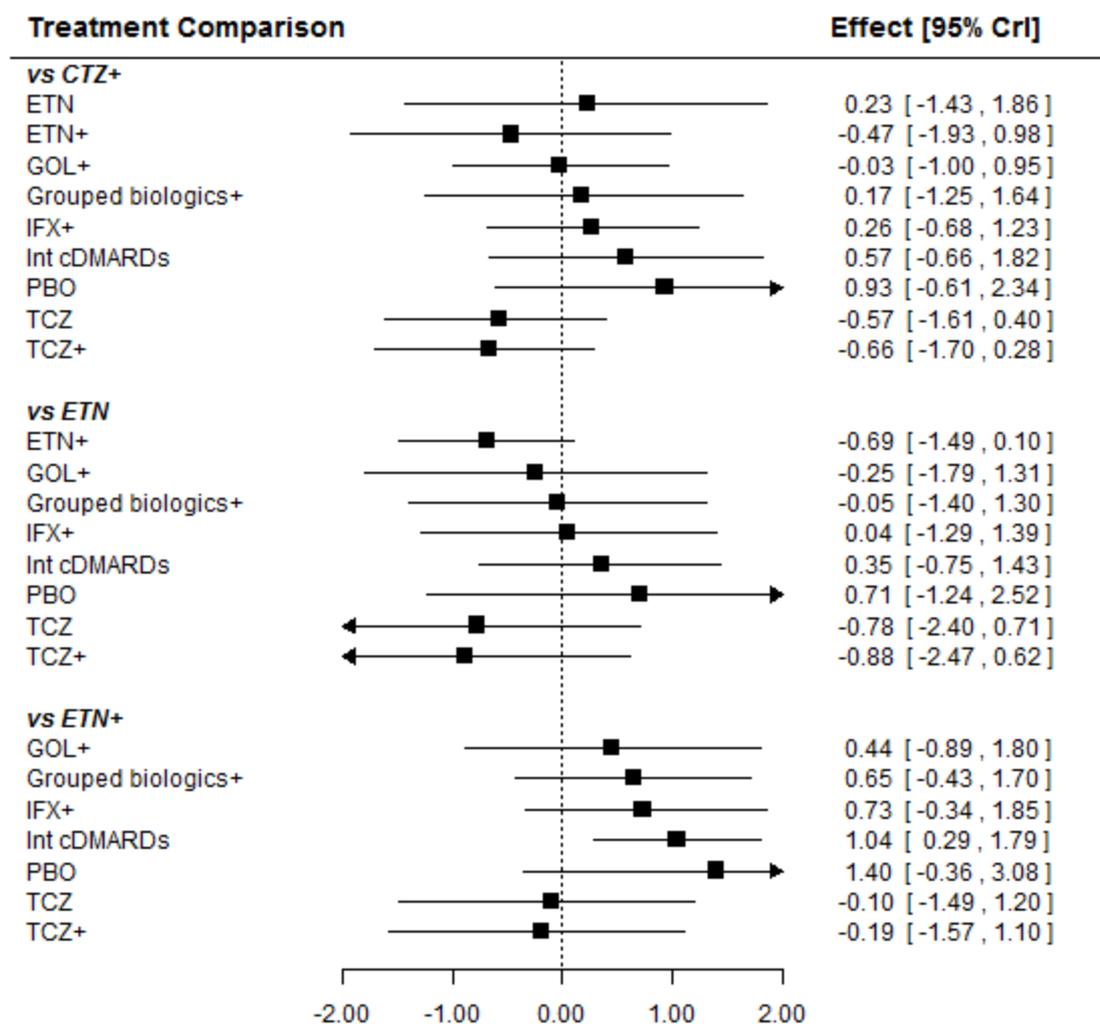


Figure 6d. Réponses EULAR : Effets des interventions – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

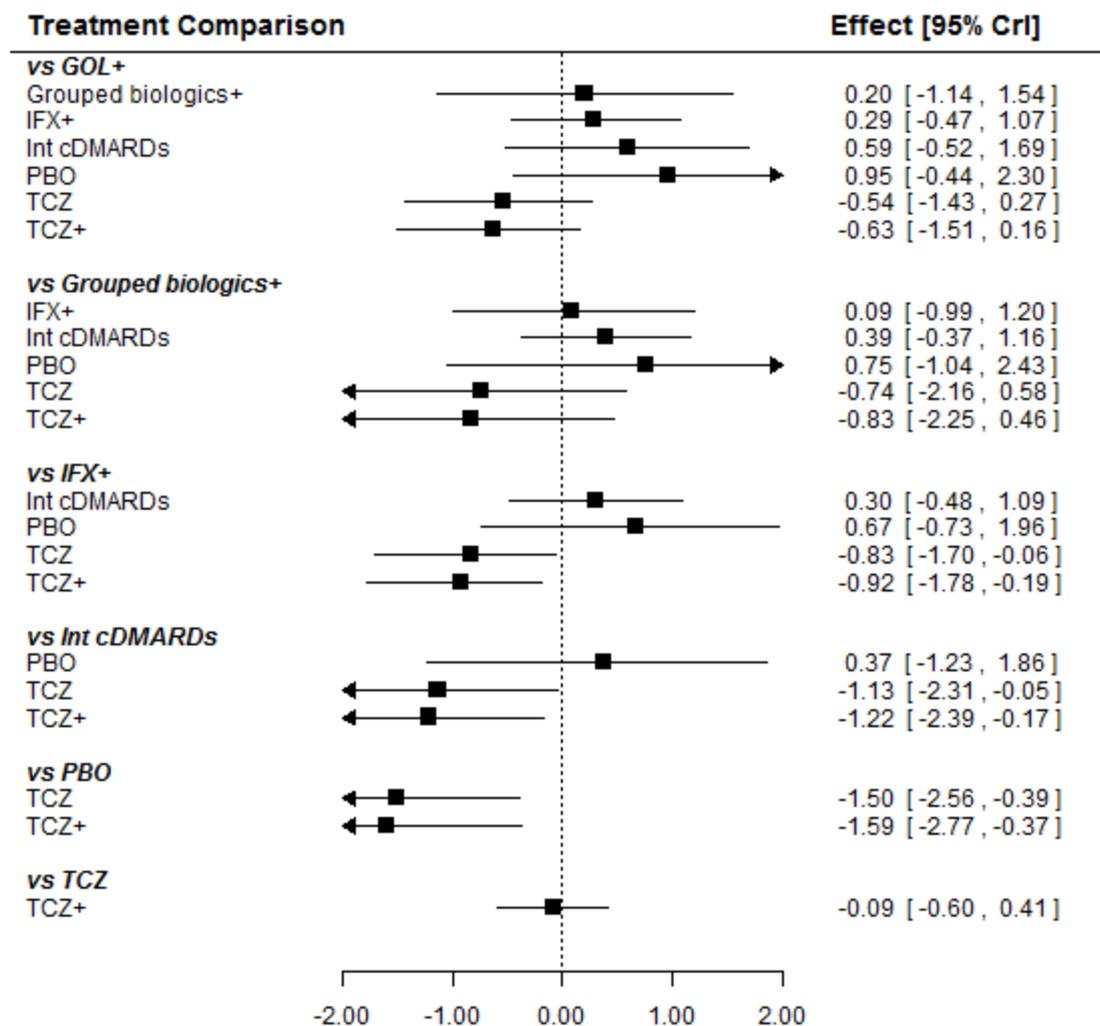


Tableau 9. Réponse EULAR : Probabilités – Patients ayant été traités par des csDMARDs (MTX) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	Au moins une réponse modérée [95 % CrI]	Au moins une bonne réponse [95 % CrI]
csDMARDs	0,45 [0,38 - 0,52]	0,12 [0,08 - 0,17]
ABT i.v.+	0,69 [0,40 - 0,89]	0,28 [0,09 - 0,57]
ADA	0,70 [0,34 - 0,94]	0,30 [0,07 - 0,70]
ADA+	0,70 [0,38 - 0,91]	0,30 [0,08 - 0,63]
CTZ+	0,77 [0,48 - 0,94]	0,38 [0,13 - 0,70]
ETN	0,70 [0,18 - 0,98]	0,30 [0,02 - 0,82]
ETN+	0,89 [0,50 - 0,99]	0,56 [0,14 - 0,92]
GOL+	0,78 [0,58 - 0,92]	0,39 [0,19 - 0,63]
Grouped biologics+	0,72 [0,26 - 0,97]	0,31 [0,04 - 0,77]
IFX+	0,69 [0,47 - 0,85]	0,28 [0,12 - 0,50]
Int cDMARDs	0,57 [0,22 - 0,87]	0,19 [0,03 - 0,53]
PBO	0,43 [0,08 - 0,87]	0,11 [0,01 - 0,53]
TCZ	0,91 [0,76 - 0,98]	0,60 [0,35 - 0,83]
TCZ+	0,92 [0,80 - 0,98]	0,63 [0,40 - 0,85]

4.3. Les analyses de sensibilité

Les analyses de sensibilité prévues dans le deuxième chapitre ont été réalisées pour explorer la variabilité des résultats si les critères d'inclusion n'ont pas été strictement vérifiés. Elles devaient être présentées dans la section analyse de l'hétérogénéité et de la cohérence des preuves (section 5). Afin de faciliter la lecture chronologique des résultats (analyses principales, analyses de sensibilité), elles ont été présentées juste après les résultats de l'analyse principale (section 4.2).

Pour les patients ayant été traités par des csDMARDs (MTX) :

- analyse de sensibilité n° 1 : ajout de réponses ACR de biosimilaires en supposant que les biosimilaires étaient différents de leurs molécules mères (études : NCT01895309, PLANETRA et NCT01936181) ;
- analyse de sensibilité n° 2 : ajout de réponses ACR provenant d'essais cliniques dont une proportion des patients avait été traitée par des bDMARDs (ACQUIRE (120), LITHE, NCT00254293, OPTION, RA0025, RAPID1, RAPID2 (121), (122), ORAL STANDARD, SUMMACTA, BREVACTA et MUSASHI) ;

- analyse de sensibilité n° 3 : ajout de réponses ACR provenant d'essais cliniques dont une proportion des patients avait été traitée par des bDMARDs (population de l'analyse de sensibilité n° 2 + patients de l'essai AMBITION qui a porté sur un « mélange » de patients traités et non traités par MTX) ;
- analyse de sensibilité n° 4 : ajout de réponses ACR provenant d'essais cliniques dont une proportion de patients était naïve de MTX (études : TEAR et TEMPO) ;
- analyse de sensibilité n° 5 : ajout de réponses EULAR provenant d'essais cliniques portant sur des biosimilaires en supposant que ces derniers étaient différents de leurs molécules mères (études : NCT01895309, PLANETRA and NCT01936181) ;
- analyse de sensibilité n° 6 : ajout de réponses EULAR provenant d'essais cliniques dont une proportion de patients a été traitée par des bDMARDs (études : OPTION, RAPID1 and Yamamoto 2011).

4.3.1. Analyse de sensibilité n° 1

L'ajout des études NCT01895309, PLANETRA et NCT01936181 ne semblait pas affecter les résultats car elles ne contribuent pas à former un réseau fermé (figures 7a-7d et tableau 10).

Les effets de traitements les plus importants étaient associés au biosimilaire de l'ETN (SB4)+, ETN+, TCZ (avec et sans MTX). Au niveau $\alpha=5\%$, ils étaient statistiquement significatifs pour toutes les interventions, sauf pour le placebo et l'ADA (l'effet de l'ADA se situait à la limite de la significativité statistique au niveau $\alpha=5\%$). Biosim ETN (SB4)+ était l'intervention qui présentait la réponse ACR la plus élevée (médiane de rang 2, probabilité d'être la meilleure 0,37) (annexe A3.9).

Figure 7a. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

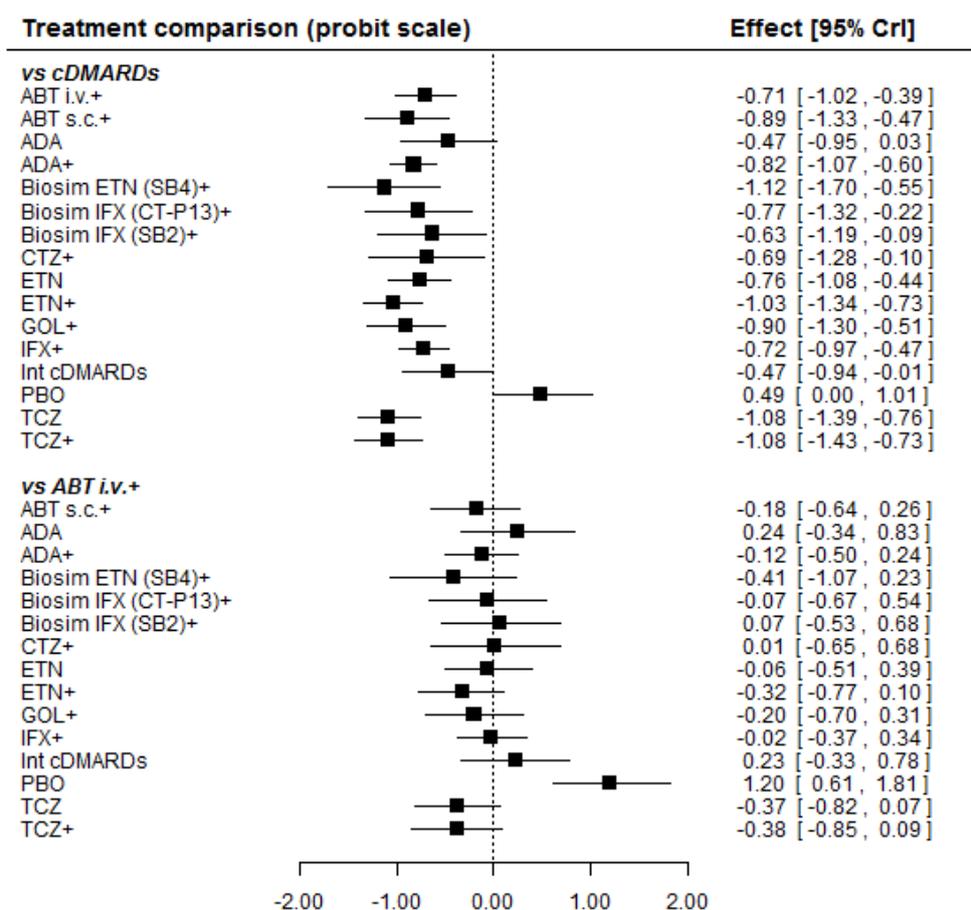


Figure 7b. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

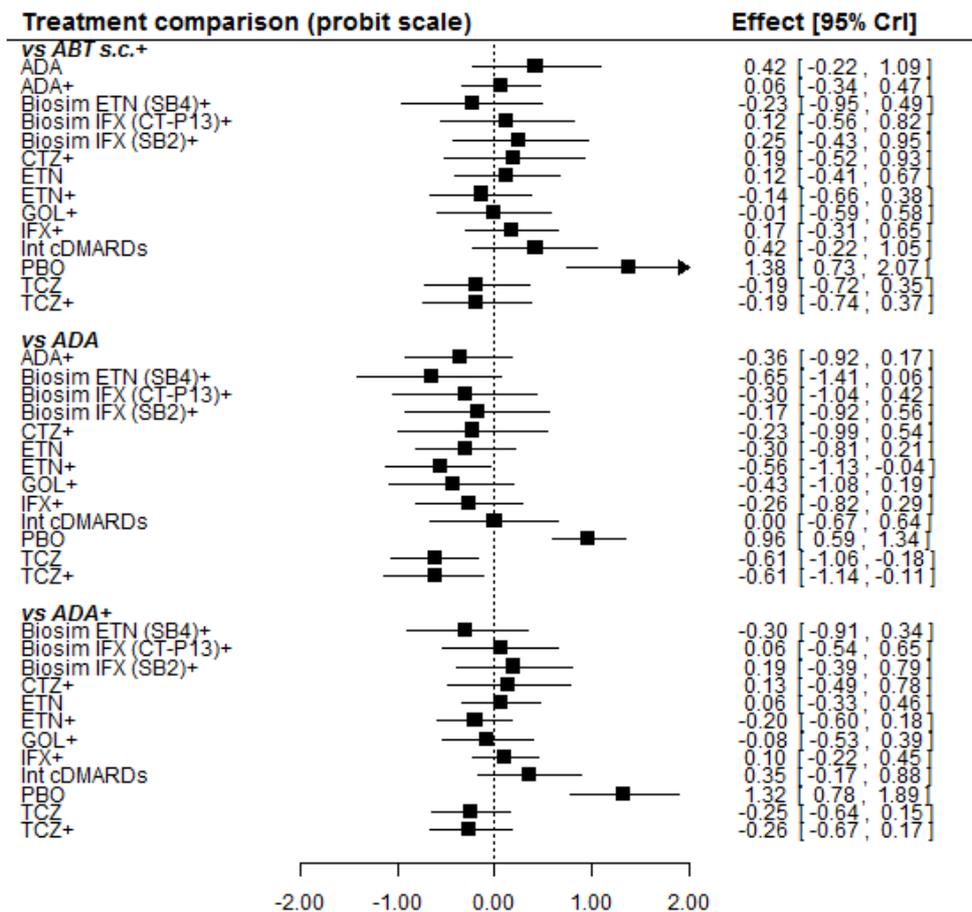


Figure 7c. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

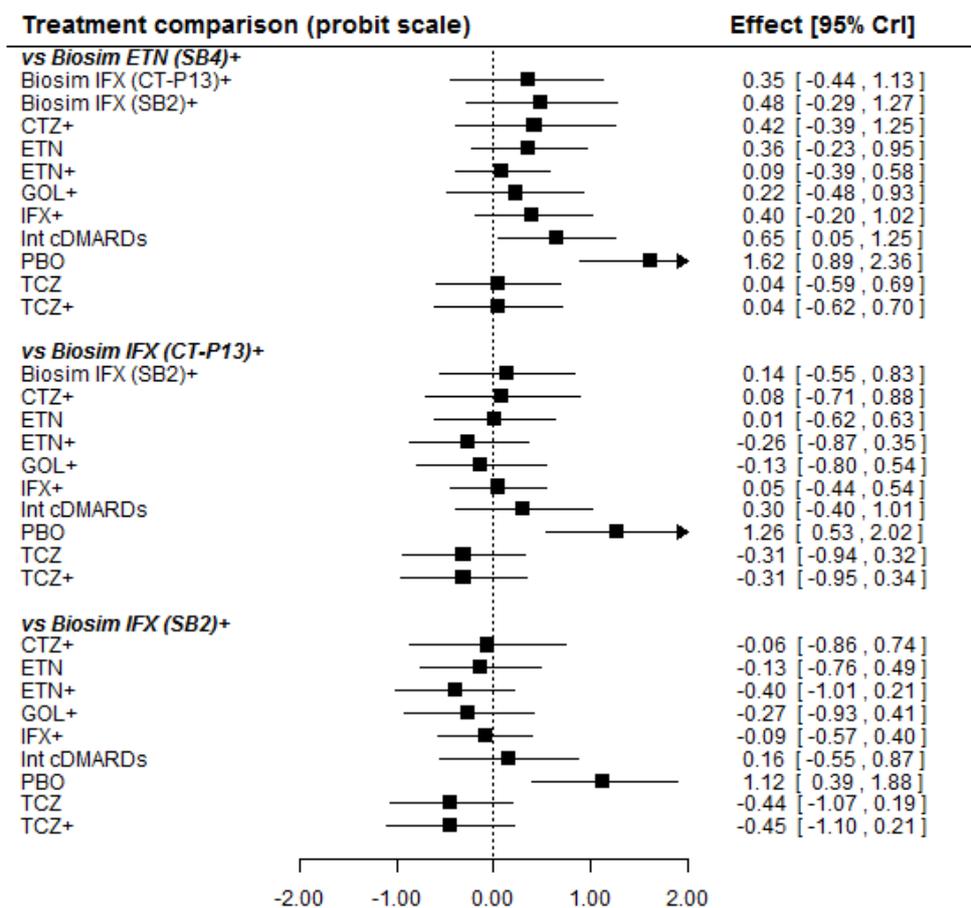


Figure 7d. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau –Université de Sheffield » (84)

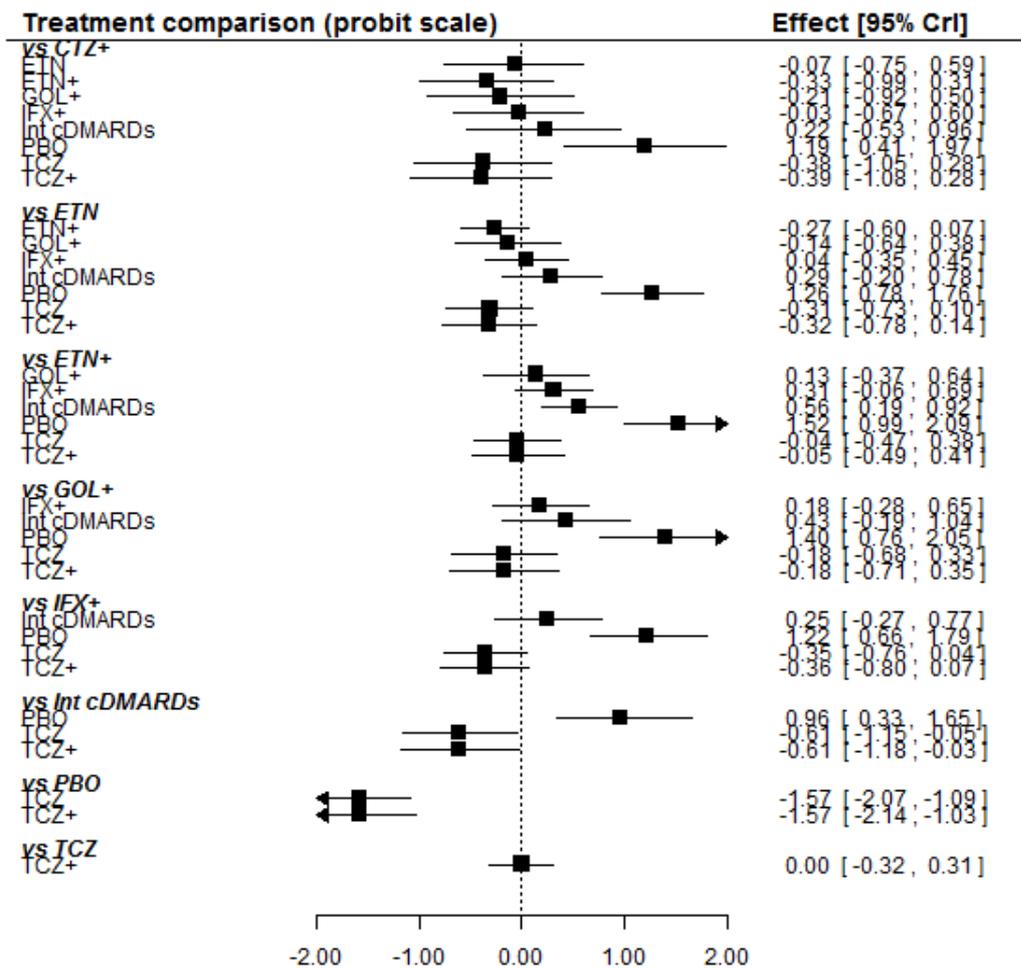


Tableau 10. Réponses ACR : Probabilités – Analyse de sensibilité – Patients ayant été traités par MTX + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	ACR20 [95 % CrI]		ACR50 [95 % CrI]		ACR70 [95 % CrI]	
csDMARDs	0,31 (0,26 - 0,36)	0,31 [0,26 - 0,36]	0,13 [0,10 - 0,16]	0,13 [0,10 - 0,17]	0,04 [0,03 - 0,06]	0,04 [0,03 - 0,06]
ABT i.v.+	0,58 [0,45 - 0,71]	0,58 [0,45 - 0,71]	0,34 [0,22 - 0,47]	0,34 [0,23 - 0,47]	0,16 [0,09 - 0,25]	0,16 [0,09 - 0,26]
ABT s.c.+	0,65 [0,48 - 0,80]	0,65 [0,48 - 0,80]	0,40 [0,25 - 0,59]	0,41 [0,25 - 0,58]	0,20 [0,10 - 0,36]	0,21 [0,10 - 0,36]
ADA	0,49 [0,29 - 0,69]	0,48 [0,29 - 0,68]	0,25 [0,12 - 0,44]	0,25 [0,12 - 0,43]	0,11 [0,04 - 0,23]	0,11 [0,04 - 0,23]
ADA+	0,63 [0,53 - 0,73]	0,63 [0,52 - 0,73]	0,38 [0,28 - 0,49]	0,38 [0,28 - 0,49]	0,19 [0,12 - 0,27]	0,19 [0,12 - 0,27]
Biosim ETN [SB4]+	0,74 [0,52 - 0,89]		0,50 [0,28 - 0,72]		0,28 [0,12 - 0,50]	
Biosim IFX [CT-P13]+	0,61 [0,39 - 0,80]		0,36 [0,18 - 0,58]		0,17 [0,07 - 0,36]	
Biosim IFX [SB2]+	0,56 [0,33 - 0,76]		0,31 [0,14 - 0,53]		0,14 [0,05 - 0,31]	
CTZ+	0,58 [0,34 - 0,79]	0,58 [0,35 - 0,79]	0,33 [0,15 - 0,57]	0,33 [0,15 - 0,57]	0,15 [0,05 - 0,34]	0,16 [0,05 - 0,35]
ETN	0,61 [0,47 - 0,73]	0,61 [0,47 - 0,73]	0,36 [0,24 - 0,49]	0,36 [0,24 - 0,50]	0,17 [0,10 - 0,28]	0,17 [0,10 - 0,28]
ETN+	0,71 [0,58 - 0,81]	0,70 [0,58 - 0,81]	0,46 [0,33 - 0,59]	0,46 [0,33 - 0,60]	0,25 [0,16 - 0,37]	0,25 [0,16 - 0,37]
GOL+	0,66 [0,50 - 0,80]	0,66 [0,50 - 0,79]	0,41 [0,26 - 0,58]	0,41 [0,26 - 0,58]	0,21 [0,11 - 0,35]	0,21 [0,11 - 0,35]
IFX+	0,59 [0,48 - 0,70]	0,59 [0,48 - 0,70]	0,34 [0,24 - 0,45]	0,35 [0,25 - 0,46]	0,16 [0,10 - 0,24]	0,17 [0,10 - 0,25]
Int cDMARDs	0,49 [0,31 - 0,68]	0,49 [0,30 - 0,68]	0,26 [0,13 - 0,43]	0,26 [0,12 - 0,44]	0,11 [0,04 - 0,23]	0,11 [0,04 - 0,23]
PBO	0,16 [0,07 - 0,32]	0,16 [0,06 - 0,32]	0,05 [0,02 - 0,13]	0,05 [0,02 - 0,14]	0,01 [0,00 - 0,05]	0,01 [0,00 - 0,05]
TCZ	0,72 [0,59 - 0,82]	0,72 [0,60 - 0,82]	0,48 [0,34 - 0,62]	0,48 [0,35 - 0,62]	0,26 [0,16 - 0,39]	0,27 [0,17 - 0,39]
TCZ+	0,72 [0,58 - 0,83]	0,72 [0,58 - 0,83]	0,48 [0,34 - 0,63]	0,48 [0,34 - 0,63]	0,26 [0,16 - 0,40]	0,27 [0,16 - 0,40]

4.3.2. Analyse de sensibilité n° 2

Quatorze études (ACQUIRE, Kremer 2012, LITHE, NCT00254293, OPTION, RA0025, RAPID1, RAPID2, van der Heijde 2013, Yamamoto 2011, ORAL STANDARD, SUMMACTA, BREVACTA et MUSASHI) ont été ajoutées à la méta-analyse en réseau principale. Ceci a contribué à diminuer l'incertitude associée aux paramètres d'hétérogénéité et a conduit à une estimation légèrement plus précise des effets des traitements relatifs aux csDMARDs (figures 8a-8d). Les « graphiques en forêt » montraient que les effets des bDMARDs ont été légèrement diminués par rapport à ceux de l'analyse principale (figures 9a-9d), à l'exception de CTZ+.

L'ajout de ces 14 études a également diminué l'effet du traitement sur les réponses ACR20, ACR50 et ACR70. Les probabilités de réponses ACR estimées pour les interventions ont diminué par rapport à l'analyse principale, exception faite pour le CTZ+ (tableau 11). Le CTZ+ était l'intervention qui présentait la réponse ACR la plus élevée (médiane de rang 1, probabilité d'être la meilleure 0,59, cf. annexe 3-9).

Figure 8a. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

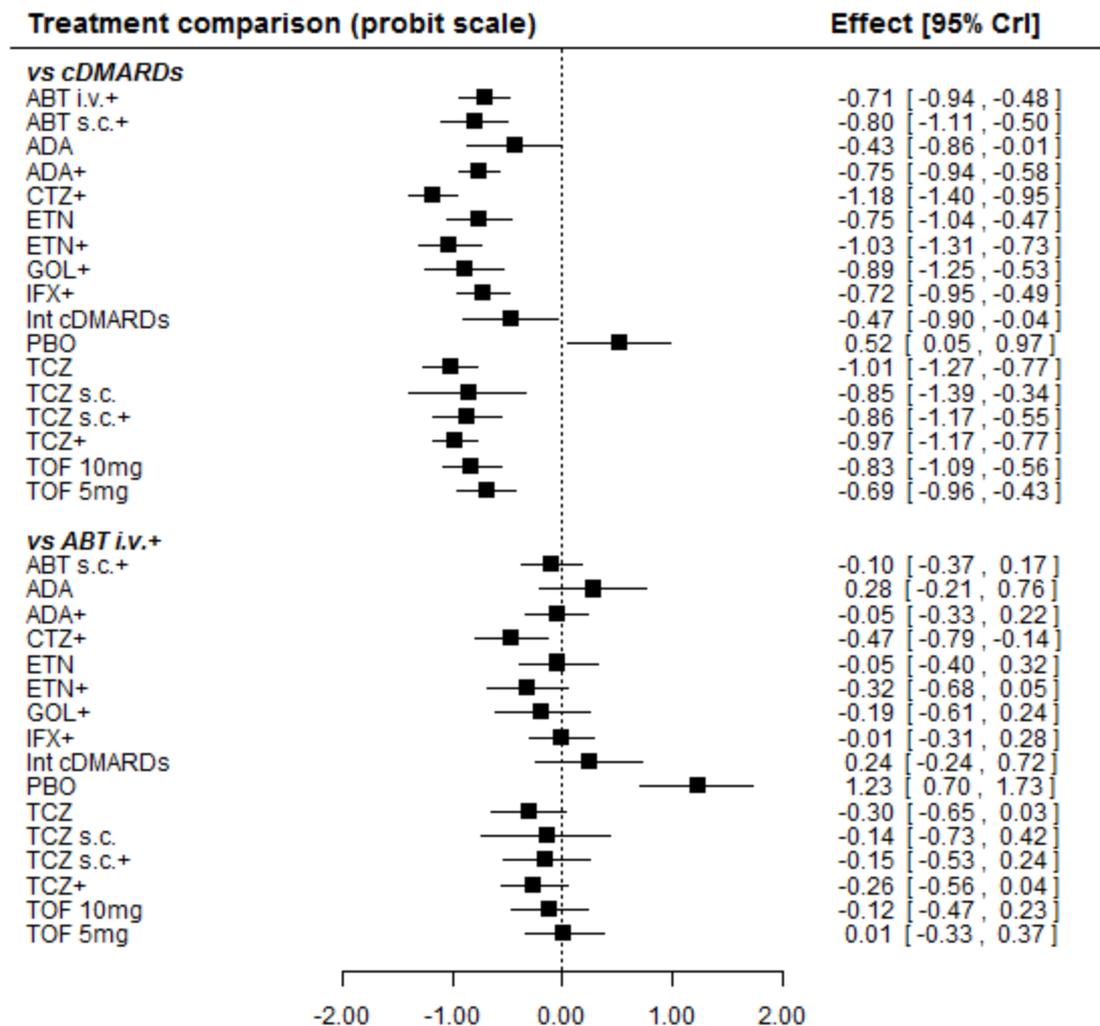


Figure 8b. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + proportion de patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

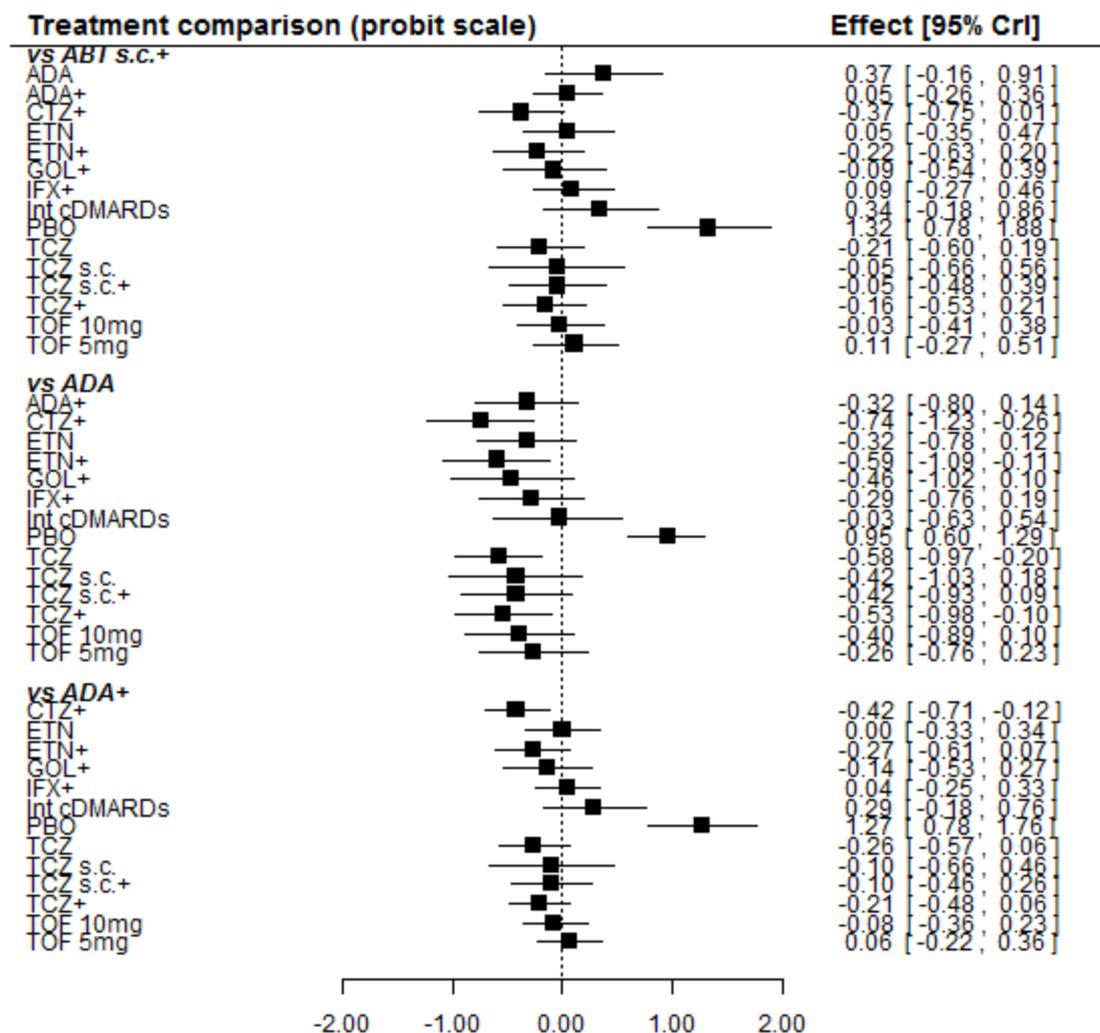


Figure 8c. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + proportion de patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

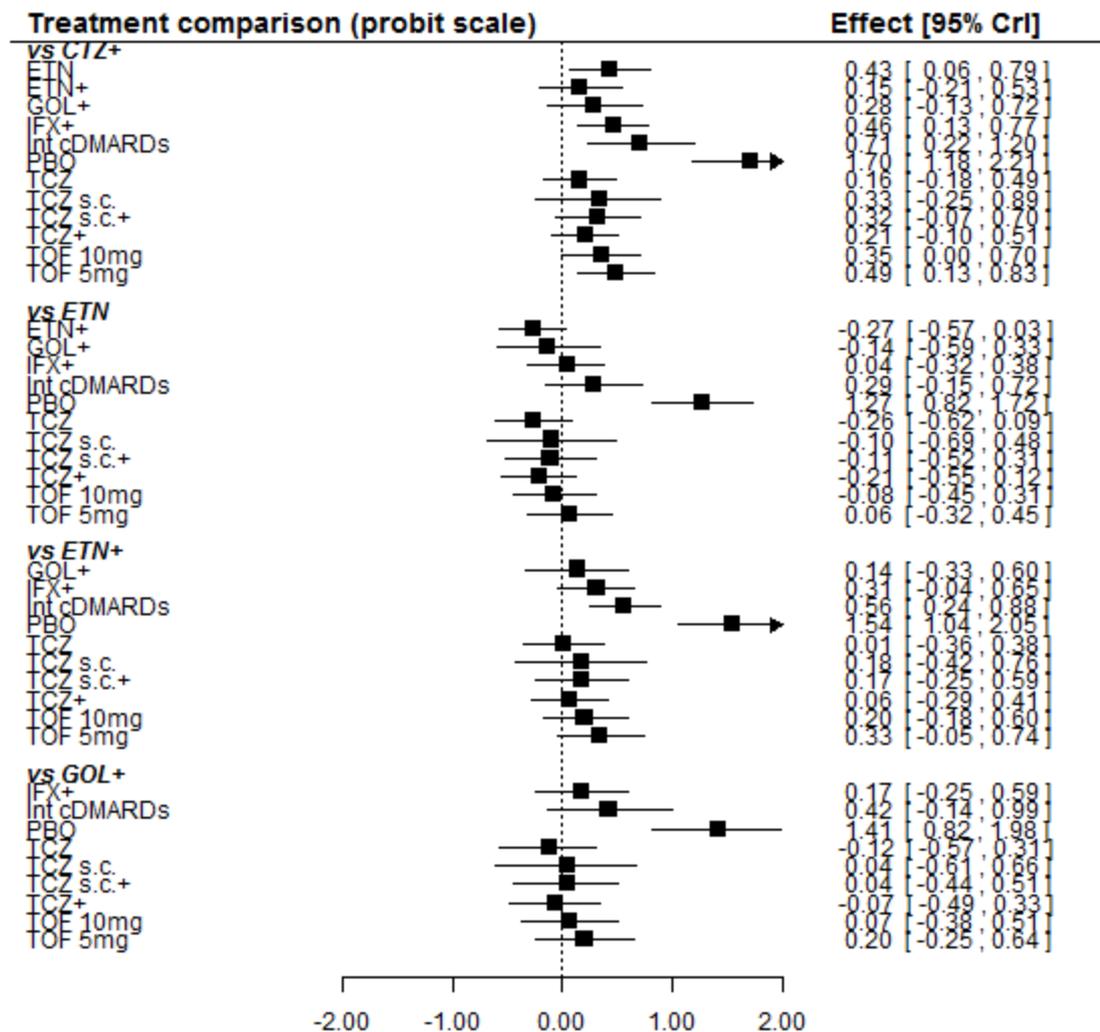


Figure 8d. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

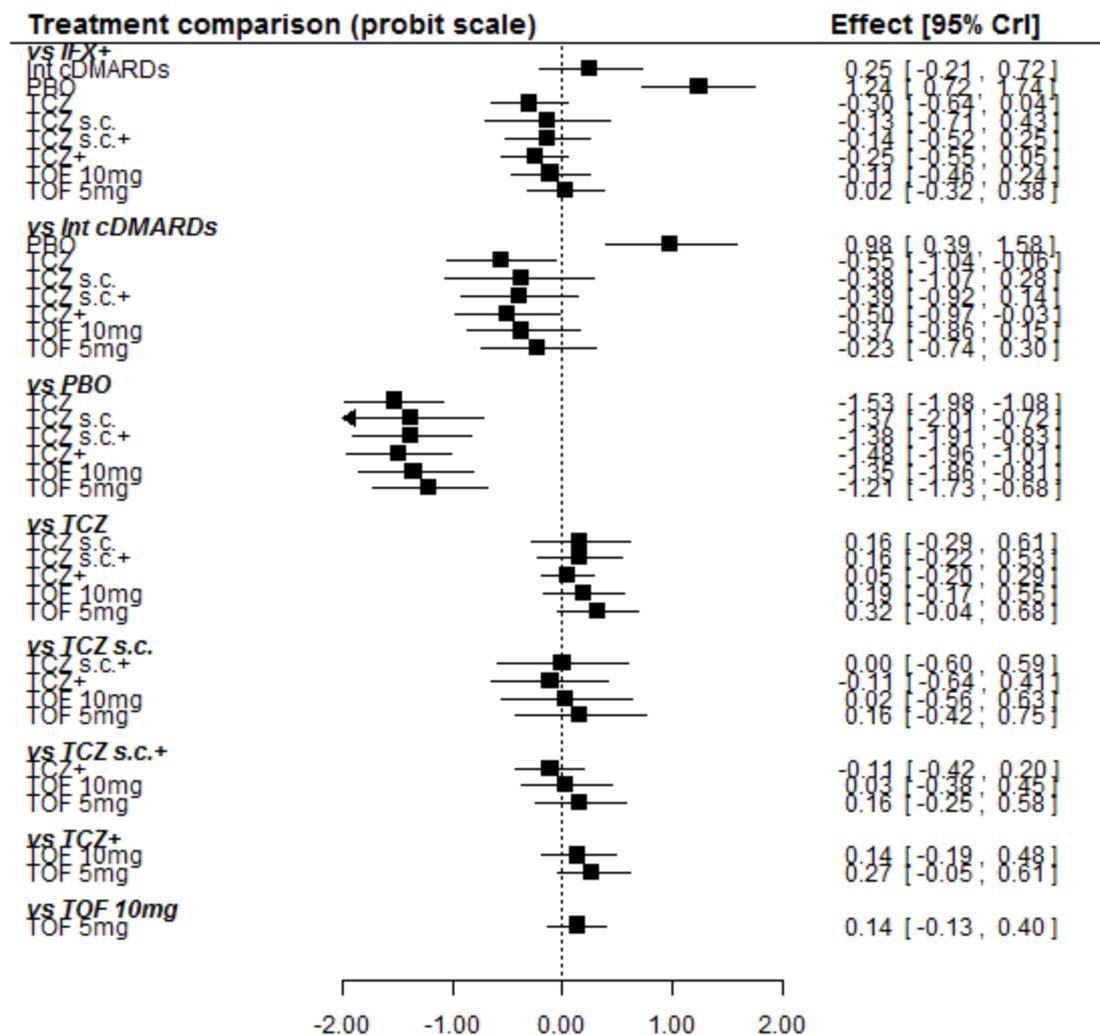


Tableau 11. Réponses ACR : Probabilités – Analyse de sensibilité – Patients ayant été traités par MTX + proportion de patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	ACR20 [95 % CrI]		ACR50 [95 % CrI]		ACR70 [95 % CrI]	
cDMARDs	0,28 [0,25 - 0,32]	0,31 [0,26 - 0,36]	0,12 [0,10 - 0,15]	0,13 [0,10 - 0,17]	0,04 [0,03 - 0,05]	0,04 [0,03 - 0,06]
ABT i.v.+	0,55 [0,45 - 0,65]	0,58 [0,45 - 0,71]	0,32 [0,23 - 0,42]	0,34 [0,23 - 0,47]	0,14 [0,09 - 0,21]	0,16 [0,09 - 0,26]
ABT s.c.+	0,59 [0,47 - 0,71]	0,65 [0,48 - 0,80]	0,36 [0,24 - 0,49]	0,41 [0,25 - 0,58]	0,17 [0,10 - 0,26]	0,21 [0,10 - 0,36]
ADA	0,44 [0,28 - 0,62]	0,48 [0,29 - 0,68]	0,23 [0,12 - 0,38]	0,25 [0,12 - 0,43]	0,09 [0,04 - 0,19]	0,11 [0,04 - 0,23]
ADA+	0,57 [0,49 - 0,66]	0,63 [0,52 - 0,73]	0,34 [0,26 - 0,42]	0,38 [0,28 - 0,49]	0,16 [0,11 - 0,22]	0,19 [0,12 - 0,27]
CTZ+	0,73 [0,64 - 0,81]	0,58 [0,35 - 0,79]	0,50 [0,40 - 0,60]	0,33 [0,15 - 0,57]	0,28 [0,20 - 0,37]	0,16 [0,05 - 0,35]
ETN	0,57 [0,45 - 0,69]	0,61 [0,47 - 0,73]	0,34 [0,23 - 0,45]	0,36 [0,24 - 0,50]	0,16 [0,09 - 0,24]	0,17 [0,10 - 0,28]
ETN+	0,68 [0,56 - 0,78]	0,70 [0,58 - 0,81]	0,44 [0,32 - 0,56]	0,46 [0,33 - 0,60]	0,23 [0,15 - 0,33]	0,25 [0,16 - 0,37]
GOL+	0,63 [0,48 - 0,76]	0,66 [0,50 - 0,79]	0,39 [0,25 - 0,54]	0,41 [0,26 - 0,58]	0,19 [0,10 - 0,31]	0,21 [0,11 - 0,35]
IFX+	0,56 [0,46 - 0,66]	0,59 [0,48 - 0,70]	0,32 [0,24 - 0,42]	0,35 [0,25 - 0,46]	0,15 [0,10 - 0,21]	0,17 [0,10 - 0,25]
Int cDMARDs	0,46 [0,29 - 0,63]	0,49 [0,30 - 0,68]	0,24 [0,12 - 0,40]	0,26 [0,12 - 0,44]	0,10 [0,04 - 0,20]	0,11 [0,04 - 0,23]
PBO	0,14 [0,06 - 0,27]	0,16 [0,06 - 0,32]	0,05 [0,02 - 0,11]	0,05 [0,02 - 0,14]	0,01 [0,00 - 0,04]	0,01 [0,00 - 0,05]
TCZ	0,67 [0,57 - 0,76]	0,72 [0,60 - 0,82]	0,44 [0,33 - 0,55]	0,48 [0,35 - 0,62]	0,23 [0,15 - 0,32]	0,27 [0,17 - 0,39]
TCZ s.c.	0,61 [0,40 - 0,80]		0,37 [0,20 - 0,59]		0,18 [0,07 - 0,36]	
TCZ s.c.+	0,61 [0,48 - 0,73]		0,38 [0,26 - 0,51]		0,18 [0,11 - 0,28]	
TCZ+	0,65 [0,57 - 0,74]	0,72 [0,58 - 0,83]	0,42 [0,33 - 0,51]	0,48 [0,34 - 0,63]	0,21 [0,15 - 0,29]	0,27 [0,16 - 0,40]
TOF 10 mg	0,60 [0,48 - 0,71]		0,37 [0,26 - 0,48]		0,17 [0,11 - 0,26]	
TOF 5 mg	0,55 [0,43 - 0,66]		0,32 [0,22 - 0,42]		0,14 [0,09 - 0,22]	

4.3.3. Analyse de sensibilité n° 3

Quinze études (ACQUIRE, Kremer 2012, LITHE, NCT00254293, OPTION, RA0025, RAPID1, RAPID2, van der Heijde 2013, Yamamoto 2011, ORAL STANDARD, SUMMACTA, BREVACTA, MUSASHI et AMBITION) ont été incluses dans la méta-analyse principale. Les résultats étaient similaires à ceux de l'analyse de sensibilité n° 2 sans l'intégration de l'étude AMBITION (figures 9a-9d et tableau 12). Le CTZ+ présentait la réponse ACR la plus élevée (médiane 1, probabilité d'être la meilleure 0,65) (cf. annexe A3-9).

Figure 9a. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + proportion de patients ayant été traités par des bDMARDs, y compris AMBITION – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

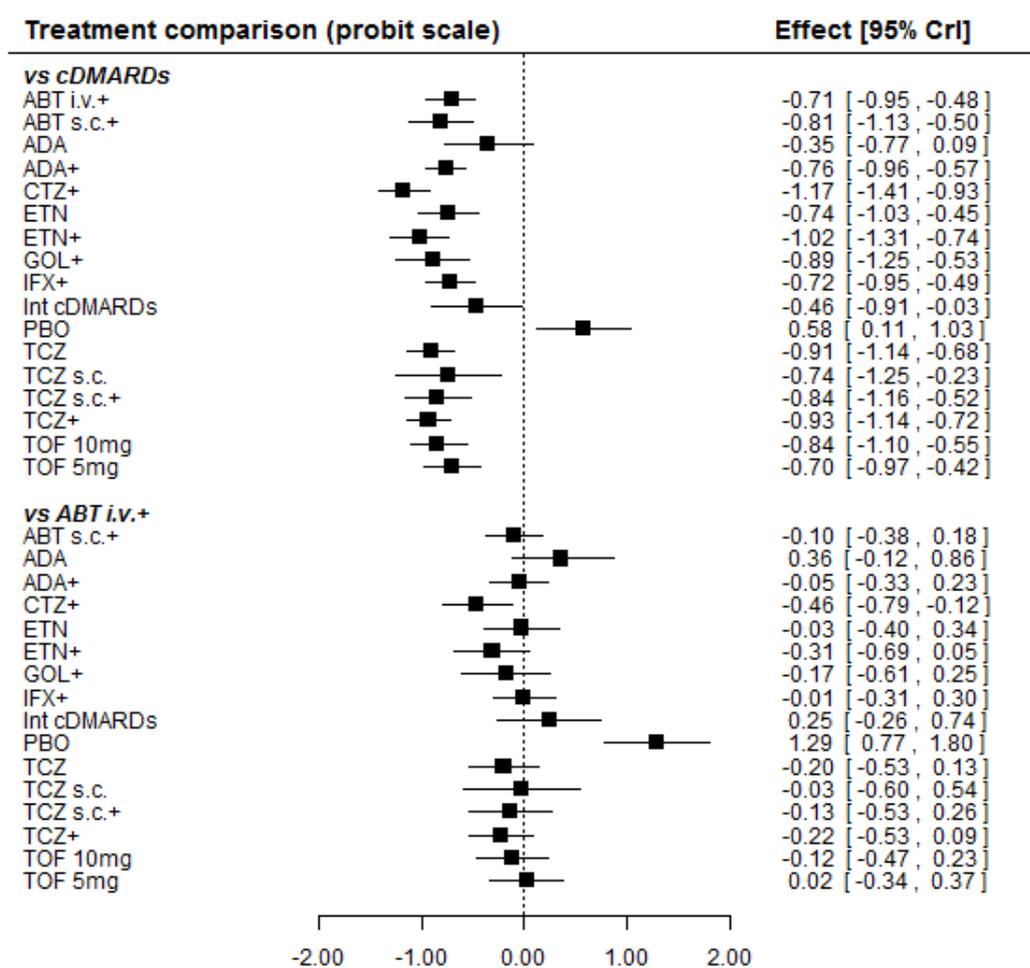


Figure 9b. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs, y compris AMBITION – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

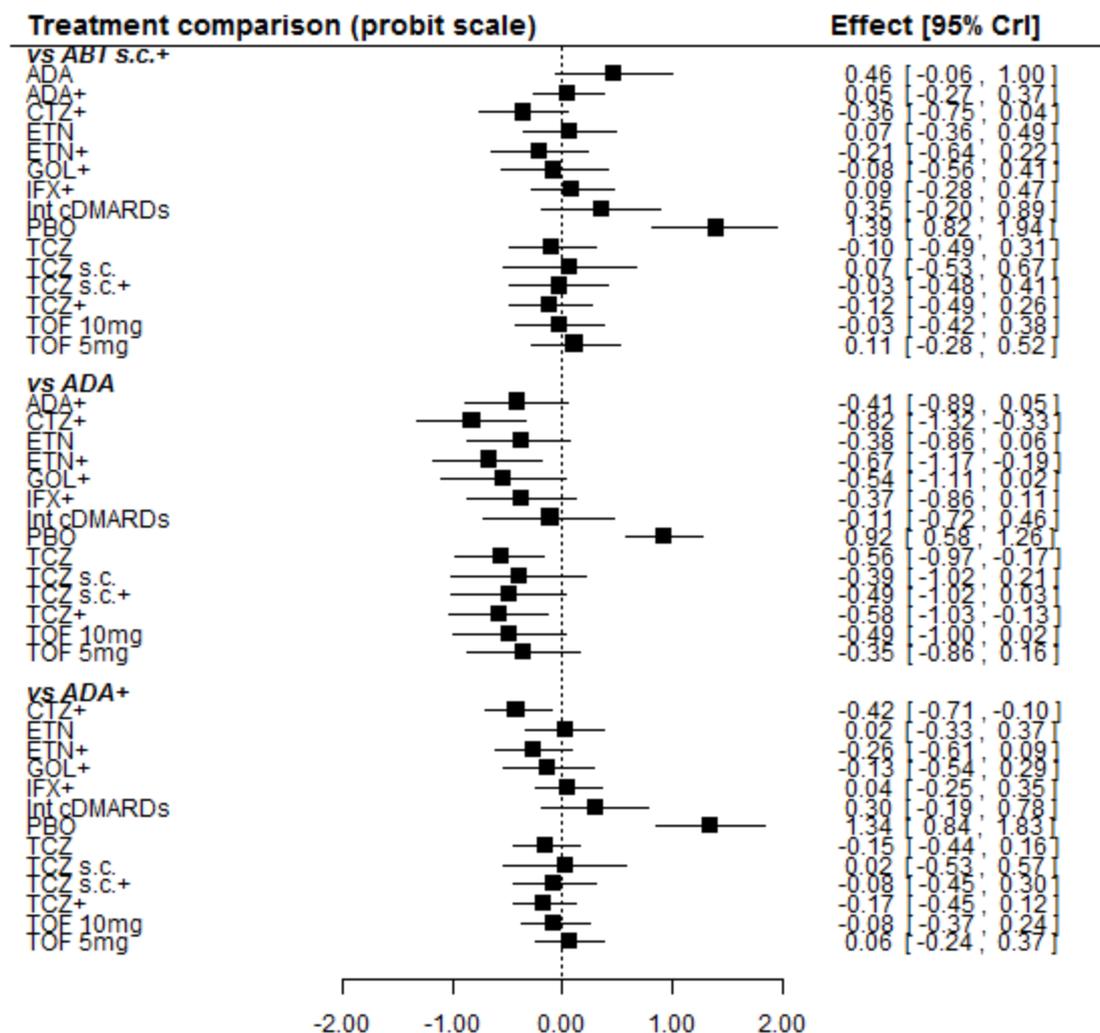


Figure 9c. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs y compris AMBITION – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

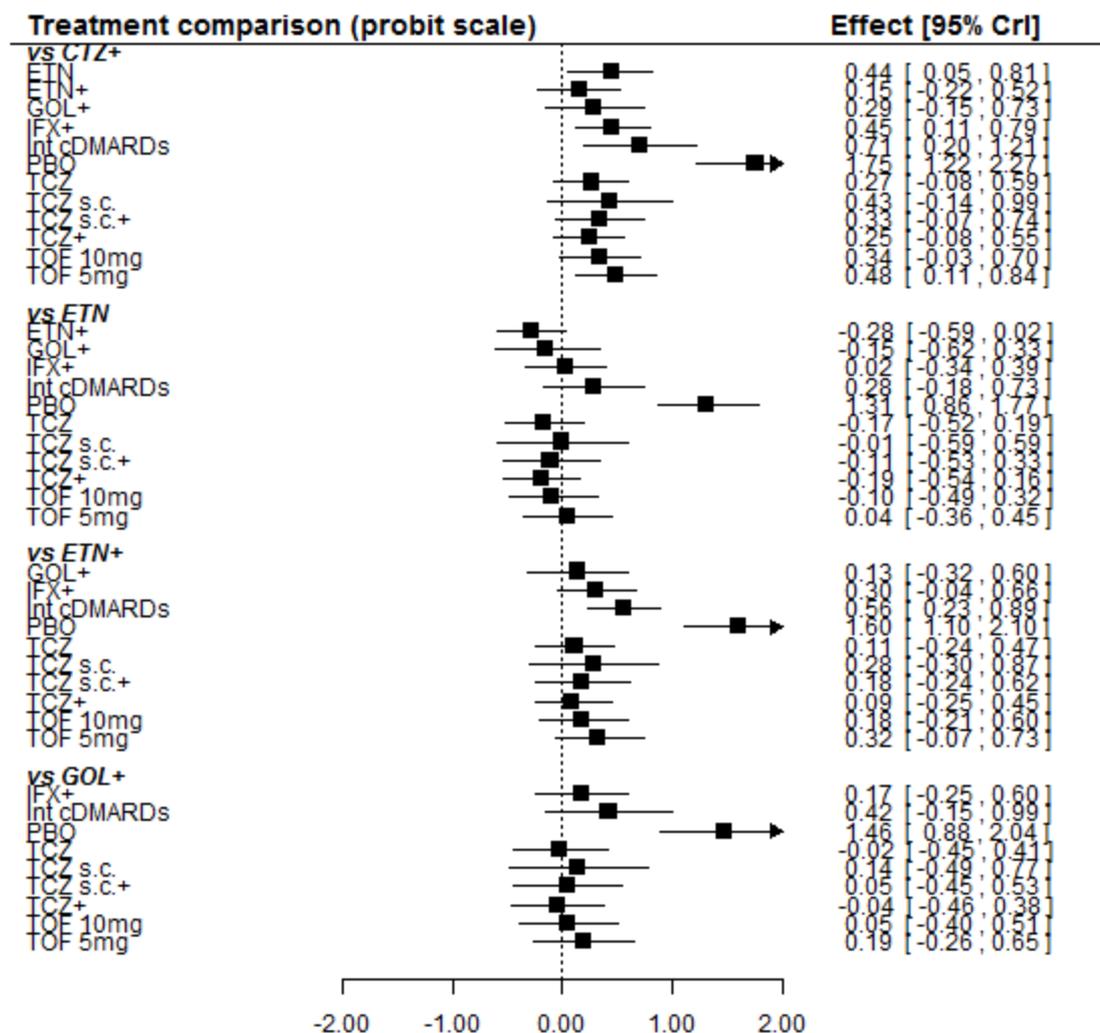


Figure 9d. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs, y compris AMBITION – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

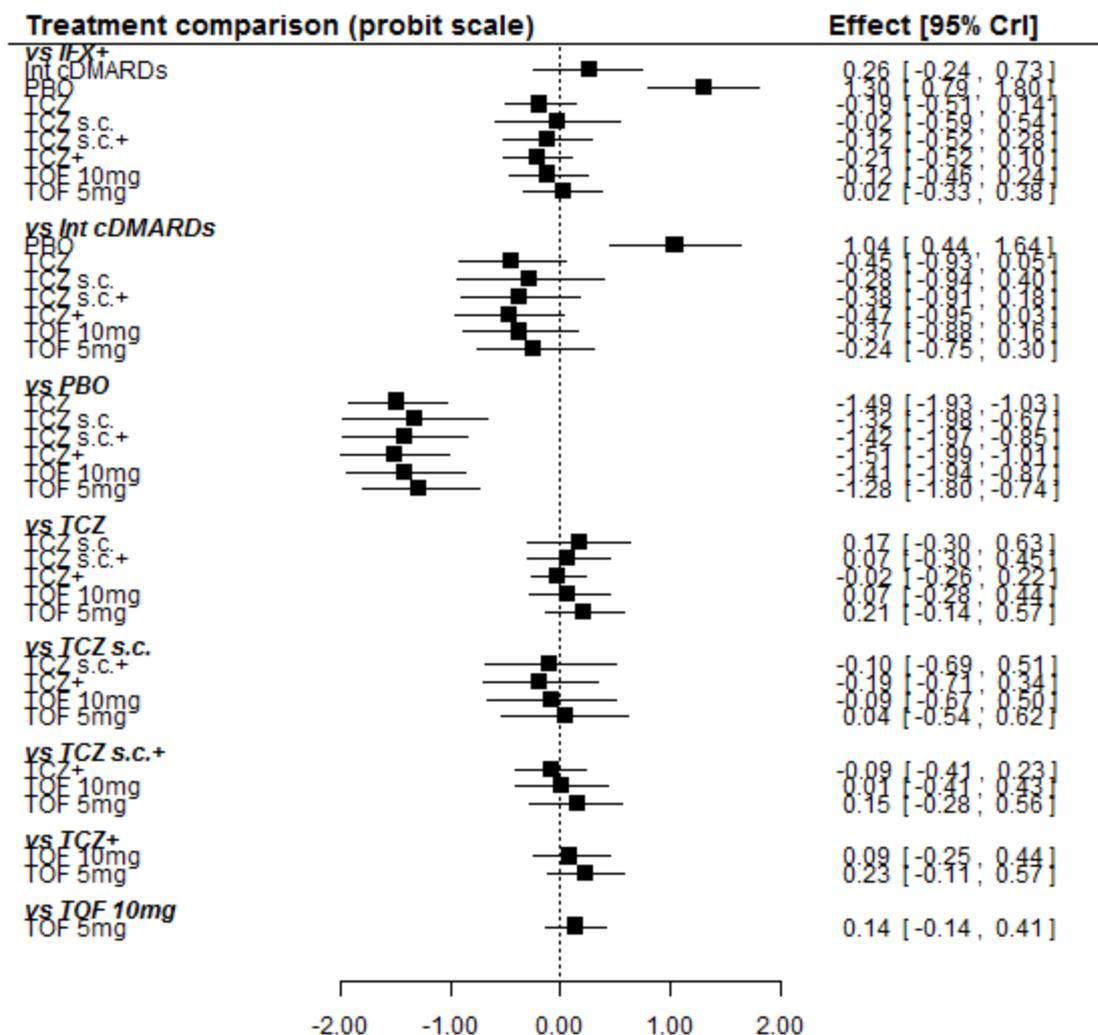


Tableau 12. Réponses ACR : Probabilités – Analyse de sensibilité – Patients ayant été traités par MTX + patients ayant été traités par des bDMARDs (y compris AMBITION) – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	ACR20 [95 % CrI]		ACR50 [95 % CrI]		ACR70 [95 % CrI]	
csDMARDs	0,29 [0,25 - 0,33]	0,31 [0,26 - 0,36]	0,12 [0,10 - 0,15]	0,13 [0,10 - 0,17]	0,04 [0,03 - 0,05]	0,04 [0,03 - 0,06]
ABT i.v.+	0,56 [0,37 - 0,74]	0,58 [0,45 - 0,71]	0,33 [0,18 - 0,51]	0,34 [0,23 - 0,47]	0,15 [0,06 - 0,29]	0,16 [0,09 - 0,26]
ABT s.c.+	0,60 [0,39 - 0,78]	0,65 [0,48 - 0,80]	0,36 [0,19 - 0,57]	0,41 [0,25 - 0,58]	0,17 [0,07 - 0,34]	0,21 [0,10 - 0,36]
ADA	0,42 [0,21 - 0,65]	0,48 [0,29 - 0,68]	0,21 [0,08 - 0,42]	0,25 [0,12 - 0,43]	0,08 [0,02 - 0,21]	0,11 [0,04 - 0,23]
ADA+	0,58 [0,40 - 0,75]	0,63 [0,52 - 0,73]	0,35 [0,20 - 0,52]	0,38 [0,28 - 0,49]	0,16 [0,07 - 0,30]	0,19 [0,12 - 0,27]
CTZ+	0,73 [0,55 - 0,86]	0,58 [0,35 - 0,79]	0,51 [0,32 - 0,69]	0,33 [0,15 - 0,57]	0,28 [0,14 - 0,46]	0,16 [0,05 - 0,35]
ETN	0,57 [0,37 - 0,76]	0,61 [0,47 - 0,73]	0,34 [0,18 - 0,54]	0,36 [0,24 - 0,50]	0,16 [0,06 - 0,31]	0,17 [0,10 - 0,28]
ETN+	0,68 [0,49 - 0,84]	0,70 [0,58 - 0,81]	0,44 [0,26 - 0,64]	0,46 [0,33 - 0,60]	0,23 [0,11 - 0,41]	0,25 [0,16 - 0,37]
GOL+	0,63 [0,41 - 0,81]	0,66 [0,50 - 0,79]	0,39 [0,21 - 0,61]	0,41 [0,26 - 0,58]	0,20 [0,08 - 0,38]	0,21 [0,11 - 0,35]
IFX+	0,57 [0,38 - 0,74]	0,59 [0,48 - 0,70]	0,33 [0,18 - 0,51]	0,35 [0,25 - 0,46]	0,15 [0,07 - 0,29]	0,17 [0,10 - 0,25]
Int cDMARDs	0,46 [0,25 - 0,70]	0,49 [0,30 - 0,68]	0,24 [0,10 - 0,46]	0,26 [0,12 - 0,44]	0,10 [0,03 - 0,25]	0,11 [0,04 - 0,23]
PBO	0,13 [0,04 - 0,31]	0,16 [0,06 - 0,32]	0,04 [0,01 - 0,13]	0,05 [0,02 - 0,14]	0,01 [0,00 - 0,04]	0,01 [0,00 - 0,05]
TCZ	0,64 [0,45 - 0,80]	0,72 [0,60 - 0,82]	0,40 [0,23 - 0,59]	0,48 [0,35 - 0,62]	0,20 [0,09 - 0,36]	0,27 [0,17 - 0,39]
TCZ s.c.	0,57 [0,32 - 0,80]		0,34 [0,14 - 0,60]		0,16 [0,05 - 0,37]	
TCZ s.c.+	0,61 [0,41 - 0,79]		0,38 [0,20 - 0,58]		0,18 [0,08 - 0,35]	
TCZ+	0,65 [0,46 - 0,80]	0,72 [0,58 - 0,83]	0,41 [0,24 - 0,59]	0,48 [0,34 - 0,63]	0,21 [0,10 - 0,36]	0,27 [0,16 - 0,40]
TOF 10 mg	0,61 [0,41 - 0,78]		0,37 [0,20 - 0,57]		0,18 [0,08 - 0,34]	
TOF 5 mg	0,56 [0,36 - 0,74]		0,32 [0,17 - 0,51]		0,15 [0,06 - 0,29]	

4.3.4. Analyse de sensibilité n° 4

L'introduction de deux études supplémentaires (TEAR et TEMPO) dans la méta-analyse en réseau principale a augmenté l'incertitude portant sur les paramètres d'hétérogénéité estimés et a conduit à une estimation moins précise des effets des traitements relatifs aux csDMARDs (figures 10a-10d). L'effet de l'ADA, de l'ETN, de l'ETN+ et des csDMARDs intensifs était moins favorable aux csDMARDs comparativement aux résultats de l'analyse de référence des figures 9a-9d. Les TCZ (avec et sans MTX) étaient encore les interventions présentant les réponses ACR les plus élevées (classement médian 2, probabilité d'être le meilleur : 0,26 pour TCZ et 0,31 pour TCZ+, respectivement).

L'intégration de ces deux études lors de l'estimation de l'effet du traitement à l'inclusion a augmenté l'effet du traitement sur les réponses ACR20, ACR50 et ACR70 pour les csDMARDs par rapport à l'analyse principale. Cela accroît l'hétérogénéité de l'effet du traitement (tableau 13). L'augmentation de l'effet du traitement à l'inclusion a conduit à une augmentation des probabilités de réalisation des réponses ACR pour toutes les interventions sauf pour l'ADA, l'ETN, l'ETN+ et les csDMARDs intensifs.

Figure10a. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients naïfs de MTX – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

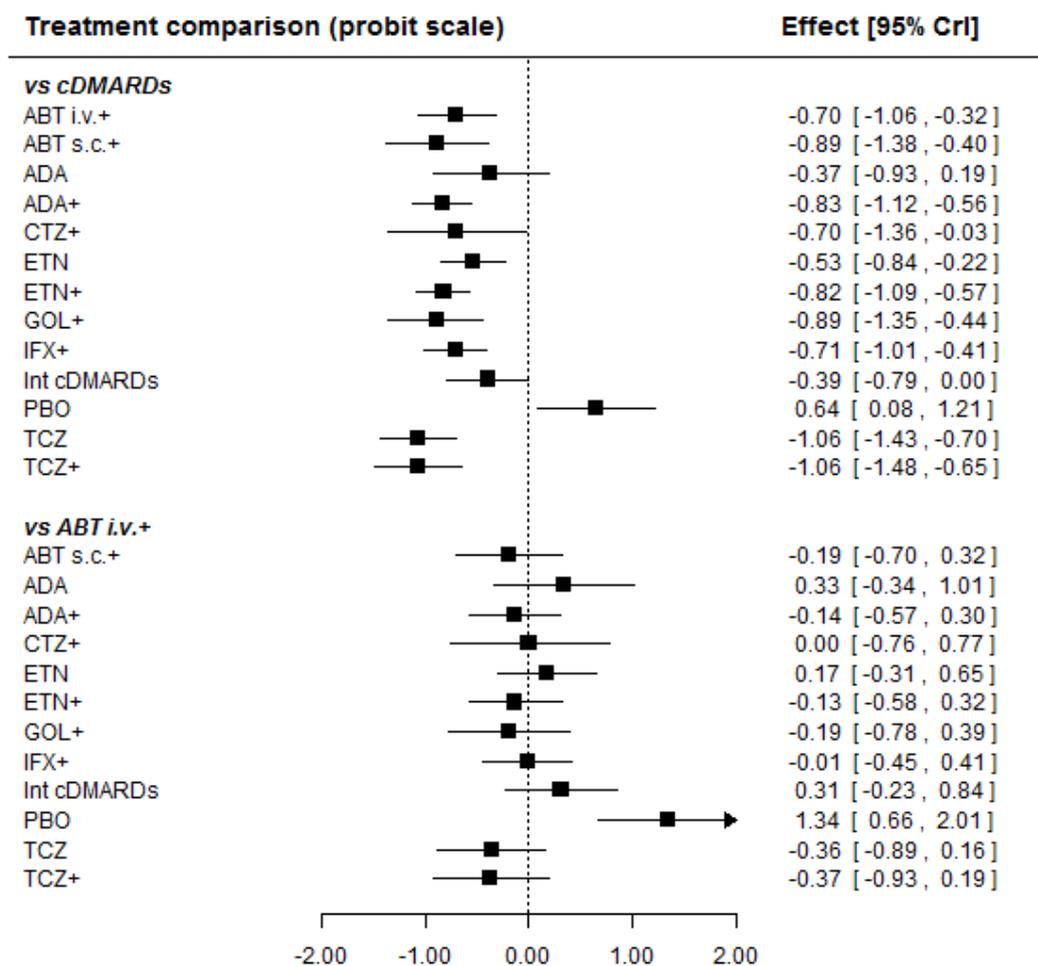


Figure 10b. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients naïfs de MTX – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

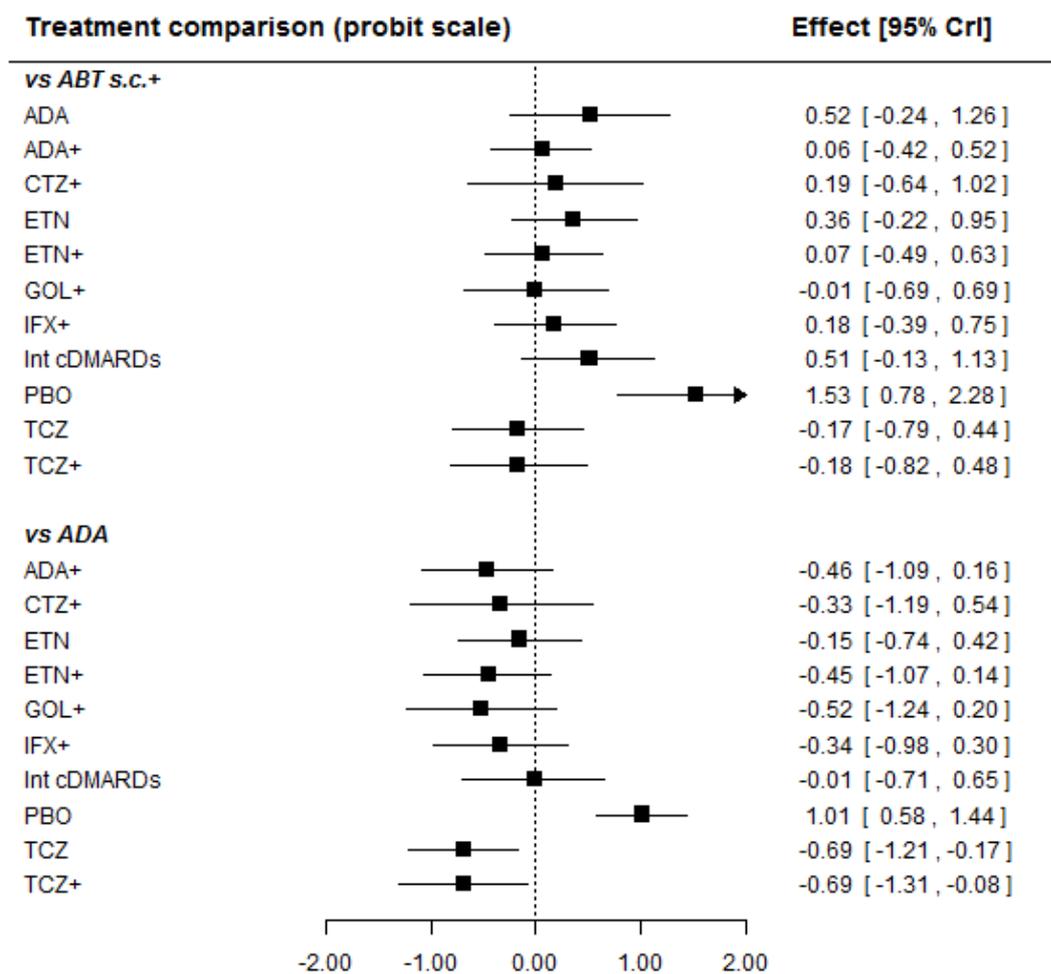


Figure 10c. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients naïfs de MTX – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

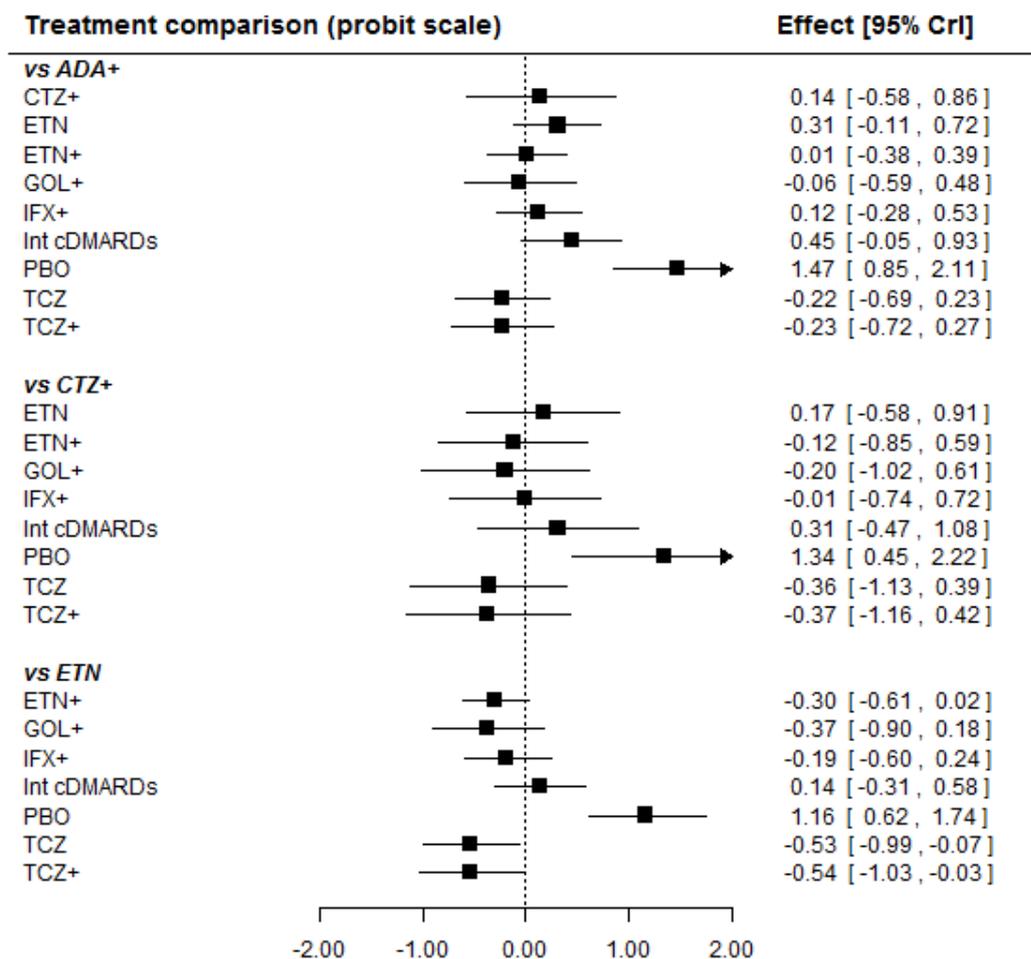


Figure 10d. Réponses ACR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par MTX + patients naïfs de MTX – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

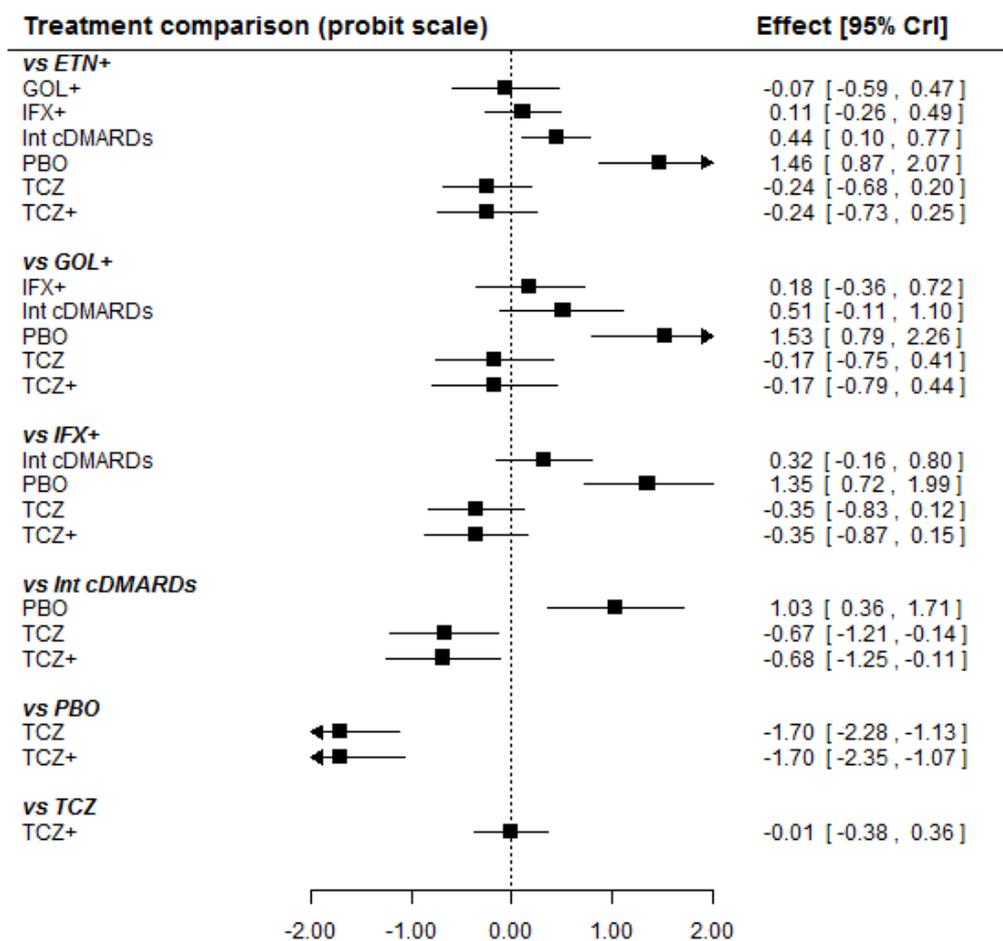


Tableau 13. Réponse ACR : Probabilités – Analyse de sensibilité – Patients ayant été traités par MTX + patients naïfs de MTX – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	ACR20 [95 % CrI]		ACR50 [95 % CrI]		ACR70 [95 % CrI]	
csDMARDs	0,33 [0,27 - 0,40]	0,31 [0,26 - 0,36]	0,14 [0,11 - 0,19]	0,13 [0,10 - 0,17]	0,05 [0,03 - 0,07]	0,04 [0,03 - 0,06]
ABT i.v.+	0,60 [0,44 - 0,75]	0,58 [0,45 - 0,71]	0,36 [0,22 - 0,51]	0,34 [0,23 - 0,47]	0,17 [0,08 - 0,29]	0,16 [0,09 - 0,26]
ABT s.c.+	0,68 [0,47 - 0,84]	0,65 [0,48 - 0,80]	0,43 [0,24 - 0,64]	0,41 [0,25 - 0,58]	0,22 [0,10 - 0,40]	0,21 [0,10 - 0,36]
ADA	0,48 [0,26 - 0,70]	0,48 [0,29 - 0,68]	0,24 [0,10 - 0,45]	0,25 [0,12 - 0,43]	0,10 [0,03 - 0,24]	0,11 [0,04 - 0,23]
ADA+	0,66 [0,53 - 0,77]	0,63 [0,52 - 0,73]	0,41 [0,29 - 0,54]	0,38 [0,28 - 0,49]	0,20 [0,12 - 0,31]	0,19 [0,12 - 0,27]
CTZ+	0,61 [0,34 - 0,83]	0,58 [0,35 - 0,79]	0,36 [0,14 - 0,62]	0,33 [0,15 - 0,57]	0,17 [0,05 - 0,39]	0,16 [0,05 - 0,35]
ETN	0,54 [0,40 - 0,68]	0,61 [0,47 - 0,73]	0,29 [0,19 - 0,43]	0,36 [0,24 - 0,50]	0,13 [0,07 - 0,22]	0,17 [0,10 - 0,28]
ETN+	0,65 [0,53 - 0,76]	0,70 [0,58 - 0,81]	0,40 [0,29 - 0,53]	0,46 [0,33 - 0,60]	0,20 [0,12 - 0,30]	0,25 [0,16 - 0,37]
GOL+	0,68 [0,49 - 0,83]	0,66 [0,50 - 0,79]	0,43 [0,25 - 0,63]	0,41 [0,26 - 0,58]	0,22 [0,10 - 0,39]	0,21 [0,11 - 0,35]
IFX+	0,61 [0,48 - 0,73]	0,59 [0,48 - 0,70]	0,36 [0,24 - 0,49]	0,35 [0,25 - 0,46]	0,17 [0,10 - 0,27]	0,17 [0,10 - 0,25]
Int cDMARDs	0,48 [0,32 - 0,65]	0,49 [0,30 - 0,68]	0,25 [0,14 - 0,40]	0,26 [0,12 - 0,44]	0,10 [0,04 - 0,20]	0,11 [0,04 - 0,23]
PBO	0,14 [0,05 - 0,32]	0,16 [0,06 - 0,32]	0,04 [0,01 - 0,13]	0,05 [0,02 - 0,14]	0,01 [0,00 - 0,04]	0,01 [0,00 - 0,05]
TCZ	0,73 [0,59 - 0,85]	0,72 [0,60 - 0,82]	0,50 [0,34 - 0,66]	0,48 [0,35 - 0,62]	0,27 [0,16 - 0,42]	0,27 [0,17 - 0,39]
TCZ+	0,74 [0,57 - 0,86]	0,72 [0,58 - 0,83]	0,50 [0,33 - 0,67]	0,48 [0,34 - 0,63]	0,27 [0,15 - 0,44]	0,27 [0,16 - 0,40]

4.3.5. Analyse de sensibilité n° 5

L'inclusion des études NCT01895309, PLANETRA et NCT01936181 portant sur les biosimilaires dans la méta-analyse en réseau réalisée sur les patients ayant été traités par MTX n'a pas affecté les résultats, car ces deux essais ne permettaient pas de contribuer à un réseau fermé (figures 11a-11d et tableau 14). Au niveau $\alpha=5\%$, les effets du traitement étaient statistiquement significatifs seulement pour ETN+, GOL+, IFX+, TCZ (avec et sans MTX). Le TCZ+ paraissait comme l'intervention présentant la réponse EULAR la plus élevée (médiane de rang 2, meilleure probabilité 0,33) (cf. annexe 3-9).

Figure 11a. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

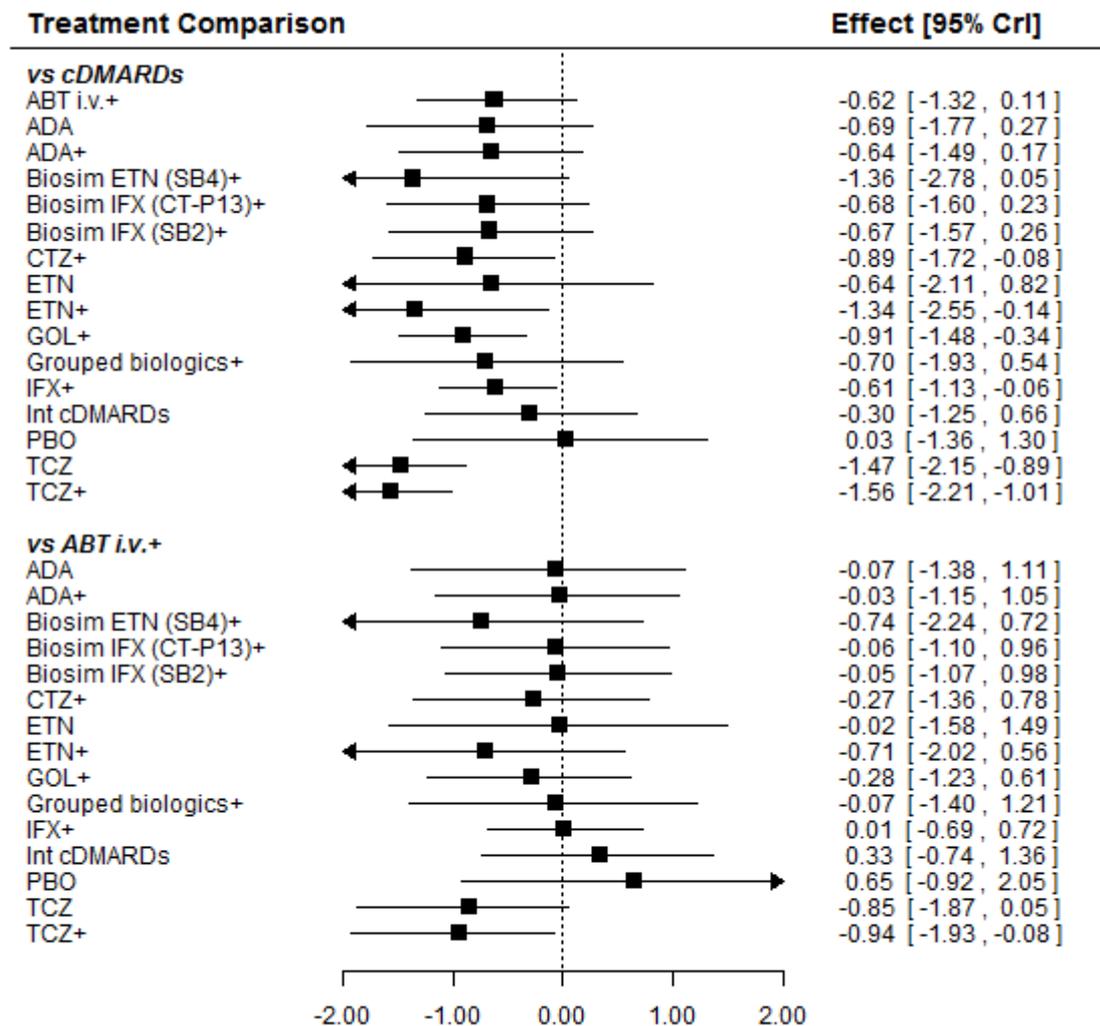


Figure 11b. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

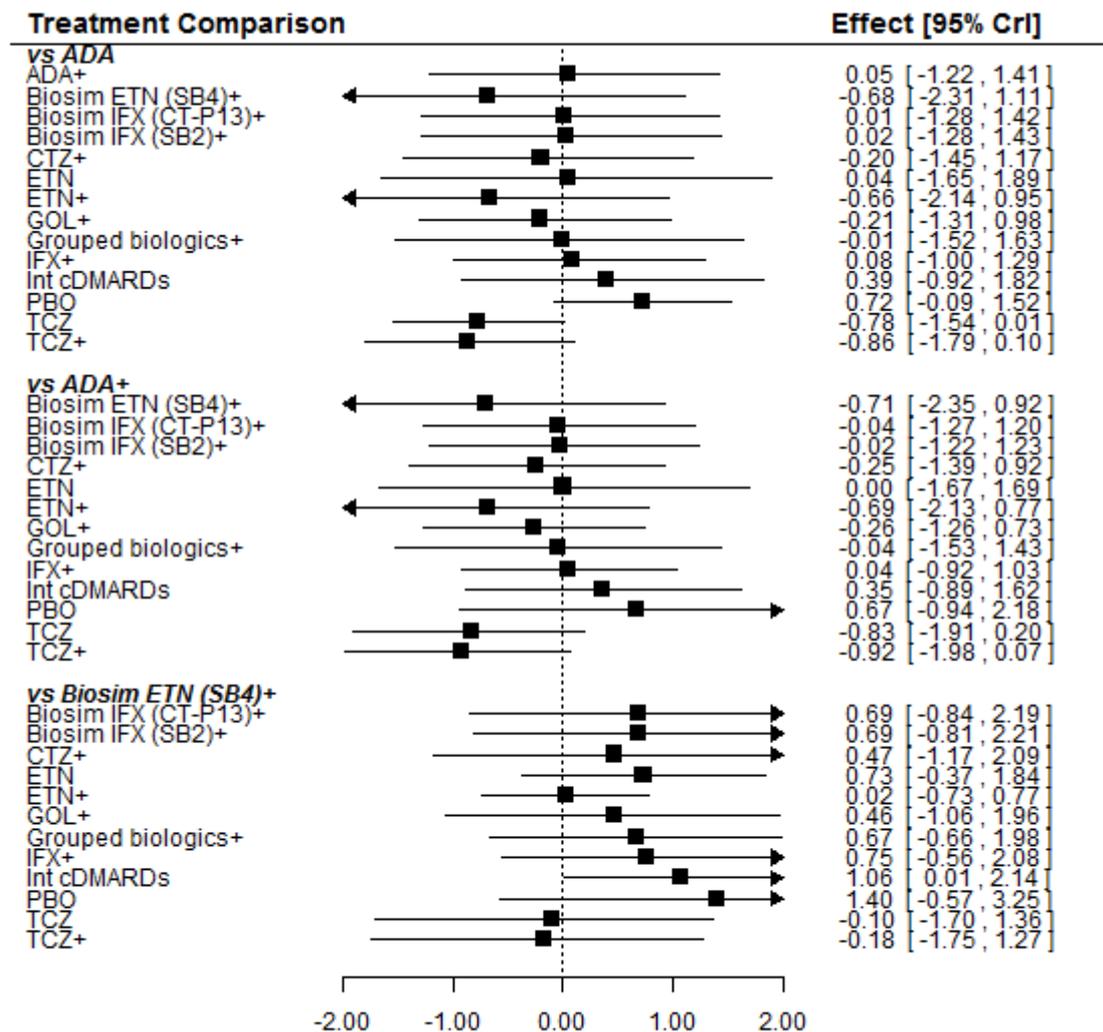


Figure 11c. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

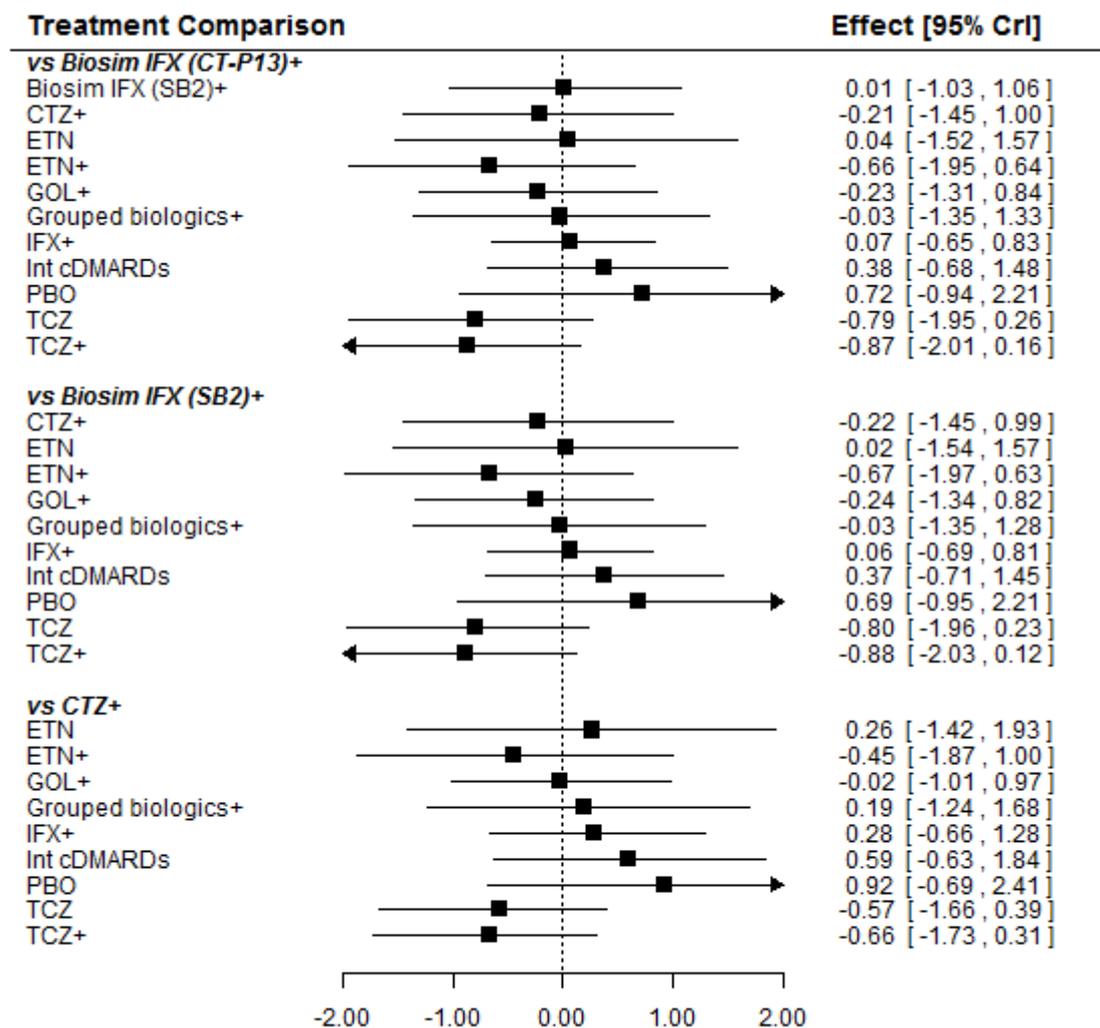


Figure 11d. Réponses EULAR : Effets des interventions (échelle probit) – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

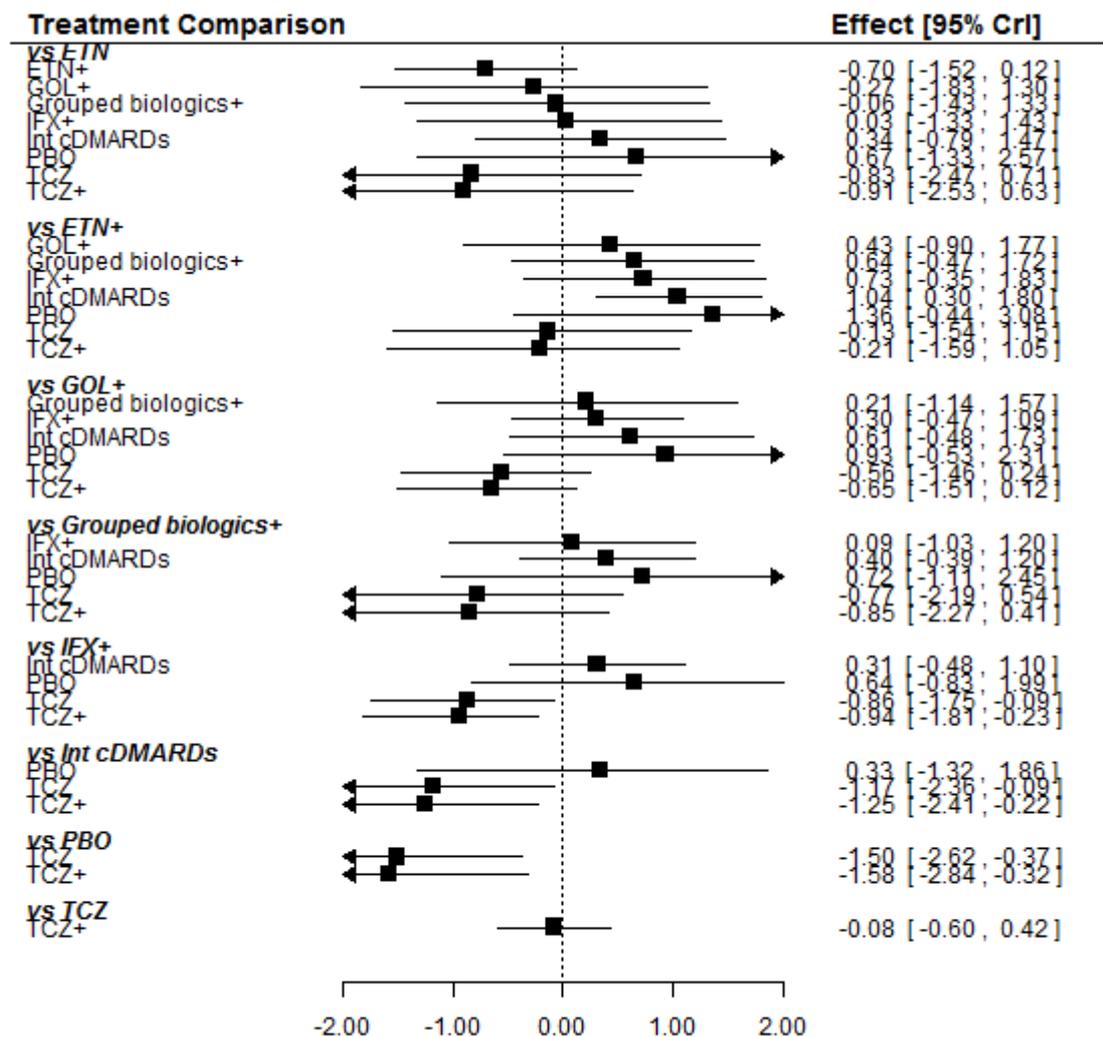


Tableau 14. Réponses EULAR : Probabilités – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + biosimilaires – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	Au moins une réponse modérée [95 % CrI]		Au moins une bonne réponse [95 % CrI]	
csDMARDs	0,45 [0,38 - 0,52]	0,45 [0,38 - 0,52]	0,09 [0,06 - 0,14]	0,12 [0,08 - 0,17]
ABT i.v.+	0,69 [0,40 - 0,89]	0,69 [0,40 - 0,89]	0,24 [0,07 - 0,52]	0,28 [0,09 - 0,57]
ADA	0,72 [0,34 - 0,95]	0,70 [0,34 - 0,94]	0,27 [0,05 - 0,68]	0,30 [0,07 - 0,70]
ADA+	0,70 [0,38 - 0,92]	0,70 [0,38 - 0,91]	0,25 [0,06 - 0,58]	0,30 [0,08 - 0,63]
Biosim ETN [SB4]+	0,89 [0,43 - 1,00]		0,52 [0,08 - 0,93]	
Biosim IFX [CT-P13]+	0,71 [0,36 - 0,93]		0,26 [0,06 - 0,62]	
Biosim IFX [SB2]+	0,71 [0,35 - 0,93]		0,26 [0,05 - 0,61]	
CTZ+	0,78 [0,47 - 0,95]	0,77 [0,48 - 0,94]	0,33 [0,10 - 0,67]	0,38 [0,13 - 0,70]
ETN	0,70 [0,17 - 0,98]	0,70 [0,18 - 0,98]	0,25 [0,02 - 0,79]	0,30 [0,02 - 0,82]
ETN+	0,89 [0,50 - 0,99]	0,89 [0,50 - 0,99]	0,51 [0,11 - 0,89]	0,56 [0,14 - 0,92]
GOL+	0,78 [0,58 - 0,92]	0,78 [0,58 - 0,92]	0,34 [0,15 - 0,58]	0,39 [0,19 - 0,63]
Grouped biologics+	0,72 [0,25 - 0,97]	0,72 [0,26 - 0,97]	0,27 [0,03 - 0,74]	0,31 [0,04 - 0,77]
IFX+	0,69 [0,47 - 0,85]	0,69 [0,47 - 0,85]	0,24 [0,10 - 0,44]	0,28 [0,12 - 0,50]
Int cDMARDs	0,57 [0,21 - 0,87]	0,57 [0,22 - 0,87]	0,15 [0,02 - 0,48]	0,19 [0,03 - 0,53]
PBO	0,44 [0,08 - 0,89]	0,43 [0,08 - 0,87]	0,09 [0,00 - 0,52]	0,11 [0,01 - 0,53]
TCZ	0,91 [0,77 - 0,98]	0,91 [0,76 - 0,98]	0,56 [0,32 - 0,81]	0,60 [0,35 - 0,83]
TCZ+	0,92 [0,80 - 0,98]	0,92 [0,80 - 0,98]	0,59 [0,36 - 0,82]	0,63 [0,40 - 0,85]

4.3.6. Analyse de sensibilité n° 6

L'ajout de trois études chez des patients ayant déjà utilisé des bDMARDs (OPTION, RAPID1 et Yamamoto 2011) dans la méta-analyse en réseau portant sur les patients ayant été traités par MTX, a augmenté l'effet du traitement par CTZ+, comparativement aux csDMARDs. CTZ+, ENT+ et TCZ (en association ou non avec le MTX) étaient les interventions associées aux effets de traitement les plus importants par rapport aux csDMARDs (figures 12a -12d). Les preuves étaient insuffisantes pour différencier ces quatre traitements. Les rangs médians pour TCZ, TCZ+, CTZ+ et ENT+ étaient respectivement 2, 3, 3 et 3. La probabilité d'être le traitement le plus efficace pour CTZ+, ENT+, TCZ et TCZ+ était de 0,26, 0,19, 0,14 et 0,34, respectivement (cf. annexe 3-9).

Par ailleurs, l'inclusion de ces trois études a réduit l'effet du traitement (réponses ACR20, ACR50 et ACR70 pour les csDMARDs). Par conséquent, les probabilités des réponses ACR

estimées pour les interventions ont diminué par rapport à l'analyse principale, exception faite pour CTZ+ (tableau 15).

Figure 12a. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + proportion de patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

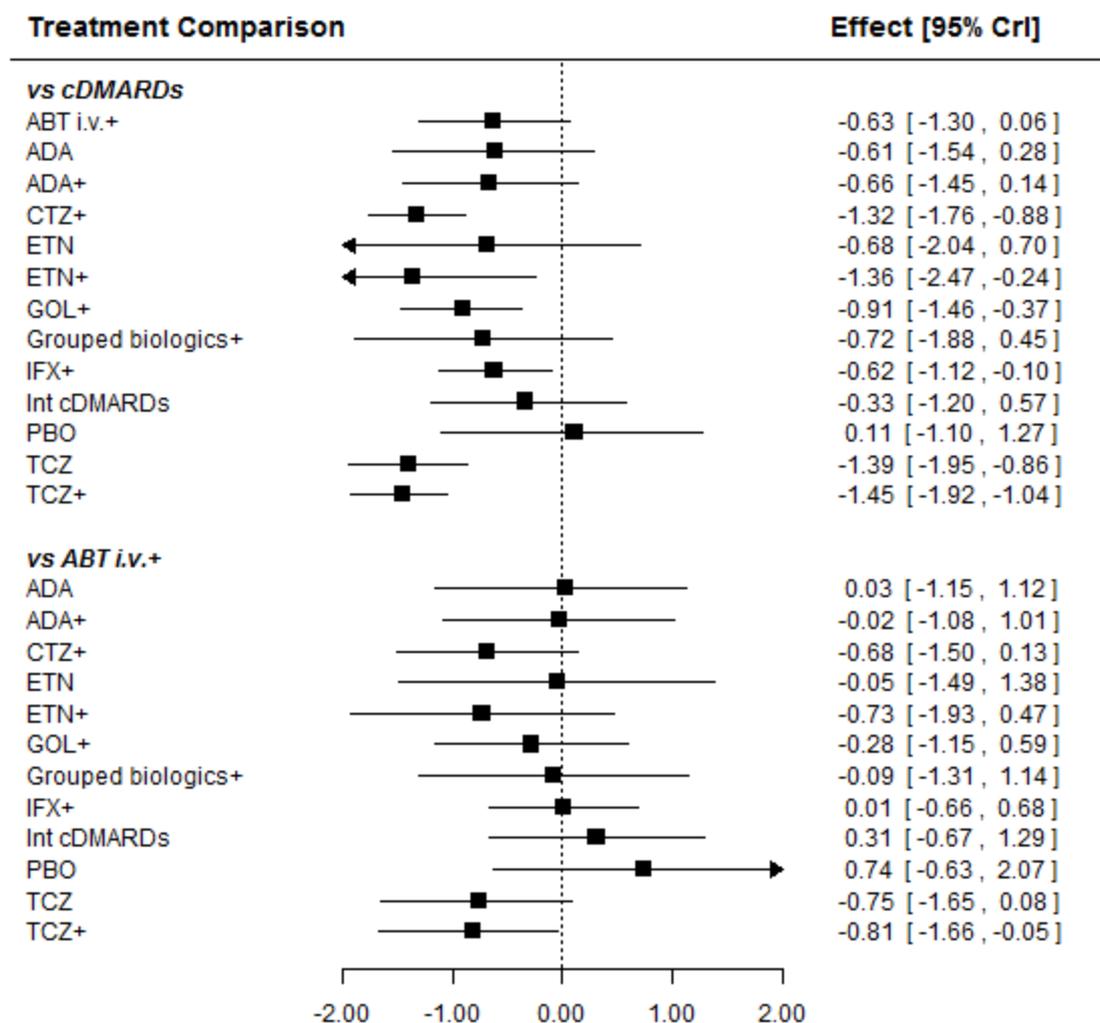


Figure 12b. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

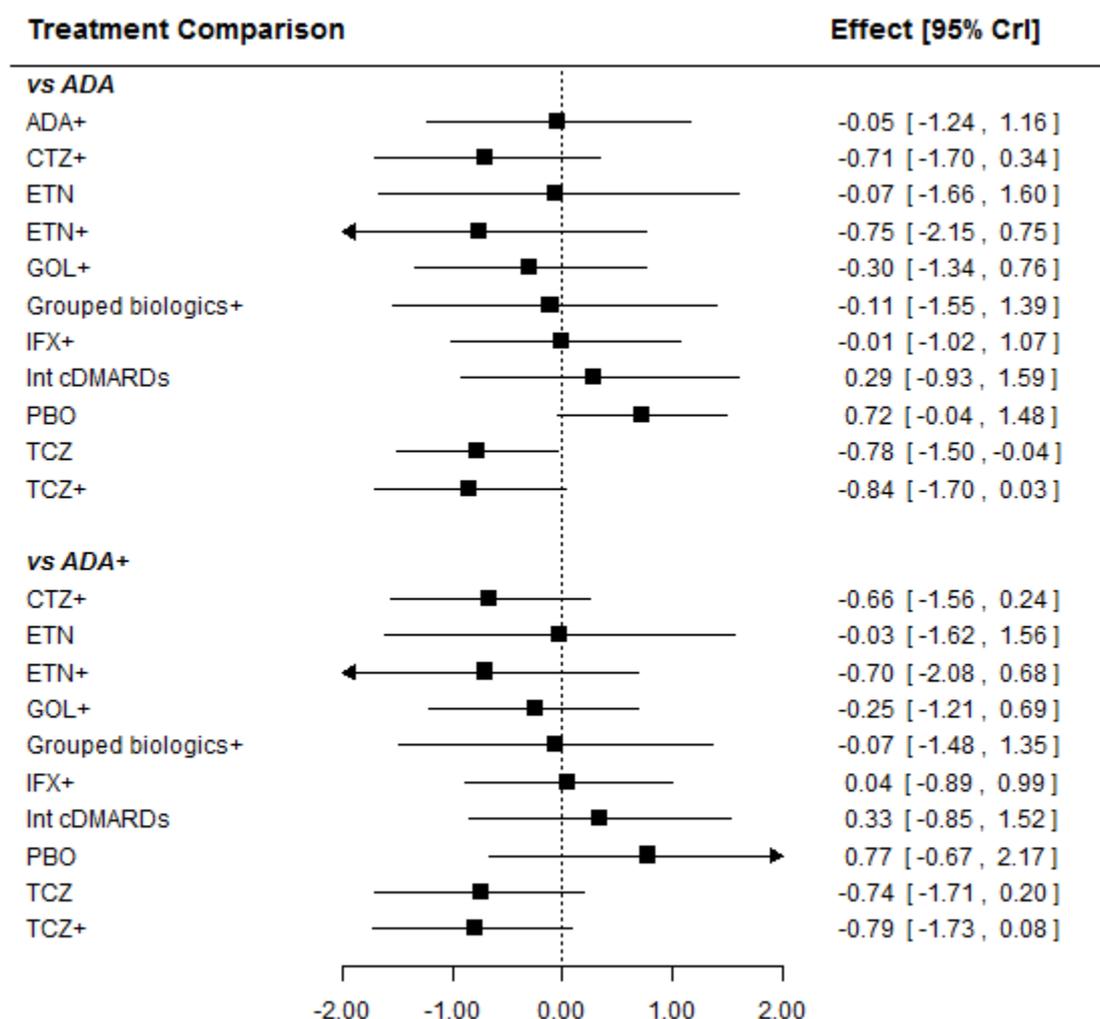


Figure 12c. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

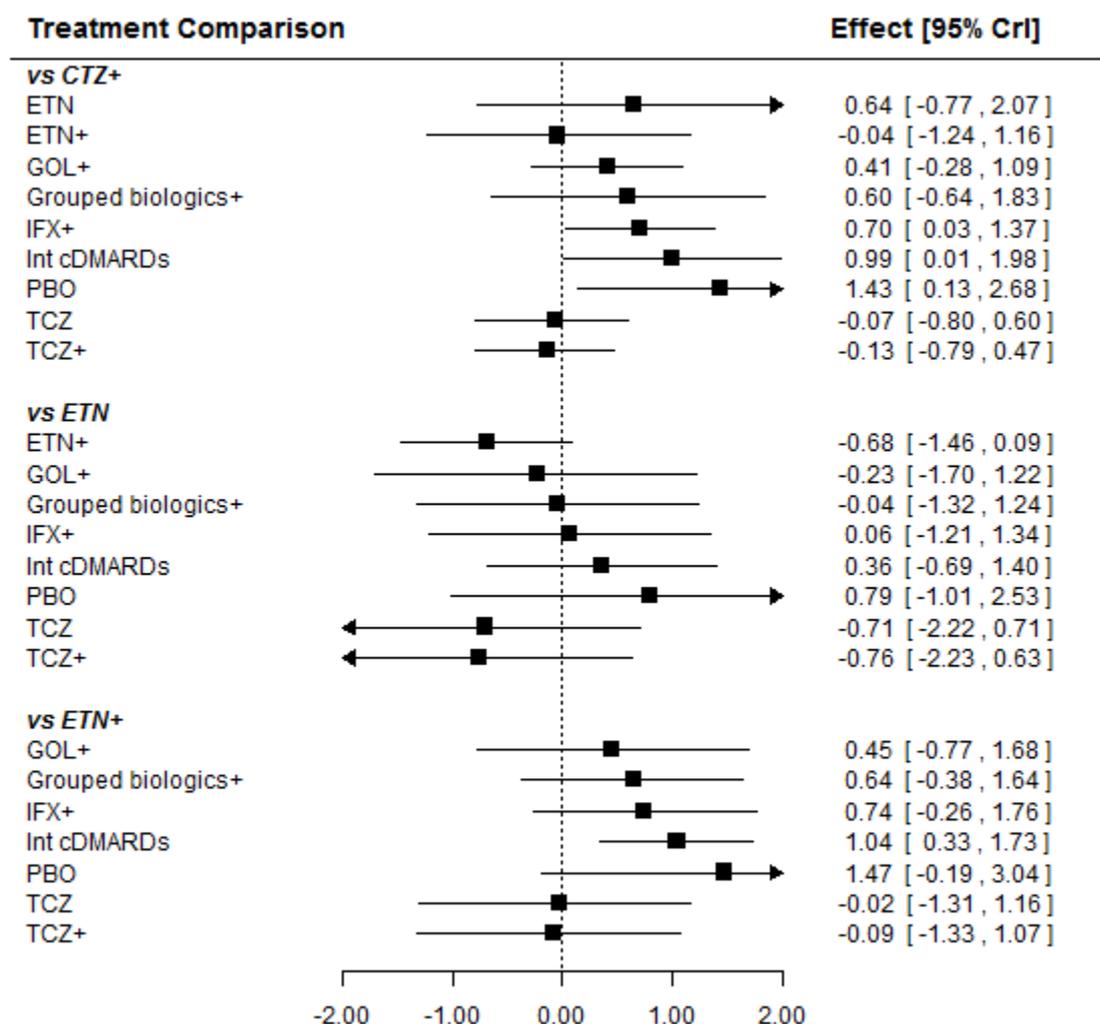


Figure 12d. Réponses EULAR : Effets des interventions – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » [\(84\)](#)

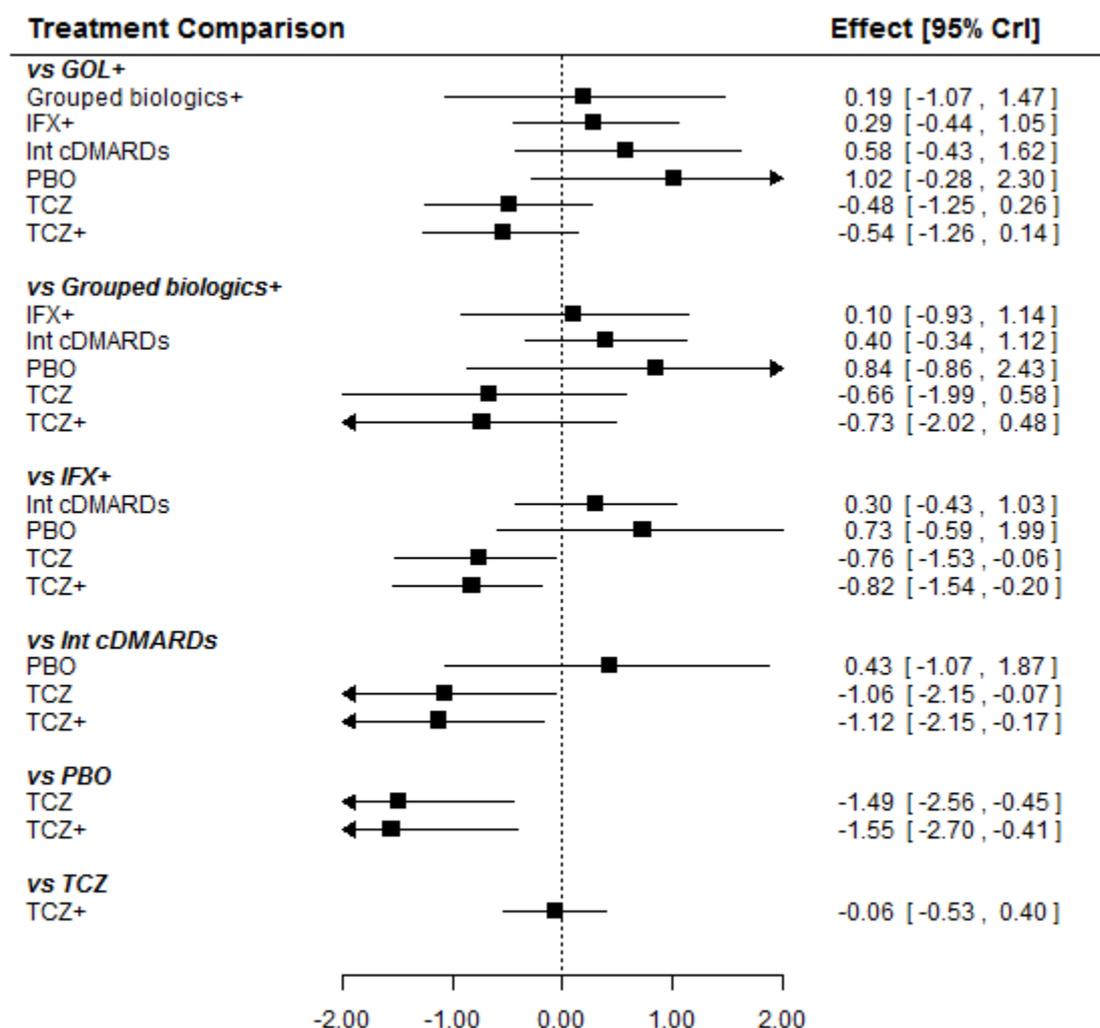


Tableau 15. Réponses EULAR : Probabilités – Analyse de sensibilité – Patients ayant été traités par des csDMARDs (MTX) + patients ayant été traités par des bDMARDs – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Intervention	Au moins une réponse modérée [95 % CrI]		Au moins une bonne réponse [95 % CrI]	
	csDMARDs	0,41 [0,34 - 0,48]	0,45 [0,38 - 0,52]	0,10 [0,06 - 0,14]
ABT i.v.+	0,66 [0,38 - 0,86]	0,69 [0,40 - 0,89]	0,25 [0,08 - 0,51]	0,28 [0,09 - 0,57]
ADA	0,65 [0,30 - 0,91]	0,70 [0,34 - 0,94]	0,25 [0,05 - 0,61]	0,30 [0,07 - 0,70]
ADA+	0,67 [0,35 - 0,89]	0,70 [0,38 - 0,91]	0,26 [0,07 - 0,57]	0,30 [0,08 - 0,63]
CTZ+	0,86 [0,73 - 0,94]	0,77 [0,48 - 0,94]	0,50 [0,31 - 0,69]	0,38 [0,13 - 0,70]
ETN	0,67 [0,17 - 0,96]	0,70 [0,18 - 0,98]	0,27 [0,02 - 0,77]	0,30 [0,02 - 0,82]
ETN+	0,87 [0,50 - 0,99]	0,89 [0,50 - 0,99]	0,52 [0,14 - 0,88]	0,56 [0,14 - 0,92]
GOL+	0,75 [0,55 - 0,90]	0,78 [0,58 - 0,92]	0,35 [0,16 - 0,58]	0,39 [0,19 - 0,63]
Grouped biologics+	0,69 [0,25 - 0,95]	0,72 [0,26 - 0,97]	0,28 [0,04 - 0,72]	0,31 [0,04 - 0,77]
IFX+	0,65 [0,44 - 0,82]	0,69 [0,47 - 0,85]	0,25 [0,11 - 0,45]	0,28 [0,12 - 0,50]
Int cDMARDs	0,54 [0,21 - 0,84]	0,57 [0,22 - 0,87]	0,16 [0,03 - 0,47]	0,19 [0,03 - 0,53]
PBO	0,37 [0,06 - 0,81]	0,43 [0,08 - 0,87]	0,08 [0,00 - 0,43]	0,11 [0,01 - 0,53]
TCZ	0,88 [0,73 - 0,96]	0,91 [0,76 - 0,98]	0,53 [0,31 - 0,75]	0,60 [0,35 - 0,83]
TCZ+	0,89 [0,78 - 0,96]	0,92 [0,80 - 0,98]	0,56 [0,37 - 0,74]	0,63 [0,40 - 0,85]

4.4. Évaluation de la qualité statistique des méta-analyses HAS & SchARR

4.4.1. Analyse de l'hétérogénéité

➤ Analyse de la variance inter-études

L'analyse de la variance inter-études estimée dans l'analyse principale des méta-analyses en réseau est présentée dans le tableau 16. L'estimation ponctuelle du paramètre associé à l'hétérogénéité suggérait :

- une hétérogénéité estimée relative aux réponses EULAR plus importante que celle des réponses ACR ;
- une hétérogénéité estimée modérée dans les analyses principales des méta-analyses en réseau portant sur les réponses ACR.

Tableau 16. Estimation de l'écart-type inter-études – Analyses principales – Source : Rapport technique « Revue systématique de la littérature et méta-analyses en réseau – Université de Sheffield » (84)

Réponses	Populations	Médiane (95 % CrI)	P(hétérogénéité faible)	P(hétérogénéité modérée)	P(forte hétérogénéité)	Modèle utilisé
Réponses ACR	Patients ayant été traités par des csDMARDs (MTX)	0,14 [0,03 - 0,48]	0,13	0,68	0,19	Modèle à effet aléatoire avec un <i>a priori</i> de référence
Réponses EULAR	Patients ayant été traités par des csDMARDs (MTX)	0,32 [0,16 - 0,53]	0	0,30	0,70	Modèle à effet aléatoire avec un <i>a priori</i> informatif

➤ Les analyses de sensibilité

Les analyses de sensibilité sur les méta-analyses principales ont été présentées dans la section 2.4. Les sept analyses réalisées ne montraient pas des différences importantes entre les méta-analyses en réseau principales et l'ajout de patients provenant d'autres types de populations. Des effets plutôt quantitatifs ont été observés au niveau des paramètres dérivant l'hétérogénéité des effets de traitement et le niveau des probabilités de réponses ACR ou EULAR estimées.

➤ Les méta-régressions

Afin d'explorer si l'hétérogénéité pouvait être expliquée par la durée de la maladie (un modificateur de l'effet du traitement), des méta-régressions portant sur la réponse ACR au sein des deux populations d'analyse ont été planifiées. Les résultats de l'annexe A3.10 suggéraient que l'ancienneté de la PR ne semblait pas être un modificateur de l'effet du traitement : le coefficient associé à cette variable n'était pas statistiquement significatif et était très proche de la valeur zéro.

Les méta-régressions concernant la réponse EULAR n'ont pas été réalisées en raison du faible nombre d'études disponibles.

4.4.2. Analyse de la cohérence des preuves

L'incohérence des preuves a été vérifiée à l'aide de trois contrastes de la réponse EULAR chez les patients qui avaient été traités par des csDMARDs (MTX) en utilisant la méthode

« *node-split* » (37). La preuve directe provenait des études ACT-RAY et SURPRISE, suggérant que l'effet thérapeutique du TCZ+ par rapport au TCZ était égal à -0,31, selon l'échelle probit. Mais les preuves indirectes provenant des études SATORI et TOWARD suggéraient que l'effet du traitement du TCZ+ par rapport au TCZ était égal à 0,98. Ceci pourrait indiquer une incohérence entre les preuves directes et indirectes pour cette comparaison de traitements (la p-valeur était égale à 0,01). L'effet du traitement estimé était de -0,09 dans la méta-analyse en réseau provenant principalement de la preuve directe. Il n'y avait pas d'éléments suggérant des incohérences dans les autres analyses principales des méta-analyses en réseau.

4.4.3. Cohérence des intervalles de confiance associés aux réponses ACR et EULAR

Les résultats des effets de traitements selon un intervalle de prédiction à 95 % sont présentés dans l'annexe A3-8. L'intervalle de prédiction à 95 % (PrI) fournit une plage de distributions d'observations prédictives de l'effet du traitement aléatoirement sélectionnées dans une population. Les PrI à 95 % fondés sur des distributions prédictives étaient toujours plus larges que les CrI à 95 % utilisant les distributions *a posteriori*.

Les effets de l'ADA, des csDMARDs intensifs et du CTZ+ comparés aux csDMARDs ne devenaient pas statistiquement significatifs. Comparativement aux csDMARDs, seuls GOL+, TCZ et TCZ+ restaient statistiquement significatifs concernant la réponse EULAR (GOL+ était proche de la significativité statistique au niveau $\alpha=5\%$).

5. Synthèse et discussion

5.1. Synthèse

Au total, 46 essais cliniques répondaient aux critères d'éligibilité de la revue systématique de la littérature HAS & SchARR. Dix ont porté sur des patients naïfs de MTX et 36 ont porté sur la population d'entrée du modèle PR France, à savoir les patients qui ont été traités par des csDMARDs (MTX) dont trois évaluaient des biosimilaires.

Les analyses principales des méta-analyses en réseau ont porté sur 33 essais cliniques, y compris celles identifiées dans l'évaluation du NICE en mai 2013, (12). Quant aux analyses de sensibilité, elles ont porté sur 20 études (14 identifiées lors de l'évaluation du NICE en mai 2013 et six nouveaux essais).

Par rapport à la revue systématique de littérature réalisée par le NICE en mai 2013, la revue systématique de la HAS & SchARR a permis d'identifier 11 nouveaux essais randomisés comparatifs, conformément au protocole d'éligibilité de la revue systématique. Parmi ceux-ci, deux ont porté sur les patients naïfs de MTX, et neuf ont porté sur des patients atteints de PR et qui avaient été traités par des csDMARDs (MTX). Tous ces nouveaux essais ont inclus des patients atteints de PR active sévère et ont colligé des réponses ACR. En revanche, les réponses EULAR étaient seulement disponibles dans cinq des 11 nouveaux essais retenus.

La majorité des études présentait un risque de biais afférant aux procédures de randomisation et de masquage des traitements attribués. Ce risque a été gradé d'un niveau faible, dans la mesure où il concernait principalement la description des éléments méthodologiques qui n'étaient pas suffisamment explicités dans plusieurs études retenues.

Le comparateur commun des méta-analyses en réseau était un csDMARD (MTX).

Concernant la réponse ACR, ABT i.v.+, ABT s.c.+, ADA, ADA+, CTZ+, ETN, ETN+, GOL+, IFX+, TCZ, TCZ+ et les csDMARDs intensifs ont présenté des effets de traitement positifs comparativement aux csDMARDs, les effets les plus importants étant associés à TCZ (en association ou non avec le MTX) et ETN+. Les effets de l'ADA, du CTZ+ et des csDMARDs intensifs n'étaient pas statistiquement significatifs selon l'intervalle de prédiction et étaient à la limite de la significativité statistique selon l'intervalle de crédibilité.

Concernant la réponse EULAR, ABT i.v.+ , ADA, ADA+, CTZ+, ETN, ETN+, GOL+, IFX+, TCZ, TCZ+, les bDMARDs+ regroupés et csDMARDs intensifs ont présenté des effets cliniques positifs, comparativement aux csDMARDs. TCZ et TCZ+ étaient les interventions associées à des résultats statistiquement significatifs selon les intervalles de crédibilité et les intervalles de prédiction. Dans les analyses de sensibilité, en supposant que les biosimilaires étaient différents du traitement parent, l'ETN biosimilaire (SB4)+ était l'intervention qui présentait la réponse ACR la plus élevée et le TCZ+ restait l'intervention qui présentait la réponse EULAR la plus élevée, avec le biosimilaire ETN (SB4)+ classé en 2^e rang. Les autres analyses de sensibilité ne changeaient pas le sens des résultats des analyses principales, mais affectaient légèrement le niveau de réponse et l'incertitude liée aux paramètres décrivant leur précision statistique.

Comparés entre eux, les résultats des méta-analyses en réseau ne détectaient généralement pas de différences significatives entre les bDMARDs.

5.2. Discussion

L'originalité de la revue systématique et des méta-analyses réalisées par la HAS & SchARR réside dans :

- la différenciation des traitements bDMARDs en première ligne de traitement en fonction des patients naïfs de MTX et ceux ayant été traités par csDMARDs. Ce choix est non seulement conforme aux indications des produits étudiés, mais facilite également les comparaisons des deux populations à l'inclusion ;
- la prise en compte des biosimilaires dans la revue systématique et les analyses de sensibilité des méta-analyses en réseau, alors que la plupart des revues précédentes, à quelques exceptions (CADTH, 2018 (32)), ont principalement inclus des anti-TNF en ne tenant pas compte de leurs médicaments biosimilaires. **L'intégration des biosimilaires permettait notamment d'estimer des réponses qui seront incluses dans le modèle PR France ;**
- la prise en compte des réponses ACR et EULAR comme critères d'efficacité définis sous forme de données ordonnées en considérant des catégories mutuellement exclusives. Ceci explique pourquoi nous avons choisi une approche bayésienne fondée sur un modèle aléatoire utilisant une fonction de lien probit (Dias *et al.*, 2013) (37) au lieu d'un lien d'odds qui est fréquemment utilisé pour estimer les probabilités à partir des données dichotomiques (Singh *et al.*, 2016, (89), ICER, 2017 (13), CADTH, 2018 (32)).

En dépit de nombreuses différences en termes de critères d'exclusion, d'inclusion de tsDMARDs (traitements de fond synthétiques ciblés tels que le baricitinib ou le tofacitinib) et du nombre d'études analysées, les analyses comparant les bDMARDs entre eux ont montré des résultats comparables aux méta-analyses récentes telles que celles de la Cochrane (Singh *et al.*, 2017) (123), de l'ICER (2017) (13) et de la CADTH (2018) (32). **Les bDMARDs (et leurs biosimilaires) comparés aux csDMARDs (i.e. MTX) ont clairement montré un bénéfice clinique supérieur ; les résultats n'ont cependant pas permis de mettre en évidence des différences d'efficacité entre bDMARDs.**

Nos résultats suggèrent que le TCZ (en monothérapie ou en association avec le MTX) se présente comme l'intervention ayant les réponses (EULAR ou ACR) les plus élevées par rapport aux csDMARDs. Ces résultats sont comparables avec les résultats d'une méta-analyse précédente (Alfonso-Cristancho *et al.*, 2017) (124) où le TCZ était, soit similaire, soit présentait des réponses supérieures aux autres bDMARDs. La supériorité du TCZ administré en monothérapie peut être expliquée par :

- un effet significatif en faveur des inhibiteurs des réactifs de phase aiguë (tels que le TCZ) par rapport aux autres bDMARDs ;
- le DAS 28 qui est susceptible de surestimer la réponse clinique EULAR, en raison du poids des réactifs de phase aiguë dans le calcul de la formule du DAS 28 (125). Il est également important de souligner que le niveau de la CRP est aussi une composante de la réponse ACR. Ainsi, il peut y avoir une inflammation résiduelle dans les articulations, même si le niveau de la CRP est faible.

Les méta-analyses réalisées présentent les limites suivantes :

- il y a peu d'essais contrôlés randomisés comparant directement les produits biologiques. La rareté des essais cliniques randomisés à comparaisons directes rend nécessaire l'utilisation des méta-analyses en réseau ;
- une limite potentielle de ce travail, et également dans d'autres méta-analyses, est la difficulté de comparer les différentes interventions selon une définition standard de la rémission acceptée de manière unanime. Les notions de rémission ou de faible activité de la maladie permettant une prise en charge ciblée de la PR sont récemment recommandées et de plus en plus utilisées dans la prise en charge des patients en pratique clinique en Europe. L'Agence européenne du médicament (EMA) (25) recommande d'utiliser la rémission comme critère principal d'évaluation dans les essais cliniques. Elle peut être définie selon les critères EULAR (DAS 28 < 2,6) ou conformément aux critères plus stricts EULAR-ACR (booléens ou indexés). **Néanmoins, il existe peu de données dans les essais cliniques randomisés et comparatifs disponibles colligeant cette variable (en tant que critère principal) et suffisantes pour réaliser des comparaisons pertinentes. De plus, les réponses ACR ou EULAR en tant que paramètres cliniques dans les modèles économiques sont fréquemment utilisées dans la modélisation économique (cf. chapitre 4). En effet, l'utilisation de la rémission n'est pas fréquente car elle présente le risque d'ignorer les gains associés à la transition d'un patient d'un état grave à un état meilleur (même si le patient n'est pas en état de rémission) (12) ;**
- les traitements de fond synthétiques ciblés (tsDMARDS) représentés par les nouveaux agents chimiques de type anti-enzymes, tels que le tofacitinib et le baricitinib, n'ont pas été inclus dans les méta-analyses en réseau principales (données non disponibles au cours de l'analyse statistique des méta-analyses en réseau principales). Ils seront intégrés dans la prochaine actualisation de ce rapport. Comme l'indiquent, par exemple, les avis de la CT ou le rapport de l'agence canadienne (CADTH, 2018) (32), ces traitements sont susceptibles d'avoir un impact sur la morbidité et la qualité de vie des patients atteints de PR. Toutefois, les avis de la CT relatifs à ces produits

(tofacitinib (30) et baricitinib (41)) ont souligné que l'impact de ces traitements devrait être limité, compte tenu :

- de l'absence de comparaison aux biothérapies de 3^e intention,
 - du faible recul en termes d'efficacité et de tolérance comparativement aux biothérapies disponibles,
 - et surtout des inquiétudes en termes de tolérance, notamment à long terme, portant en particulier sur les risques infectieux et les risques potentiels cardio-vasculaires et carcinogènes ;
- par ailleurs, bien que l'analyse systématique de la littérature clinique ait constaté un risque de biais faible afférent à la description de la procédure de randomisation et ait statistiquement montré que l'ancienneté de la PR n'était pas un modificateur de l'effet du traitement concernant la réponse ACR, d'autres limites statistiques devraient être prises en compte dans de futures méta-analyses :
- le nombre limité d'essais cliniques comparatifs randomisés portant sur la réponse EULAR dans les deux populations étudiées n'a pas permis de réaliser des méta-régressions susceptibles de tester si l'ancienneté de la PR représente un facteur modificateur du traitement. Ceci s'ajoute aux difficultés de comparaison des études en raison notamment de l'utilisation de différents critères de réponses, périodes d'inclusion, des doses des traitements biologiques et des traitements de fond synthétiques concomitants, ainsi que des *designs* des études,
 - seuls trois essais portant sur les produits biosimilaires répondant aux critères d'inclusion ont été identifiés. Comme ces derniers ont été comparés uniquement à leurs traitements biologiques, leur inclusion dans la méta-analyse en réseau n'a pas affecté les résultats des autres interventions dans la mesure où ils ne permettaient pas de former une boucle fermée dans le réseau.

6. Synthèse d'efficacité comparative sur les tsDMARDs et le sarilumab

Deux médicaments inhibiteurs sélectifs et réversibles des Janus kinases (enzymes impliquées dans la physiopathologie de la PR à travers la transduction de signaux de nombreuses cytokines) ont été évalués par la CT en 2017 : le baricitinib (OLUMIANT[®]) et le tofacitinib (XELJANZ[®]). Ces deux traitements appartiennent à une nouvelle classe thérapeutique définie par l'EULAR en 2016 et nommée « traitements de fond synthétiques ciblés » (tsDMARDs). Ils sont administrés par voie orale alors que les traitements de fond biologiques sont administrés par voie parentérale. Le schéma posologique du tofacitinib est un comprimé à 5 mg deux fois par jour ; celui du baricitinib est de 4 mg par jour. L'indication de l'AMM, correspondant à la prescription de ces deux traitements, est « *le traitement de la polyarthrite active modérée à sévère chez les patients adultes qui ont présenté une réponse inadéquate, ou une tolérance insuffisante à plusieurs traitements de fond (DMARDs)* ».

Un deuxième médicament appartenant à la classe des anticorps monoclonaux inhibiteurs de l'interleukine 6 (même classe thérapeutique que le tocilizumab) a été évalué par la CT en 2018. Il s'agit du sarilumab (KEVZARA[®]). Il est administré par voie sous-cutanée une fois toutes les 2 semaines. Il est indiqué dans le traitement de la PR après échec (réponse inadéquate ou tolérance insuffisante) d'un ou plusieurs traitements de fond, c'est-à-dire en 2^e intention (après échec des traitements de fond synthétiques conventionnels tels que le MTX) et en 3^e intention (échec d'un bDMARD), voire plus.

6.1. Synthèse des évaluations cliniques de la HAS (avis CT)

➤ Les tsDMARDs (baricitinib et tofacitinib)

L'évaluation de l'efficacité du baricitinib repose sur quatre essais cliniques randomisés en double aveugle de phase III. Le baricitinib a été comparé à un placebo et/ou à un comparateur actif (adalimumab). Trois essais (RA-BEAM, RA-BUILD et RA-BEACON) ont évalué l'efficacité et la tolérance du baricitinib en association au MTX conformément à l'indication de l'AMM, et une étude (RA-BEGIN) a comparé le baricitinib en monothérapie (ou en association au MTX) sur une population naïve aux traitements de fond (hors AMM). Les résultats de ces essais cliniques ont démontré une supériorité par rapport au MTX et à l'adalimumab.

L'efficacité du tofacitinib a été évaluée à l'aide de sept études cliniques, contrôlées, randomisées et en double aveugle. Plusieurs co-critères principaux de jugement d'efficacité ont été analysés (réponses ACR 20, 50 ou 70, HAQ) selon une approche hiérarchique afin

de tenir compte de la multiplicité des analyses effectuées. Quatre études (ORAL SCAN, ORAL SYNC, ORAL STEP et ORAL STANDARD) ont porté sur l'indication de l'AMM. Elles ont comparé le tofacitinib en association avec le MTX, après échec d'un csDMARD (ex. : MTX) ou d'un bDMARD (ex. : les anti-TNF).

Les études portant sur le tofacitinib démontrent une supériorité par rapport au MTX et sa non-infériorité par rapport à l'adalimumab.

Selon les avis de la CT relatifs à ces deux produits (41) et (80), le mode d'administration par voie orale du baricitinib et du tofacitinib pourrait avoir un impact favorable sur l'organisation du parcours de soins et permettre d'alléger les contraintes liées à l'injection pour les patients. Cependant, l'appréciation de leur impact sur la morbidité et la qualité de vie des patients devrait tenir compte des limites suivantes :

- l'absence de comparaison de ces produits avec les bDMARDs de 3^e ligne ;
- le faible recul quant à ces deux traitements en termes d'efficacité et de tolérance comparativement aux bDMARDs actuels ;
- et notamment les inquiétudes en termes de tolérance à long terme portant sur les risques infectieux et les risques cardio-vasculaires et carcinogènes.

➤ **L'inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 (sarilumab)**

L'efficacité et la tolérance du sarilumab ont été principalement évaluées dans trois études randomisées en double aveugle, dont deux *versus* placebo (études MOBILITY et TARGET) et une étude *versus* comparateur actif, l'adalimumab (étude MONARCH).

Ces études ont montré la supériorité du sarilumab :

- en monothérapie par rapport à l'adalimumab uniquement chez les patients ayant une contre-indication au MTX ;
- en association au MTX par rapport au placebo+MTX chez des patients en échec du MTX ;
- en association à un csDMARD par rapport au placebo+csDMARD chez des patients en échec d'au moins un anti-TNF.

Selon l'avis de la CT, le sarilumab enrichit l'arsenal thérapeutique des traitements de fond biologiques dans la prise en charge de la PR. Cependant, l'examen des données cliniques relatives à ce produit a permis de souligner les limites suivantes :

- l'absence de comparaison directe au tocilizumab (produit appartenant à la même classe) alors qu'elle était faisable ;
- l'absence de comparaison directe aux autres alternatives disponibles en 2^e ligne ou plus, alors qu'elles étaient pour certaines faisables.

6.2. Synthèse des méta-analyses récentes

6.2.1. Revue systématique et méta-analyses de l'Agence canadienne des technologies de santé (CADTH)

Une revue systématique a inclus 91 études dont la moitié présentait un risque de biais jugé acceptable. La population d'analyse visait les patients atteints de PR modérée à sévère et présentant une réponse inadéquate ou une tolérance insuffisante au MTX. Des méta-analyses en réseau ont été réalisées. Le critère principal d'analyse était la réponse ACR50 et, selon la disponibilité, les critères secondaires étaient le HAQ, le DAS 28, les arrêts de traitements et les critères de qualité de vie. Une méthode bayésienne a été retenue pour la réalisation des méta-analyses en réseau. Elle a permis d'estimer les odds ratio avec leurs intervalles de crédibilité à 95 %. Les méta-analyses en réseau distinguaient deux types de comparateurs communs principaux : le MTX et les traitements de fond synthétiques (csDMARDs).

Les principales conclusions des méta-analyses en réseau étaient :

- dans la population de patients inclus dans la revue (c'est-à-dire ceux dont la réponse au MTX est inadéquate), les associations des traitements de fond (bDMARDs ou tsDMARDs) avec le MTX étaient généralement plus efficaces que le MTX en monothérapie ;
- les traitements de fond biologiques ou synthétiques ciblés ne présentaient pas de différences importantes en termes de bénéfice clinique entre eux. Ces traitements n'avaient pas été évalués sur l'ensemble des critères d'intérêt et il n'y avait généralement pas de différence statistiquement significative dans les comparaisons directes entre ces traitements.

6.2.2. Revue systématique et méta-analyses en réseau de l'*Institute for Clinical and Economic Review (ICER)*

Une revue systématique a inclus 132 rapports de 67 essais cliniques et 17 études d'observation portant sur des patients atteints de PR modérée à sévère présentant une réponse inadéquate ou une tolérance insuffisante à un csDMARD (ex. : MTX) ou partiellement à un traitement de fond biologique (bDMARD). Les 67 essais cliniques ont fourni des données sur plus de 28 000 patients. Parmi ces essais, 60 portaient sur des associations de MTX avec des bDMARDs ou quelques tsDMARDs. Dix-neuf essais cliniques étaient des comparaisons directes dont huit étaient des comparaisons d'un bDMARD ou d'un tsDMARD par rapport à un autre, et onze comparaient des biosimilaires à leurs molécules mères.

Les données ont été analysées par des méta-analyses en réseau fondées sur une approche bayésienne estimant les odds ratios avec leurs intervalles de crédibilité à 95 %. Le critère d'analyse principal était la réponse ACR (réponse symptomatique de 20 %, 50 % et 70 %) et le critère secondaire était la progression structurale (évaluée par le score de Sharp) sur un nombre d'études limité.

Les principales conclusions des méta-analyses en réseau étaient :

- concernant l'ACR20, tous les traitements de fond (biologiques ou synthétiques ciblés) en association avec le MTX (ou en monothérapie) étaient plus efficaces que le MTX. Ce résultat était également observé pour les patients naïfs de traitements de fond synthétiques ou ayant été partiellement traités, notamment par des traitements de fond biologiques ;
- les comparaisons entre les bDMARDs et les tsDMARDs n'étaient pour la plupart pas statistiquement significatives (les intervalles de crédibilité contenaient la valeur de l'odds ratio 1,0).

6.2.3. Revues systématiques et méta-analyses en réseau de la Cochrane

Les revues systématiques et les méta-analyses de la Cochrane ont porté sur des populations similaires à celles des revues systématiques décrites dans les paragraphes 6.2.1 et 6.2.2. Les travaux de Singh *et al.* (2016) et Singh *et al.* (2017) ([89](#), [92](#)) incluant le tofacitinib en plus des bDMARDs aboutissaient à des résultats sur l'efficacité et la tolérance assez comparables à ceux des méta-analyses en réseau décrites dans les paragraphes 6.2.1 et 6.2.2. En l'occurrence, les résultats qui paraissaient globalement assez robustes étaient ceux qui démontraient un bénéfice clinique (défini en termes d'amélioration de la réponse ACR50) au profit des bDMARDs et du tofacitinib, comparativement au MTX (ou un autre csDMARD) en monothérapie.

6.3. Synthèse des critiques des évaluations du NICE

➤ Les tsDMARDs (baricitinib et tofacitinib)

À notre connaissance, une actualisation de l'évaluation complète de Stevenson *et al.* (2016) ([12](#)) sur les bDMARDs n'a pas été réalisée à ce jour. Cependant, des synthèses de la Commission indépendante de l'évaluation des avis du NICE sur les soumissions des industriels commercialisant les tsDMARDs (tofacitinib, baricitinib) ont récemment été publiées. Sans remettre en question les principales conclusions des méta-analyses en réseau de la CADTH et de l'ICER décrites dans la section 6.2, Uttley *et al.* (2018) ([126](#))

et Ren *et al.* (2018) (127) ont rapporté les principales critiques méthodologiques de l'ERG sur les méta-analyses en réseau relatives au tofacitinib et au baricitinib.

➤ **L'inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 (sarilumab)**

Une synthèse de la Commission indépendante de l'évaluation des avis du NICE de la soumission de l'industriel commercialisant le sarilumab a été récemment publiée. Sans remettre en question les principales conclusions des méta-analyses en réseau de la CADTH et de l'ICER décrites dans le paragraphe 6.3, Bermejo *et al.* (2018) a rapporté les principales critiques méthodologiques de l'ERG sur les méta-analyses en réseau portant sur le sarilumab.

- **L'association du MTX avec les tsDMARDs et l'association du MTX avec le sarilumab étaient plus efficaces que les csDMARDs en monothérapie (ex. : MTX).**
- **Aujourd'hui, il y a peu de preuves pour différencier en termes d'efficacité et de tolérance un traitement de fond biologique (ou biosimilaire) ou synthétique ciblé (associé au MTX).**

7. Conclusion

Les deux dernières décennies ont été marquées par une évolution de la prise en charge de la PR résultant de l'apparition de nouveaux traitements de fond (les traitements de fond biologiques et les traitements de fond ciblés) et de nouvelles stratégies thérapeutiques (ex. : les stratégies de traitement ciblé « *treat to target* »). Ces traitements de fond concernent notamment les traitements de fond synthétiques (csDMARDs) ou une combinaison des csDMARDs dite « csDMARDs intensifs » (deux ou trois csDMARDs), les traitements de fond biologiques (bDMARDs), en monothérapie ou en association avec MTX, et leurs biosimilaires, et récemment les tsDMARDs (les traitements de fond synthétiques ciblés).

Après avoir réalisé une revue systématique portant sur ces traitements, nos méta-analyses ont noté que chez les patients ayant une PR modérée à sévère et ayant été traités par des csDMARDs, les biosimilaires étaient aussi efficaces que leur traitement biologique de référence. Les associations de MTX avec les bDMARDs et les bsDMARDs ont été plus efficaces que les csDMARDs (en monothérapie). **Comparativement au MTX (comparateur commun le plus fréquent)** dans les méta-analyses en réseau réalisées dans ce projet, les associations MTX+ biosimilaire de l'ETN, MTX+ ETN, et MTX+TCZ présentaient les réponses ACR et EULAR les plus élevées. Cependant, les résultats des méta-analyses en réseau ne détectaient généralement pas de différences significatives entre les biothérapies (bDMARDs).

La synthèse des évaluations sur les nouveaux produits analysés par la HAS en 2017 et en 2018, l'agence canadienne en 2018, l'institut américain sur l'efficacité comparative (ICER) en 2017 a complété les résultats des méta-analyses réalisées dans ce projet par les constats suivants :

- les associations du MTX avec les tsDMARDs et le nouvel inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 (sarilumab) sont plus efficaces que les cDMARDs en monothérapie (ex. : MTX) ;
- il est généralement difficile de différencier en termes de bénéfice clinique un traitement de fond biologique par rapport à un autre, une classe thérapeutique de traitements biologiques par rapport à une autre, ainsi que des bDMARDs par rapport à des tsDMARDs.

En résumé, l'évaluation de l'efficacité comparative de cet arsenal thérapeutique de traitements de fond de la PR est non seulement une étape préalable et indispensable de l'évaluation économique des traitements de fond dans la prise en charge de la PR en France, mais fournit aussi des résultats permettant d'éclairer les choix des rhumatologues, des patients et autres praticiens afin de faciliter une prise en charge efficace et optimisant un contrôle strict de l'activité de la maladie, et de consolider *in fine* un processus de décision partagée entre le rhumatologue et le patient, la prise en charge globale (médicamenteuse et non médicamenteuse) de la PR, la nécessité de l'introduction la plus précoce possible d'un traitement de fond et la pertinence d'une évaluation régulière de l'activité de la maladie par des scores composites permettant de vérifier l'atteinte de la rémission.

Selon les recommandations de l'EULAR (Smollen *et al.*, 2017) (6) et de la SFR (Daien *et al.*, 2019) (7), le but du traitement devrait être d'atteindre la rémission ou une faible activité de la maladie, ainsi qu'une absence de progression radiologique. En même temps, l'EULAR souligne que les coûts à prendre en compte dans le choix de l'initiation d'un traitement incluent des coûts individuels, sociaux et médicaux (Smollen *et al.*, 2017) (6). Dans cette double perspective, **il est essentiel de comparer différents choix thérapeutiques en tenant compte de leur efficacité clinique et de leur rapport coût-efficacité/utilité incrémental. Ce dernier aspect sera développé dans le chapitre 5.**

Chapitre 4. Revue systématique de la littérature économique portant sur les traitements de fond biologiques

Depuis les années 2000, une littérature assez abondante a porté sur des évaluations économiques des traitements de fond (DMARDs) dans la prise en charge de la polyarthrite rhumatoïde (PR). Un certain nombre de ces évaluations économiques a été réalisé par des agences d'évaluation des technologies de santé. À titre d'exemple, le *National Institute for Health and Care Excellence* (NICE) a réalisé plusieurs évaluations dont la plus récente a été publiée en avril 2016 (Stevenson *et al.*, 2016) (128) et concernait les traitements de fond biologiques (bDMARDs) suivants : l'adalimumab, l'étanercept, l'infliximab, le certolizumab pegol, le golimumab, le tocilizumab et l'abatacept. Aux États-Unis, l'*Institute for Clinical and Economic Review* (ICER) (13), organisme à but non lucratif et indépendant américain, a publié une étude économique sur les traitements de fond biologiques et les traitements de fond ciblés (i.e. tofacitinib, baricitinib) dans la prise en charge de la PR, en avril 2017.

En France, des travaux ont porté sur les aspects économiques associés à la prise en charge de la PR. En particulier, certaines études ont abordé la part des coûts des bDMARDs dans le coût total de la prise en charge de la PR par patient : Maravic *et al.*, 2010 (14, 15) ; Kobelt *et al.*, 2008 (16) ; Fautrel *et al.*, 2006 (129) ; Fautrel *et al.*, 2012 (17) ; Fautrel *et al.*, 2016 (8). Deux études ont analysé le ratio coût-efficacité des bDMARDs (adalimumab, infliximab et rituximab) après une réponse inadéquate à au moins un anti-TNF α (Sany *et al.*, 2009 ; Sarraux *et al.*, 2010) (18, 19) ; et une étude de type coût-utilité, fondée sur un essai clinique français, STRASS « *Spacing of TNF-blocker Injections in Rheumatoid Arthritis Study* » (Vannier *et al.*, 2017) (20), a comparé la stratégie d'espacement de doses des anti-TNF α administrés par voie sous-cutanée à la stratégie d'administration standard.

Ce chapitre présente une revue systématique de la littérature économique portant sur les analyses coût-efficacité/utilité des séquences de traitements de fond biologiques (bDMARDs) dans la prise en charge de la PR. **Les objectifs étaient de :**

- **synthétiser la littérature économique existante portant sur les patients atteints de PR et éligibles à un traitement de fond biologique en première ligne de traitement biologique (ou en 2^e ligne de traitement biologique et plus) ;**
- **identifier les aspects de l'évaluation économique communs à la plupart des études et ceux suscitant des interrogations méthodologiques. Ce travail devait permettre d'alimenter la discussion sur la transposabilité des rationnels de certains choix structurants au contexte français de la prise en charge de la PR.**

Le choix des séquences des bDMARDs est conforme à l'objectif de la saisine. Il répond à un besoin de modélisation économique, rarement exploré dans la littérature économique. Celle-ci est en effet fréquemment centrée sur la comparaison des médicaments prescrits en première ligne de traitement, comme l'illustrent plusieurs études économiques de la PR depuis les années 2000. **La prise en compte de cette problématique est aujourd'hui très utile dans l'analyse du parcours médicamenteux dans les maladies chroniques ; elle pourrait être adaptée à d'autres projets d'évaluation économique de la HAS portant sur des aires thérapeutiques comparables à la PR et bénéficiant notamment d'un recul sur les données d'efficacité et de tolérance en vie réelle.**

La revue systématique de la littérature a porté sur la période de janvier 2000 à mai 2018. Elle a été limitée aux études publiées en anglais ou en français, mais le périmètre de l'évaluation économique a couvert 16 pays répartis sur trois continents (Amérique du Nord, Europe et Asie). Cette revue systématique a porté sur :

- l'indication principale à l'entrée du modèle PR France, à savoir les patients atteints de PR modérée à sévère, présentant une réponse inadéquate ou une tolérance insuffisante associée à l'utilisation des csDMARDs (i.e. le MTX). Cette indication n'est pas toujours formulée de la même façon dans la littérature, compte tenu de l'évolution de la prise en charge de la PR et de l'éventuelle différence de formulation des indications entre les agences réglementaires. À cet égard, l'indication « *patients présentant une PR active et présentant une réponse inadéquate ou une tolérance insuffisante à un csDMARD (i.e. le MTX)* », mentionnée dans plusieurs publications sélectionnées, a également été retenue. Ainsi, pour assurer la cohérence avec l'indication de l'entrée du modèle PR France, tous les patients candidats à un bDMARD en première ligne de traitement biologique et ayant présenté une réponse inadéquate ou une tolérance insuffisante à des csDMARDs ont été retenus. En revanche, conformément aux champs retenus dans ce projet (cf. chapitre 1, section Champs exclus de l'évaluation), l'indication relative aux patients atteints d'une PR sévère n'ayant jamais été précédemment traités par des csDMARDs (les patients « naïfs » de MTX) n'a pas fait l'objet d'analyse dans ce rapport ;
- au-delà de l'indication principale et puisque le modèle PR France simule le parcours médicamenteux en tenant compte de plusieurs lignes biologiques (trois lignes biologiques, suivies par un quatrième traitement biologique identique dans toutes les séquences retenues (cf. chapitre 5)), les évaluations économiques portant sur les patients atteints d'une PR active et présentant une réponse inadéquate à un anti-TNF α ont été retenues.

La première section présente les méthodes pour l'élaboration de la revue systématique de la littérature économique. La deuxième section présente les résultats de la recherche documentaire, décrit les résultats de la revue systématique de littérature, ainsi que son analyse critique. Le processus et les résultats de la revue systématique de la littérature économique seront discutés dans la troisième section.

1. Méthodes pour l'élaboration de la revue systématique de la littérature économique

1.1. Les questions spécifiques de la revue systématique de la littérature économique

La revue systématique de la littérature économique réalisée visait à répondre à trois catégories de questions :

- 1. des questions principales spécifiques à la modélisation économique** (i.e. études coût-efficacité/utilité) **des séquences des bDMARDs** : par exemple, quelles sont les structures de modèles identifiées dans la littérature ? Quels sont les critères de résultats de santé retenus ? Comment les résultats de santé tels que le QALY ont-ils été estimés ? Les réponses à ces dernières questions permettent-elles de savoir si **le rationnel de la définition des séquences de traitements** (ex. : changement ou *switch* de traitements), **la progression de l'activité de la maladie et la prise en compte des événements indésirables ont été suffisamment analysés dans la littérature ?** ;
- 2. des questions complémentaires** et spécifiques aux critères de résultats économiques complétant la modélisation économique : **quels sont les aspects méthodologiques liés aux résultats de santé** (ex. : années de vie ajustées sur la qualité de vie, QALY) **portant sur les bDMARDs** ? Cette question privilégie notamment les éléments méthodologiques concernant les études françaises portant sur les coûts et les questionnaires de qualité de vie (ex. : EQ-5D et SF-6D) ;
- 3. des questions complémentaires** portant sur les aspects contextuels épidémiologiques (ex. : les études portant sur la mortalité liée à la PR) et cliniques (ex. : les recommandations de prise en charge de la PR formulées par les sociétés savantes telles que la SFR, l'EULAR et l'ACR). Les éléments épidémiologiques relatifs à la mortalité sont indispensables pour justifier éventuellement l'ajustement de la mortalité des patients atteints d'une PR (extraite de la mortalité de la population générale française) par un taux de surmortalité associée à la PR.

La stratégie de recherche documentaire est présentée dans l'annexe A4.1.

1.2. Identification des sources de données

La recherche bibliographique a porté sur la période de janvier 2000 à octobre 2018. Son utilisation dans la description des résultats de la littérature économique a été fondée sur une

approche méthodologique en partie similaire à celle proposée par Mauskopf *et al.* (2014) (130) dans la revue systématique des traitements de fond biologiques dans la prise en charge du psoriasis.

Des recherches systématiques sur des bases de données électroniques ont été réalisées pour identifier les évaluations économiques portant sur les traitements de fond synthétiques et biologiques des patients atteints de PR. Afin d'optimiser la revue systématique (i.e. cibler le maximum de publications portant sur la problématique d'analyse), des mots clés (cf. tableau 1) ont été choisis selon des termes économiques standards couramment utilisés dans l'évaluation économique de la PR, au même titre que l'évaluation clinique des traitements de fond synthétiques ou biologiques.

Dans les bases de données interrogées (par exemple, les bases de données internationales (Medline, Central) et nationales (Pascal et Lissa)), une recherche générique sur les termes de la pathologie et sur le type d'études a été effectuée (ex. : utilisation des termes « polyarthrite rhumatoïde », « coûts », « qualité de vie », « années de vie ajustées sur la qualité de vie », « efficience », « anti-TNF », « recommandations »). Les bases de données électroniques interrogées sont présentées dans le tableau 2 ; la stratégie de recherche et la liste des sites interrogés, dans l'annexe A4.1.

Tableau 1. Mots clés figurant dans la recherche documentaire

Items	Mots clés
Pathologie et interventions	Arthritis rheumatoid, polyarthrite rhumatoïde, Tumor necrosis-alpha/antagonists and inhibitors, infliximab, golimumab, abatacept, adalimumab, certolizumab pegol, rituximab, tocilizumab, abatacept, rituximab, anti-TNF α , tocilizumab, etanercept, infliximab, baricitinib, tofacitinib, sarilumab.
Aspects économiques	Analyse des coûts, cost-effectiveness, health economics, efficiency, quality of life, adjusted quality of life, QALY, HAQ cost of illness, cost sharing, cost allocation, cost analysis, Health care cost, health expenditure, health care sector, budget, financial support, health insurance, burden disease, ESPOIR.

Tableau 2. Extrait simplifié des bases de données électroniques interrogées dans la revue systématique de la littérature économique

Bases de données interrogées	Période de recherche
Medline	Janvier 2000 – Décembre 2018
Central	Janvier 2010 – Décembre 2018
Cochrane library	Janvier 2010 – Décembre 2018
Pascal	Janvier 2010 – Décembre 2018
Lissa	Janvier 2010 – Décembre 2018

En plus de la recherche menée sur les bases de données électroniques, des publications ont été identifiées *via* :

- les sites Internet publiant des rapports d'évaluation technologique/économique et des recommandations des agences (ou organismes) d'évaluation des technologies de santé ;
- les sites Internet des sociétés savantes spécialisées en rhumatologie (ex. : la Société française de rhumatologie (SFR), la Société britannique de rhumatologie (BSR), le Collège américain de rhumatologie (ACR)). Le site de la SFR a notamment permis de réaliser un contrôle qualité sur toutes les publications scientifiques portant sur la cohorte ESPOIR ;
- les publications identifiées dans le rapport de Stevenson *et al.*, 2016 ([12](#)) portant sur l'évaluation du NICE réalisée en 2014, ainsi que celles proposées lors des différents échanges avec les experts du groupe de travail PR et l'expert en modélisation collaborant à notre travail (Pr Jaime Caro, *London School of Economics*).

1.3. Critères de sélection des études éligibles à la revue systématique de la littérature

1.3.1. Critères d'inclusion

Populations éligibles

Les populations éligibles étaient :

- *première ligne biologique (population principale considérée à l'entrée du modèle PR France)* : les patients adultes atteints d'une PR active modérée à sévère ayant présenté une réponse inadéquate ou une tolérance insuffisante à un csDMARD (ex. : le MTX) ;
- *lignes au-delà de la première ligne biologique* : les patients atteints d'une PR active et présentant une réponse inadéquate à un anti-TNF α .

Les populations éligibles ont été étendues aux traitements de fond synthétiques ciblés qui sont également prescrits pour les patients appartenant à ces deux populations.

Les interventions

Il s'agit des csDMARDs et des bDMARDs prescrits conformément aux indications délivrées par les autorités réglementaires telles que l'Agence européenne du médicament (EMA), les agences fédérales telles que l'Agence américaine du médicament (*Food and Drug Administration* (FDA)), l'agence canadienne (Agence canadienne des médicaments et des technologies de la santé (CADTH)) et les agences nationales des médicaments (ex. : les agences japonaise, chinoise et iranienne).

Les bDMARDs retenus étaient :

- l'adalimumab (ADA) ;
- l'étanercept (ETA) ;
- l'infliximab (INF) ;
- le golimumab (GOL) ;
- le certolizumab pegol (CERT) ;
- l'abatacept : formes IV et SC (ABA) ;
- le tocilizumab : formes IV et SC (TOC) ;
- le sarilumab (appartenant à la même classe thérapeutique du tocilizumab, à savoir les inhibiteurs de la transmission du signal médié par les récepteurs de l'IL-6).

Les csDMARDs retenus étaient les suivants :

- le méthotrexate (MTX) ;
- la sulfasalazine (SULF) ;
- le léflunomide (LEF) ;

- l'hydroxychloroquine (HYD).
- Les traitements de fond synthétiques ciblés : baricitinib et tofacitinib.

Les interventions retenues pouvaient également inclure des combinaisons de csDMARDs prescrites selon les recommandations cliniques en vigueur des pays relatifs aux publications analysées : deux csDMARDs (ex. : MTX + sulfasalazine) et trois csDMARDs (ex. : MTX + sulfasalazine + hydroxychloroquine).

Les comparateurs (stratégies comparées)

Il s'agissait des bDMARDs ou des séquences de traitements définies en fonction des bDMARDs prescrits en monothérapie ou en association avec un csDMARD (MTX ou sulfasalazine ou hydroxychloroquine), conformément aux indications délivrées par l'Agence européenne du médicament (EMA). **L'attention a été portée à la comparaison des séquences de bDMARDs ou de tsDMARDs.**

Le type d'analyse économique

Toutes les études incluses concernaient des analyses coût-efficacité et/ou coût-utilité. Les critères de résultats économiques permettant de calculer le ratio différentiel coût-résultat (RDCR) étaient :

- les années de vie, les années de vie ajustées sur la qualité de vie (QALY) ou le temps de faible activité de la PR (ou temps en rémission). **Les années de vie gagnées ne représentent pas le résultat de santé le plus important, car le lien de causalité entre la survie et la prise des bDMARDs n'est pas démontré.** Aujourd'hui, l'amélioration clinique de l'état de santé des patients atteints d'une PR est évaluée selon des critères d'évolution de la maladie (DAS 28), d'incapacité fonctionnelle (HAQ), de réponses (ACR ou EULAR) ou des critères décrivant l'érosion osseuse ;
- les coûts exprimés en termes de coûts directs médicaux et non médicaux. Les études prenant en compte les coûts indirects n'ont pas été exclues de la revue systématique.

1.3.2. Les critères d'exclusion

Les études suivantes ont été exclues :

- les études coût-efficacité/utilité portant sur les patients atteints de PR sévère, n'ayant jamais été traités par des csDMARDs tels que le MTX (les patients « naïfs » de csDMARDs). Comme précisé dans le premier chapitre (section 1.4, Champs d'analyse exclus dans le rapport), l'évaluation économique de cette population n'a pas été retenue dans le projet ;
- les études coût-efficacité/utilité portant seulement sur les csDMARDs ;

- les études portant sur l'utilisation des bDMARDs dans d'autres aires thérapeutiques (ex. : psoriasis, maladie de Crohn) ;
- les études portant seulement sur des coûts relatifs à des pays étrangers ;
- les études de minimisation des coûts, les études d'impact budgétaire ou les études des contrats de performance et/ou partage de risques portant sur les bDMARDs (conformément à l'objectif de la saisine).

1.3.3. Critères supplémentaires

Afin d'enrichir la description et la discussion des questions et des problèmes de modélisation économique, des études supplémentaires ont été retenues :

- des études sur des questions méthodologiques portant sur les utilités. Ces questions permettant, par exemple, de discuter du calcul des années de vie ajustées sur la qualité de vie (QALY) ;
- des études sur l'éventuel risque de surmortalité associée à la PR pouvant être exploré dans des analyses de sensibilité ;
- les études de coûts françaises apportant des informations complémentaires aux études françaises de type coût-efficacité/utilité ;
- les populations éligibles ont été élargies en intégrant les publications disponibles sur les nouveaux traitements de fond synthétiques ciblés (tofacitinib et baricitinib) et le sarilumab disponibles au cours de la période de la recherche documentaire.

1.4. Processus de sélection, d'extraction et d'évaluation des résultats

1.4.1. Sélection, extraction et évaluation des résultats

Le processus de la revue systématique de la littérature mis en place s'articulait autour de trois étapes principales :

- la première étape de sélection des études a été fondée sur un processus similaire à celui de l'analyse systématique de la littérature clinique. L'approche *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) (Moher *et al.*, 2009) (36) a ainsi été adaptée à la revue systématique de la littérature économique ;
- la deuxième étape a concerné la synthèse de l'extraction des résultats des publications retenues. Dans un premier temps, une grille simplifiée d'évaluation de la littérature, adaptée à la PR et créée à partir de la grille CHEERS *Consolidated Health Economic Evaluation Reporting Standards* (Husereau *et al.*, 2013) (38), a été élaborée et validée par trois lecteurs indépendants. Elle a permis de réaliser des résumés de toutes les études retenues (cf. annexe 4-2). Dans un deuxième temps, à partir de ces résumés,

quatre grilles de lecture des caractéristiques des études retenues ont été élaborées. Ces quatre grilles (cf. annexe 4-2) étaient fondées sur les grilles de lecture validées lors de l'élaboration du guide de l'analyse d'impact budgétaire à la HAS et publiées dans Ghabri *et al.*, 2017 (131) :

- la première grille d'analyse résume les caractéristiques globales des études (indication, objectif, prise en compte des séquences et stratégies comparées),
 - la deuxième grille d'analyse synthétise les sources de données mobilisées dans l'extraction des entrées des modèles (efficacité, tolérance/événements indésirables, mortalité, coûts et résultats de santé),
 - la troisième grille d'analyse décrit les choix structurants : type d'étude (coût-utilité ou coût-efficacité), caractéristiques de la population à l'entrée du modèle, horizon temporel et actualisation, rationnel et méthodes utilisés dans l'estimation des coûts, rationnel et méthodes utilisés dans l'estimation des utilités ou des critères d'efficacité (utilisés dans les analyses coût-efficacité),
 - la quatrième grille d'analyse décrit les résultats des modèles (RDCR, types d'analyses de sensibilité effectuées, facteurs influençant la variabilité du RDCR, les conclusions/perspectives/limites), le financement de l'étude et les conflits d'intérêt déclarés.
- La troisième étape concerne l'analyse critique de la littérature portant sur les séquences de traitements. Elle a été réalisée en deux étapes :
- 1) l'analyse critique de la qualité de la description des études fondée sur la grille CHEERS adaptée à la PR (cf. annexe 4-3),
 - 2) une analyse critique en combinant la grille CHEERS avec les *items* de la grille de Drummond (Drummond *et al.*, 2015) (39), la grille de l'*Outcome Measures in Rheumatology* (OMERACT) (Gabriel *et al.*, 2002) (40) et notamment les différentes questions méthodologiques soulevées lors des réunions du groupe de travail PR (juin 2017 et juillet 2018) et des présentations lors des congrès internationaux tels que le 19^e congrès annuel de l'ISPOR (*International Society for Pharmacoeconomics and Outcomes Research*) (Vienne, 2017) et le 14^e congrès annuel de la HTAi (*Health Technology Assessment International*) (Rome, 2017). Des exemples de ces grilles sont donnés en annexe 4-4..

1.4.2. Contrôle qualité

Les trois étapes décrites dans la section 2.4.1 ont fait l'objet d'une double vérification indépendante. Un troisième évaluateur intervenait dans le cas de discordances dans les

extractions et l'analyse critique de la littérature réalisée par les deux évaluateurs. Les deux grilles relatives aux étapes 2 et 3 ont été réalisées par le coordinateur du projet PR, testées par trois analystes (un pharmacien, un médecin de santé publique et un économiste de la santé) et approuvées par le groupe de travail PR. Aucune remarque sur des problèmes de reproductibilité des grilles utilisées n' a été faite par le groupe de lecture PR et la CEESP.

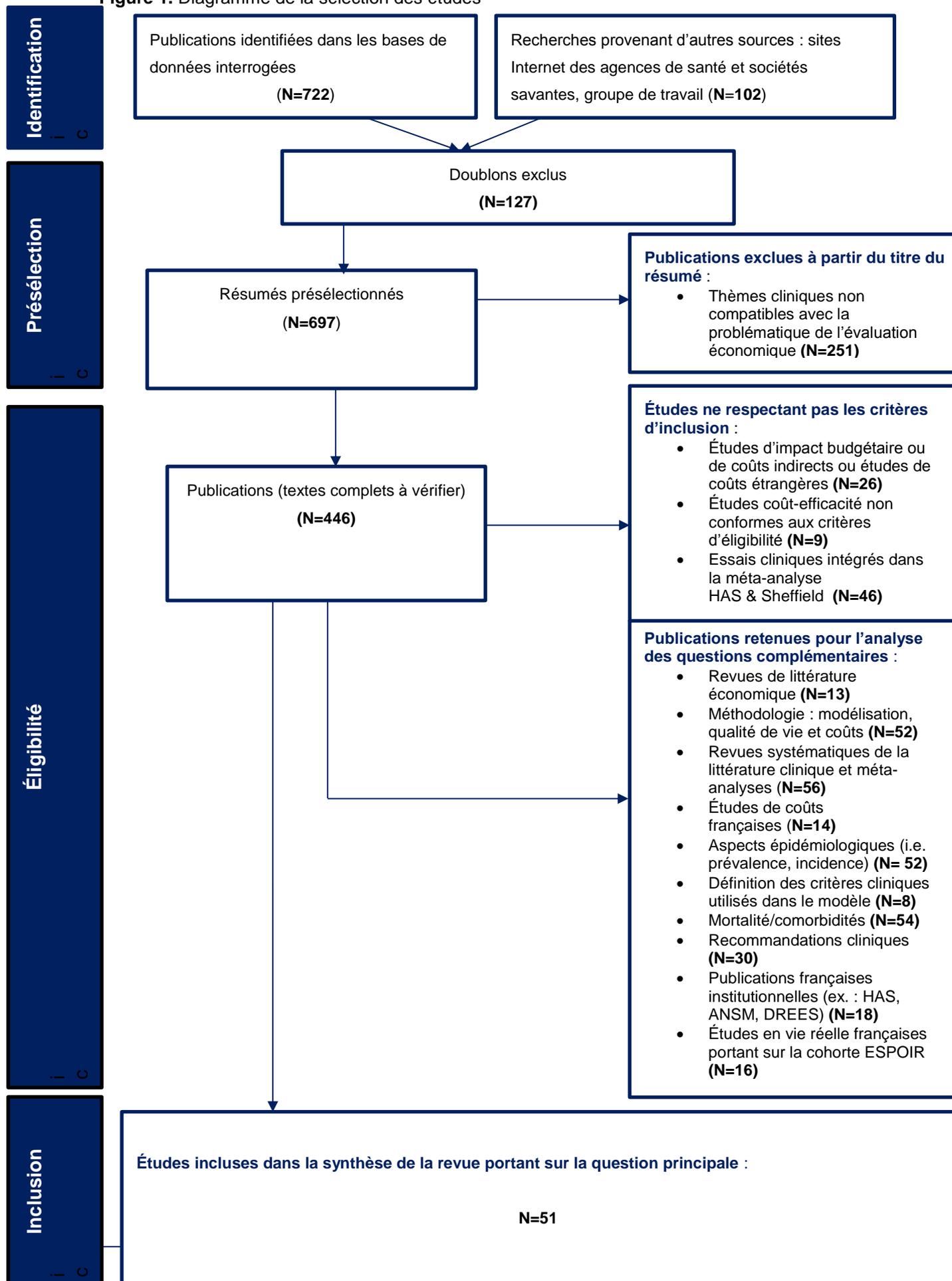
2. Résultats de la revue systématique de la littérature économique

2.1. Sélection des références

Le processus de sélection des études est présenté dans la figure 1. La base de données électronique a identifié 824 références. Après suppression des doublons, il restait 697 références présélectionnées à partir de leur résumé. Parmi ces dernières, 251 résumés n'ont pas rempli les critères d'éligibilité et 446 ont fait l'objet d'un examen complet. Quarante-neuf publications, ne respectant pas les critères d'inclusion, ont été exclues. Trois cent treize publications ont été retenues pour analyser les questions complémentaires de la revue systématique de la littérature économique (ex. : 54 études ont été retenues pour analyser la mortalité associée à la PR).

Au final, 51 études ont été retenues dans la revue systématique de la littérature économique.

Figure 1. Diagramme de la sélection des études



2.2. Synthèse des résultats de la revue systématique de la littérature économique

Les annexes 4-5 à 4-8 décrivent les caractéristiques des études coût-utilité/efficacité retenues selon les grilles 1, 2, 3 et 4. Au total, 51 études ont été sélectionnées.

2.2.1. Caractéristiques des études

➤ Répartition des études par indication et pays

Les études sélectionnées se répartissaient de la façon suivante :

- 40 études ont porté sur les patients traités par un bDMARD en première ligne de traitement biologique et présentant une réponse inadéquate ou une tolérance insuffisante à des csDMARDs. Parmi ces études, 22 ont été décrites en termes de « *traitement des patients adultes atteints de PR modérément à sévèrement active lorsque la réponse aux traitements de fond (ou leur tolérance), y compris le MTX, est inadéquate (ou insuffisante)* » ;
- 11 études ont porté sur le traitement des patients adultes atteints de PR active, sévère, présentant une réponse inadéquate ou une intolérance aux traitements de fond, dont au moins un anti-TNF α .

Les études sélectionnées ont été réalisées dans 15 pays¹¹ : 14 (soit 27 % des études) scandinaves (huit suédoises, cinq finlandaises, et une norvégienne) ; 11 britanniques (21 % des études) et dix américaines (21 % des études). Une seule étude française, portant sur des patients atteints de PR active modérée à sévère et ayant une réponse insuffisante à au moins un anti-TNF α , a été identifiée (Saraux *et al.*, 2010).

➤ Financement des études et conflits d'intérêts

La majorité des études (57 % ; 29 /51) a été au moins partiellement financée par l'industrie pharmaceutique commercialisant les bDMARDs analysés. Parmi les 22 % (11/51) d'études coordonnées par des organismes publics, cinq étaient des évaluations de technologies de santé (HTA réalisées au Royaume-Uni dans le cadre des programmes de recherche initiés sous l'égide du NICE et publiées en 2002, 2004, 2006, 2011 et 2014). Par ailleurs, l'ICER ([13](#)) a réalisé l'évaluation des traitements de fond biologiques aux États-Unis en 2017.

Au moins 59 % (30/51) des études présentaient un conflit d'intérêts ; 18 % (9/51) des études ne mentionnaient ni l'existence de conflit d'intérêt, ni le financement de l'étude en question.

¹¹ Ces pays sont : France, Espagne, Italie, Pays-Bas, Grèce, Suède, Norvège, Royaume-Uni, Allemagne, États-Unis, Chine, Iran, Corée du Sud et Japon.

2.2.2. Données mobilisées pour l'extraction des paramètres de la modélisation économique

➤ L'efficacité, la tolérance et la mortalité

- Trente-et-un pour cent (16/51) des études publiées à partir de 2010 ont mentionné des méta-analyses (comparaisons indirectes ou des méta-analyses en réseau) portant majoritairement sur les probabilités de réponses ACR (i.e. ACR20, ACR50 et ACR70). Seules les méta-analyses en réseau récentes réalisées par le NICE (Stevenson *et al.*, 2016) ([12](#)) ont porté sur des réponses ACR et des réponses EULAR.
- Les sources de données sur la tolérance (concernant principalement les événements indésirables) ont été mentionnées dans 23 % (12/51) des études ; 8 % (4/51) mentionnaient des méta-analyses dans l'extraction des données sur les événements indésirables. Les autres sources étaient très hétérogènes (ex. : publications, données d'essais cliniques).
- Concernant la survie, les tables de mortalité spécifiques aux pays d'analyse (ou rarement les tables de mortalité issues des statistiques internationales telles que celles de l'OMS) ont été utilisées dans 74 % (38/51) des études.

➤ Les ressources et les coûts

- Les types de sources de données portant sur les coûts étaient comparables entre la majorité des études et dépendaient du mode de prise en charge de la PR dans le pays de l'étude en question. Elles ont été clairement mentionnées dans 96 % (49/51) des études. Ces sources de données étaient des bases de données médico-administratives, des listes nationales de tarifs des ressources ou des prix des médicaments, des études observationnelles ou même des essais cliniques, notamment dans le cas des évaluations économiques où l'horizon temporel était inférieur ou égal à 2 ans.
- Les données macroéconomiques, telles que les statistiques de l'emploi ou portant sur le revenu national, ont été utilisées afin d'estimer les pertes de productivité, les retraites liées à une incapacité due à la PR et les arrêts de travail. Ce type de données a été exploité dans des études scandinaves et japonaises représentant 16 % (8/51) des études analysées.

➤ Les résultats de santé retenus dans les évaluations économiques

- Trente-sept pour cent (19/51) des études se fondaient sur la littérature (ex. : Bansback *et al.*, 2005 ; Wailoo *et al.*, 2008 ; Hernandez *et al.*, 2013) en ce qui concerne l'extraction des équations de prédiction (ou *mapping*) des utilités en fonction du HAQ et d'autres co-variables telles que le sexe et l'échelle visuelle analogique (EVA) de la douleur.

2.2.3. Les choix structurants des évaluations économiques

➤ Le type d'analyse économique :

Quatre-vingt-huit pour cent des études (45/51) étaient des études coût-utilité (les résultats de santé étaient exprimés en années de vie ajustées sur la qualité de vie, QALY). Douze pour cent [6 /51] des études étaient des analyses de type coût-efficacité dont cinq ont porté sur la population des patients atteints de PR active (ou modérée à sévère) et présentant une réponse insuffisante à au moins un anti-TNF α . La seule étude française respectant les critères de sélection était une étude de type coût-efficacité fondée sur le coût d'une rémission ou d'une faible activité de la maladie (Saraux *et al.*, 2010).

➤ La population d'analyse dans le modèle économique

Les caractéristiques de la population d'analyse (en particulier l'âge, le sexe, le HAQ à l'entrée du modèle économique) ont été mentionnées dans 76 % (39/51) des études.

➤ La perspective

Trente-cinq pour cent (18/51) des études étaient fondées sur une perspective sociétale et 65 % (33/51), sur celle du payeur ou « principal financeur de soins ». Les études reposant sur une perspective sociétale étaient principalement scandinaves et/ou japonaises (16 %, 8/51) et prenaient en compte les coûts indirects liés à la PR (ex. : les pertes de productivité et les arrêts de travail).

➤ Comparateurs/séquences de traitement

- Soixante-et-un pour cent (31/51) des études présentaient des stratégies exprimées en termes de séquences de traitements et 39 % (20/51) comparaient des traitements.
- Le MTX en monothérapie était le comparateur le plus fréquent dans les études comparant des traitements (55 %, 11/20).
- Les séquences de csDMARDs intégrant le MTX en monothérapie ou en combinaison avec les autres csDMARDs étaient les stratégies standards les plus fréquemment retenues dans les études comparant des stratégies définies *via* des séquences de traitements (45 %, 14/31).

➤ L'horizon temporel et l'actualisation

- Cinquante-et-un pour cent (26/51) des études reposaient sur un horizon temporel vie entière, 25 % (13/51) sur un horizon inférieur ou égal à 2 ans, 10 % (5/51) retenaient un horizon de 10 ans, 8 % (4/51) un horizon de 5 ans et 4 % (2/51) un horizon de 20 ans.
- Les coûts et les QALYs ont été actualisés dans 80 % (41/51) des études. Les taux d'actualisation utilisés dans les études médico-économiques analysées variaient entre 1,5 % et 6 % et dépendaient des recommandations économiques et de leur mise à jour dans le pays de l'étude (ex. : Royaume-Uni).

➤ **L'estimation et la valorisation des coûts**

- La typologie des coûts directs liés à l'utilisation des bDMARDs (coûts d'acquisition, coûts d'administration, coûts de suivi et coûts d'hospitalisation) et leur valorisation étaient mentionnées dans près de 98 % (50/51) des publications selon les recommandations en vigueur du pays de l'étude.
- La méthode d'estimation des coûts directs relatifs aux hospitalisations (en particulier les jours d'hospitalisation) et/ou les ressources consommées durant le suivi du patient différaient selon les études : dans 37 % (19/51) des études, l'estimation de ces coûts était stratifiée sur la catégorie du score HAQ (six niveaux de gravité de la PR) et dans 10 % (5/51) sur l'activité de la maladie (catégories du DAS 28). Les études restantes ne mentionnaient pas clairement si une stratification en fonction d'autres critères avait été effectuée.
- Les coûts associés aux événements indésirables étaient retenus dans 16 % (8/51) des études. Il s'agissait principalement des infections graves (ex. : les pneumonies).

➤ **L'estimation et la valorisation des résultats de santé**

- Le résultat de santé le plus fréquent était le QALY (88 %, 45/51) des études. La rémission et la faible activité ont été retenues dans 12 % (6/51) des études.
- Le QALY a été estimé en pondérant la survie par les scores d'utilité associés aux états (ou aux équations du modèle d'analyse) :
 - la mortalité a été majoritairement (75 %, 38/51) estimée à l'aide des tables de mortalité nationales (couramment stratifiées par âge et sexe) et pondérée, dans 39 % (20/51) des études analysées, par un risque de surmortalité. Ce dernier est souvent calculé en fonction de la variation initiale du HAQ (ex. Michaud *et al.*, 2012 ([132](#))). Dans certaines études, la mortalité a été simplement pondérée par un surrisque de mortalité associée à la PR (ex. : Brenan, 2007) ;
 - dans la majorité des études (78 %, 40/51), les utilités (scores EQ-5D ou HUI3) ont été obtenues selon une prédiction (*mapping*) réalisée en fonction du HAQ et d'autres variables (âge, sexe, échelle visuelle analogique de la douleur) tout en supposant généralement l'existence d'une relation linéaire entre la variation du HAQ et l'utilité. Quatorze pour cent (7/51) des études présentaient des scores d'utilité fondés uniquement sur un instrument générique (ex. : EQ-5D).

- Les désutilités liées aux événements indésirables et pouvant être liées à l'utilisation des bDMARDs (ex. : tuberculose, pneumonie, infection urinaire, infection respiratoire) étaient seulement intégrées dans 10 % (5/51) des analyses.

2.2.4. Aspects liés à la modélisation économique

➤ Le type de modélisation

Quarante-trois pour cent (22/51) des publications analysées étaient fondées sur des modèles de micro-simulations : simulations à événements discrets « *Discrete-Event Simulation* » (DES) ou des simulations markoviennes individuelles. Trente-trois pour cent (17/51) des études reposaient sur des modèles markoviens de cohorte multi-états, 12 % (6/51) étaient des arbres de décision et 12 % (6/51) étaient issus d'un calcul statistique (ex. : régression linéaire ou calcul par la méthode de *bootstrap*). Les modèles de micro-simulation étaient de plus en plus utilisés depuis les années 2010. Les modèles retenus dans les évaluations réalisées par les organismes d'évaluation des technologies de santé (ex. : NICE et ICER) étaient des modèles de micro-simulation. Des comparaisons de ces modèles par rapport au modèle PR France sont présentées dans la section 4 du chapitre 6 et leurs principales caractéristiques sont décrites l'annexe 6.4.

➤ Les séquences

Soixante-et-un pour cent (31/51) des structures de modèles reposaient sur des séquences de traitements. La majorité d'entre elles comparaient seulement les choix pour la première ligne biologique, toujours suivis par les mêmes séquences. Deux études (ICER, 2017 ([13](#)) ; Jansen *et al.*, 2017 ([133](#))) proposaient des séquences différentes au niveau de la 1^{re} et la 2^e ligne de traitement.

➤ Hypothèses principales de modélisation

- Quarante-cinq pour cent (23/51) des études analysées (ex. : les modèles markoviens de cohortes multi-états ou individuels) se fondaient sur des états de santé ou utilisaient des probabilités de transition reposant sur les réponses aux traitements selon la réponse ACR (ACR20 ou ACR50 ou ACR70). Dix pour cent (5/51) considéraient des états de santé et/ou utilisaient des probabilités fondées sur la réponse EULAR, alors que l'indice composite d'activité de la PR (DAS 28), seul ou en combinaison avec le HAQ, a été retenu dans 31 % (16/51) des études.
- Immédiatement à l'arrêt du traitement, la dégradation du score de l'incapacité fonctionnelle HAQ a été supposée égale à l'amélioration du HAQ précédemment observée. Ce phénomène connu sous le nom d'effet « rebond » indique le retour au score HAQ avant la mise en place du traitement, et ce, jusqu'à l'initiation d'une nouvelle ligne de traitement. Cette hypothèse a été utilisée dans plusieurs études anciennes et récentes (ex. : Brennan *et al.*, 2004 ; Diamantopoulous *et al.*, 2012 et Stevenson *et al.*, 2016) ([12](#), [134](#), [135](#)).

- Un surrisque de mortalité associé à la PR a été retenu dans la majorité des études sans forcément supposer un effet direct des DMARDs (synthétiques ou biologiques) sur la mortalité. La majorité des études utilisaient un ratio égal à 1,33^{HAQ} (Wolfe *et al.*, 2003) (136). En revanche, l'évaluation de Stevenson *et al.* (2016) (12) considérait des odds ratios correspondant aux classes du HAQ telles qu'elles étaient définies dans l'étude de Michaud *et al.* (2012) (132).
- Le changement de traitement a été systématiquement justifié soit par une réponse inadéquate, soit par une tolérance insuffisante à un csDMARD (indication principale de l'entrée du modèle) ou à un bDMARD (utilisation d'un bDMARD en 2^e ligne biologique ou plus). Les préférences des patients et l'adhésion au traitement n'ont pas été directement spécifiées comme raisons de changement de traitement.
- La première génération des modèles de transformation (*mapping*) du HAQ en scores d'utilités reposait sur des régressions linéaires reliant les utilités au HAQ et aux variables socio-démographiques telles que l'âge et le sexe. À partir de 2013, des modèles non linéaires « *mixture models* » adaptés aux spécificités des distributions statistiques des utilités calculées à partir de l'EQ-5D ont été utilisés dans trois évaluations : Jansen *et al.*, 2017 (133), Athanasakis *et al.*, 2015 (137) et Stevenson *et al.*, 2016 (12).

2.2.5. Les résultats, les analyses de sensibilité, les conclusions et les limites

➤ Les résultats

- Les résultats reflètent une forte variabilité des RDCRs se manifestant au niveau des dimensions suivantes : l'indication analysée, la variabilité inter-temporelle, la variabilité intra-pays et la variabilité inter-pays.
- D'une manière générale, dans 33 % (17/51) des études, le RDCR comparant les bDMARDs (ou leurs séquences) aux csDMARDs (ou leurs séquences) était inférieur à 30 000 €/QALY. Les RDCRs étaient compris entre 30 000 et 100 000 €/QALY dans 33 % (17/51) des études, alors qu'ils étaient supérieurs à 100 000 €/QALY dans 18 % (9/51) des études.
- En termes absolus, les RDCRs variaient de 9 000 €/QALY à 1 948 919 €/QALY pour les patients traités par un bDMARD en première ligne de traitement biologique après une réponse inadéquate ou une tolérance insuffisante à au moins un csDMARD (ex. : MTX). Comparativement aux csDMARDs (ou leurs séquences), ils étaient supérieurs à 50 000 €/QALY ou à 100 000 €/QALY dans les deux tiers des études coût-utilité portant sur des patients ayant une réponse inadéquate à un anti-TNF α (soit 4/6 de ces dernières études).

- Au sein du même pays, pour les patients traités par un bDMARD en première ligne de traitement biologique après une réponse inadéquate ou une tolérance insuffisante à au moins un csDMARD, le RDCR comparant les bDMARDs (ou leurs séquences) aux csDMARDs (ou leurs séquences) variait de 2 000 à 1 948 919 €/QALY en Suède, de 16 000 à plus de 40 000 £/QALY au Royaume-Uni, de 7 925 \$ à plus de 150 000 \$ US/QALY aux États-Unis.
- Le calcul du coût de la rémission ou de l'activité faible a été réalisé dans 12 % (6/51) des études (y compris l'étude française de Saraux *et al.*, 2010). Dans toutes les études de type coût-efficacité portant sur des patients ayant une réponse inadéquate à un anti-TNF α , les auteurs suggéraient que la stratégie intégrant l'abatacept était « dominante » (en 2^e ou 3^e ligne), quelle que soit la séquence comparée (ETA->RTX->ADA ou ETA->ADA->INF).

➤ **Analyses de sensibilité et exploration de l'incertitude**

- Les analyses de sensibilité déterministes (notamment des analyses de scénarios) ont été réalisées dans 71 % (36/51) des études. Les analyses de sensibilité probabilistes (i.e. des courbes d'acceptabilité) ont été intégrées dans presque la moitié des études (49 %, 25/51) ;
- Par ordre croissant, les facteurs affectant la variabilité du RDCR et la caractérisation de l'incertitude et de l'hétérogénéité étaient : les hypothèses relatives à la progression et/ou la sévérité de la maladie définies en fonction de la variation initiale et à long terme du HAQ [39 %, 20/51], les variations du prix d'un bDMARD (25 %, 13/51), les utilités (24 %, 12/51) et les probabilités des réponses au traitement telles que les réponses ACR20, ACR50 ou ACR70 (14 %, 7/51).

➤ **Les conclusions, les limites et les perspectives**

- La plupart des études concluaient à un ratio coût-efficacité/utilité comparant les bDMARDs (ou leurs séquences) aux csDMARDs (ou leurs séquences) favorable au traitement étudié (ou la séquence débutant par le traitement étudié).
- La référence à un seuil de disposition à payer a été explicitée dans les études portant sur les programmes britanniques. Au Royaume-Uni, pour les patients traités par un bDMARD en première ligne de traitement biologique après une réponse inadéquate ou une tolérance insuffisante à au moins un csDMARD (ex. : MTX) et comparativement à ceux traités par csDMARDs (i.e. MTX), les valeurs de RDCR étaient supérieures au seuil retenu dans ce pays.

- Les principales limites soulevées dans les études étaient : le manque de données portant sur des comparaisons directes entre les bDMARDs (27 %, 14/51 des études) et le manque de données de long terme portant sur l'incapacité fonctionnelle et l'activité de la PR concernant les csDMARDs ou les bDMARDs et la prise en compte de l'adhésion des patients et de la persistance des traitements.

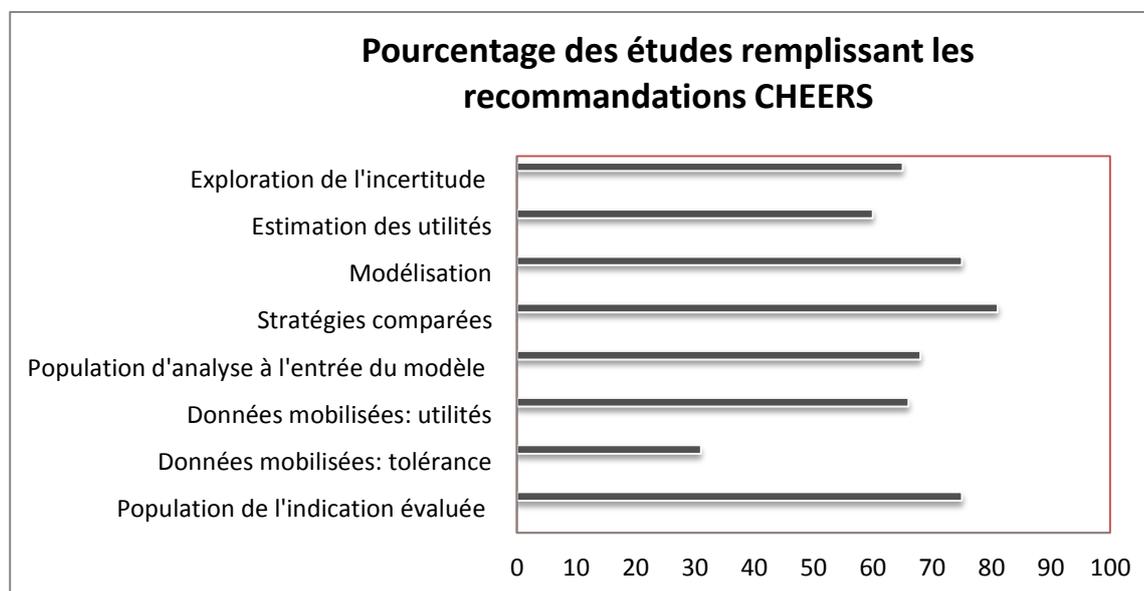
3. Analyse critique de la revue systématique de littérature économique

L'analyse critique de la qualité des études a été réalisée sur les 32 études comparant des stratégies définies en termes de séquences de traitements (cf. annexe 4-9).

3.1. Évaluation selon les recommandations CHEERS adaptées à la PR

L'analyse de la description des items selon la liste CHEERS adaptée à la PR a porté sur les 32 études. L'annexe 4-10 synthétise tous les items de la grille CHEERS adaptée à la PR. Les descriptions des sources de données portant sur l'efficacité, l'approche globale d'estimation des coûts et les résultats de la modélisation étaient satisfaisants (les scores relatifs à ces items étaient proches ou égaux à 100 %). La figure 2 inclut les scores des principaux items n'indiquant pas une qualité de description satisfaisante selon les recommandations de la grille CHEERS. Il s'agit de la population de l'indication évaluée, des données de tolérance et d'utilités utilisées dans le modèle, des stratégies ou séquences comparées, des aspects techniques spécifiques à la modélisation et de l'exploration de l'incertitude *via* les analyses de sensibilité déterministes (ex. : analyses de scénarios) et probabilistes.

Figure 2. Pourcentage des études remplissant les recommandations CHEERS



Les pourcentages des items associés aux sources de données de tolérance et des utilités mobilisées dans la modélisation ainsi que l'approche d'estimation des utilités étaient les plus faibles (respectivement 31 %, 66 % et 60 %). Quelle que soit la période étudiée de la prise en charge de la PR (années 2000 ou années 2010), ces aspects de l'évaluation économique posent problème. D'une part, le choix des instruments génériques appropriés dans l'estimation du QALY ne fait pas

l'objet d'un consensus entre les experts cliniques (Gaujoux-Viala *et al.*, 2013 (26), Beresniak *et al.*, 2013 (138)). À travers ces divergences, on retrouve indirectement les critiques classiques de l'utilisation du QALY dans l'évaluation économique des maladies chroniques (faible sensibilité des instruments pour capter la sévérité de la PR et ses différentes dimensions impactant la qualité de vie des patients). D'autre part, il existe de réelles lacunes méthodologiques dans l'utilisation des données de tolérance et dans la disponibilité des données comparatives par classe de bDMARDs (Heather *et al.*, 2014) (139).

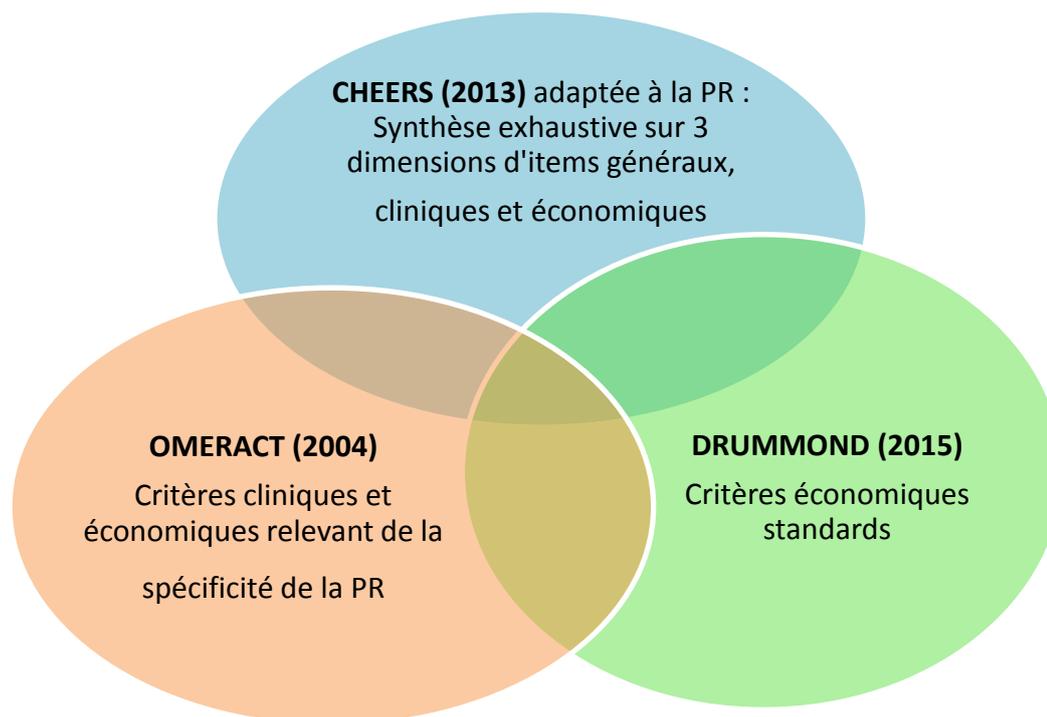
Les pourcentages relatifs aux caractéristiques de la population d'analyse de l'indication évaluée (74 %, 23/31), les stratégies comparées et les modèles économiques utilisés (cohorte multi-états ou micro-simulation) (75 %, 23/31) étaient assez élevés. La description de ces items s'est améliorée après les années 2010. Cette amélioration s'explique par la clarté des indications des bDMARDs dans la prise en charge de la PR, les exigences réglementaires de réalisation des analyses de sensibilité probabilistes, et la tendance à l'utilisation des modèles de micro-simulations (permettant de tenir compte de la variabilité stochastique, des profils complexes des patients atteints de PR et des caractéristiques du système de prise en charge dans le pays de l'étude en question).

3.2. Évaluation critique selon une liste de questions critiques provenant de plusieurs grilles

La grille CHEERS a permis d'apprécier la qualité de la description des items synthétiques des évaluations économiques, d'identifier les éléments communs dans la littérature faisant l'objet d'un usage standardisé (ex. : typologie des coûts et leurs approches de calcul) et de mettre l'accent sur des problématiques méthodologiques (ex. : les approches de mesure de la qualité de vie des patients et la collecte des données nécessaires à l'estimation des utilités) ou à la fois des contraintes de disponibilité de données et des problèmes méthodologiques (ex. : l'intégration des événements indésirables dans les modélisations à travers l'évolution de la prise en charge de la PR). Toutefois, comme toute grille de synthèse critique générale centrée sur le degré d'explicitation des items analysés, la grille CHEERS ne permet pas d'alimenter l'analyse critique de certains éléments complémentaires et spécifiques à l'évaluation économique de la prise en charge de la PR en France.

Afin d'optimiser l'identification de ces éléments, conformément aux discussions au sein du groupe de travail PR, et d'étoffer (et/ou compléter) les items de la grille CHEERS, cette dernière a été croisée avec deux autres grilles (grille de l'*Outcome Measures in Rheumatology* « OMERACT », Gabriel *et al.*, 2002 (40) ; et grille de Drummond *et al.*, 2015 (39)) utilisées dans la littérature économique sur l'évaluation des stratégies des traitements de fond biologiques dans la prise en charge de la PR.

L'annexe 4-3 présente le contenu des grilles de l'OMERACT et de Drummond. La grille CHEERS adaptée à la PR est présentée dans l'annexe 4-2. Ces trois grilles d'évaluation de la qualité des études économiques présentent des points communs. En particulier, la grille OMERACT tient compte des spécificités cliniques de la PR ; la grille de Drummond se focalise sur les questions économiques incluses pour la plupart dans la grille CHEERS. La grille CHEERS est néanmoins la plus récente et la plus exhaustive sur les aspects cliniques liés à la modélisation au même titre que les aspects strictement économiques (ex. : choix structurants, aspects techniques relatifs à la modélisation).



Les principaux éléments résultant du croisement de ces trois grilles sont résumés dans le **tableau 3**. Leur niveau d'appréciation a été réalisé selon une échelle ordinale (-- : non pris en compte ; - : non satisfaisant ; +/- : moyen ; + : assez bon ; ++ : bon) en se fondant sur des éléments d'évaluation économique discutés par le GT PR.

Neuf évaluations parmi les 31 études présentant des séquences de traitements ont été retenues. Il s'agit des études fondées sur des méta-analyses d'efficacité comparative, à horizon vie entière, présentant des détails techniques sur l'approche de l'estimation des utilités et dont les séquences de traitements étaient clairement présentées.

Tableau 3. Évaluation critique d'une courte liste d'études selon des critères émanant des trois grilles : CHEERS, OMERACT et DRUMMOND (Drummond *et al.*, 2015) (39)

Étude /année / pays	Sévérité de la maladie	Progression /activité de la maladie	Revue systématique/ méta-analyses	Rationnel séquences et des hypothèses structurelles	Rationnel estimation des résultats de santé	Rationnel et intégration des EI spécifiques aux DMARDs	Références par rapport à un guide de prise en charge	Stratification de la population en groupes (selon risque de comorbidités)
Barton <i>et al.</i> 2004 Royaume-Uni (23)	-	++	(+/-)	+	-	-	++	--
Chen <i>et al.</i> 2006 Royaume-Uni (140)	-	++	(+/-)	+	-	-	++	--
Merkesdal <i>et al.</i> 2010 Allemagne (141)	+	+	(+/-)	(+/-)	+	-	+	--
Mattloki <i>et al.</i> 2011 Royaume-Uni (142)	+	++	+	+	+	-	++	--
Diamantopoulos <i>et al.</i> 2012 Royaume-Uni (135)	+	+	(+/-)	+	+	+	+	--
Stevenson <i>et al.</i> 2016 Royaume-Uni (12)	++	++	++	++	++	+	++	--
ICER 2017 États-Unis (13)	+	++	++	++	+	+	+	--
Jansen <i>et al.</i> 2017 États-Unis (133)	+	+	(+/-)	+	+	+	+	--
Claxton <i>et al.</i> 2018 États-Unis (143)	+	+	(+/-)	+	+	+	+	--

3.3. Les aspects présentant une amélioration par rapport aux évaluations économiques identifiées dans la littérature

- Les composantes du rationnel des séquences de traitements étaient clairement argumentées dans les évaluations britanniques (Barton *et al.*, 2004 ; Malotki *et al.*, 2011 ; Stevenson *et al.*, 2016). Ces dernières évaluations ont notamment proposé : 1) des séquences de traitements explicites dans des simulations à événements discrets modélisant dans le temps l'arrêt (et le changement) de traitement, à la suite d'une réponse

inadéquate ou d'une tolérance insuffisante à un csDMARD ; 2) un arsenal de différentes combinaisons de csDMARDs (en monothérapie, en combinaison de deux à trois csDMARDs) utilisées dans des séquences de csDMARDs servant comme des comparateurs « standards » aux séquences incluant les bDMARDs. Ces possibilités de comparaisons étaient adaptées au contexte clinique de prise en charge de la PR au Royaume-Uni (et dans une moindre mesure dans les pays scandinaves), pays où l'utilisation des combinaisons de csDMARDs semble être plus fréquente que dans le contexte français (cf. chapitre 5).

L'utilisation de la réponse EULAR (peu fréquente dans les évaluations économiques) dans le modèle de simulations à événements discrets a été bien argumentée dans la dernière évaluation du NICE (Stevenson *et al.*, 2016), en fonction des pratiques de prise en charge de la PR au Royaume-Uni.

- **L'évaluation de l'ICER (2017) (13) et l'étude de Janssen *et al.* (2017) (133) se sont particulièrement distinguées par leur rationnel du choix de séquences de traitements incluant trois lignes de traitements bDMARDs différentes. Même si l'évaluation de l'ICER avait pour objectif (final) de comparer des bDMARDs et des DMARDs ciblés, cette spécification des stratégies de séquences est assez proche de celles retenues dans notre évaluation économique.** Quant aux évaluations du NICE (notamment celle de Stevenson *et al.*, 2016 (12)), elles intégraient des séquences de traitements mais différenciaient uniquement les séquences au niveau de la première ligne de traitement. **Ceci s'explique par le fait que l'objectif des évaluations était d'analyser le ratio coût-utilité des médicaments en première ligne de traitement biologique.**
- Les rationnels des méta-analyses en réseau utilisées dans Stevenson *et al.* (2016) (12) ou l'ICER (2017) (13) ont montré le poids des comparaisons indirectes dans les estimations des effets d'efficacité *via* le choix d'un comparateur commun à partir des csDMARDs. Ceci a concerné les réponses ACR au même titre que les réponses EULAR. Les estimations CODA (fichier aléatoire des probabilités de réponses issues de l'estimation bayésienne de la méta-analyse en réseau) des effets des traitements issues des méta-analyses en réseau de Stevenson *et al.* (2016) (12) permettent de fournir une spécification statistique adéquate des paramètres d'efficacité comparative dans les analyses de sensibilité probabilistes en tenant compte de leur corrélation.
- **Le rationnel décrivant la progression de l'activité de la PR dans Stevenson *et al.* (2016) (12) se fonde sur le HAQ.** Il a été appliqué sur les données à long terme d'un registre britannique sur la PR (registre BSR de la Société britannique de rhumatologie) et a

été spécifié à l'aide d'une modélisation statistique permettant de prédire la variation du HAQ dans le temps, et ce, quel que soit le traitement pris par le patient (un csDMARD ou un bDMARD). Cette modélisation permet de dépasser des hypothèses alternatives utilisées dans d'autres évaluations économiques fondées sur des extrapolations linéaires simples qui ne sont pas plausibles d'un point de vue clinique.

- **Amélioration de la modélisation des équations de prédiction des utilités en fonction du HAQ.** Dans la littérature, des hypothèses portant sur la linéarité de la relation entre le HAQ et l'utilité ont été couramment utilisées dans les modèles économiques. La plausibilité d'un point de vue clinique de ces hypothèses a été également critiquée. Afin de pallier cette limite, la dernière version du modèle Sheffield RA s'est fondée sur un modèle économétrique non linéaire pour estimer l'utilité en fonction des caractéristiques des patients (y compris leur HAQ). Ce modèle (Hernandez *et al.*, 2013, (27)) a été appliqué à des données américaines contenant plus de 100 000 observations.

3.4. Les limites actuelles de l'évaluation économique des séquences de traitement dans la littérature

- Au-delà de l'intégration implicite des événements indésirables dans l'arrêt de traitement et le passage d'une ligne de traitement à une autre, la prise en compte des événements indésirables spécifiques aux bDMARDs dans l'estimation des coûts et les éventuelles pertes de qualité de vie demeure insatisfaisante. En dépit des efforts fournis pour intégrer les coûts et les pertes de qualité de vie liées aux événements aigus (ex. : infections graves) dans la plupart des modélisations retenues dans certaines études (cf. **tableau 3**), un argumentaire clinique et/ou économique expliquant l'impact de l'absence de ces éléments sur les résultats de santé observés fait défaut. La modélisation de **l'occurrence des événements indésirables et l'estimation des conséquences des événements indésirables aigus (infections graves) et chroniques (ex. : cancers) sur les résultats de santé sont loin d'être réalisées d'une manière satisfaisante, et les tentatives méthodologiques actuelles ne sont pas non plus jugées satisfaisantes par les experts participant à ces différentes évaluations économiques.**
- Le très faible pourcentage des études explorant l'hétérogénéité au sein de la population d'analyse (3 %, 1/31), identifié dans l'analyse critique issue de la liste CHEERS adaptée à la PR, reflète très clairement l'absence d'évaluations économiques de populations atteintes de PR selon des sous-populations stratifiées sur le risque de comorbidités. Ceci s'explique par le manque de données d'essais cliniques ou de données de registres renseignant sur ces facteurs de risque.

- Seule l'évaluation de Barton *et al.* (2004) distingue et compare l'évaluation économique des PR débutantes et tardives.
- Que l'on soit ou non en accord avec l'utilisation d'une équation de transformation d'un critère d'incapacité fonctionnelle en score d'utilité (*mapping*), il est important de constater que la prédiction de l'utilité en fonction du HAQ et d'autres variables s'est techniquement améliorée dans la nouvelle version du modèle de Sheffield (12). Toutefois, le degré de pertinence de la transposabilité de cette modélisation à d'autres contextes cliniques (ex. : le contexte français) dépendrait non seulement de l'utilisation du HAQ en pratique clinique courante, mais également d'autres critères cliniques pouvant décrire l'activité de la maladie tels que l'indice composite DAS 28 et le critère de détérioration structurale des articulations (indicateur renseignant sur l'érosion osseuse). Par exemple, l'étude française de Saraux *et al.* (2010) (19) et d'autres études telles que l'étude canadienne de Russel *et al.* (2009) (144) ont privilégié l'utilisation du DAS 28 dans la caractérisation de l'activité de la PR.
- Aucune méta-analyse portant sur la réponse EULAR et permettant d'extraire les réponses de traitement en deuxième ligne biologique et plus n'a été identifiée dans les évaluations économiques retenues. Ceci s'explique principalement par la rareté des comparaisons directes qui concerne aussi bien la première ligne que la deuxième ligne de traitements biologiques.

4. Discussion de la revue systématique de la littérature économique

La revue systématique de la littérature économique a été réalisée selon une méthodologie faisant appel à :

- un processus de contrôle qualité reposant sur deux vérifications indépendantes des extractions et des analyses des publications retenues ;
- des grilles de lecture reposant sur les résumés de la grille CHEERS appliquée à la PR ;
- une grille critique adaptée à la PR intégrant des items provenant du croisement de trois grilles d'analyse critique : Drummond *et al.* (2015) (39), OMERACT (2002) (40) et CHEERS (2013) (38).

Les résultats de cette revue systématique de la littérature économique étaient en cohérence avec ceux de Heather *et al.* (2014) (139) portant sur l'intégration des événements indésirables dans l'évaluation économique des anti-TNF α dans la prise en charge de la PR. De même, les résultats portant sur les modèles utilisés et les choix structurants étaient concordants avec la revue de littérature technique de Scholz *et al.* (2014) (145) portant sur les structures des modèles utilisées dans la prise en charge de la PR et celle de Joensuu *et al.* (2015) (146) portant sur les études coût-utilité dans la PR.

La plupart des études concluaient à un ratio coût-efficacité/utilité favorable de la séquence analysée ou du traitement étudié. À titre informatif, plus de 61% (31/51) des études financées par l'industrie pharmaceutique et/ou présentant un conflit d'intérêts aboutissaient à des conclusions favorables au bDMARD (ou à la stratégie du bDMARD) analysé(e).

Ce travail se distingue de la plupart des travaux de la littérature économique analysés dans les sections précédentes tels que l'évaluation du NICE (Stevenson *et al.*, 2016) (12) dans la mesure où :

- il ne s'est pas limité à la revue de la littérature économique des traitements biologiques utilisés en première ligne « biologique » mais a analysé également les études évaluant les stratégies des traitements biologiques considérés en deuxième ligne et plus ;
- il a pris en compte l'ensemble des « dimensions » de l'évaluation économique *via* quatre grilles d'analyse fondées sur des résumés CHEERS ;
- il a utilisé l'ensemble des grilles actuellement disponibles dans l'analyse critique de la littérature.

Ce travail souligne les points suivants :

- une seule étude française portant sur les patients atteints de PR modérée à sévère et présentant une réponse inadéquate à au moins un anti-TNF α respectait les critères d'inclusion de la revue systématique de littérature économique. Les autres études

françaises ne respectaient pas les critères d'éligibilité (Vanier *et al.*, 2017 (20) ; Sany *et al.*, 2009 (18)) ou étaient des analyses de coûts telles que les études de Fautrel *et al.* (2016) (8) et Chevreul *et al.* (2014) (147) ;

- comparativement à l'analyse économique des traitements de fond biologiques en deuxième ligne de la PR, le nombre d'études concernant la troisième ligne et plus était faible. L'extrapolation des analyses et conclusions associées à cette dernière indication doit être lue avec prudence : les études étaient entachées de nombreuses limites liées aux spécificités des études économiques telles que le grand nombre d'hypothèses dont la pertinence est discutable, la variabilité d'estimation des utilités conduisant à des résultats hétérogènes et non reproductibles.

Ce travail présente « les limites » suivantes :

- il n'a pas inclus les évaluations économiques des traitements de fond portant sur d'autres pays et non publiées en français ou en anglais (même si le périmètre de la revue systématique inclut des pays appartenant à trois continents) ;
- comparativement à la revue systématique de la littérature clinique, il n'a pas été possible d'utiliser une grille critique « graduant » et définissant d'une manière standard les risques de biais dans les évaluations économiques. Afin de pallier l'absence de cet instrument d'analyse, le travail s'est fondé sur une grille d'analyse critique combinant des items issus de trois grilles fréquemment utilisées dans la revue systématique de la littérature : la grille CHEERS (grille d'analyse critique de *reporting*), la grille OMERACT (grille implicite d'analyse critique des biais) et la grille de Drummond (ne séparant pas clairement les items de *reporting* des items critiques de l'analyse économique). Comparativement à la grille de lecture CHEC « *Consensus on Health Economic Criteria* » (Evers *et al.*, 2005) (148), les items de l'analyse critique utilisés dans ce rapport concernaient les limites relatives aux rationnels des données (notamment les données de tolérance et de qualité de vie), des hypothèses structurelles utilisées dans la modélisation et l'adéquation de ces dernières à l'évolution de la prise en charge de la PR en vie réelle selon le contexte du pays étudié. La grille ECOBIAS « *Bias in Economic Evaluation* » (Adarkwah *et al.*, 2015) (149) est une piste méthodologique récente à explorer dans l'analyse des biais des revues systématiques de littérature économique. Cette grille propose une liste d'items décrivant les risques de biais liés aux choix structurants de l'évaluation économique et aux caractéristiques des modèles économiques dans les revues de littérature économique, mais, à notre connaissance, elle n'est pas encore validée par un panel d'experts (autre que celui des auteurs de la grille) et ne couvre pas les problèmes pratiques liés à la validation des modèles économiques (Ghabri *et al.*, 2018) (150) ;

- le nombre de publications économiques portant sur les nouveaux traitements de fond synthétiques ciblés (tsDMARDs) et le sarilumab était plus faible que celui portant sur les bDMARDs (soit 6 % du nombre total des études retenues). Ceci s'explique notamment par l'arrivée récente de ces produits (2017 et 2018) et la disponibilité réduite de données sur l'efficacité et la tolérance de ces produits en vie réelle. Il est à noter que des synthèses des avis du NICE sur les nouveaux tsDMARDs (tofacitinib, baricitinib) et le deuxième produit appartenant à la classe des inhibiteurs de la transmission du signal médié par les récepteurs de l'IL-6 (sarilumab) ont été récemment publiées : Ren *et al.* (2018) ([127](#)), Utley *et al.* (2018) ([126](#)), Bermejo *et al.* (2018) ([151](#)). Ces publications ont inclus des appréciations méthodologiques générales sur les évaluations de ces produits par les industriels, mais contrairement à l'évaluation complète de Stevenson *et al.* (2016) ([12](#)) sur les bDMARDs, elles n'ont pas fourni d'informations suffisantes (ex. : choix structurants des évaluations économiques étudiées) permettant de les intégrer dans les grilles d'analyse proposées dans ce projet. En revanche, leurs aspects méthodologiques ont été pris en compte afin de les comparer aux évaluations spécifiques des bDMARDs.

5. Conclusion

La revue systématique de la littérature économique a abouti à la sélection de 52 études dans 15 pays faisant partie de trois continents (Amérique, Europe et Asie) : 41 études ont porté sur les patients traités par un bDMARD en première ligne de traitement biologique et présentant une réponse inadéquate ou une tolérance insuffisante à des csDMARDs, et 11 études ont porté sur le traitement des patients adultes atteints de PR active, sévère, présentant une réponse inadéquate ou une intolérance aux traitements de fond, dont au moins un anti-TNF α . Quarante-huit pour cent des études ont été réalisées dans les pays scandinaves et au Royaume-Uni. Une seule étude française portant sur la PR active modérée à sévère et ayant une réponse insuffisante à au moins un anti TNF α a été identifiée (Saraux *et al.*, 2010) ([19](#)).

L'extraction des données s'est fondée sur un processus de contrôle qualité similaire à celui des méta-analyses en réseau décrites dans le chapitre 3. L'analyse de la littérature a reposé sur des résumés issus de la grille CHEERS et sur les grilles de lecture utilisées dans l'élaboration du guide méthodologique de l'impact budgétaire de la HAS (Ghabri *et al.*, 2018 ([131](#))) et adaptées à la PR. Quant à l'analyse critique de la littérature économique, elle s'est fondée sur le recensement des items de l'évaluation respectant les recommandations CHEERS et a été complétée par une liste courte de critères provenant du croisement de trois grilles d'analyse critique : Drummond *et al.* (2015) ([39](#)), OMERACT (2002) ([40](#)) et CHEERS (2013) ([38](#)). Ces critères ont été validés par le groupe de travail PR.

La qualité des études sélectionnées concernant la description générale des études portant sur les coûts, les sources de données d'efficacité mobilisées dans les modélisations était bonne. Les modèles agrégés (modèles de cohortes multi-états markoviens), au même titre que les modèles de micro-simulation (modèles markoviens individuels ou les simulations à événements discrets), peuvent être utilisés dans la modélisation économique de la prise en charge de la PR. Cependant, les modèles retenus dans les organismes évaluant les technologies de santé étaient plutôt des modèles de micro-simulation (ex. : simulations à événements discrets). La prise en compte des séquences de traitements est assez fréquente dans les utilisations récentes des modèles de micro-simulation, mais reste majoritairement adaptée à l'évaluation des bDMARDs en première ligne de traitement biologique plutôt qu'à l'évaluation d'un parcours médicamenteux fondée sur au moins trois lignes différentes de traitements biologiques.

Peu d'informations concernant les sources de données portant sur la tolérance des bDMARDs (i.e. les événements indésirables) étaient disponibles dans les études. De même, peu d'études intégraient les événements indésirables dans l'estimation des coûts et des éventuelles pertes de qualité de vie des patients. Par ailleurs, à part les évaluations récentes du NICE et de l'ICER,

l'utilisation des équations de prédiction des utilités en fonction du HAQ était peu argumentée et loin d'être acceptée par l'ensemble des experts cliniciens et économistes.

D'une manière générale, les résultats reflètent une forte hétérogénéité et variabilité des RDCRs mises en évidence dans les dimensions suivantes : l'indication analysée, la variabilité inter-temporelle, la variabilité intra-pays (recommandations en vigueur concernant le traitement de la PR) et la variabilité inter-pays. Toutefois, dans des évaluations récentes des organismes d'évaluation de technologies de santé (i.e. NICE, ICER), l'efficience n'a pas été établie : le RDCR était supérieur à 40 000 £/QALY pour l'indication de la PR sévère à modérée au Royaume-Uni ([12](#)) et supérieur à 100 000 \$ US/QALY pour la même indication aux États-Unis (ICER, 2017) ([13](#)).

Les résultats de cette revue de littérature sont certainement très utiles pour renseigner les aspects économiques communs à la majorité des évaluations économiques (ex. : typologie de coûts, horizon temporel, actualisation, présentation des résultats) mais n'apportent pas aujourd'hui de réponses précises à un certain nombre de questions : mobilisation des données de tolérance et de qualité de vie dans la modélisation, approche d'estimation des utilités dans la PR, intégration d'au moins trois lignes de traitements biologiques différentes dans les séquences de traitements biologiques. De plus, les données recueillies ne portent pas ou très peu sur le contexte français. Ceci montre clairement l'intérêt de ce travail.

Chapitre 5. Modélisation pour l'évaluation économique de la prise en charge de la PR en France

L'évaluation économique proposée dans ce rapport ne répond pas uniquement à la rareté des études coût-efficacité/utilité portant sur la prise en charge de la PR en France, mais propose également d'analyser le parcours médicamenteux des patients atteints de PR modérée à sévère. À ce propos, elle compare des séquences de traitements de fond biologiques entre elles et non des médicaments *via* une ligne de traitement préalablement définie, comme dans le cas des comparaisons en première ligne de traitements biologiques évalués par le NICE (Stevenson *et al.*, 2016) ou l'ICER (ICER, 2017) (13).

Ce chapitre est dédié à la présentation de la méthodologie et des données utilisées dans l'évaluation économique des traitements de fond biologiques dans la prise en charge de la PR en France. Cette évaluation s'appuie sur un nouveau modèle de décision économique dans la prise en charge de la PR, appelé « PR France ». **PR France compare les stratégies de prise en charge de la PR en termes de séquences de traitements biologiques conformément à la saisine de ce projet** (cf. section 1, chapitre 1). Il est fondé sur une approche méthodologique appropriée à la prise en charge de la PR en France et est en accord avec :

- les recommandations du groupe de travail PR ;
- les conclusions des avis d'évaluation et de réévaluation de la Commission de transparence (CT) de la HAS portant sur les bDMARDs ;
- les recommandations récentes des sociétés de rhumatologie (ex. : la Société française de rhumatologie, SFR) et de la Ligue européenne contre les maladies rhumatismales (*European League Against Rheumatism*, EULAR) ;
- les données françaises disponibles en vie réelle ;
- les recommandations des sociétés professionnelles (ISPOR, SMDM) sur les bonnes pratiques en modélisation économique.

La première section explicite le choix du modèle PR France, présente sa structure, son implémentation technique et le processus de sa validation. Les choix structurants de l'évaluation économique sont décrits dans la deuxième section. La troisième section présente les sources de données mobilisées dans l'évaluation économique, incluant les données d'efficacité, de tolérance, de mortalité, de coûts et des résultats de qualité de vie/utilité. La quatrième section présente les résultats des équations de survie du modèle PR France. La cinquième section synthétise les autres entrées du modèle PR France.

1. Le modèle PR France

Depuis les années 2000, des modèles markoviens agrégés (i.e. cohorte multi-états) ou des modèles de micro-simulations (ex. : modèles markoviens individuels ou simulations à événements discrets) ont été proposés pour estimer le rapport différentiel coût-résultat (RDCR) des traitements pour la PR (Scholz *et al.*, 2014) (145). Aujourd'hui, les modèles de micro-simulations sont de plus en plus utilisés dans les évaluations économiques (Stevenson *et al.*, 2016 (12) ; ICER, 2017 (13)). Ce type de modélisation permet notamment de simuler l'évolution du parcours de patients atteints d'une PR modérée à sévère en fonction des recommandations de bonnes pratiques cliniques.

1.1. Les raisons du choix de modèle PR France

Un modèle de micro-simulation « modèle à conditions et événements discrètement intégrés » (CEDI) (ou *Discretely Integrated Condition Event*, DICE) (Caro, 2016 (33) ; Caro & Moller, 2018 (34)) a été retenu dans le cadre de cette évaluation. Ce choix se justifie par les éléments suivants :

1. ce type de modélisation simule les caractéristiques individuelles des patients et permet de tenir compte de l'hétérogénéité des parcours médicamenteux des patients ;
2. il tient compte des séquences de traitements de fond de la PR (*Disease-Modifying Antirheumatic Drugs*) incluant une combinaison de traitements csDMARDs (méthotrexate + sulfasalazine + hydroxychloroquine) appelée « trithérapie » (TT) ou des traitements de fond biologiques (bDMARDs) et leurs biosimilaires ou des traitements de fond synthétiques ciblés (tsDMARDs), en association ou non avec le méthotrexate (MTX) ;
3. il repose sur la distinction entre des éléments du modèle qui persistent dans le temps, appelés « conditions » et d'autres qui surviennent à un moment donné, appelés « événements » (cf. paragraphe 2.2) ;
4. il offre des possibilités en termes de « flexibilité » de modélisation (introduction ou modification des entrées du modèle et de sa structure selon le choix de l'utilisateur), de transparence et de documentation du processus de modélisation (interface DICE).

1.2. Structure du modèle PR France

Le modèle PR France est fondé sur une dimension temporelle qui repose sur **la réalisation de l'événement d'intérêt** « *time-to-event* ». Cette technique de modélisation est de plus en plus utilisée dans les évaluations des technologies de santé (Caro *et al.*, 2016) (33). Selon cette technique, le devenir d'un patient est modélisé dans le temps en considérant les événements qui peuvent se produire (ex. : le développement d'une érosion) ou ceux qui peuvent se répéter plusieurs fois (ex. :

le changement de ligne de traitement de fond). Tous les patients atteints de PR modérée à sévère et présentant une réponse inadéquate (ou une tolérance insuffisante) à un traitement de fond synthétique (ex. MTX) « entrent » dans le modèle au moment de la décision d'initiation du traitement par un bDMARD ou un tsDMARD ou par la trithérapie (TT). Les séquences de traitements de fond (de la 1^{re} à la 3^e ligne) après échec du MTX sont simulées, et le bDMARD est sélectionné en fonction de la séquence pré-spécifiée. Pour chaque patient, une date est choisie pour chaque événement d'intérêt du modèle. Le modèle procède ensuite au « déclenchement » de chaque événement selon son occurrence dans le temps. Certains événements peuvent, à leur tour, modifier les temps d'occurrence d'autres événements. Par exemple, l'événement « Réponse EULAR » peut déclencher à nouveau l'événement « Initier traitement », si la réponse est jugée insuffisante. Lors de la survenue d'un événement, en particulier lors d'une consultation médicale (ex. : consultation d'un rhumatologue), le tableau clinique est mis à jour et la décision de changer de traitement peut être prise, conformément aux traitements spécifiés selon les séquences prédéfinies. Les valorisations en termes de qualité de vie et de coûts sont appliquées au moment de l'occurrence de chaque événement, le cas échéant, et elles sont accumulées en tant que sorties du modèle. Ce processus continue jusqu'à la fin de l'horizon temporel.

Ce modèle de micro-simulation a été implémenté en utilisant la simulation DICE (Caro, 2016) (33). Cette technique s'articule autour de deux concepts principaux, les événements et les conditions :

- **un événement** dans un modèle DICE reflète tout aspect du problème qui se produit à un moment donné et entraîne des conséquences (ex. : le développement d'une érosion peut entraîner un changement de traitement de fond) ;
- **une condition** exprime un aspect du problème qui persiste au cours du temps (ex. : l'activité de la PR calculée selon le score DAS 28).

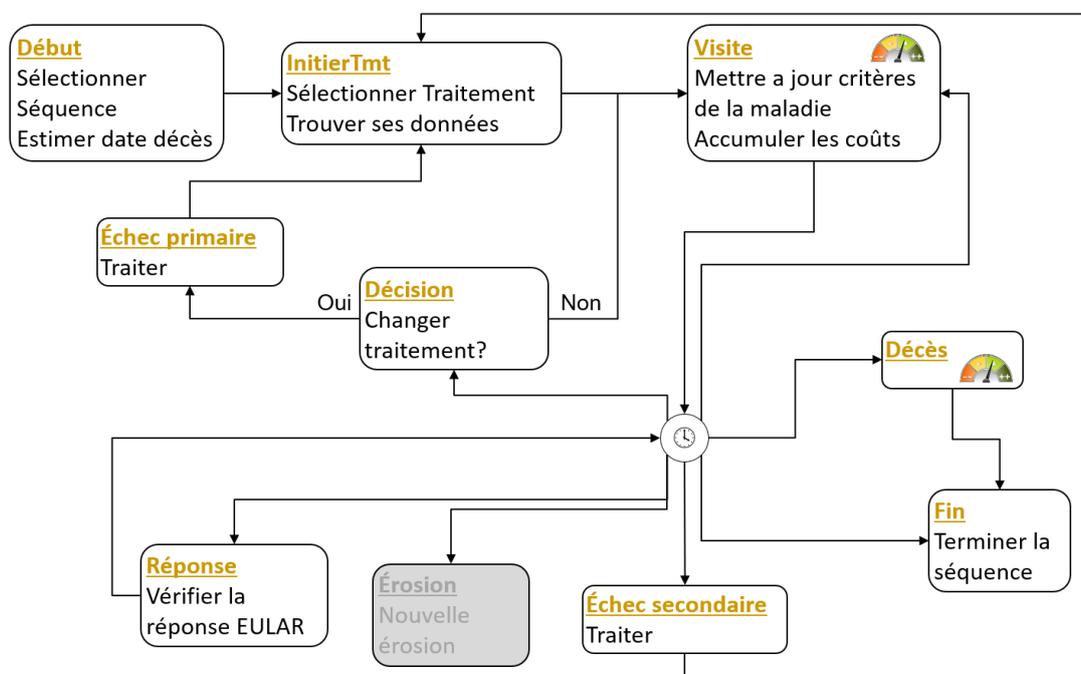
1.2.1. Les événements

Le modèle inclut 11 événements : **DÉBUTER**, **ÉVALUER**, **INITIER TRAITEMENT**, **RÉPONSE EULAR**, **DÉCISION**, **VISITE**, **ÉCHEC PRIMAIRE**, **ÉCHEC SECONDAIRE**, **ÉROSION**, **DÉCÈS** et **FIN**. Ces événements sont liés dans le modèle pour créer la structure logique (**Figure 1**). L'événement « **DÉBUTER** » correspond à la sélection de la séquence à simuler, l'estimation du temps jusqu'au décès (i.e. la survie du patient). L'événement suivant est « **INITIER TRAITEMENT** » où le premier traitement est choisi (bDMARD ou TT) lors d'une consultation et ses caractéristiques (ex. : dose, prix) sont retenues. L'événement « **VISITE** » est celui au cours duquel les critères spécifiques de la maladie (activité, incapacité fonctionnelle) et du traitement (dose) sont mis à jour et les coûts de la visite et des examens sont accumulés. L'événement « **ÉVALUER** » est toujours appelé immédiatement après un autre événement pour accumuler les résultats. À ce stade, plusieurs événements peuvent se réaliser selon ces différentes dates. L'événement suivant peut être « **RÉPONSE EULAR** », qui est défini selon une durée par défaut égale à 6 mois. Lors de la survenue de cet événement, la réponse au traitement est examinée selon les critères EULAR et le changement résultant de l'indice DAS 28 est calculé¹². L'événement « **DÉCISION** » est appelé immédiatement, ou après un certain délai, selon le choix de l'utilisateur. Si l'événement est reporté à une date ultérieure, le changement dans le DAS 28 est projeté de façon linéaire jusqu'à ce que la décision soit prise de maintenir le traitement ou de passer en deuxième ligne. Si la décision est de passer tout de suite en deuxième ligne de traitement, le patient procède à l'événement « **ÉCHEC PRIMAIRE** » où les conséquences sont comptabilisées, et ensuite, il revient à l'événement « **INITIER TRAITEMENT** ». Sinon, un délai d'échec secondaire est défini. Le patient continue alors ses visites jusqu'à l'échec secondaire qui le conduit à l'événement « **ÉCHEC SECONDAIRE** » ; l'événement « **INITIER TRAITEMENT** » est alors appelé afin que le patient passe à la ligne de traitement suivante. Le processus de vérification de la réponse EULAR et, par conséquent, la prise d'une nouvelle décision sont réactivés. L'événement « **FIN** » se produit quand le « **DÉCÈS** » survient (selon la date choisie au début) ou quand l'horizon temporel est atteint. L'événement « **ÉROSION** » n'est pas actif dans la version courante du modèle en raison de l'indisponibilité des données nécessaires (taux de développement et efficacité des bDMARDs pour le prévenir ou le « retarder »).

¹² Le médecin évalue le DAS 28 et à partir de cela il détermine la réponse. Nous la simulons « à l'envers » parce que la méta-analyse nous a fourni la réponse mais pas la valeur du DAS 28.

Tableau 1. Les événements dans le modèle PR France

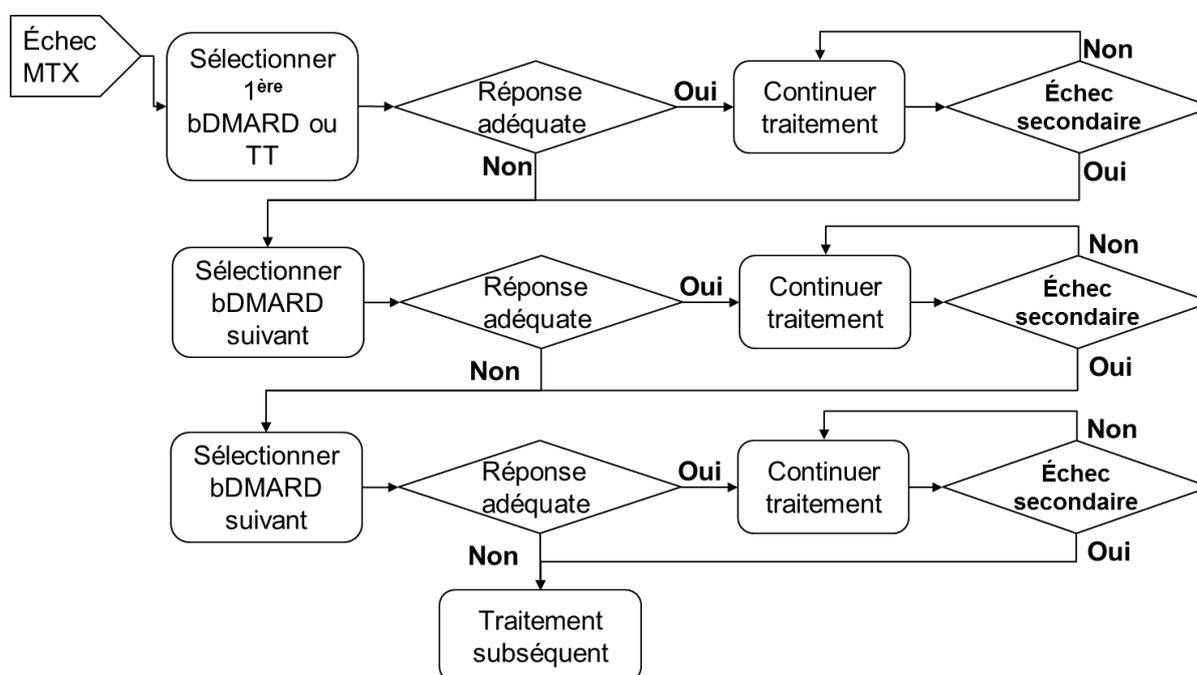
Nom	Objectif
Débuter	Initialiser l'exécution pour la séquence pour chaque patient
Évaluer	Appliquer la valorisation, accumuler les résultats
Initier Tmt	Initialiser le traitement (n'importe quelle ligne)
Réponse EULAR	Vérifier la réponse au traitement
Décision	Prendre la décision de changer ou continuer le traitement
Échec Primaire	Traiter l'échec primaire
Échec Secondaire	Traiter l'échec secondaire
Visite	Mettre à jour les critères spécifiques de la maladie, accumuler les coûts des visites
Érosion	Processus de développement d'une nouvelle érosion
Décès	Traiter le décès
Fin	Fin d'exécution de la séquence

Figure 1. Diagramme du parcours du patient dans le modèle PR France

NB : L'icône 🌈 indique qu'un événement « Évaluer » est immédiatement déclenché. Cet événement est spécialement conçu pour cumuler les coûts et les QALYs. La structure du modèle PR France reste inchangée dans les analyses de scénarios complémentaires dans lesquels les tsDMARDs sont inclus.

Le modèle considère les séquences contenant jusqu'à trois lignes de traitement après échec du MTX (cf. **Figure 2**).

Figure 2. Les lignes de traitement considérées dans le modèle PR France



1.2.2. Les conditions

Il existe plusieurs types de conditions dans le modèle PR France. Certaines sont **statiques** (leurs valeurs ne changent pas dans le temps), tandis que d'autres sont **dynamiques** (leurs valeurs peuvent varier dans le temps). Les principales caractéristiques du patient sont simulées en définissant une condition pour chacune d'elles. Les valeurs à l'état initial du modèle ne changent pas dans le temps et sont stockées pour chaque patient dans la table **Profil de patients** (cf. tableau 2).

Tableau 2. Profil des patients

ID	Sexe	Âge	Poids	DAS 28	HAQ	Érosion	Activité	Utilité EQ-5D	Utilité SF-6D
1	1=H 2=F	A la v0	kg à la v0	Score à la vB	Catégorie à la vB	Oui/Non	Calculée à partir du DAS 28	A la vB	A la vB
2									
3									
...									

Légende : ID = numéro séquentiel du profil ; v0 = visite à l'inclusion dans la cohorte ESPOIR ; vB = visite de base dans le modèle (début des bDMARDs) ; **catégories du HAQ** (1:<0,6, 2:[0,6-1,1[, 3:[1,1-1,6[, 4:[1,6-2,1[,5:>2,1) ; **activité** : DAS 28<2,6 : rémission ; 2,6<=DAS 28<3,1 : activité faible ; 3,1<=DAS 28<5,1 : activité modérée ; DAS 28>5,1 : activité sévère.

D'autres conditions permettent de contrôler l'attribution du traitement de fond en cours et ses caractéristiques, les caractéristiques spécifiques de la PR (ex. : activité de la PR mesurée à l'aide du DAS 28), les paramètres des fonctions de survie, les divers types de coûts, les événements indésirables (EI). L'ensemble des conditions (tableau 3) est stocké dans une table **Conditions** qui contient deux colonnes : la première affiche le nom de chaque condition et la seconde contient sa valeur initiale (zéro si elle n'est pas connue au départ).

Tableau 3. Conditions incluses dans le modèle PR France

Noms	Définitions et valeurs initiales
Actif	Condition obligatoire qui indique que le patient en cours de simulation est vivant
Activité	Niveau actuel de l'activité de la PR (selon DAS 28 : 1=rémission <2,6 ; 2=faible 2,6-3,1 ; 3=modéré 3,2-5,1 ; 4=élevé >5,1)
Utilité EI	Perte de qualité de vie due aux EI, comptée une seule fois pour chaque ligne de traitement
BsActivité	Niveau d'activité au début, à partir de la table Profils
BsÂge	Âge au démarrage du modèle, extrait à partir de la table Profils
BsDAS	DAS 28 à l'initiation du modèle, extrait à partir de la table Profils
BsÉrosion	Statut initial des modifications radiologiques, extrait à partir de la table Profils
BsHAQ	HAQ à l'état initial du modèle (1 : <0,6 ; 2 : [0,6-1,1[; 3 : [1,1-1,6[; 4 : [1,6-2,1[; 5 : >2,1) extrait à partir de la table Profils
TempsDécision	Temps pour déterminer s'il faut changer le traitement, valeur initiale selon l'utilisateur, mais au moins 6 mois
ChgDAS	Variation en cours du DAS 28 fondée sur un ensemble possible de valeurs en utilisant une distribution uniforme entre les deux limites
ChgDASInf	Limite inférieure du changement possible du DAS 28 calculée en fonction de la réponse
ChgDASSup	Limite supérieure du changement possible du DAS 28 calculée en fonction de la réponse
ChgTmt	"Oui" ou "Non" selon l'activité ou la réponse ou l'apparition d'une érosion
Classe	Classe thérapeutique du médicament en cours
ClasseAnt	Classe thérapeutique du médicament antérieur dans la séquence
CoûtDose	Coût unitaire d'une dose, calculé en fonction du poids, le cas échéant, extrait de la table Médicaments
CoûtLab	Coût des examens de laboratoire relatifs aux visites effectuées
CoûtLabInit	Coût supplémentaire des examens biologiques lors de la première consultation
CoûtMG	Coût d'une consultation auprès d'un médecin généraliste
CoûtRad	Coût de l'imagerie lors d'une consultation
CoûtRH	Coût d'une consultation auprès d'un rhumatologue
CoûtAdmin	Coût unitaire d'administration selon le lieu d'administration, extrait à partir de la table Médicaments
CoûtEI	Coût de la prise en charge des événements indésirables (par classe thérapeutique) compté une seule fois pour chaque ligne de traitement
CoûtDoseCharge	Coût total des doses de chargement, extrait de la table Médicaments
Âge	Âge actuel

Noms	Définitions et valeurs initiales
TmtAct	Index du traitement en cours à partir de la table Médicaments
DAS	DAS 28 actuel
Décédé	“Oui” si décès, “Non” si vivant
DoseParPoids	Dosage du traitement en fonction du poids : Oui=1, Non=0
DosesDues	Nombre de doses qui auraient dû être prises depuis la dernière visite
DosesDonnées	Nombre total de doses déjà prises jusqu’à la visite précédente
DosesAdonner	Nombre total de doses à prendre depuis la visite précédente
EffPertePGOther	Perte proportionnelle d'efficacité (bonne réponse) lorsqu'un médicament autre qu'un anti-TNF est prescrit en 2 ^e ligne biologique
EffMultpB	Multiplicateur permettant d'appliquer la perte d'efficacité relative (bonne réponse)
EffMultpM	Multiplicateur permettant d'appliquer la perte d'efficacité relative (réponse bonne ou modérée)
EffPertePBAntiTNF	Perte proportionnelle d'efficacité (bonne réponse) lorsqu'un anti-TNF est prescrit en 2 ^e ligne biologique après un premier anti-TNF
EffPertePMAntiTNF	Perte proportionnelle d'efficacité (réponse bonne ou modérée) lorsqu'un anti-TNF est prescrit en 2 ^e ligne biologique après un premier anti-TNF
EffPertePMOther	Perte proportionnelle d'efficacité (réponse bonne ou modérée) lorsqu'un médicament autre qu'un anti-TNF est prescrit en 2 ^e ligne biologique
GErosion	Paramètre Gamma pour l'équation de la survenue d'une érosion
GSurv	Paramètre Gamma dans l'équation de survie
HAQ	Catégorie du HAQ actuel
HRHAQ	Hazard ratio appliqué à la mortalité associée au HAQ
ID	Nombre séquentiel associé à chaque profil
InterventionNum	Numéro de la séquence
LErosion	Paramètre Lambda de l'équation relative à l'apparition d'une érosion
Ligne	Ligne du traitement dans la séquence
LigneTemps	Temps depuis l'initiation d'une ligne de traitement
PeriodDoseCharge	Durée des doses de chargement
LSurv	Paramètre Lambda dans l'équation de survie
MaxLigne	Nombre maximal de lignes
MortAdj	Hazard ratio appliqué à la mortalité associée à la PR
MTXcoût	Coût annuel du méthotrexate
PAug	Probabilité selon laquelle l'activité augmente d'une visite à la suivante
PBonne	Probabilité d'une bonne réponse EULAR, extraite de la table Médicaments
PDim	Probabilité selon laquelle l'activité diminue d'une visite à la suivante
Periode	Temps entre les doses (en jours), extrait de la table Médicaments
PMod	Probabilité d'une réponse EULAR modérée ou bonne, extraite de la table Médicaments
RandChgActivité	Numéro aléatoire utilisé pour déterminer si l'activité change
RandÉchec	Numéro aléatoire utilisé pour déterminer s'il y a un échec
RandHAQ	Numéro aléatoire permettant de sélectionner la catégorie HAQ à partir de la distribution de l'activité de la PR.
RandRéponse	Numéro aléatoire utilisé pour assigner la réponse EULAR
RandSurv	Numéro aléatoire utilisé pour sélectionner le temps de décès
RandSwitch	Numéro aléatoire utilisé pour sélectionner le temps de changement de traitement dû à un échec secondaire
Réponse	Réponse EULAR (non, modérée, bonne)
RepNum	Numéro de la réitération en cours
Séquence	Séquence en cours
Sexe	Sexe (1=Homme, 2=Femme), extrait de la table profils

Noms	Définitions et valeurs initiales
SwitchK	Paramètre Kappa relatif au changement de traitement
SwitchMu	Paramètre associé à la constante dans l'équation du temps jusqu'au changement de traitement dû à un échec secondaire
SwitchQ	Paramètre Lambda dans l'équation relative au changement secondaire de traitement
SwitchModbD	Paramètre identifiant les bDMARDs dans l'équation relative au changement secondaire de traitement
SwitchModL2	Paramètre diminuant le taux d'échec secondaire en ligne 2
SwitchModL3	Paramètre diminuant le taux d'échec secondaire en ligne 3
SwitchRéduc	Paramètre qui résume la réduction totale du taux d'échec
TempsInit	Enregistre la date et l'heure du début de chaque ligne de traitement
Temps	Horloge de la simulation
HorizTemp	Durée de la simulation
TmpCoût	Condition temporaire utilisée dans le calcul des coûts
TmpSwitch	Condition temporaire utilisée dans le calcul du temps jusqu'au changement de traitement
TmpUtil	Condition temporaire utilisée dans l'accumulation des QALYs
UtilEQ-5D	Utilité fondée sur l'EQ-5D
Utiliser EQ-5D	1 si Oui ou 2 si Utiliser SF-6D pour les utilités
Utilité	Score d'utilité actuel
UtilSF-6D	Score d'utilité estimé à partir du SF-6D
VisitNum	Nombre séquentiel de visites dans la ligne en cours
Poids	Poids en kg, extrait de la table Profils

1.3. Exécution du modèle

L'exécution de PR FRANCE (qui se traduit par la lecture des tables d'événements et la réalisation de chaque équation) est implémentée dans **MS Excel®** en utilisant une macro écrite selon le langage « *Visual Basic* pour applications » (VBA). La macro parcourt chaque ligne de chaque table d'événements et exécute tout ce que l'expression spécifie. La macro réalise ce travail en convertissant chaque expression de texte en une formule qui suit la syntaxe MS Excel® et l'exécute avec VBA. La boucle continue jusqu'à ce que l'événement **Fin** soit réalisé de telle sorte que les résultats accumulés dans les conditions de sortie soient affichés. Il est à noter que la macro est générique pour tous les modèles car les spécificités du modèle sont indiquées dans les tables **Conditions et Événements**, et non dans le code VBA. Comme le langage VBA ne permet d'interpréter que des instructions rédigées en anglais selon le format américain, les instructions sont conformes à cette restriction.

L'application VBA, interagissant avec une feuille de calcul MS Excel®, est lente. Ceci représente un problème lors de l'exécution des analyses de sensibilité probabilistes (ASP) où plusieurs milliers d'exécutions du modèle sont requises. Toutefois, pour accélérer l'exécution, la macro lit toutes les conditions dans la mémoire et exécute les instructions tabulées dans les événements sans

utiliser la feuille de calcul elle-même pour les calculs. Cette version de la macro (appelée ROLL) est programmée en C#¹³ et est appelée chaque fois que le fichier du modèle est ouvert.

Une itération est une réalisation du modèle pour toutes les séquences. Plusieurs itérations peuvent être exécutées pour constituer une exécution de modèle pour les ASP. Pour chaque exécution, la macro vérifie que toutes les tables requises existent ; toutes les expressions sont dans la syntaxe MS Excel® appropriée ; toutes les **conditions référencées** dans les tables **d'événements** sont mentionnées dans la table **Conditions** ; et toutes les sorties appelées dans les tables **d'événements** existent dans la table **des sorties**.

1.4. Validation du modèle

Le processus de validation du modèle PR France se fonde sur les bonnes pratiques développées dans les recommandations de l'ISPOR sur la conceptualisation et la validation des modèles (Roberts *et al.*, 2012 ([152](#)) ; Eddy *et al.*, 2012 ([153](#))), les recommandations méthodologiques de la HAS pour l'évaluation économique ([30](#)) et Ghabri *et al.*, 2018 ([131](#)).

1.4.1. Adéquation des résultats avec l'intuition : *face validity*

L'adéquation des résultats avec l'intuition permet de vérifier si la structure du modèle et ses hypothèses sont compatibles avec l'histoire naturelle de la pathologie étudiée, ses caractéristiques et sa prise en charge en France. Ce processus de validation a reposé principalement sur :

- les échanges relatifs à la structure du modèle avec les membres du groupe de travail PR lors de ses première et deuxième réunions (1^{er} juillet 2016 et 22 juin 2017) ;
- l'élaboration d'un algorithme de changement de traitements (*switch treatments*) et un calendrier de gestion de la pathologie fondé sur les recommandations pour la prise en charge de la PR en France et sur les données en vie réelle disponibles. Ces aspects ont été approuvés par les rhumatologues du groupe de travail PR et vérifiés d'une manière indépendante par deux pharmaciens du Service évaluation économique et santé publique (SEESP) ;
- l'examen du modèle par des experts indépendants spécialisés dans la modélisation de la PR (ex. : Pr Pelham Barton, Université de Birmingham et Franck Maunoury, Stasesia, CNAM et Université du Mans) et les membres du groupe de travail PR (rhumatologues, économistes, pharmaciens et un représentant d'une association de patients français trai-

¹³ C'est le nom du langage de programmation : [https://en.wikipedia.org/wiki/C_Sharp_\(programming_language\)](https://en.wikipedia.org/wiki/C_Sharp_(programming_language)).

tant de la PR). Les noms et les affiliations des membres du groupe de travail sont donnés dans l'annexe 1-3 du chapitre n° 1.

1.4.2. Validation interne

La validité interne concerne la vérification **complète et indépendante** de la programmation du modèle, l'intégration des calculs (i.e. équations du modèle PR France), la plausibilité des valeurs (entrées et obtenues). Elle a été réalisée d'une manière indépendante et selon un processus à deux étapes par un expert de modélisation non impliqué dans la conception du modèle (Dr Jorgen Moller, Université de Lund, Suède, et Evidera, Londres). Une synthèse des éléments techniques vérifiés et validés est présentée dans l'annexe 6.2 du chapitre n° 6.

1.4.3. Validation externe

La validation externe du modèle PR France est conditionnée par l'existence de données « indépendantes » portant sur la trajectoire du DAS 28, les échecs de traitements et les réponses EULAR obtenues en vie réelle. Le modèle de Sheffield utilisé dans l'évaluation du NICE ([12](#)) se fondait sur la réponse EULAR mais ne reposait pas sur l'évolution de l'activité de la PR selon le DAS 28. **De ce fait, il n'était pas adapté à la modélisation de la prise en charge de la PR en France : ses simulations reposaient essentiellement sur l'évolution de l'indicateur d'incapacité fonctionnelle** (le HAQ). Au même titre que le modèle de Sheffield, le modèle de l'ICER (2017) ([13](#)) repose également sur l'évolution du HAQ.

Une validation fondée sur des données françaises « non indépendantes » telles que les données de la cohorte ESPOIR (Ghabri *et al.*, 2018) ([150](#)) est prévue dans les mises à jour du modèle PR France.

2. Choix structurants de l'évaluation économique

Les choix méthodologiques et de présentation des résultats de cette évaluation économique ont été principalement fondés sur les recommandations du guide sur les « *Choix méthodologiques pour l'évaluation économique à la HAS* » (30). Ils ont été complétés à partir d'autres recommandations méthodologiques telles que les bonnes pratiques de l'ISPOR sur la modélisation (Roberts *et al.*, 2012 (152), Caro *et al.*, 2012 (154)), la validation des modèles (Eddy *et al.*, 2012) (153) et la prise en compte de l'incertitude (Briggs *et al.*, 2012) (155).

2.1. Type d'analyse

Le type d'évaluation économique retenue dans le cadre de l'analyse de référence est l'analyse coût-utilité. En effet, la qualité de vie liée à la santé a été identifiée comme une conséquence importante associée à l'activité et la sévérité de la PR. La PR altère la qualité de vie du patient, conduit à une incapacité fonctionnelle et à la difficulté d'exercer une activité professionnelle. Dans certaines situations, elle nécessite le recours à la chirurgie (arthroplastie, arthrodèse, synovectomie). L'analyse coût-utilité est accompagnée d'une analyse coût-efficacité (ex. : résultats exprimés en temps passé avec faible activité de la PR). Le temps passé en faible activité est un résultat de santé calculé en fonction du DAS 28 qui a été identifié dans la revue systématique de la littérature économique et qui est cliniquement considéré comme l'un des objectifs des bDMARDs (en plus de la rémission).

2.2. Stratégies à comparer, stratégies de l'analyse de référence, méthode d'analyse

Le nombre de séquences étant initialement trop élevé (352 séquences, annexe A5.1), il a fait l'objet d'une optimisation mathématique après avoir réduit le nombre initial de séquences théoriquement possibles (758 séquences) en faisant les choix suivants :

1. on considère 10 produits (objectif de la saisine) : trithérapie (TT), adalimumab (ADA), étanercept (ETA), infliximab (INF), certolizumab (CER), golimumab (GOL), abatacept (ABA), tocilizumab (TOC), rituximab (RTX), Benepali (BEN, biosimilaire de l'ETA). Les produits concernant les traitements de fond ciblant les JAKs (ex. : baricitinib, tofacitinib) et le deuxième produit de la classe « inhibiteur de la transmission du signal médié par les récepteurs de l'IL-6 » (sarilumab) sont intégrés dans des analyses de scénarios complémentaires selon la disponibilité des données¹⁴. Tous les produits biologiques sont considérés en association avec le MTX ;

¹⁴ La prochaine mise à jour du modèle PR France intégrera les résultats de la mise à jour de la méta-analyse incluant les traitements de fond synthétiques ciblés.

2. on considère **trois lignes de traitements** différentes **après échec du MTX** ;
3. après échange avec le groupe de travail PR, les produits considérés en première ligne après échec du MTX sont :
 1. les anti-TNF, sauf l'infliximab,
 2. le biosimilaire de l'étanercept : Benepali.

Le rituximab peut être utilisé comme médicament de secours (quatrième ligne biologique et plus). La trithérapie (en première ligne après échec du MTX) est considérée dans une analyse de scénario. De même, l'abatacept et le tocilizumab sont également considérés en première ligne biologique après échec du MTX dans une analyse complémentaire modifiant le rationnel du modèle PR France (cf section 2.3, chapitre 6) ;

4. **la stratégie (séquence) de « référence » était la séquence la moins coûteuse.** La stratégie de référence est, par définition, la moins chère quand on analyse plusieurs stratégies. De cette façon, il est possible de calculer les RDCRs pour les stratégies qui ne sont pas dominées.

Les stratégies non cliniquement pertinentes ont été exclues, telles que les séquences présentant à la fois la molécule mère et son biosimilaire (ex. : Étanercept → Infliximab → Remsima). **Au final, 180 séquences ont été retenues** (cf. annexe 5-1). En outre, il y avait 54 autres séquences qui commencent par la TT dans la première ligne. Une simulation d'une séquence dure 1,185 secondes par patient. Ainsi, il faut compter environ 3,5 minutes pour un patient pour les 180 séquences. Pour une population de 1 000 patients, cela reviendrait à 59 heures. Même si le nombre de patients était réduit à 500, il faudrait attendre 29,5 heures sans prendre en compte la réalisation des analyses de sensibilité. En raison de ces contraintes techniques qui augmentent considérablement le temps de calcul des sorties du modèle PR France, au même titre que la quantité d'informations à traiter, la question de l'utilisation d'un logiciel différent s'est posée. Le logiciel VBA inclus avec Excel est bien connu, mais son exécution s'est avérée très lente pour ce type de simulation. En revanche, même si le modèle reste dans Excel pour conserver sa facilité d'utilisation et sa transparence, l'utilisation d'un langage comme C# (pour la macro qui exécute les instructions) permet d'atteindre des vitesses d'exécution entre 20 et 30 fois plus élevées. Une simulation de 1 000 patients avec 180 séquences pourrait durer 2,9 heures en utilisant la version du DICE¹⁵ programmée en C# (0,06 secondes par séquence et par patient). Pour information, ROLL (nom du logiciel DICE programmée en C#) était en phase de test, lors de la validation interne du modèle PR France.

¹⁵ DICE (modèle à conditions et événements discrètement intégrés) est le type de simulation utilisé dans le modèle PR France.

Afin de prendre en compte d'autres analyses des séquences qui ne sont pas exactement fondées sur le même rationnel des stratégies à comparer dans l'analyse de référence et qui peuvent correspondre à d'autres stratégies médicamenteuses de prise en charge de la PR, deux analyses complémentaires ont été approuvées par le groupe de travail PR et suggérées par certains membres du groupe de lecture. L'objectif de la première analyse complémentaire était d'intégrer certaines séquences de bDMARDs qui n'ont pas été incluses dans le rationnel de l'analyse de référence : par exemple, des séquences débutant la première ligne biologique par un non-anti-TNF α (ex. : tocilizumab ou abatacept). L'objectif de la deuxième était de comparer des stratégies de séquences incluant **les tsDMARDs (tofacitinib et baricitinib) et les bDMARDs (sarilumab et IMRALDI, biosimilaire de l'adalimumab)**. En effet, ces produits n'ont pas été inclus dans l'analyse de référence des méta-analyses en réseau réalisées puisqu'ils ont été évalués par la CT fin 2017 et en 2018. **Le groupe de travail PR (et des membres du groupe de lecture) ont proposé d'explorer l'évaluation économique de ces produits tout en reconnaissant notamment la disponibilité réduite des données concernant les tsDMARDs.**

Les descriptions complètes de ces deux analyses complémentaires qui ont été planifiées après la réalisation de l'analyse de référence sont données dans la section 3 « Analyses complémentaires » du chapitre 6.

2.3. Population d'entrée du modèle

La population d'entrée du modèle inclut les patients adultes atteints d'une PR active (définie par un critère d'activité de la maladie, DAS 28 \geq 5,1) ou modérée à sévère (définie par un score de DAS entre 3,2 et 5,1), qui ont été traités par des csDMARDs (ex. : le MTX) et présentant une réponse inadéquate ou une tolérance insuffisante à la suite de l'administration de ces derniers.

2.4. Horizon temporel et prise en compte de la dimension temporelle

Comme la PR est une maladie chronique, un horizon temporel « vie entière » a été retenu afin d'intégrer l'ensemble des différentiels de coûts et des résultats attendus. Une durée de simulation de 40 ans a été choisie dans l'analyse de référence. L'âge moyen (49 ans) à l'entrée du modèle PR France a été extrait d'une cohorte française « Étude et suivi des polyarthrites indifférenciées récentes » (ESPOIR). La durée de simulation choisie est compatible avec l'espérance de vie en France [16] et donc l'horizon temporel « vie entière » retenu : dans les conditions de mortalité de 2015, l'espérance de vie à la naissance était de 85 ans pour une femme et 78,9 ans pour un homme.

La dimension temporelle a été intégrée dans le modèle PR France *via* sa structure DICE fondée sur la technique *time to event*, permettant de définir un ensemble d'événements d'intérêt (cf. para-

graphe structure DICE). L'occurrence de ces événements est fondée sur des équations et plus particulièrement sur des distributions paramétriques de survie permettant de générer les simulations. Par exemple, l'équation du changement de ligne de traitement faisait appel à une distribution de survie paramétrique (ex. : distribution de gamma généralisée).

2.5. Perspective

Le guide HAS recommande généralement l'adoption d'une perspective de l'ensemble des parties prenantes concernées par les interventions étudiées dans le contexte du système de santé français (assurance maladie, reste à charge, autres payeurs tels que les assurances complémentaires).

Dans le cadre de cette évaluation, la perspective repose autant que possible sur les éléments de la perspective recommandée par la HAS et privilégie la valorisation la plus proche des pratiques réelles. La perspective adoptée dans cette évaluation n'intègre pas les items liés aux éventuels autres payeurs que l'assurance maladie obligatoire AMO (ex. : éventuels restes à charge pour les patients et autres coûts financés par d'autres payeurs). En effet :

- les patients « ciblés » dans cette évaluation appartiennent à la catégorie des patients en ALD (« Affections de longue durée ») ;
- à l'exception des coûts d'acquisition, d'administration des bDMARDs et des coûts d'hospitalisation relatifs aux événements indésirables sévères et/ou graves par classe, les autres coûts sont comparables entre les DMARDs. Ce choix a été validé par les rhumatologues du GT PR.

La perspective de l'assurance maladie est complètement applicable dans ce travail à l'exception de la valorisation des hospitalisations liées aux événements indésirables. Cette dernière a été fondée sur les données du site de l'ATIH dans le secteur public en 2016 à partir de l'étude nationale de coûts (ENC).

2.6. Actualisation

Conformément aux recommandations du guide de la HAS sur les choix d'évaluation économique, l'analyse de référence mobilisait un taux d'actualisation public fixé à 4 % depuis 2005 pour des horizons temporels inférieurs à 30 ans et un taux de 2 % au-delà. Ce taux est applicable pour les coûts, au même titre que pour les résultats de santé.

2.7. Les coûts et leur valorisation

En accord avec la littérature économique sur la prise en charge de la PR et le guide méthodologique de la HAS, les catégories de coûts pris en compte dans la prise en charge de la PR sont les coûts directs : les coûts d'acquisition et d'administration des traitements, les coûts de suivi (consultations médicales, examens biologiques et imagerie) et, selon leur pertinence, les coûts de prise en charge des événements indésirables sévères et/ou graves susceptibles d'entraîner une hospitalisation du patient. Compte tenu de la perspective retenue, l'analyse a privilégié la valorisation la plus proche des pratiques réelles. La valorisation des ressources repose ainsi autant que possible sur des coûts et tarifs relatifs à l'organisme dont on adopte la perspective (l'assurance maladie obligatoire à titre principal).

2.8. Résultats de santé et leur valorisation

Les résultats de santé ont été exprimés en années de vie ajustées sur la qualité de vie (QALY), en temps (les années) de faible activité de la PR et en années de vie gagnées.

- La survie est fondée sur les données de l'Insee portant sur l'espérance de vie en France stratifiées par âge et sexe et leur extrapolation par une distribution statistique (ex. : une distribution de Gompertz). Les années de vie gagnées ne représentent pas un critère pertinent à la différence du QALY ou du temps de faible activité de la PR, car il n'existe pas d'éléments de preuves montrant clairement l'impact de l'utilisation des traitements de fond biologiques sur la survie. Ce choix était en accord avec les résultats de la revue systématique de littérature présentés dans le chapitre n° 4. En revanche, dans l'analyse de référence, quelle que soit la stratégie comparée, la mortalité a été ajustée sur un risque relatif tenant compte d'une éventuelle surmortalité associée à la PR (Van den Hoek *et al.*, 2017), (55). La section 3.2 du chapitre n° 1 explicite les raisons d'ajustement de la mortalité par un surrisque de mortalité associé à la PR qui n'est pas affecté par l'absence d'un traitement par un bDMARD.
- Les scores d'utilités ont été calculés à l'aide de deux questionnaires : EQ-5D-3L et SF-6D, extraits de données françaises (cohorte ESPOIR « Étude et suivi des polyarthrites indifférenciées récentes ») :
 - **dans l'analyse de référence**, les utilités ont été estimées à l'aide de l'EQ-5D-3L portant sur la population d'analyse (extraite à partir de la cohorte ESPOIR). Elles ont été valorisées à l'aide des tarifs français (Chevalier *et al.*, 2013) (156) ;
 - **dans les analyses secondaires**, les utilités ont été calculées à l'aide du SF-6D portant sur la population d'analyse (extraite à partir de la cohorte ESPOIR). Elles ont été valori-

sées à l'aide de l'algorithme de la matrice de pondération britannique fondée sur les travaux de Brazier (Brazier *et al.*, 2002 (157), Brazier *et al.*, 2004 (158)).

2.9. Présentation des résultats

Les interventions de santé constituant la frontière d'efficacité ont été identifiées et un ratio différentiel coût-résultat (RDCR) a été calculé pour chacune d'elles, avec le détail des différentiels de coûts et de résultats de santé.

Le ratio différentiel coût-résultat (RDCR) a été calculé en suivant la procédure suivante :

1. les stratégies sont ordonnées en termes de coûts (de la moins chère à la plus chère) ;
2. les stratégies dominées, au sens de la dominance stricte, sont identifiées : si une stratégie est plus coûteuse mais moins efficace que la précédente, elle est dominée ;
3. les stratégies dominées au sens de la dominance généralisée sont identifiées : les RDCRs sont calculés pour chaque intervention par rapport à l'intervention non strictement dominée qui la précède en termes de coût. Si le RDCR d'une séquence est supérieur à celui de la séquence juste au-dessus d'elle en termes d'efficacité, cette séquence est exclue au titre de la dominance généralisée ;
4. les RDCRs sont alors recalculés avec les stratégies restantes.

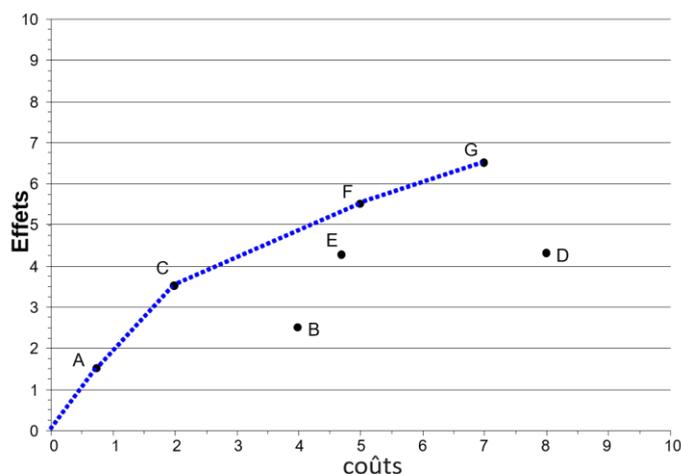
Les coûts et les résultats des stratégies sont synthétisés dans un tableau afin de mettre en évidence toutes les situations dans lesquelles une situation de dominance stricte ou de dominance généralisée existe (**tableau 4**).

Tableau 4. Exemple de tableau synthétisant les résultats de l'analyse de référence

Séquences	Coûts actualisés (€)	QALY actualisé	Coût par QALY gagné (€)
Séquence A	XXXX	YYYY	Référence
Séquence B	XXXX	YYYY	Dominée (Dominance généralisée)
Séquence C	XXXX	YYYY	454 200
Séquence D	XXXX	YYYY	Dominée
Séquence E	XXXX	YYYY	Dominée
Séquence n	XXXX		

Ce tableau est complété par la représentation graphique dans le plan coût-résultat des points correspondant aux différentes stratégies ainsi que la frontière d'efficacité correspondant à l'ensemble des situations non dominées (figure 3).

Figure 3. Représentation graphique de la frontière d'efficacité



Afin de valider une stratégie d'interprétation des résultats de l'analyse de référence utiles pour le décideur et les recommandations de la HAS :

- l'interprétation en termes de frontière d'efficacité a été retenue ;
- l'utilisation de la métrique du bénéfice monétaire a été jugée discutable car les différences entre les stratégies étaient très faibles (cf. chapitre 6).

2.10. Analyses de sensibilité

Les sources d'incertitude identifiées étaient l'incertitude sur les paramètres (entrées) du modèle (incertitude paramétrique) et l'incertitude sur la structure du modèle (incertitude structurelle) dont les sources pourraient être liées aux hypothèses structurelles du modèle.

Compte tenu de la nature de la problématique (comparaisons de séquences de traitements composées par plusieurs lignes de traitements) et du nombre important des séquences à comparer (180), des choix adaptés ont été privilégiés :

- des analyses de scénarios déterministes, susceptibles d'influencer la variabilité du RDCR, ont été choisies. Les diagrammes de Tornado n'ont pas été retenus. Dans notre cas, ils ne permettent pas des interprétations pertinentes : ils multiplient le nombre de comparaisons deux à deux. Le cas échéant, ils auraient pu être utiles si deux stratégies avaient été comparées ;

- une analyse de sensibilité probabiliste portant sur un nombre de séquences raisonnable. En modifiant les choix des lignes de traitements à partir de la séquence de référence, dix séquences représentatives de la prise en charge de la PR ont été identifiées ;
- il a été convenu de réaliser une analyse de sensibilité probabiliste sur les paramètres d'intérêt entachés d'incertitude statistique en considérant, notamment :
 - les valeurs d'efficacité comparative (les probabilités de réponse EULAR) générées par le processus CODA,
 - des décompositions de Cholesky réalisées sur les paramètres d'intérêt qui sont corrélés, et ce, après avoir effectué :
 - des estimations de modèles mixtes à effet aléatoire régressant les scores d'utilité sur l'activité de la PR, le HAQ et leurs interactions,
 - des estimations de modèles linéaires généralisés à effet aléatoire régressant les catégories du HAQ sur l'activité de la PR.

Les résultats complets, permettant de spécifier les distributions statistiques des paramètres d'intérêt, sont donnés dans le chapitre 6.

3. Sources de données mobilisées dans le modèle PR France

3.1. Les sources de données de l'efficacité comparative extraites à partir d'essais cliniques

3.1.1. Données en 1^{re} ligne de traitements biologiques

Les données d'efficacité comparative portaient sur les probabilités de réponses EULAR qui représentaient une **entrée principale du modèle PR France**. Ces données s'appliquaient aux patients présentant une réponse inadéquate ou une tolérance insuffisante à un traitement de fond synthétique (MTX). Ce choix a été également adopté dans la récente évaluation du NICE (12) et s'explique par le fait que la réponse EULAR en tant que critère d'efficacité clinique est plus adaptée/spécifique que la réponse ACR dans le contexte clinique européen et français de la prise en charge de la PR.

Les données d'efficacité comparative relatives à la première ligne de traitement biologique étaient issues des résultats de méta-analyses en réseau réalisées par la HAS & SchARR (*School of Health and Related Research*) de l'Université de Sheffield. Les résultats complets de la méta-analyse en réseau sont présentés dans le chapitre n° 3.

En résumé, les probabilités de réponses EULAR à court terme (i.e. 6 mois) issues des méta-analyses en réseau réalisées par la HAS & Sheffield portaient sur la population d'entrée du modèle PR France. Il s'agissait d'estimations ponctuelles des probabilités de réponses EULAR et des sorties CODA (*Convergence Diagnosis and Output Analysis*).

3.1.2. Données en 2^e ligne et plus de traitements biologiques

Aujourd'hui, le faible nombre d'essais cliniques comparant directement les bDMARDs en 2^e ligne de traitement de fond biologique (ou plus) et la disponibilité réduite des réponses EULAR (ou ACR) dans ces essais semblent affecter considérablement la pertinence (i.e. la qualité de niveau de preuve) des méta-analyses en réseau, soit en termes de puissance statistique, soit en termes de pertinence des réseaux entre les traitements. Par conséquent, des essais cliniques randomisés comparatifs ont été retenus. Le choix de ces essais cliniques a été validé par le groupe de travail PR. Les éléments de ce choix se fondent sur :

1. la qualité méthodologique de ces essais cliniques, mentionnée dans les avis d'évaluation et de réévaluation de ces produits par la Commission de transparence de la HAS (CT) ;
 2. la comparabilité des populations d'analyse de ces essais cliniques à la population française dans la prise en charge de la PR.
- **Données en 2^e et 3^e lignes de traitements biologiques après réponse inadéquate ou tolérance insuffisante à un traitement de fond biologique**

Ces données ont été extraites d'un essai clinique français appelé « *Rotation or Change* » (ROC) (Gottenberg *et al.*, 2016) (159). ROC est un essai pragmatique, randomisé, multicentrique (47 centres français) et comparatif. Il compare des anti-TNF α et des non-anti-TNF α en 2^e ligne et porte sur des patients atteints de PR et présentant une réponse insuffisante à un anti-TNF α .

L'essai ROC a permis d'extraire des estimations de perte d'efficacité et de risques relatifs associés aux probabilités de réponses EULAR pour des patients atteints de PR et présentant une réponse inadéquate ou une tolérance insuffisante à un bDMARD (ex. : un anti-TNF α).

Les données utilisées sont synthétisées dans le **tableau 5**.

Tableau 5. Efficacité comparative en 2^e ligne des bDMARDs – Source : Gottenberg *et al.*, 2016 (159)

Anti-TNF	N	(%)	Réponse EULAR bonne	Réponse EULAR modérée
Adalimumab	85	29,2	0,3	0,7
Étanercept	157	54,0	0,56	0,89
Infliximab	41	14,1	0,28	0,69
Certolizumab	5	1,7	0,38	0,77
Golimumab	3	1	0,39	0,78
	291		44,0 %	80,3 %
Moyenne			21 %	52 %
Anti-TNF en 2^e ligne biologique				
Perte			23,0 %	28,3 %
Risque relatif			0,478	0,647
Non-anti-TNF				
Abatacept	33	22,9	28,0 %	69,0 %
Rituximab	41	28,5	15,0 %	50,0 %
Tocilizumab	70	48,6	63,0 %	92,0 %
	144		41,3 %	74,8 %
Moyenne			39,0 %	69,0 %
Non-anti-TNF en 2^e ligne biologique				
Perte			5,0 %	11,3 %
Risque relatif			0,944	0,923

➤ **Des estimations de probabilités de réponses EULAR issues d'un essai clinique randomisé comparatif portant sur l'efficacité du rituximab**

Dans le modèle PR France, les séquences de traitements sélectionnées intégraient le rituximab après une réponse inadéquate ou une tolérance insuffisante à un 3^e bDMARD. Les estimations de la réponse EULAR provenaient d'un essai comparatif randomisé « *Randomized Evaluation of Long-term Efficacy of Rituximab* » (REFLEX) (Cohen *et al.*, 2006) (160). Cet essai, dont la durée est égale à 2 ans, a comparé l'association rituximab+MTX à celle MTX+placebo.

Les estimations des réponses EULAR issues de REFLEX étaient assez similaires à celles d'un registre français (AIR-PR) portant sur des patients atteints de PR et traités en pratique courante par rituximab. Les estimations de la réponse EULAR issues du registre AIR-PR n'ont pas été retenues. En effet, il a été mentionné dans l'avis CT du 22 mars 2017 (rituximab) : « *concernant les données d'efficacité, elles sont à interpréter avec prudence en raison d'un risque de biais inhérent à la méthodologie de l'étude et notamment du fait de nombreuses données manquantes. Les données d'efficacité ne portent que sur 55 % de patients de la cohorte* ».

Les données utilisées sont présentées dans le tableau 6.

Tableau 6. Réponses EULAR relatives au rituximab – Source : Cohen *et al.* (2006) (160)

Groupes à comparer	N	%	Réponse EULAR au moins modérée	Réponse EULAR bonne
Rituximab+MTX	298	60 %	65 % (149/298)	15 % (545/298)
MTX+Placebo	201	40 %	22 % (40/201)	2 % (10/201)

3.2. Les données en vie réelle

En plus des données d'efficacité fondées sur la méta-analyse en réseau HAS & SchARR, le modèle PR France repose sur **des données françaises en vie réelle** portant sur des patients atteints de PR. Ces données ont été utilisées dans l'extraction **des profils des patients** conformément aux caractéristiques de la population d'analyse, aux **estimations des données d'utilité**, aux paramètres **des équations du modèle** (ex. : **changement de ligne de traitement**) et **certaines données décrivant les coûts de suivi de la PR** (ex. : la moyenne des consultations des médecins généralistes et des rhumatologues).

3.2.1. Sources des données françaises en vie réelle portant sur les patients atteints de la PR

Les registres et cohortes français portant sur des patients atteints de PR en vie réelle ont été identifiés à partir :

1. des descriptions des registres issues des avis CT sur l'évaluation et la réévaluation des traitements de fond biologiques et si nécessaire des compléments d'informations fournis par les industriels au Service d'évaluation des médicaments (SEM) de la HAS sur les études post-inscriptions ;
2. des informations mises sur les registres issues du site de la SFR (www.larhumatologie.fr) ;
3. des publications proposant des descriptions globales des données françaises en vie réelle (efficacité et tolérance) telles que :

- l'étude de Mariette *et al.*, 2011 ([161](#)) décrivant trois registres français utilisés dans l'analyse de long terme de la tolérance et de l'efficacité des bDMARDs,
 - la méta-analyse récente sur les données de tolérance des bDMARDs issues des registres portant sur la PR (Divonne *et al.*, 2017) ([162](#)) ;
4. des informations issues de la base de données de la cohorte ESPOIR dans le cadre de l'accord entre la SFR et la HAS sur l'utilisation de la base de données ESPOIR dans le projet de la HAS portant sur l'évaluation des traitements de fond biologiques dans la prise en charge de la PR.

3.2.2. Registres et cohorte français identifiés

Une cohorte et quatre registres français portant sur des patients atteints de PR ont été identifiés. Il s'agissait de la cohorte ESPOIR (Étude et suivi des polyarthrites indifférenciées récentes) et des registres : ORA (*Orencia and Rheumatoid Arthritis*) ; RATIO (Recherche sur les anti-TNF et infections opportunistes) ; AIR-PR (*AutoImmunity and Rituximab*) et REGATE (*REGistry-RoAcTEmra*).

AIR-PR, ORA et REGATE portaient respectivement sur la tolérance et l'efficacité en vie réelle du rituximab, de l'abatacept et du tocilizumab, alors que le registre RATIO rapportait uniquement la tolérance de certains anti-TNF (ex. : certains types d'événements indésirables associés à l'utilisation des anti-TNF tels que les infections et les lymphomes). En revanche, ESPOIR représente une cohorte multicentrique d'arthrites indifférenciées débutantes évoluant vers une PR et servant de base de données à des études diagnostiques (i.e. diagnostic de PR afin de permettre un traitement précoce), pronostiques (identifier les patients qui vont développer des PR sévères), médico-économiques (i.e. évaluer les coûts de la PR) et pathogéniques (i.e. constitution d'une biothèque).

Les caractéristiques globales de ces registres sont résumées dans le tableau 7.

Tableau 7. Caractéristiques des registres français sur les patients atteints de PR.

Registre	Critères d'inclusion	csDMARDs /bDMARDs	Population	Nombre de patients inclus	Comparateur	Période de suivi prévue	Année d'inclusion
ESPOIR	Patients de 18 à 70 ans, au moins deux arthrites depuis au moins 6 semaines, depuis moins de 6 mois, pas de traitement de fond ou corticoïdes au cours des 6 derniers mois*	csDMARDs (ex. : MTX, SUL) et différentes classes de bDMARDs (ex. : anti-TNF α)	Arthrites débutantes évoluant vers une PR	813	Pas de comparateur	Tous les 6 mois pendant 2 ans puis tous les ans pendant 20 ans	2002
ORA	Patients traités avec abatacept (hors essais cliniques) après échec ou contre-indication d'un anti-TNF	Abatacept	PR et autres maladies auto-immunes	1017 (*)	Pas de comparateur	7 années	2008
RATIO	Les cas d'infections opportunistes, les infections bactériennes graves et les lymphomes survenant chez les patients traités ou ayant été traités par des anti-TNF α	Anti-TNF α	PR et autres maladies auto-immunes	(Étude de cas)	Population générale	3 années	2004
AIR	Tout patient porteur d'une PR chez qui on débute un traitement par rituximab	Rituximab	PR	1709 (*)	Pas de comparateur	Suivi 2 fois par an pendant 5 ans	2009
REGATE	Patients traités avec tocilizumab en pratique réelle	Tocilizumab	PR	1500 (*)	Pas de comparateur	5 années	2010

(*) : Source : Divonne *et al.* (2017) (163).

3.2.3. Pourquoi la cohorte ESPOIR a-t-elle été sélectionnée ?

Le choix de l'utilisation des données en vie réelle a porté sur la cohorte ESPOIR pour les raisons suivantes :

- comparativement aux quatre autres registres décrits ci-dessus, ESPOIR se distingue par un suivi longitudinal suffisamment long : le suivi des patients s'effectue tous les 6 mois pendant 2 ans (2002-2007), puis tous les ans pendant 20 ans. Le suivi des patients est décrit dans l'annexe 5.2. Cette dimension temporelle permet de collecter les dates et les raisons d'un arrêt de traitement nécessaires à la définition des événements de changement de ligne de traitement liés à une insuffisance d'efficacité ou de tolérance et permet *in fine* d'estimer les paramètres des équations de survie « *time to event* » sur lesquelles se fonde le modèle PR France ;
- elle collige des informations sur des csDMARDs au même titre que les bDMARDs sans se focaliser sur la description d'un DMARD particulier comme c'était le cas des registres ORENCIA, AIR-PR et REGATE. L'utilisation des données de la cohorte ESPOIR a permis la constitution de lignes de traitements synthétiques et/ou biologiques. Cette dimension a été très utile pour spécifier les équations de changement de traitement en fonction d'une réponse inadéquate ou d'une tolérance insuffisante ;
- elle permet de colliger les caractéristiques (profil des patients) qui sont compatibles avec l'indication de la population d'analyse du modèle PR France (cf. tableau 1) ;
- contrairement aux trois registres mentionnés plus haut, elle permet à la fois de colliger des critères décrivant :
 - l'activité de la PR : le DAS 28,
 - l'incapacité fonctionnelle : le HAQ,
 - la qualité de vie : EQ-5D et SF-6D.

Les instruments décrivant la qualité de vie tels que l'EQ-5D et le SF-6D ont permis de calculer les scores d'utilité selon les tarifs français pour l'EQ-5D (Chevalier *et al.*, 2013) et les tarifs anglais pour le SF-6D (Brazier *et al.*, 2004) ;

- environ 90 publications scientifiques ont porté sur divers aspects de la cohorte ESPOIR, y compris les aspects économiques tels que les coûts (ex. : Chevreul *et al.*, 2014) et les comparaisons des utilités dérivées de l'EQ-5D et du SF-6D (ex. : Gaujoux-Viala *et al.*, 2012) ;
- au-delà de ce projet et dans les années à venir, la perspective d'appariement des données de la cohorte ESPOIR avec les données du SNIIRAM est prometteuse. Elle permettrait :

- d'approfondir l'analyse médico-économique en vie réelle de la prise en charge des patients atteints de PR et de réaliser une validation du modèle PR France avec des données « non indépendantes » permettant d'évaluer les prédictions sur les coûts,
- d'intégrer d'autres stratégies de traitement,
- d'évaluer l'incidence des comorbidités et les coûts de prise en charge des patients atteints de PR en utilisant le DCIR (Datamart consommation inter-régimes).

3.2.4. La cohorte ESPOIR

La disponibilité d'une cohorte française sur la PR (cohorte ESPOIR) a permis d'extraire un échantillon de patients dont le profil était compatible avec la population d'analyse. La population d'analyse a été décrite dans les choix structurants de l'évaluation économique. L'algorithme d'extraction de cette population est décrit dans un protocole dédié à l'analyse de cette cohorte pour compléter les événements et les conditions du modèle PR France (« protocole HAS d'analyse de la cohorte ESPOIR », cf. annexe 5-3).

➤ **Profil des patients issus de la cohorte ESPOIR**

La cohorte a permis d'extraire un échantillon de patients atteints de PR et respectant les critères de la population d'analyse. Les variables constituant cet échantillon sont : l'âge, le sexe, la valeur du DAS 28 et la catégorie de l'activité correspondante, la valeur du HAQ et la catégorie correspondante, le poids, les utilités calculées à l'aide de l'EQ-5D ou le SF-6D (cf. tableau 1).

➤ **Les événements associés aux changements de traitement**

Les analyses portent sur les événements de changement (de lignes) de traitement de fond en raison d'une insuffisance d'efficacité ou de tolérance conformément à la population d'analyse et à la structure du modèle PR France. Les événements relatifs aux changements de traitement concernent particulièrement le temps jusqu'au deuxième changement de ligne de traitement. Le protocole HAS d'analyse de la cohorte ESPOIR présente la méthodologie de ces analyses (cf. annexe 5-3).

➤ **L'érosion définie en termes de modification radiologique selon les critères de l'ACR 1987**

Afin de tenir compte de l'évolution de l'érosion d'une manière qualitative, la variable binaire « modifications radiologiques » fondée sur les critères de l'ACR 1987 a été retenue. L'analyse du développement d'une érosion est décrite dans le protocole HAS d'analyse de la cohorte ESPOIR (cf. annexe 5-2).

➤ **DAS 28 et HAQ exprimés en termes de catégories**

Les données actuelles de la cohorte ESPOIR permettaient de fournir un suivi des patients sur 10 ans. La variable DAS 28 a été colligée à chaque visite de suivi de la cohorte ESPOIR. Conformément à la structure du modèle PR France, le DAS 28 et le HAQ sont analysés sous formes catégorielle et continue.

- Pour le **DAS 28**, les catégories suivantes ont été définies :
 - catégorie 1 : $DAS\ 28 < 2,6$ (rémission) ;
 - catégorie 2 : $2,6 \leq DAS\ 28 < 3,2$ (faible activité) ;
 - catégorie 3 : $3,2 \leq DAS\ 28 < 5,1$ (activité modérée) ;
 - catégorie 4 : $DAS\ 28 \geq 5,1$ (activité sévère).
- Pour le **HAQ**, les catégories suivantes ont été définies :
 - catégorie 1 : $HAQ < 0,6$;
 - catégorie 2 : $0,6 \leq HAQ < 1,1$;
 - catégorie 3 : $1,1 \leq HAQ < 1,6$;
 - catégorie 4 : $1,6 \leq HAQ < 2,1$;
 - catégorie 5 : $HAQ > 2,1$.

Un score HAQ global de « 0 » signifie l'absence d'incapacité, alors qu'un score à « 3 » correspond à une incapacité maximale (cf. annexe A1.7).

➤ Les données d'utilité : EQ-5D, SF-6D

Les questionnaires EQ-5D et SF-6D ont été colligés dans la cohorte ESPOIR. Le calcul des utilités pour la population d'analyse du modèle PR France est décrit dans le protocole HAS d'analyse de la cohorte ESPOIR (cf. annexe 1). L'algorithme de la valorisation des utilités du SF-6D a été programmé sous SAS® (version 9.4) par la HAS et validé par Nathalie Rincheval (INSERM, Montpellier). Cet algorithme a été fondé sur les travaux de Brazier *et al.* (2002 ; 2004) ([157](#), [158](#)) et le support méthodologique, fourni par l'équipe SchARR de l'Université de Sheffield.

3.3. Les données portant sur la tolérance

3.3.1. Types de données de tolérance utilisées dans le modèle PR France

Les raisons de changement de traitement ont été collectées à partir de la cohorte ESPOIR et utilisées dans l'équation du changement de traitement *via* l'identification de l'événement changement de traitement en raison d'une réponse inadéquate ou d'une tolérance insuffisante (cf. « protocole HAS d'analyse de la cohorte ESPOIR », annexe 5-3).

Quant aux données décrivant les événements indésirables susceptibles d'entraîner des hospitalisations, elles ont fait l'objet d'un protocole d'étude nommé « Profil des événements indésirables et valorisation de leurs coûts » (cf. annexe 5.4).

Ce protocole a été réalisé au sein du SEESP pour analyser les profils et les coûts des événements indésirables relatifs à l'administration des traitements de fond qui étaient susceptibles d'entraîner des hospitalisations dans la prise en charge de la PR en France.

3.3.2. Sources des données décrivant les événements indésirables

Le protocole « Profil des événements indésirables et valorisation de leurs coûts » se fondait sur les sources de données suivantes :

1. les avis de la Commission de la transparence (CT) de la HAS ;
2. les rapports EPAR (*European Public Assessment Reports*) de l'Agence européenne des médicaments (EMA) et les résumés des caractéristiques des produits (RCP) ;
3. les archives des dossiers médico-techniques déposés par les industriels auprès de la HAS depuis l'apparition des bDMARDs (début des années 2000) ;
4. la base de données européenne du système de gestion et d'analyse des informations sur les effets indésirables de l'EMA (EudraVigilance) ;
5. la littérature récente portant sur la tolérance des traitements de fond :
 - a. des méta-analyses récentes et des articles scientifiques portant sur la tolérance des traitements de fond,
 - b. des rapports d'évaluation de technologies de santé.

L'annexe 5-4 concernant le protocole « Profil des événements indésirables et valorisation de leurs coûts » présente les détails sur les sources utilisées dans l'analyse des profils des événements indésirables selon la classe thérapeutique des traitements de fond synthétiques.

3.4. Les données portant sur la mortalité

Les données utilisées dans cette évaluation économique ont reposé sur :

- les statistiques de l'Insee portant sur la mortalité en population générale française stratifiée par âge et par sexe (Source : tableau 68 – table de mortalité des années 2013-2015, données arrêtées à fin décembre 2015 (https://www.insee.fr/fr/statistiques/fichier/2045139/irsocsd2014_t68_f.xls) ;
- le risque de surmortalité associé à la PR a été extrait d'une étude de mortalité (Van den Hoek *et al.*, 2017) (55) réalisée sur une cohorte hollandaise longitudinale (1997-2012). La mortalité observée chez les patients atteints de PR était supérieure de 54 % à celle obser-

vée dans la population générale après ajustement sur l'âge, le sexe et l'année civile. La méthodologie et les résultats de cette étude étaient en accord avec la revue systématique et la méta-analyse (Dadoun *et al.*, 2013) (51) effectuée sur des patients atteints de PR et réalisée entre 1966 et 2010.

3.5. Les estimations des désutilités associées aux événements indésirables

Les données de la cohorte ESPOIR n'ont pas permis d'estimer des désutilités en raison du nombre très faible des événements indésirables collectés et également en raison du fait que ces derniers n'ont pas été nécessairement associés à la PR et à ses traitements. De même, les autres registres français ORA, RATIO et REGATE colligeaient des événements indésirables portant seulement sur les bDMARDs étudiés, et leurs *designs* et objectifs ne proposaient pas d'estimer des pertes d'utilité associées aux événements indésirables spécifiques aux bDMARDs analysés.

À défaut de données de meilleure qualité, des données de désutilité (perte de qualité de vie) associées aux infections graves et faisant consensus au sein des groupes de travail des trois organismes d'évaluation de technologies de santé (américain, britannique et français) ont été retenues. Ainsi, des estimations de désutilités sur les infections graves, issues des rapports récents des évaluations des technologies de santé du NICE (12) et de l'ICER (13), ont été extraites. Ces estimations ont été fondées sur des hypothèses validées par les comités scientifiques de ces deux institutions. Ensuite, les désutilités retenues, telles que celles relatives aux infections graves, ont été appliquées sur une période de 30 jours. En l'absence d'une littérature portant sur les épisodes et sur **l'occurrence des événements sévères et/ou graves entraînant des hospitalisations**, le choix de 30 jours a été appliqué dans la littérature dans le cas des événements aigus tels que les infections graves.

D'autres estimations de désutilités correspondant à d'autres événements chroniques (ex. : cancers) ont été extraites de la littérature (cf. tableau 24). Ceci illustre une nouvelle fois que la prise en compte des événements indésirables chroniques dans l'évaluation économique est loin d'être conceptuellement et méthodologiquement résolue et constitue une des recherches méthodologiques dans l'évaluation économique des technologies de santé (Heather *et al.*, 2014) (139).

3.6. L'estimation des coûts

D'une manière générale, les mêmes types de sources sur les coûts et les ressources identifiées dans le chapitre 4 ont été identifiés dans le contexte français. Les coûts ont été actualisés aux prix de 2017.

3.6.1. Les coûts d'acquisition des traitements de fond synthétiques et biologiques

Les coûts d'acquisition des traitements ont été calculés à partir des prix publics des produits toutes taxes comprises (PPTTC) issus des sources de données suivantes :

- la base de données nationale des médicaments dans sa version de février 2018 (<http://base-donnees-publique.medicaments.gouv.fr/>) ;
- la banque de données sur les médicaments Thériaque (<http://www.theriaque.org/apps/contenu/accueil.php>) ;
- les doses et durées de traitement ont été collectées à partir des résumés des caractéristiques des produits (RCP). Ce choix est compatible avec l'esprit du modèle PR France qui propose une modélisation selon les bonnes pratiques de prise en charge fondées sur les recommandations européennes et françaises. Les réductions des doses et durées de traitement en vie réelle (ex. : données extraites à partir du Système national d'information inter-régimes de l'assurance maladie (SNIIRAM) ou de l'Échantillon généraliste des bénéficiaires (EGB)) ont été également envisagées pour tenir compte des stratégies ciblées (ex. : les stratégies *treat to target*, T2T), mais l'absence de données dans ces bases décrivant les critères cliniques, tels que les variables renseignant sur l'activité de la maladie (ex. : DAS 28) à 3 mois, n'a pas permis de modéliser ce type de stratégies. Si des données relatives à ces stratégies sont disponibles, elles pourront être incluses dans une version future du modèle PR France en mettant à jour les équations portant sur les doses, les visites et le changement de traitement.

3.6.2. Les coûts d'administration

La valorisation des coûts d'administration des traitements a été réalisée selon les tarifs nationaux.

- *Les coûts d'administration en ville*

Les coûts d'administration infirmiers ont été extraits de l'étude de Chevreul *et al.* (2014) et validés à partir de l'analyse des données de la cohorte française ESPOIR.

- *Les coûts d'administration à l'hôpital*

Les coûts d'administration à l'hôpital ont été issus des données de l'ATIH (Agence technique de l'information sur l'hospitalisation) sur les tarifs des GHS en date de mars 2017 (dernière version en vigueur au 7 juillet 2017).

3.6.3. Les coûts de suivi : consultations, examens biologiques

- Les coûts des différents examens biologiques ont été extraits de la Table nationale de codage de biologie (<http://www.codage.ext.cnamts.fr/>) avec la valeur de référence unitaire de chaque examen (le B) actualisée en 2017.
- Les coûts des examens d'imagerie ont été extraits de l'étude de Chevreul *et al.* (2014) ([147](#)) et fondés sur les données économiques de la cohorte ESPOIR. Tous ces examens ont été regroupés sous forme de forfaits, appliqués selon des dates de suivi de la PR.

3.6.4. Coûts des événements indésirables

Une fois les profils des événements indésirables sélectionnés et leur fréquence par classe thérapeutique synthétisée, les sources pour valoriser le coût moyen des hospitalisations imputables aux EI ont été extraites des données du site de l'ATIH dans le secteur public en 2016 à partir de l'étude nationale de coût (ENC).

Le protocole « Profil des événements indésirables et valorisation de leurs coûts » (annexe 5.2) présente les détails du calcul des coûts des événements indésirables susceptibles d'entraîner une hospitalisation (cf. annexe 5.4). Des analyses de sensibilité fondées sur les coûts d'hospitalisation estimés dans des études françaises récentes sur la cohorte ESPOIR (Chevreul *et al.*, 2014) ([147](#)) et sur l'Échantillon généraliste des bénéficiaires « EGB » (Fautrel *et al.*, 2016) ([8](#)) ont été discutées mais n'ont finalement pas été retenues car, d'une part, ces sources de données ne différenciaient pas les différents bDMARDs analysés et, d'autre part, il n'était pas clairement indiqué dans la méthode de description de ces coûts d'hospitalisation si les coûts d'administration des bDMARDs étaient exclus des calculs.

4. Les équations du modèle PR France

L'exécution du modèle PR France a été réalisée à l'aide d'un ensemble d'équations. Ces dernières ont été incorporées dans une table d'équations et identifiées avec un nom unique. Ceci facilite la vérification des équations et permet également d'intégrer des mises à jour (par exemple, ajouter d'autres équations dans les analyses de sensibilité structurelles).

Pour implémenter les équations « *time to event* » (tableau 8), une fonction de survie intégrée a été utilisée : *Sltte*. Cette fonction sélectionne un temps pour un événement selon la distribution paramétrique associée à cet événement en appelant ses paramètres génériques (ex. paramètre lambda pour une distribution gamma généralisée). Elle s'écrit de la manière suivante :

Sltte (Distribution, IP, HR, Temps, Numéro aléatoire, Gamma, Lambda).

Tableau 8. Paramètres et métriques des distributions “*time to event*” – Source : Équations programmées dans les modules du modèle PR France

Distribution	Metric	Survivor function	Parameterization	Ancillary parameters
Exponential	PH	$\exp(-\lambda_j t_j)$	$\lambda_j = \exp(\mathbf{x}_j \beta)$	
Exponential	AFT	$\exp(-\lambda_j t_j)$	$\lambda_j = \exp(-\mathbf{x}_j \beta)$	
Weibull	PH	$\exp(-\lambda_j t_j^p)$	$\lambda_j = \exp(\mathbf{x}_j \beta)$	p
Weibull	AFT	$\exp(-\lambda_j t_j^p)$	$\lambda_j = \exp(-p \mathbf{x}_j \beta)$	p
Gompertz	PH	$\exp\{-\lambda_j \gamma^{-1} (e^{\gamma t_j} - 1)\}$	$\lambda_j = \exp(\mathbf{x}_j \beta)$	γ
Lognormal	AFT	$1 - \Phi\left\{\frac{\log(t_j) - \mu_j}{\sigma}\right\}$	$\mu_j = \mathbf{x}_j \beta$	σ
Loglogistic	AFT	$\{1 + (\lambda_j t_j)^{1/\gamma}\}^{-1}$	$\lambda_j = \exp(-\mathbf{x}_j \beta)$	γ
Generalized gamma				
if $\kappa > 0$	AFT	$1 - I(\gamma, u)$	$\mu_j = \mathbf{x}_j \beta$	σ, κ
if $\kappa = 0$	AFT	$1 - \Phi(z)$	$\mu_j = \mathbf{x}_j \beta$	σ, κ
if $\kappa < 0$	AFT	$I(\gamma, u)$	$\mu_j = \mathbf{x}_j \beta$	σ, κ

L'exécution du modèle PR France a été réalisée à l'aide d'un ensemble d'équations (**Tableau 9**).

Tableau 9. Équations incluses dans le modèle PR France

Nom	Équation
Activité	SI(DAS>5.1, 4, SI(DAS>3.2, 3, SI(DAS>2.6, 2, 1)))
ChgActivité	SI(RandChgActivité<pDiminuer, MAX(1,Activité-1),SI(RandChgActivité<pAugmenter, Min(4,Activité+1), Activité))
Doses	INT(MAX((TempsLigne-PeriodeCharge)/MAX(0.00001,Periode),0))
ChgTmt	SI(OU(Activité=4,ET(Activité=3, OU(ChgDAS<1.2, ET(ChgDAS>=1.2, RandEchec<iPEchec))),ET(Activité=2, "Réponse"="Aucune", Ligne=1),FAUX,FAUX),"Oui", "Non")
HAQ	MATCH(RandHAQ, CHOISIR(Activité,rgRemission,rgFaible,rgMod,rgSévère) ,1)
LimiteInf	SI("Réponse"="Aucune", 0, SI("Réponse"="Mod", SI(DAS<=6.3, MAX(0.7, MIN(DAS, DAS - 5.1)), 1.3), SI(DAS<=4.5, 1.3, DAS-3.2)))
RelEffpG	SI(ET(Ligne>1,"Classe"<>"cDMARD"), SI("Classe"="PrevClass", syEffLossPGAntiTNF, syEffLossPGOther),1)
RelEffpM	SI(ET(Ligne>1,"Classe"<>"cDMARD"), SI("Classe"="PrevClass", syEffLossPMAntiTNF, syEffLossPMOther),1)
ChoisirTmt	SI(Ligne>MaxLigne, EndMed, INDEX(tblSéquences, Séquence, SI(Ligne>1, Ligne + SeqSkip,Ligne)))
Survie	Sltte("Gomp", MortAdj,HRHAQ, AgeAct, RandSurv, GSurv, LN(LSurv))-AgeAct
Switch	DaysToYr*SLtte("GG",1,1,0,RandSwitch,SwitchQ,SwitchSigma,SwitchMu+SwitchReducer)
SwitchReduction	SI(ET("Classe"="cDMARD", UtiliserbDMARD=2), 0, SwitchModbD + SI(Ligne=2, SwitchModL2, SI(Ligne=3, SwitchModL3, 0)))
LimiteSup	SI("Réponse"="Aucune", SI(DAS>5.8, MIN(1.2,DAS-5.2),0.6), SI("Réponse"="Mod", SI(DAS<4.5, MIN(1.2,DAS), DAS-3.3), DAS))
Utilité	INDEX(CHOISIR(iUseEQ5D, rgUtilEQ5D, rgUtilSF6d), HAQ, Activité)
VisiteCoûtMG	SI(OU(visitNum >= iVisitGPAfter, SIERROU(MATCH(visitNum, iVisitGP, 0), 0) > 0), syCostGP,0)
VisiteCoûtLab	SI(OU(visitNum >= iVisitLabAfter, SIERROU(MATCH(visitNum, iVisitLab, 0),0) > 0), syCostLab + (visitNum=1)*syCostLabInit, 0)
VisiteCoûtRad	SI(OU(visitNum >= iVisitRadAfter, SIERROU(MATCH(visitNum, iVisitRad, 0), 0) > 0), syCostRad, 0)
VisiteCoûtRH	SI(OU(visitNum >= iVisitRHAfter, SIERROU(MATCH(visitNum, iVisitRH, 0), 0) > 0), syCostRH, 0)
VisiteSchedule	SI(LineTime < 0.5, timeYearsInMonth, SI(LineTime < 24, 3*timeYearsInMonth, 1))

4.1. La survie

La survie a été estimée pour chaque patient *via* un échantillonnage aléatoire de la distribution du temps de décès générée à partir de la table de mortalité française de 2015 issue des statistiques de l'Insee. Les distributions pour les hommes et les femmes ont été obtenues en ajustant plusieurs fonctions de survie paramétriques. La distribution Gompertz donnait le meilleur ajustement (figures 4 et 5). Un échantillonnage a été réalisé en utilisant cette distribution en fonction de l'âge du patient au début de la simulation. La survie a été ajustée par un facteur de surmortalité de la PR extrait de l'étude de Van den Hoek *et al.* (2017) (55).

Figure 4. Survie stratifiée par sexe obtenue à l'aide de la table de l'espérance de vie en France

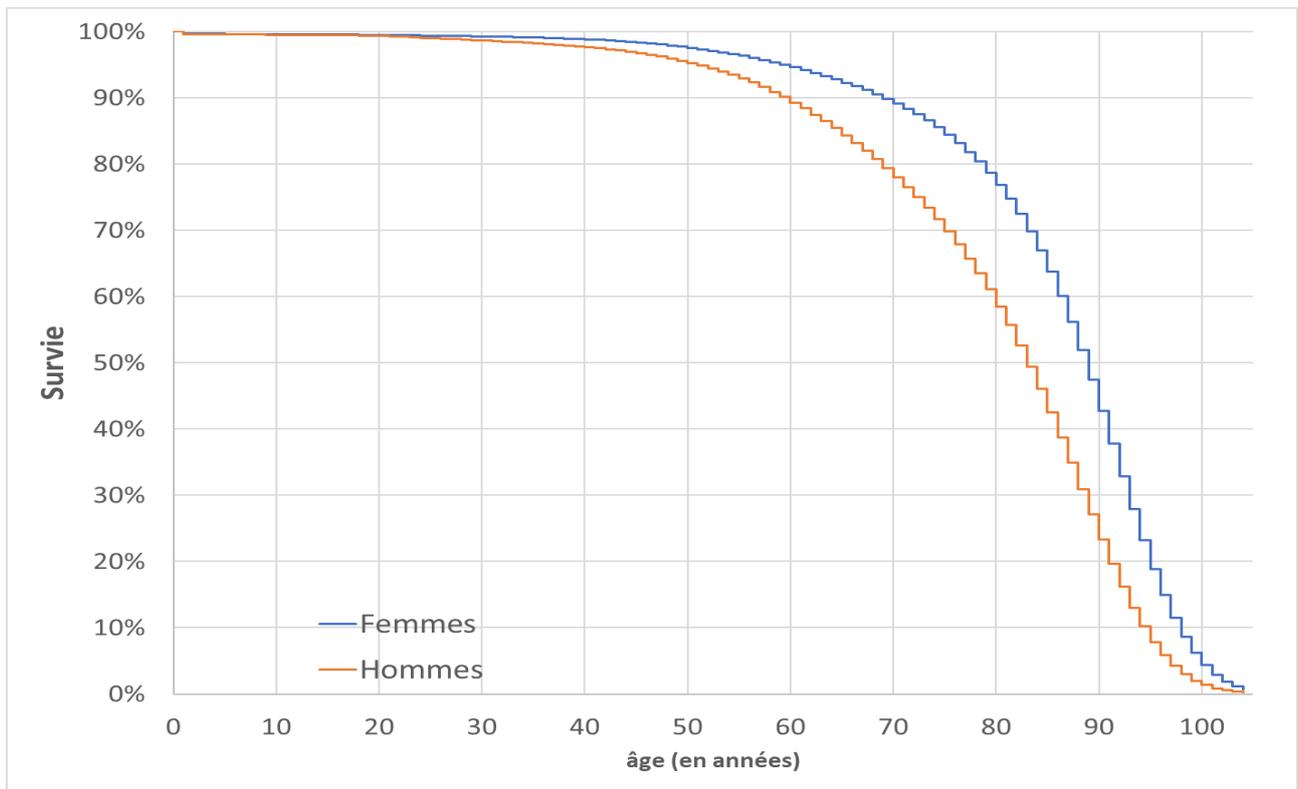


Tableau 10. Estimations des paramètres des distributions de mortalité fondées sur la table de mortalité française – Données Insee (2015)

Analyse	Constante	Écart-type	Paramètre d'échelle	Écart-type	Paramètre de forme	Écart-type	Critère d'information d'Akaike AICC	Critère d'information de Bayes BIC
Femmes								
Weibull	4,5038	0,0004	0,1146	0,0003	8,7223	0,0236	787229,931	787248,957
Log-normale	4,4256	0,0010	0,3226	0,0007			936490,173	936509,199
Log-logistique	4,4659	0,0005	0,0930	0,0003			824677,113	824696,138
Exponentielle	4,4581	0,0032	1,0000	0,0000	1,0000	0,0000	1083560,486	1083569,999
Gamma généralisée	4,5496	0,0005	0,0791	0,0003	2,1740	0,0132	767541,970	767570,508
Gompertz	12,5361	0,0254			0,1137	0,0003	756741,127	756760,153
Hommes								
Weibull	4,4417	0,0005	0,1490	0,0004	6,7099	0,0180	822701,038	822720,064
Log-normale	4,3405	0,0011	0,3632	0,0008			947385,153	947404,179
Log-logistique	4,3889	0,0006	0,1206	0,0003			858955,773	858974,799
Exponentielle	4,3784	0,0032	1,0000	0,0000	1,0000	0,0000	1073239,875	1073249,388
Gamma généralisée	4,5053	0,0006	0,1006	0,0004	2,2339	0,0125	803554,230	803582,769
Gompertz	10,3786	0,0193			0,0930	0,0002	795748,637	795767,663

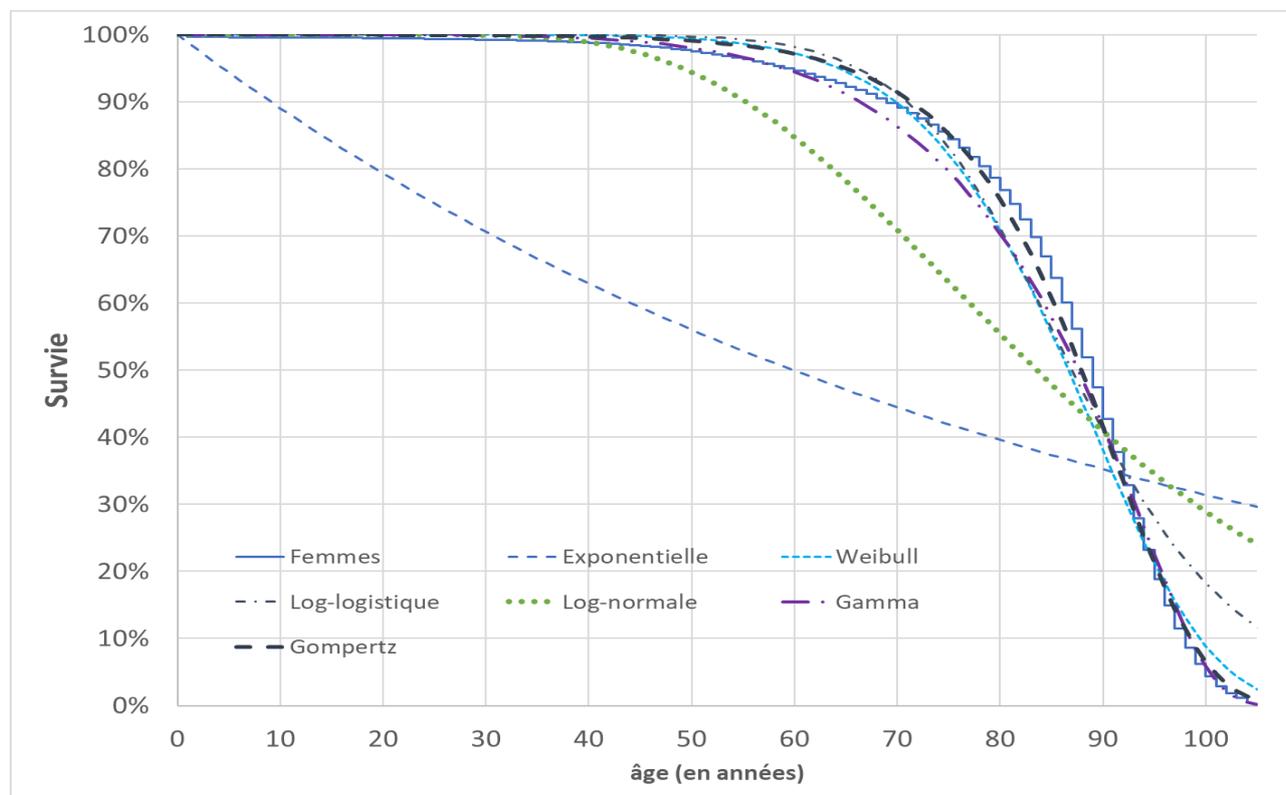
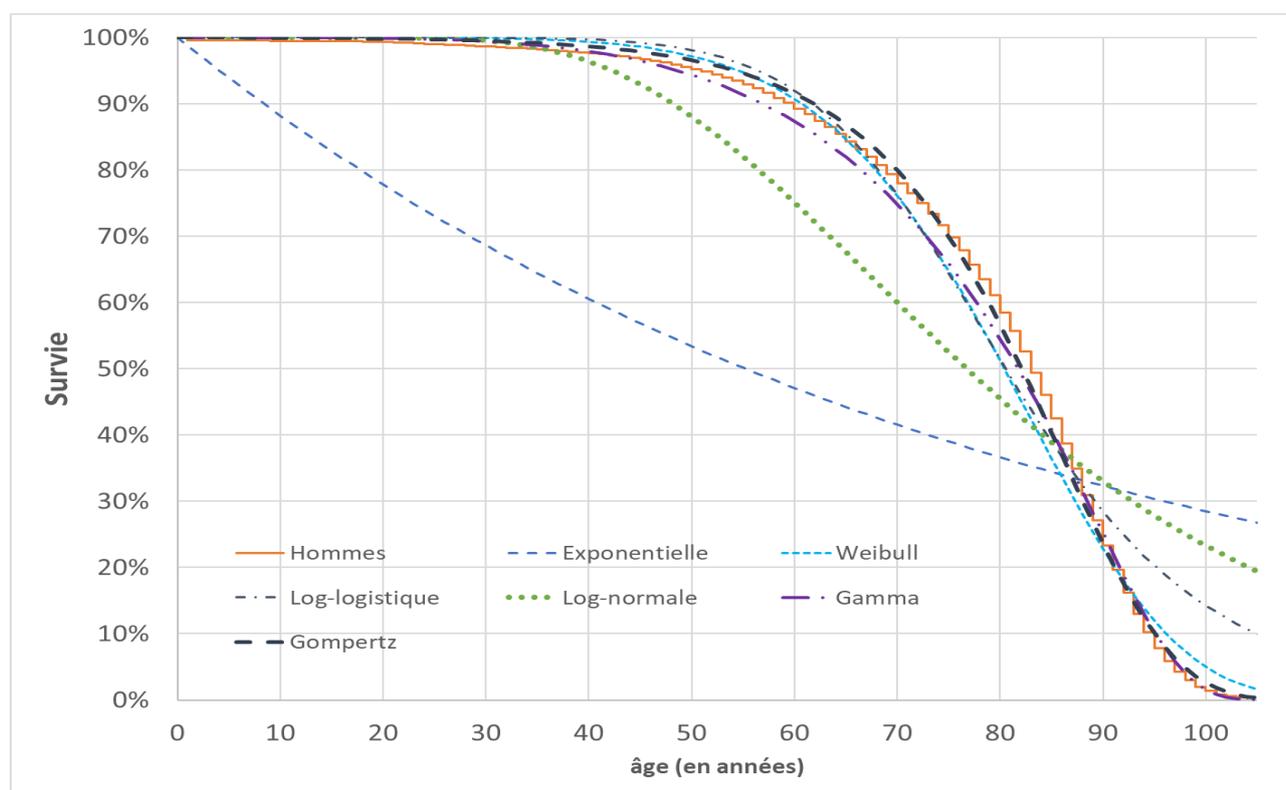
Figure 5. Distribution observée vs ajustée pour les femmes

Figure 6. Distribution observée vs ajustée pour les hommes

4.2. Temps pour développer une érosion

Le temps pour développer une érosion chez les patients (ne présentant pas d'érosion articulaire lors de l'initiation du modèle) a été estimé pour chaque patient selon un échantillonnage aléatoire de la distribution des nouveaux temps d'érosion obtenus à l'aide des analyses de la cohorte ESPOIR. L'équation du temps pour développer une érosion n'a pas été activée dans la version actuelle du modèle parce que les estimations relatives au développement d'une érosion dans la cohorte ESPOIR présentaient une tendance à être associées à un changement de traitement. En effet, ceci risquait d'entraîner un scénario de double comptage vis-à-vis du changement de traitement « *switch* ».

4.3. Les coûts relatifs aux consultations (rhumatologue, médecin généraliste), examens biologiques et imagerie

Le calendrier des visites (tableau 11), des examens de laboratoire et de l'imagerie a été validé par les rhumatologues du groupe de travail PR. Comme l'utilisation des ressources variait dans le temps, des équations ont été créées pour permettre au modèle d'appliquer correctement le coût lors d'une visite quand la ressource correspondante était utilisée. Pour faciliter l'exécution, un ensemble de quatre équations, correspondant aux planifications de visites, a été défini.

Tableau 11. Calendrier des examens biologiques et imagerie. Après 24 mois, les visites sont annuellement planifiées

Temps (mois)	Rhumatologue	Généralistes	Examens biologiques	Imagerie
0	X		X	X
1	X		X	
2			X	
3	X	X	X	
4			X	
5			X	
6	X	X	X	X
7				
8				
9	X	X	X	
10				
11				
12	X	X	X	X
13				
14				
15		X	X	
16				
17				
18	X	X	X	
19				
20				
21		X	X	
22				
23				
24	X	X	X	X

4.4. Calendrier des visites

Cette équation implémentait le calendrier des visites en se fondant sur l'avis du groupe de travail PR. Le calendrier utilisé était mensuel pour les 6 premiers mois, puis trimestriel jusqu'à la fin de la deuxième année et annuel au-delà de la deuxième année.

4.5. Calendrier des doses

Comme le calendrier de dosage et celui des visites ne coïncidaient pas nécessairement, une équation distincte a été mise en œuvre à chaque visite pour calculer le nombre de doses qui auraient dû être consommées depuis la dernière visite. Elle permettait de « remettre le compteur à zéro » lorsque la ligne de traitement suivante était activée et d'appliquer la période recommandée entre les doses. Un exemple est donné dans le tableau 12 pour un médicament avec une période de dosage tous les 56 jours (0,15332 an) et une période de dose de charge de 48 jours (0,131417 an).

Tableau 12. Exemple de calcul de doses

	Temps (années)	Doses Dues	Doses déjà prises	Doses à donner
1	0	0	0	0
2	0,083333	0	0	0
3	0,166667	0	0	0
4	0,25	0	0	0
5	0,333333	1	0	1
6	0,416667	1	1	0
7	0,5	2	1	1
8	0,75	4	2	2
9	1	5	4	1
10	1,25	7	5	2
11	1,5	8	7	1
12	1,75	10	8	2
13	2	12	10	2
14	3	18	12	6
15	4	25	18	7
16	5	31	25	6

4.6. Bornes inférieure et supérieure de la variation du DAS 28

Étant donné que l'efficacité des différents traitements a été appliquée en termes de probabilité d'obtention d'une réponse EULAR (nulle, bonne, modérée), et que ces dernières se présentaient sous forme catégorielle, la nouvelle valeur ponctuelle du DAS 28 a été obtenue en calculant un ensemble possible de valeurs compatibles avec le niveau précédent de DAS 28 et la réponse EULAR obtenue. Par exemple, si initialement, un patient a un DAS 28 de 5,5 (activité de la PR sévère) et que sa réponse au 1^{er} bDMARD est modérée, il devrait avoir une amélioration dans son DAS28 d'au moins 0,7 et d'au plus 2,2 points (tableau 13). Ainsi, l'étendue atteinte du DAS 28 sera de 3,3 (5,5 - 2,2) à 4,8 (5,5 - 0,7). Il faut noter que si l'amélioration avait été inférieure à 0,7, la réponse aurait été classée comme « aucune » et si elle avait été supérieure à 2,2, la réponse aurait été classée « bonne ». Les équations réalisant ce calcul sont fondées sur les définitions des trois catégories de réponses.

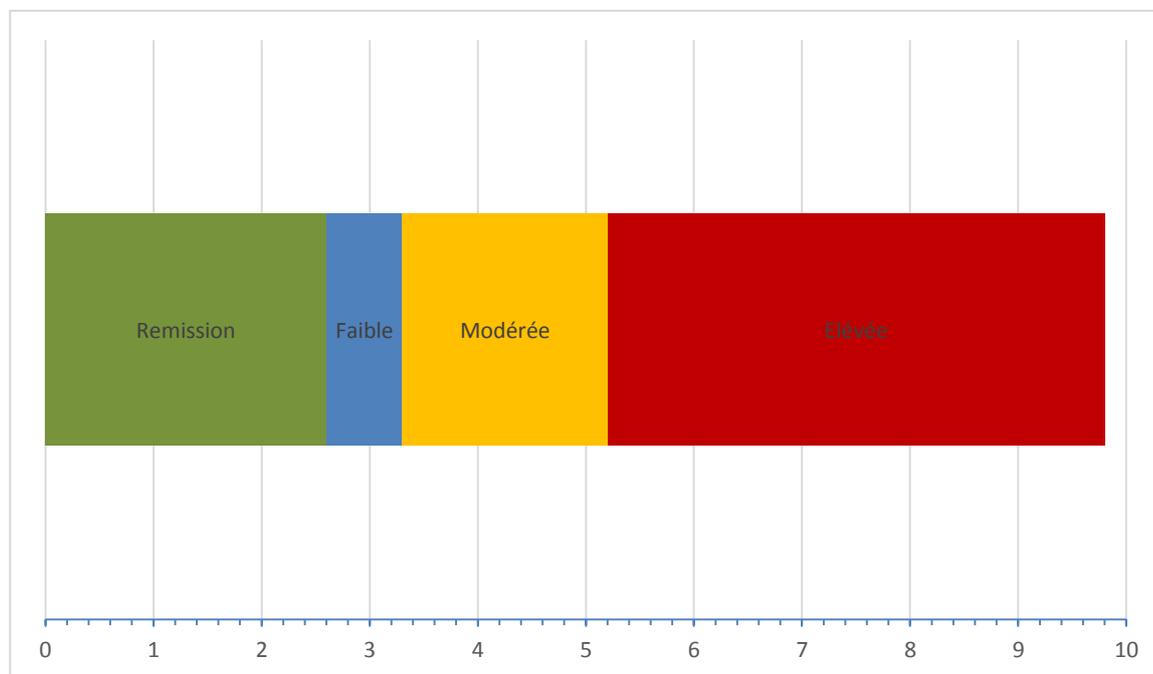
Tableau 13. Définition de la réponse EULAR selon le niveau d'amélioration et le niveau atteint de l'activité

Niveau atteint DAS 28	Amélioration dans le DAS 28		
	≤0,6	>0.6 to ≤1,2	>1,2
≤3,2	Nulle	Modérée	Bonne
>3,2 to ≤5,1	Nulle	Modérée	Modérée
>5,1	Nulle	Nulle	Modérée

4.7. L'activité (DAS 28)

L'activité de la maladie a été classée selon le DAS 28. Ceci a été réalisé à l'aide d'une équation simple qui situait le niveau de l'indice dans l'une des plages définies par l'EULAR (figure 7).

Figure 7. Niveau d'activité fondé sur le score du DAS 28

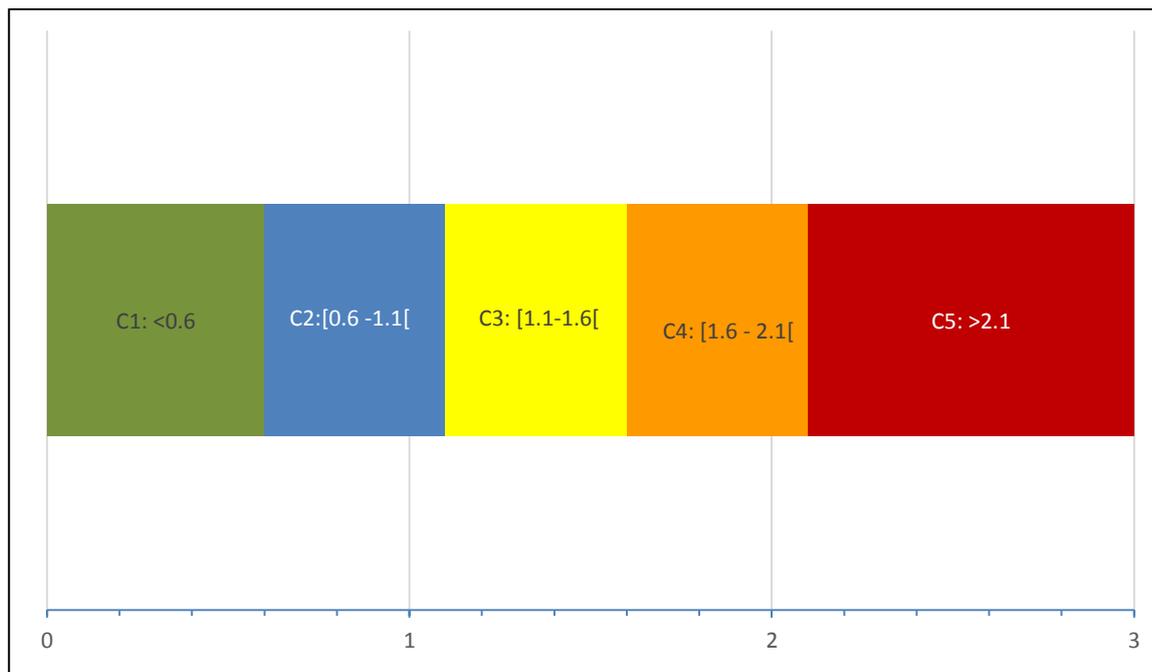


4.8. Changement de l'activité

L'équation de changement d'activité s'applique à chaque visite pour simuler les changements observés dans le temps chez les patients extraits de la cohorte ESPOIR. La proportion de visites pour lesquelles une augmentation de l'activité a été observée (21 %) et pour lesquelles l'activité a diminué (25 %) a été simulée en considérant un nombre aléatoire pour appliquer les probabilités exprimées sous la forme d'une distribution cumulative.

4.9. L'incapacité fonctionnelle (le HAQ)

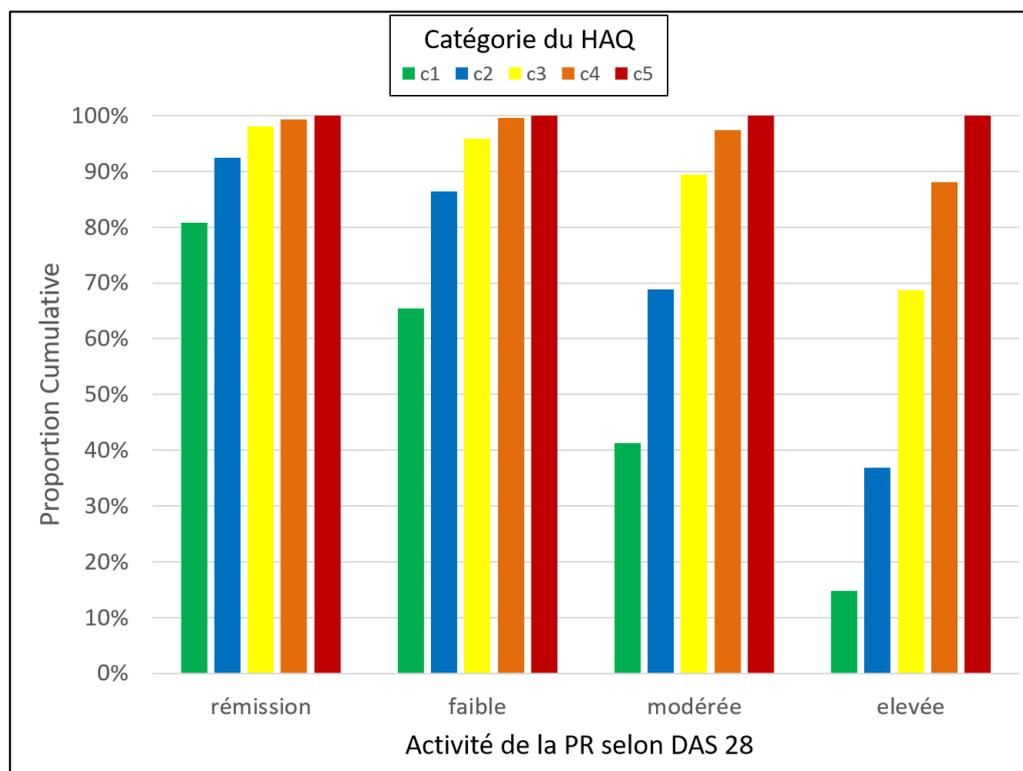
Le score HAQ a été utilisé sous forme catégorielle dans le modèle. Les catégories ont également été utilisées dans les travaux suédois de Kobelt *et al.*, 2014 ([164](#)) et Lekander *et al.*, 2013 ([165](#)) (figure 8).

Figure 8. Score d'incapacité fonctionnelle (HAQ) exprimé en catégories

L'utilisation du HAQ sous forme catégorielle plutôt que sous la forme d'une distribution continue s'explique par la très forte variabilité du score HAQ observée dans l'analyse de l'évolution dans le temps du HAQ dans la cohorte ESPOIR. Ce constat a également été fait par Wolfe *et al.* (2000) (166) et avait conduit ce dernier à considérer que la modélisation de la PR, en fonction de l'évolution des valeurs individuelles du HAQ, ne semblait pas optimale.

Les catégories du HAQ ont été stratifiées sur les catégories de l'activité de la PR (DAS 28) en utilisant l'équation HAQ et en échantillonnant de manière aléatoire les distributions cumulatives spécifiques aux catégories de l'activité de la PR extraites à partir de l'analyse de la cohorte ESPOIR (figure 9).

Figure 9. Distributions cumulatives des catégories du HAQ selon le niveau d'activité observée dans la cohorte ESPOIR



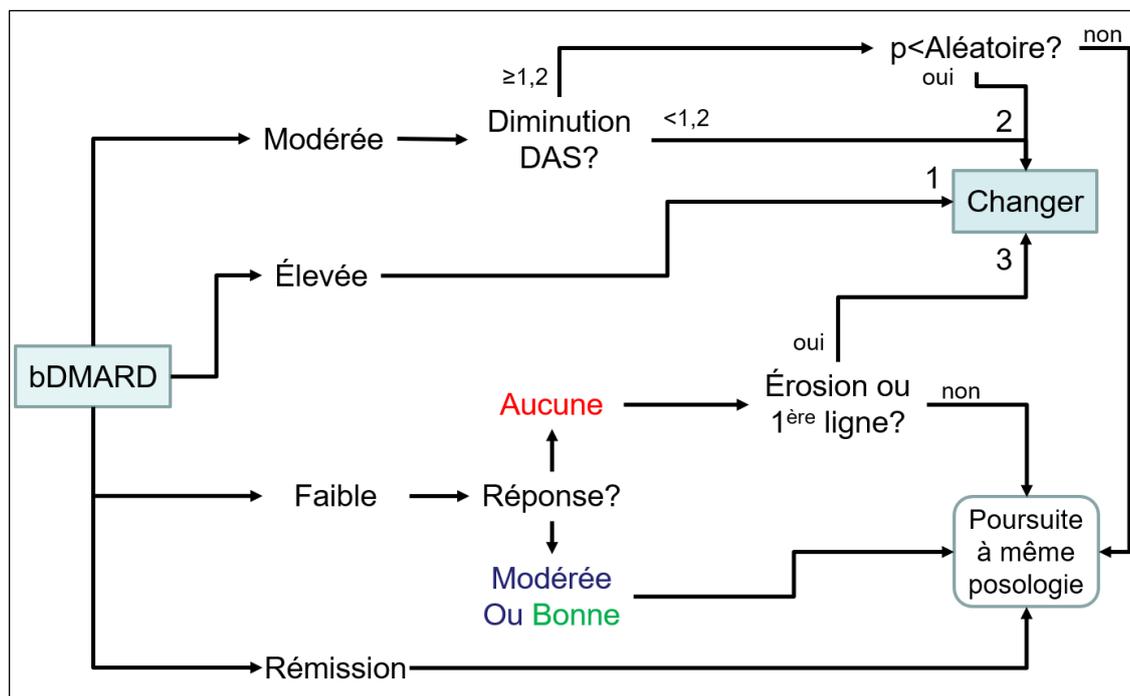
4.10. Changer le traitement

L'équation « changer le traitement » est utilisée lorsque l'événement **DÉCISION** est déclenché afin d'appliquer les recommandations françaises de prise en charge qui déterminent si le traitement doit être poursuivi ou s'il faut passer à la ligne suivante de la séquence en cours. L'algorithme pour déterminer s'il faut changer le traitement consiste en trois règles selon la prise en charge dans le contexte français et il a été validé par les experts rhumatologues du groupe de travail PR. L'équation tient compte des raisons pour lesquelles le traitement doit continuer :

1. l'activité reste élevée ;
2. si l'activité devient modérée mais la diminution du DAS 28 est de moins de 1,2 ou si la diminution a été $\geq 1,2$, mais l'activité reste modérée, on applique une probabilité de changer (50 % dans l'analyse de référence) ;
3. l'activité est devenue faible mais la réponse EULAR est « aucune » et le patient a été en ligne 1 ;

Sinon, le traitement est poursuivi (figure 10).

Figure 10. Algorithme de changement de bDMARD. Les trois règles indiquant un changement de traitement sont numérotées.



Si la DÉCISION n'est pas prise au même temps que la vérification de la réponse EULAR, le changement dans le DAS 28 est projeté de façon linéaire jusqu'à la date de la décision : Changement DAS (à 6 mois) multiplié par le temps de la décision par rapport au début de la ligne divisé par 0,5. Par exemple, si le changement dans le DAS 28 à 6 mois (0,5 ans) est une diminution de 3 points et la décision est reportée à 9 mois (0,75 ans), le DAS 28 va diminuer de 4,5 (égal à $3 \times 0,75/0,5$).

4.11. Temps de changement de traitement

Le temps de changement de traitement en raison d'une perte d'efficacité secondaire ou d'une tolérance insuffisante a été simulé pour chaque patient qui n'était pas en échec primaire. Pour ce faire, la distribution du temps du changement de traitement a été obtenue à l'aide des données extraites de l'analyse de la cohorte ESPOIR en ajustant plusieurs fonctions de survie paramétriques. Dans cet ajustement, les données au second temps de changement de traitement ont été considérées. Le premier changement de traitement n'a pas été utilisé car 98 % des cas de ce changement correspondaient à un changement de csDMARD ; il ne correspond pas, par conséquent, aux caractéristiques de la population d'analyse.

Le temps relatif au second changement de traitement est la différence entre le début de la deuxième ligne de traitement (qui commence immédiatement après le premier changement) et la date de fin du traitement de deuxième ligne en raison de la perte d'efficacité ou d'une tolérance insuffisante. Comme les échecs primaires dans le modèle PR France ont été déterminés à 6 mois pour

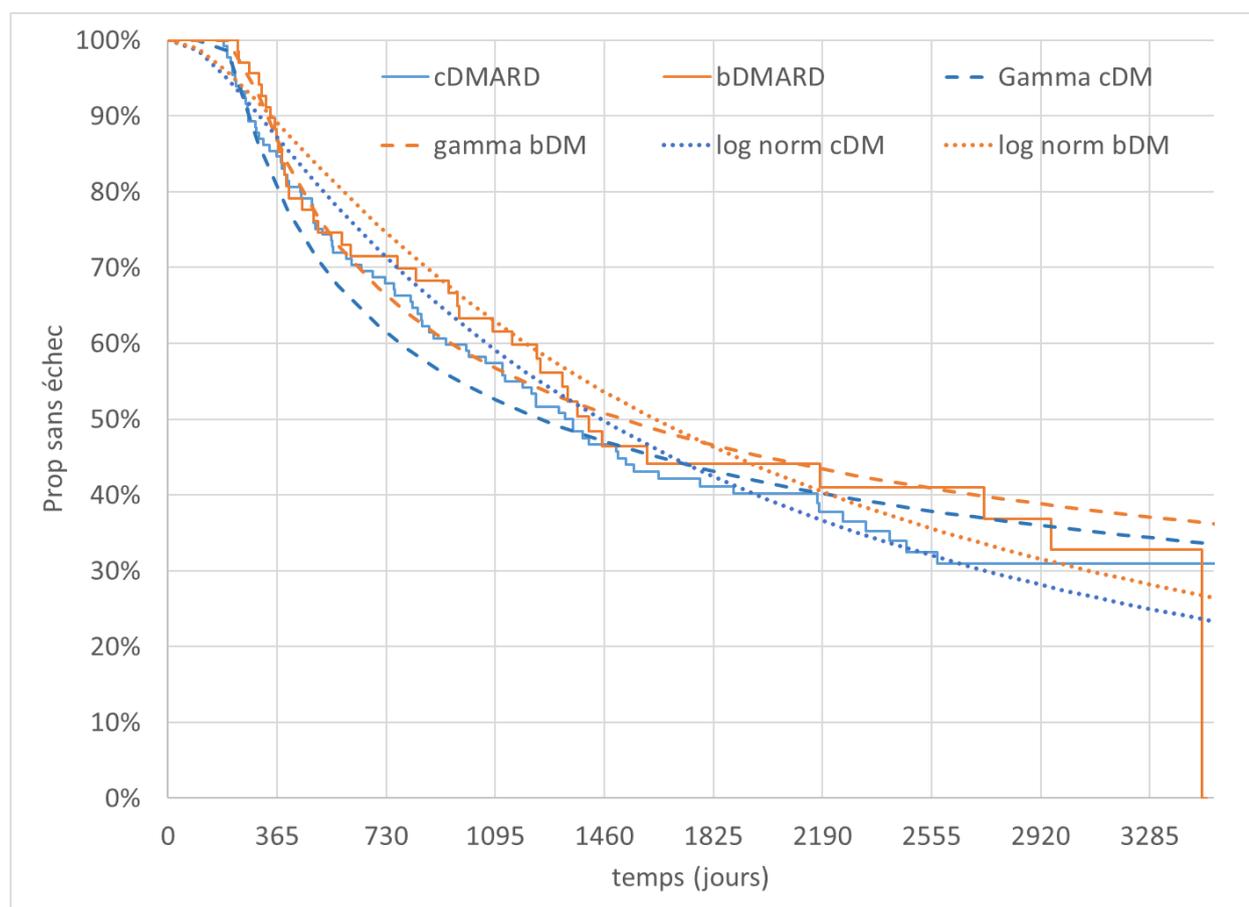
correspondre aux temps d'évaluation des essais cliniques inclus dans la méta-analyse en réseau, les données ESPOIR ont été limitées aux échecs survenant au moins 180 jours après le début du traitement (i.e. censurées à gauche). Ceux-ci devraient correspondre mieux aux échecs secondaires. Les observations ont été censurées à droite si un deuxième changement de traitement n'a pas eu lieu :

- si un 2^e changement de traitement se produit, le délai avant le deuxième changement de traitement = [date de changement de traitement] - [date de début de la 2^e ligne de traitement] ;
- si un 2^e changement de traitement ne se produit pas, le délai avant d'être censuré = [date de la dernière visite] - [date de début de la 2^e ligne de traitement].

La distribution gamma généralisée a présenté le meilleur ajustement selon le critère AIC (tableau 14) et a été retenue (figure 11).

Tableau 14. Temps de deuxième changement de traitement (trithérapie et bDMARDs) : distributions et ajustements pour tous les traitements

Distributions paramétriques	Log (maximum de vraisemblance)	Critère d'information d'Akaike (AIC)
Exponentielle	-1044,225	2092,451
Weibull	-104,196	2094,391
Gompertz	-141,6951	2089,390
Log normale	-1032,051	2070,101
Log logistique	-1036,469	2078,938
Gamma généralisée	-1018,587	2045,179

Figure 11. Temps de deuxième changement de traitement : distributions observées et distributions paramétriques ajustées

Pour appliquer la distribution gamma généralisée dans le modèle, il a fallu faire quelques ajustements. Les paramètres de l'équation sont donnés dans le tableau 15.

Tableau 15. Temps de deuxième changement de traitement (trithérapie et bDMARDs) : paramètres de la distribution retenue (Gamma généralisée)

Paramètre	Estimation	BI 95 % IC	BS 95 % IC	Écart-type estimé
Mu	5,6168	5,0300	6,2206	0,3081
Sigma	0,5299	0,1993	1,4088	0,2644
Q	-4,8904	-10,3927	0,6120	2,8074
SwitchRed	0,1980	0,0423	0,3537	0,0794

L'analyse permettant la réalisation de l'équation a été effectuée sur toutes les données concernant tous les échecs secondaires dans l'étude ESPOIR. Les échecs survenant après un traitement biologique ont été un peu moins fréquents qu'après un traitement synthétique. Cette différence est reflétée dans un facteur SwitchModbD qui est incorporé dans le facteur **SwitchRed**. SwitchRed ajustait l'équation pour être en accord avec l'opinion des experts rhumatologues selon lesquels la tendance à changer de traitement concernant les 2^e et 3^e lignes diminue par rapport à la 1^{re} ligne. Il est incorporé dans l'équation gamma généralisée en modifiant le paramètre *mu*. Pour tenir compte

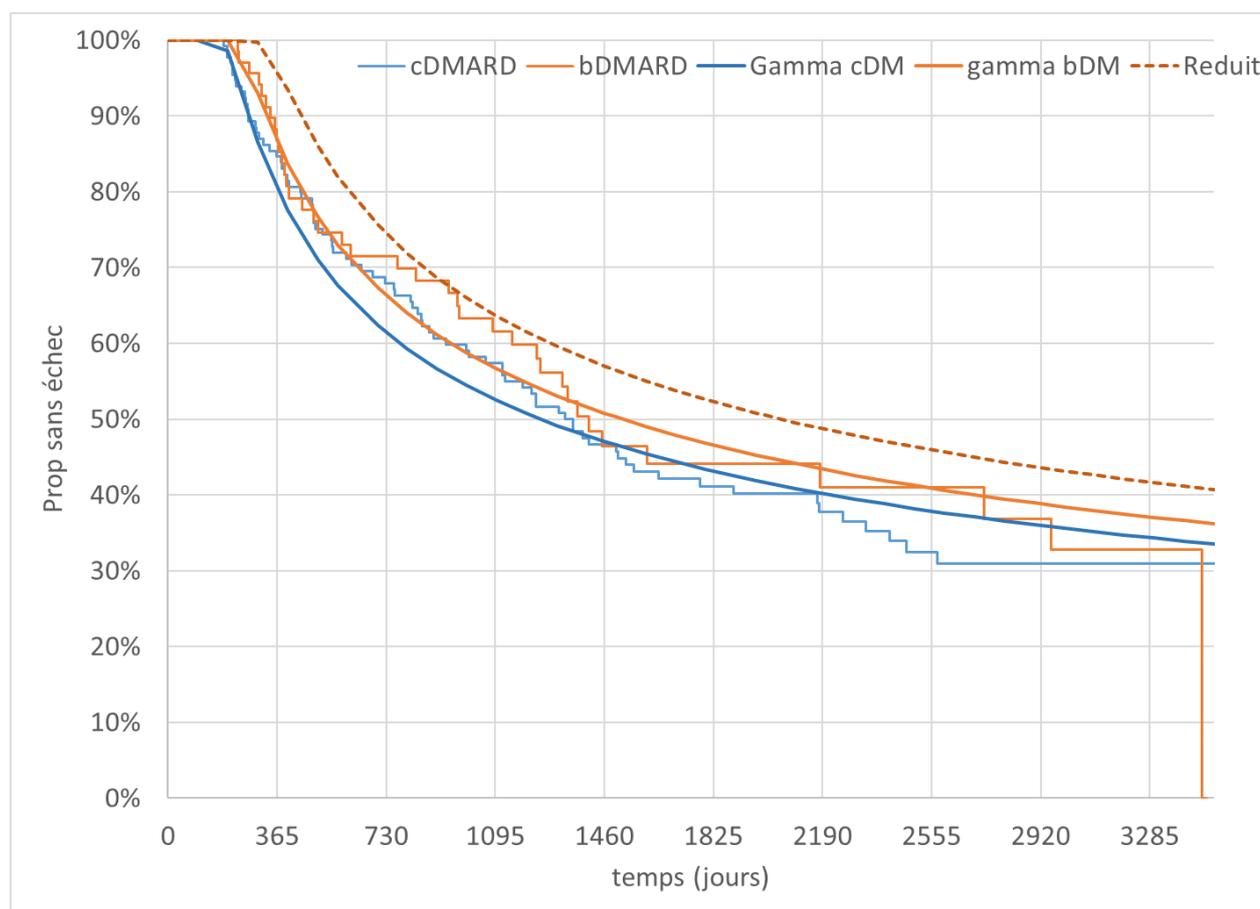
de la possible adhésion un peu inférieure des csDMARDs comparativement à celle des bDMARDs, le facteur SwitchModbD doit être mis à zéro quand la TT est utilisée (c'est le cas dans les analyses de référence). Le facteur SwitchRed incorpore aussi un facteur « **SwitchModL2** » permettant une diminution du taux de changement en 2^e ligne et un facteur « **SwitchModL3** » permettant une diminution du taux de changement en 3^e ligne :

$$\text{SwitchRed} = \text{SI}(\text{ET}(\text{Classe}=\text{cDMARD}, \text{cDMARDmoin}), 0, \text{SwitchModbD}) + \text{SI}(\text{Ligne}=2, \text{SwitchModL2}, \text{SI}(\text{Ligne}=3, \text{SwitchModL3}, 0))$$

- si le traitement pris est la TT et l'analyse permet une diminution de l'adhérence par conséquent, le facteur est égal à **zéro** et *mu* reste inchangé ;
- si le traitement pris est le premier bDMARD, le facteur devient ***mu* + SwitchModbD** ;
- si le traitement de la 2^e ligne est déjà pris, il devient ***mu* + SwitchModbD + SwitchModL2** ;
- si le traitement de la ligne 3 est déjà pris, il devient ***mu* + SwitchModbD + SwitchModL3**.

Dans la **figure 12**, la ligne en pointillé montre l'effet du facteur de réduction SwitchModL2.

Figure 12. Temps de deuxième changement de traitement : distributions Gamma généralisée ajustées sur les courbes Kaplan-Meier



L'impact de ces facteurs intégrés dans la distribution gamma généralisée sur la réduction du taux de changement des traitements peut être expliqué en termes de comparaisons des médianes : l'équation donne 570 jours pour la TT quand une adhésion diminuée est incorporée, 1 461 pour un premier bDMARD et 1 972 pour un bDMARD en 2^e ligne.

4.12. Les utilités

Les scores d'utilités associés à un temps donné ont été estimés selon les catégories du HAQ et de l'activité de la PR. Les données extraites à partir de l'analyse de la cohorte ESPOIR ont montré qu'une bonne prédiction des utilités était obtenue en **utilisant conjointement les catégories des deux dimensions : activité de la maladie (DAS 28) et incapacité fonctionnelle (HAQ) plutôt que le HAQ seul, comme cela a été simulé dans d'autres modèles** (ex. : le modèle de Sheffield de Stevenson *et al*, 2016 ; le modèle de l'ICER, 2017). Comme attendu, l'utilité diminuait avec l'augmentation du score HAQ (incapacité fonctionnelle plus prononcée) mais elle diminuait encore plus lorsque l'activité de la maladie était plus sévère, et ce, quel que soit l'instrument utilisé pour générer les scores d'utilités (EQ-5D ou SF-6D). Il faut noter que les différences sont beaucoup plus prononcées avec les utilités EQ-5D qu'avec les SF-6D.

Dans le modèle PR France, l'utilisateur a le choix d'utiliser les scores d'utilités obtenus selon l'EQ-5D ou le SF-6D. Quel que soit l'instrument choisi, le modèle sélectionne l'utilité moyenne stratifiée sur les catégories du DAS 28 et celles du HAQ.

5. Les entrées du modèle PR France

Le modèle PR France examine les entrées de l'utilisateur et, le cas échéant, les convertit selon le format approprié (par exemple, tous les temps sont exprimés dans le modèle en fractions d'une année) pour leur utilisation pendant l'exécution. Certaines entrées étaient des valeurs uniques (tableau 16), tandis que d'autres correspondaient à une plage possible de valeurs (tableaux 17-24).

Tableau 16. Les entrées principales du modèle PR France

Nom	Définitions	Valeur	Source
Horizon temporel	Durée de la simulation (années)	40	Entrée utilisateur
Perspective	Perspective de l'évaluation économique	collective	Saisine PR
Utiliser EQ-5D	Choix de l'utilisateur de l'instrument des utilités	Oui	Choix structurant
Quand décider si changer tmt	Temps pour décider s'il faut changer le traitement (en années)	0,5	Entrée utilisateur
TA santé	Taux d'actualisation pour les résultats de santé	0,04	Guide HAS 2011
TA coûts	Taux d'actualisation pour les résultats sur les coûts	0,04	Guide HAS 2011
Ajustement Mortalité PR	Ajustement de la mortalité associée à la PR	1,54	Choix structurant
Ajustement Mortalité HAQ	Hazard ratio de la mortalité conditionnelle au HAQ	1	Choix structurant
Lambda décès Femme	Paramètre Lambda dans l'équation de survie	3,59452E-6	Analyses données INSEE
Gamma décès Femme	Paramètre Gamma dans l'équation de survie	0,1137	Analyses données INSEE
Lambda décès Homme	Paramètre Lambda dans l'équation de survie	3,109076E-5	Analyses données INSEE
Gamma décès Homme	Paramètre Gamma dans l'équation de survie	0,930	Analyses données INSEE
Kappa	Paramètre Kappa dans l'équation de changement de traitement	-4,890	Analyses cohorte ESPOIR
Sigma	Paramètre Sigma dans l'équation de changement de traitement	0,53	Analyses cohorte ESPOIR
Mu	Paramètre Mu dans l'équation de changement de traitement	5,617	Analyses cohorte ESPOIR
SwitchbDMARD D	Coefficient dans l'équation de changement de traitement identifiant les bDMARDs	0,198	Analyses cohorte ESPOIR
Appliquer param bDMARD à la TT	Devrait-on utiliser la distribution de changement des bDMARDs pour la TT	Oui	Choix structurant
Réduction L2-L3	Devrait-on permettre une réduction dans les échecs en ligne 2 et en ligne 3		Choix structurant
Réduction L2	Facteur de réduction dans l'équation de changement de traitement pour la ligne 2	0,3	Hypothèse
Réduction L3	Facteur de réduction dans l'équation de changement de traitement pour la ligne 3	0,5	Hypothèse
Perte pB 2^e anti	Perte relative d'efficacité (pBonne) si le 2 ^e traitement est aussi un anti-TNF	47,8 %	Gottenberg <i>et al.</i> (2016)
Perte pM 2^e anti	Perte relative d'efficacité (pBonne ou modérée) si le 2 ^e traitement est aussi un anti-TNF	64,7 %	
Perte pB non	Perte relative d'efficacité (pBonne) si le 2 ^e	94,4 %	

Nom	Définitions	Valeur	Source
anti	traitement n'est pas un anti-TNF		
Perte pM non anti	Perte relative d'efficacité (pBonne ou modérée) si le 2 ^e traitement n'est pas un anti-TNF	92,3 %	
max Ligne	Nombre maximal de lignes de traitements dans une séquence	3	Choix structurant
P augmenter	Probabilité que l'activité de la PR augmente	0,213	Hypothèse
P diminuer	Probabilité que l'activité de la PR diminue	0,248	Hypothèse
Séquences à simuler	Simuler toutes les séquences ou une sélection (jusqu'à 10)	Toutes	Choix structurant
Revenir au DAS initial	Devrait-on revenir au DAS initial lors d'un échec	Non	Choix structurant
Coût Rad	Coût de l'imagerie par visite	101,18	Calcul
Coût Lab	Coût des examens biologiques par visite	84,78	Calcul
Coût Lab Init	Coût des examens à réaliser pendant la première visite	122,85	Calcul
Coût MG	Tarif de consultation d'un médecin généraliste	23 €	CNAM
Coût RH	Tarif de consultation d'un rhumatologue	28,27 €	CNAM
Coût adm hop	Coût de l'administration du médicament à l'hôpital	363,36 €	ATIH – Décembre 2017
Coût adm ville	Coût de l'administration du médicament en ville	3,39 €	Chevreur <i>et al.</i> 2014
Coût MTX	Coût par mg du méthotrexate	0,03 €	Calcul fondé sur les RCP*

(*) : RCP = résumé sur les caractéristiques du produit.

Tableau 17. Tableau de valeurs utilisées dans le modèle PR France

Nom	Définition	Source
Médicaments	Données portant sur les médicaments analysés	- Banque de donnée sur les médicaments : http://www.theriaque.org - Base de données nationale des médicaments : URL : http://base-donnees-publique.medicaments.gouv.fr/
HAQ rémission	Probabilités cumulatives correspondant aux catégories du HAQ pour une rémission de la PR	Analyses cohorte ESPOIR
HAQ faible	Probabilités cumulatives correspondant aux catégories du HAQ pour une PR faiblement active	Analyse cohorte ESPOIR
HAQ mod	Probabilités cumulatives correspondant aux catégories du HAQ pour une PR modérément active	Analyse cohorte ESPOIR
HAQ élevé	Probabilités cumulatives correspondant aux catégories du HAQ pour une PR très active	Analyse cohorte ESPOIR
Util EQ-5D	Utilités EQ-5D stratifiées sur les catégories du HAQ et du DAS 28	Analyse cohorte ESPOIR
Util SF-6D	Utilités SF-6D stratifiées sur les catégories du HAQ et du DAS 28	Analyse cohorte ESPOIR

Tableau 18. Distributions cumulatives des catégories du HAQ stratifiées sur l'activité de la PR

Catégories HAQ	Activité (selon DAS 28)			
	Rémission : <2,6	Faible : [2,6-3,2[Modérée : [3,2-5,1[Élevée : >=5,1
C1 : <0,6	0,811	0,658	0,413	0,149
C2 : [0,6-1,1[0,932	0,86	0,691	0,358
C3 : [1,1-1,6[0,982	0,96	0,898	0,687
C4 : [1,6-2,1[0,996	0,99	0,974	0,901
C5 : >2,1	1	1	1	1

Tableau 19. Scores moyens d'utilités EQ-5D stratifiés sur les catégories du HAQ et du DAS 28

Catégories HAQ	Activité (selon DAS 28)			
	Rémission : <2,6	Faible : [2,6-3,2[Modérée : [3,2-5,1[Élevée : >=5,1
C1 : <0,6	0,833	0,789	0,749	0,663
C2 : [0,6-1,1[0,709	0,665	0,625	0,538
C3 : [1,1-1,6[0,562	0,518	0,478	0,391
C4 : [1,6-2,1[0,396	0,353	0,312	0,226
C5 : >2,1	0,243	0,199	0,159	0,072
N	2539	909	1807	1030

Tableau 20. Scores moyens d'utilités SF-6D stratifiés sur les catégories du HAQ et du DAS 28

Catégories HAQ	Activité (selon DAS 28)			
	Rémission : <2,6	Faible : [2,6-3,2[Modérée : [3,2-5,1[Élevée : >=5,1
C1 : <0,6	0,773	0,732	0,697	0,703
C2 : [0,6-1,1[0,666	0,632	0,622	0,597
C3 : [1,1-1,6[0,605	0,597	0,570	0,550
C4 : [1,6-2,1[0,515	0,549	0,524	0,487
C5 : >2,1	0,484	0,462	0,454	0,425
N	2528	901	1798	1022

Tableau 21. Données portant sur les médicaments

Classe	DCI	Nom commercial	P bonne	P modérée-bonne	Coût d'une dose (mg)	Ajust Poids	Coût dose de charge	Coût d'administration	Période (année)	Période de Charge ment
csDMARD	Trithérapie	NA	0,19	0,57	0,8	0	1,53	0	0,002738	0,057495
Anti-TNF α	Infliximab	REMICADE	0,28	0,69	8,90	1	26,7	364,67	0,15332	0,11499
Anti-TNF α	Adalimumab	HUMIRA	0,3	0,7	377,68	0	0	3,39	0,03833	0
Anti-TNF α	Étanercept	ENBREL	0,56	0,89	184,16	0	0	3,39	0,019165	0
Anti-TNF α	Golimumab	SIMPONI	0,39	0,78	816,6	0	0	3,39	0,082136	0
Anti-TNF α	Certolizumab	CIMZIA	0,38	0,77	358,79	0	1076,37	3,39	0,03833	0,07666
Anti-TNF α	InfliximabInf	INFLECTRA	0,28	0,69	8,9	1	26,7	364,67	0,15332	0,11499
Anti-TNF α	ÉtanerceptBe	BENEPALI	0,52	0,69	146,56	0	0	3,39	0,019165	0
CTLA4-Ig	Abatacept	ORENCIA	0,28	0,69	230,76	0	0	3,39	0,07666	0,07666
Anti-IL6	Tocilizumab	ROACTEMRA	0,63	0,92	230,76	1	0	3,39	0,07666	0
Anti-CD20	Rituximab	MABTHERA	0,15	0,65	1696	0	3392	364,67	0,459959	0,03833
csDMARD	Sulfasalazine	SALAZOPYRINE	NA	NA	0,51	0	1,53	0	0,002738	0,057495
csDMARD	Hydroxychloroquine	PLAQUENIL	NA	NA	0,29	0	0	0	0,002738	0
csDMARD	Méthotrexate	plusieurs	NA	NA	0,03	1	0	0	0,019165	0

Tableau 22. Coûts de l'imagerie

Item	Coût unitaire	Nombre	Source
Main	19,95	1	http://www.codage.ext.cnamts.fr/cgi/nabm
Pied	39,95	1	
Poitrine	21,28	1	
Ciblée	20	1	
Total	101,18		Somme produits (coût unitaire, nombre)

Tableau 23. Coûts des examens biologiques lors de la visite d'initiation du traitement

Item	Coût unitaire	Nombre	Source
Forfait de prise en charge	4,32	1	http://www.codage.ext.cnamts.fr/cgi/nabm
Forfait de sécurité	1,35	1	
Forfait de prélèvement	3,78	1	
NFS, plaquettes	7,29	1	
Ionogramme	6,21	1	
CRP	2,43	1	
VS	1,89	1	
Urée et créatinine	1,89	1	
Transaminases	2,70	1	
GGT	1,62	1	
Facteur rhumatoïde	9,45	1	
Anti-CCP	16,20	1	
AAN	10,80	1	
Bandette urinaire	0,54	1	
Électrophorèse des protéines	14,31	1	
Total visite standard	84,78	1	Somme des produits
Quantiféron	54,00		http://www.codage.ext.cnamts.fr/cgi/nabm
Sérologie VHB	40,50	1	
Sérologie VHC	14,85	1	
Sérologie VIH	13,50	1	
Additionnel 1^{re} visite	122,85		Somme des produits

Tableau 24. Coûts et désutilités des effets indésirables

Profil EI	Coût GHM	Désutilité (1 mois)	Anti-TNF α	Anti-CD20	CTLA4-Ig	Anti-IL6	TT
Pyélonéphrite aiguë	3071	0,371	1,48%	6,20%	5,50%		
Réaction allergique	1311	0,01	5,50%	26,00%	0,55%		0,55%
Asthme	2752	0,136	5,50%	0,55%			
Cellulite	3303	0,371	1,70%	6,20%		2,00%	
Hémorragie GI	4075	0,254	5,50%				0,06%
Cancer sang	3540	0,371	0,10%	7,46%			
Tumeur maligne	2721	0,371	0,76%		0,40%	0,65%	0,55%
Mélanome	1902	0,371	0,55%	4,48%			
Neutropénie	3114	0,15	2,27%	5,50%	0,55%	4,80%	5,50%
Cancer peau hors mélanome	1902	0,1	0,70%				
Pneumonie	3744	0,156	6,60%	8,50%	5,50%	3,00%	1,80%
Sepsis	6812	0,156	5,05%		0,55%		0,06%
Pneumopathie interstitielle	3640	0,156	0,51%	0,055%			1,35%
Perforation intestinale	4256	0,371		0,005%		0,70%	
Coût pondéré			1277,32	1592,25	447,51	375,33	316,2
Désutilité pondérée			0,0621	0,1153	0,0322	0,0243	0,0141

6. Discussion

L'évaluation économique proposée dans ce projet se fonde sur un nouveau modèle appelé PR France. Il s'agit d'un modèle de micro-simulation « à conditions et événements discrètement intégrés » (CEDI) ou « *Discretely Integrated Condition Event* » (DICE), qui simule le parcours médicamenteux des patients atteints de PR modérée à sévère *via* les séquences des bDMARDs.

Cette évaluation est la première étude réalisée par une agence d'évaluation des technologies de santé (*Health Technology Assessment*, HTA) qui propose d'intégrer la trithérapie (TT) et de simuler le parcours médicamenteux de la prise en charge de la PR au-delà d'une seule ligne de traitement de fond biologique.

En accord avec les recommandations cliniques actuelles en France et les données d'essais cliniques randomisés et comparatifs ainsi que les données françaises en vie réelle, les stratégies de comparaison retenues incluaient toutes les séquences de bDMARDs pouvant être observées en vie réelle et validées par le groupe de travail PR. Trois lignes de traitements biologiques et un traitement en 4^e ligne biologique indifférencié constituaient chaque séquence de traitement. La trithérapie (TT) était incluse aux premières lignes de certaines séquences de traitements qui seront considérées dans les analyses de sensibilité de scénarios, conformément aux recommandations du groupe de travail PR.

La souplesse de la structure du modèle PR France reposant à la fois sur des concepts de **conditions** et **d'événements** a permis d'inclure un nombre important de séquences de traitements de bDMARDs.

Il est important de souligner que le modèle PR France reposait principalement sur les pratiques françaises de prise en charge de la PR considérant le DAS 28 comme le critère principal caractérisant l'activité de la PR plutôt que le HAQ (qui traduit les conséquences de l'évolution de l'activité définies en termes d'incapacité fonctionnelle). À la différence des évaluations du NICE (Stevenson *et al.*, 2016) et de l'ICER (ICER, 2017) qui approximaient l'évolution de l'activité de la PR à long terme par l'incapacité fonctionnelle (le HAQ), l'algorithme de simulation de PR France repose sur la variation du DAS 28.

À l'instar de l'évaluation du NICE ([12](#)), PR France repose sur des données d'efficacité comparative à court terme issues des probabilités de réponses EULAR extraites de la méta-analyse en réseau HAS & Sheffield (cf. chapitre 3).

Afin de « pallier » en partie l'absence de données de méta-analyses en 2^e ligne biologique et plus comparant les probabilités de réponses EULAR des bDMARDs, des données issues d'un essai

clinique français pragmatique randomisé et comparatif ont été utilisées pour estimer les probabilités de réponses EULAR.

Les choix structurants de l'évaluation économique présentée dans ce rapport présentaient des points communs avec les évaluations économiques identifiées dans la revue systématique de la littérature économique : l'horizon temporel, la perspective, l'actualisation, le type d'analyse économique (coût-utilité) et la typologie des coûts (coûts d'acquisition, coûts d'administration, coûts de suivi (examens biologiques et radiologiques) et coûts d'hospitalisation générés par les événements indésirables graves (ex. : les infections graves)). Néanmoins, comme le modèle PR France simule et analyse des séquences de traitements, notre évaluation suggère des problématiques méthodologiques qui diffèrent de celles identifiées dans la revue systématique de la littérature : ces problématiques concernent la présentation et l'interprétation des résultats au même titre que les analyses de sensibilité et leur interprétation, compte tenu du nombre important des séquences comparées (i.e. nombre de séquences supérieur à 100).

À l'exception des données extraites de la méta-analyse en réseau concernant les probabilités de réponses EULAR pour la première ligne de traitement biologique et les probabilités de réponses EULAR du rituximab extraites d'un essai clinique comparatif randomisé, les données restantes étaient françaises.

L'utilisation des données en vie réelle a fait l'objet de deux protocoles d'analyse : le premier a porté sur l'analyse de la cohorte ESPOIR, le second s'est focalisé sur l'analyse des événements indésirables colligés à partir de plusieurs sources de données et en accord avec les évaluations et les réévaluations des bDMARDs réalisées par la CT.

Notre modélisation présente également des points forts au niveau de :

- l'utilisation des données françaises en vie réelle (cohorte ESPOIR) dont l'aspect longitudinal est adapté à la structure du modèle PR France et à la réalisation des équations de survie ;
- la disponibilité des instruments portant sur la qualité de vie (EQ-5D et SF-6D) permettant de calculer deux types de scores d'utilité (respectivement en analyse de référence et en analyse de sensibilité) ;
- les estimations des paramètres des équations de survie (supposées similaires pour toutes les classes des traitements de fond biologiques) ne posaient pas de problème : les calculs de décès étaient fondés sur une distribution (distribution de Gompertz) fréquemment retenue pour estimer la mortalité à partir des statistiques nationales de mortalité stratifiées

sur l'âge et le sexe (i.e. statistiques de l'Insee sur la mortalité française de 2015). Quant aux estimations des paramètres de l'équation de changement de traitement (figure 9) qui contrôlaient « les échecs secondaires », entre 25 % et 30 % des patients continuaient leur traitement de fond biologique après 10 ans. Ce constat (validé par le groupe de travail PR) était assez similaire aux études d'extension des anti-TNF α dans lesquelles presque 30 % des patients étaient encore dans l'étude à 10 ans (Favalli *et al.*, 2016) (167) ;

- l'équation de la variabilité du DAS 28 reposait sur les proportions d'augmentation et de baisse de l'activité relatives à chacune des visites de la cohorte ESPOIR. Cette dernière a été marquée par une forte variabilité observée dans la cohorte ESPOIR où l'activité augmentait, restait plate, puis baissait, etc. Contrairement à l'hypothèse se fondant sur une absence de variabilité de l'activité, la variabilité observée à partir de la cohorte ESPOIR a été considérée par le groupe de travail PR comme une caractéristique réaliste de l'activité d'une PR en vie réelle. Il est vraisemblable (cf. chapitre 6) que l'élimination de la variabilité dans l'activité de la PR entraîne un effet important sur les résultats (les séquences les plus efficaces « gagnent » des QALYs grâce à cette absence de variabilité et prolongent les effets), mais cette élimination artificielle de la variabilité ne rend pas le modèle fidèle à la réalité.

Les données utilisées présentent des limites qui ont été prises en compte ou, le cas échéant, discutées et présentées comme des limites à pallier dans la prochaine actualisation du modèle PR France :

- bien qu'elle soit adaptée à notre problématique, la cohorte ESPOIR ne permet pas d'analyser des sous-populations stratifiées selon le risque de comorbidités. La cohorte ESPOIR permet un suivi d'une population atteinte de PR débutante (sans ancienneté de la PR). Cependant, l'ancienneté de la PR ne semblait pas créer un problème méthodologique majeur dans la mesure où elle ne représentait pas un facteur de modification statistiquement significatif dans la méta-régression effectuée sur les réponses EULAR (et ACR) correspondant à l'indication de la population d'entrée du modèle PR France (cf. annexe 3-10, chapitre 3). Cet argument était en accord avec celui de l'évaluation économique des traitements de fond biologiques et ciblés (Jansen *et al.*, 2017) (133) proposant une explication comparable en fonction de l'ancienneté de la maladie ;
- la périodicité annuelle des visites de la cohorte ESPOIR ne prend pas en compte des éventuels changements de traitement survenus au cours de l'année ;

- la proportion des patients traités avec des bDMARDs a été relativement faible durant les 3 premières années suivant la mise en place de l'étude. Toutefois, ce constat semble être réaliste : il est en accord avec l'évolution de la prescription des traitements de fond biologiques en France et en phase avec les autorisations sur le marché des différentes classes des produits biologiques au cours de la période 2000-2012 ;
- certaines données économiques de la cohorte ESPOIR sont caractérisées par une proportion de données manquantes assez élevée sur le suivi des patients et des hospitalisations ;
- le faible nombre d'essais cliniques pour l'efficacité clinique des traitements en 2^e et 3^e lignes est une limite importante de l'étude. Nous ne disposons pas de méta-analyses en réseau (*a minima* de comparaisons indirectes) sur les traitements de fond en 3^e ligne et plus de la PR. Le choix des données issues de l'essai clinique français ROC randomisé et comparatif représente un choix pragmatique utilisé pour la première fois dans la littérature sur l'évaluation économique des séquences des bDMARDs. Malgré ces limites (ex. : effectifs assez faibles, non prise en compte de tous les bDMARDs tels que le golimumab), cet essai a permis de fournir une mesure relative sur l'utilisation des anti-TNF vs non-anti-TNF en 2^e ligne de traitement biologique (3^e ligne de la PR). Dans le chapitre 6, les limites de l'utilisation de l'essai ROC sont palliées par des analyses de sensibilité en scénarios ;
- à l'instar des modèles du NICE, de l'ICER et d'autres évaluations économiques analysées dans la littérature, le modèle PR France dans sa version actuelle ne tient pas compte de l'adhérence au traitement et des événements indésirables « chroniques ».

7. Conclusion

Ce chapitre a permis de présenter les supports de l'évaluation économique des séquences de traitements biologiques dans la prise en charge de la PR en France :

- ✓ le choix et la structure d'un nouveau modèle de micro-simulation « à conditions et événements discrètement intégrés » décrivant le parcours médicamenteux des patients atteints de PR modérée à sévère en France selon l'indication de la population d'analyse, ses composantes en termes de conditions et d'événements et son processus de validation ;
- ✓ les choix structurants de l'évaluation économique menée : les points communs et les aspects méthodologiques suscitant un traitement spécifique aux stratégies définies en termes de séquences de traitements. Ce dernier aspect est argumenté à partir d'exemples issus des résultats présentés dans le chapitre 6 ;
- ✓ les sources de données mobilisées (efficacité, tolérance et qualité de vie). La méthodologie adoptée pour utiliser les données françaises en vie réelle a fait l'objet d'un protocole d'étude de la cohorte ESPOIR. Parmi les quatre sources françaises de données en vie réelle, la cohorte ESPOIR était la plus adaptée à la structure du modèle PR France grâce à la fois à sa structure longitudinale et la disponibilité des données sur l'activité de la PR, la modification structurale, l'incapacité fonctionnelle et la qualité de vie ;
- ✓ la mise en place d'une étude proposant une méthodologie de traitement des événements indésirables susceptibles d'entraîner des hospitalisations et colligés dans plusieurs sources de données réglementaires et en vie réelle.

La méthodologie adoptée présente des points forts tels que la capacité du modèle PR France à simuler plus de 100 séquences de traitements dans la prise en charge de la PR et la qualité des données mobilisées dans cette évaluation. Toutefois, des aspects méthodologiques tels que la prise en compte de l'adhérence, les événements indésirables « chroniques » tels que les cancers et les modifications structurales (i.e. érosions) peuvent indéniablement améliorer la modélisation si des données appropriées s'avéraient disponibles.

Chapitre 6. Résultats de l'évaluation économique des séquences de traitements de fond biologiques dans la prise en charge de la PR en France

Comparativement à la littérature anglo-saxonne, les études économiques des traitements de fond biologiques réalisées en France sont peu nombreuses et assez hétérogènes. Parmi les 17 études françaises recensées lors de la sélection intermédiaire de la revue systématique de littérature économique, une seule (Saraux *et al.*, 2010) (19) remplissait les critères d'éligibilité. Elle portait sur des patients atteints de PR active modérée à sévère, présentant une réponse inadéquate à au moins un anti-TNF α , et comparait la rémission (ou la faible activité de la PR) selon quatre stratégies de traitements de fond biologiques incluant l'éta nercept, l'adalimumab, l'infliximab et le rituximab. L'étude de Sany *et al.*, en 2009 (18), était non comparative et évaluait la réponse avant et après administration d'infliximab dans l'indication de la PR sévère. L'étude de Vanier *et al.*, en 2017 (20), a évalué à partir d'un essai clinique français (STRASS), l'espacement des injections des anti-TNF α pour des patients présentant une rémission établie, tandis que Fagnani *et al.*, en 2016 (8), ont évalué l'utilisation du certolizumab selon une stratégie de traitement jusqu'à l'atteinte de la cible (« *treat to target strategy* ») avec et sans un schéma de contrat de partage de risque avec le Comité économique des produits de santé (CEPS). Les études restantes (Beck *et al.*, 2015 (15) ; Beresniak *et al.*, 2010 (138) ; Bertin *et al.*, 2015 (168) ; Chevreul *et al.*, 2014 (147) ; Fautrel *et al.*, 2005 (169) ; Fautrel *et al.*, 2012 (17) ; Fautrel *et al.*, 2016 (8) ; Guillemain *et al.*, 2004 (2) ; Kobelt *et al.*, 2008 (16) ; Maravic *et al.*, 2010 (14) ; Rat *et al.*, 2004 (170) ; Julliard-Candat, 2007 (171) ; Julliard-Candat, 2008 (172)) étaient des analyses des coûts directs traitant principalement l'utilisation des traitements de fond biologiques dont les sources de données, les populations et les méthodologies de calcul des coûts étaient hétérogènes.

Notre évaluation ne constitue pas seulement une réponse à la saisine de la Direction de la Sécurité sociale. Elle répond au besoin d'une évaluation économique française portant sur les séquences de bDMARDs et qui est fondée notamment sur des données de méta-analyses et des choix structurants recommandés par la HAS.

Ce chapitre a pour objectif de présenter les résultats de l'évaluation économique des séquences de traitements de fond biologiques dans la prise en charge de la PR en France. La première section présente les résultats de l'analyse de référence. Les résultats des analyses de sensibilité (scénarios déterministes et probabilistes) sont synthétisés dans la deuxième section. Les discussions sont présentées dans la troisième section.

1. Les résultats de l'analyse de référence

Pour obtenir les sorties de la modélisation économique, le modèle PR France compile les résultats (tableau 1) relatifs aux coûts, à la survie, aux QALYs, au temps passé dans chaque ligne de traitement, au temps durant lequel l'activité de la PR était en rémission et aux proportions des patients subissant des échecs primaires et secondaires à chaque ligne.

Tableau 1. Sorties du modèle PR France

Nom	Objectif
Coût Dose Chg	Coût moyen de toutes les doses de chargement
Coût Visite RH	Coût moyen des visites de spécialistes (rhumatologues)
Coût Visite MG	Coût moyen des visites de médecins généralistes
Coût Visite Lab	Coût moyen des examens de laboratoire
Coût Visite Rad	Coût moyen de l'imagerie radiologique
Coût Visite	Somme des coûts : rhumatologues, généralistes, examens de laboratoire, coûts imagerie
Coût Tmt	Coût cumulé des traitements utilisés pour traiter la polyarthrite rhumatoïde
Coût MTX	Coût cumulé du méthotrexate de fond
Coût Admin	Coût cumulé de l'administration de médicaments
Coût IE	Coût de la prise en charge des événements indésirables
Perte QALY IE	Perte d'utilité à cause des événements indésirables
QALY	Année de vie ajustée sur la qualité
AV	Années de vie
Dur Ligne 1	Durée de traitement en 1 ^{re} ligne
Temps en rémission	Temps associé à une rémission
Échecs	Nombre d'échecs par patient
Décès	Proportion décédant pendant l'horizon temporel
Dernière Ligne	Proportion dépassant la 3 ^e ligne
Switch Prim L1	Proportion qui fait un échec primaire en 1 ^{re} ligne
Switch Prim L2	Proportion qui fait un échec primaire en 2 ^e ligne
Switch Prim L3	Proportion qui fait un échec primaire en 3 ^e ligne
Switch Sec L1	Proportion qui fait un échec secondaire en 1 ^{re} ligne
Switch Sec L2	Proportion qui fait un échec secondaire en 2 ^e ligne
Switch Sec L3	Proportion qui fait un échec secondaire en 3 ^e ligne
Raison L1	Raison pour laquelle le traitement a été changé en 1 ^{re} ligne
Raison L2	Raison pour laquelle le traitement a été changé en 2 ^e ligne
Raison L3	Raison pour laquelle le traitement a été changé en 3 ^e ligne
Coût L1	Coût moyen correspondant à 1 ^{re} ligne
Coût L2	Coût moyen correspondant à 2 ^e ligne
Coût L3	Coût moyen correspondant à 1 ^{re} ligne
Coût >L3	Coût moyen après la 3 ^e ligne
Temps L1	Temps moyen dans la 1 ^{re} ligne
Temps L2	Temps moyen dans la 2 ^e ligne
Temps L3	Temps moyen dans la 3 ^e ligne
Temps >L3	Temps moyen après la 3 ^e ligne

L'analyse de référence a été réalisée conformément aux choix structurants de l'évaluation économique (section 3, chapitre 5) et aux recommandations du groupe de travail PR. Les points suivants ont été retenus lors de la 3^e réunion du 17 juillet 2018 :

- les échecs secondaires (i.e. réponse inappropriée à un traitement de fond synthétique) survenus lors des 180 premiers jours dans la cohorte ESPOIR ne doivent pas être pris en compte dans la courbe de changement de traitement puisque le modèle PR France démarre avec une réponse au traitement à 6 mois issue de la méta-analyse en réseau ;
- un risque de mortalité associé à la PR tel que l'estimation du risque relatif de Hoek *et al.*, (2017) (55), qui était similaire à celle issue de la méta-analyse de Dadoun *et al.*, (2013) (51), doit être appliqué ;
- conformément à l'objectif de la saisine et à l'indication de la population à l'entrée du modèle, les séquences débutant par la trithérapie (TT) seront considérées seulement dans les analyses de scénarios. Dans la prise en charge actuelle de la PR en France, les séquences de traitements débutant par la TT ne représentent pas des stratégies courantes en vie réelle. La proportion des patients traités par la TT apparaît très faible : dans la cohorte ESPOIR, le pourcentage des patients traités par une combinaison de trois DMARDs était inférieur à 1 %.

1.1. Les résultats de santé

Le tableau 2 synthétise les statistiques descriptives des résultats de santé. Il montre que les différences de QALY (actualisées) entre les séquences de traitements de fond biologiques étaient faibles : les valeurs moyennes, maximales et minimales du QALY présentaient une différence de l'ordre de 0,14 an (soit à peine 2 mois).

La différence de temps moyenne en rémission de la PR était légèrement supérieure à un peu plus de 6 mois (0,58 an).

Le critère des années de vie gagnées ne montrait aucune différence entre les séquences de traitements : en l'absence de preuves montrant un effet des bDMARDs sur la survie des patients atteints de PR, l'hypothèse selon laquelle l'utilisation de ces derniers n'affectait pas la mortalité a été retenue.

Tableau 2. Résultats de santé : statistiques descriptives pour les 180 séquences analysées

Résultats	Moyenne	Minimum	Maximum	Écart-type
Années de vie	28,55	28,55	28,55	0
Années de vie actualisées	16,58	16,58	16,58	0
QALYs	19,12	19,02	19,20	0,0463
QALYs actualisés	11,13	11,06	11,20	3.54E-02
Faible activité de la PR	10,12	9,80	10,38	0,146
Temps passé après la 3 ^e ligne	9,79	7,21	11,99	1,1

1.2. Les coûts

L'impact du choix du traitement de fond biologique sur le différentiel des coûts entre les séquences était plus important et les séquences différaient donc beaucoup plus que pour les QALYs : alors que la différence dans les QALYs **actualisés n'était en moyenne que de 1 %, les coûts actualisés diffèrent de plus de 35 %**. Indépendamment du bDMARD initial (1^{re} ligne biologique), les séquences les moins coûteuses étaient celles contenant de l'abatacept en 2^e ou en 3^e ligne (coût total actualisé inférieur à 150 000 €) et les plus coûteuses (plus de 150 000 €) étaient celles intégrant du tocilizumab ou trois anti-TNF α . En général, les séquences commençant par le biosimilaire de l'étanercept (ETABepali®) étaient moins coûteuses que les séquences correspondantes débutant par d'autres anti-TNF α . Par exemple, le coût moyen de la séquence la moins coûteuse **ETAB --> ABA --> INF** est estimé à 116 000 € (coûts actualisés), contre 123 000 € à 128 000 € pour les autres **anti-TNF α --> ABA --> INF**.

La majeure partie du coût total était expliquée par les coûts d'acquisition des traitements.

Quant aux consultations et aux examens biologiques/radiologiques, ils représentaient un poids beaucoup moins important que les coûts d'acquisition des médicaments biologiques. Le différentiel entre les coûts des visites variait de 16 299 € à 16 484 €. Le tableau 1 de l'annexe 6A.1 décrit un exemple des composantes de coûts d'une séquence sélectionnée aléatoirement parmi les 180 séquences analysées dans l'analyse de référence et illustre, comparativement aux coûts des événements indésirables et du suivi de la PR, le poids important des coûts d'acquisition dans le coût total de la séquence. Le graphique 1 de l'annexe 6A.2 décrit la répartition du coût total de la séquence par ligne de traitement et montre que la dernière ligne de traitement (4^e ligne de traitement biologique ou 5^e ligne de traitement de la PR) n'est pas nécessairement celle qui contribue le plus dans le coût total de la séquence. En effet, le coût dépend essentiellement de l'efficacité du bDMARD associé à la ligne du traitement et de son prix. Le graphique 2 de l'annexe 6A.1 montre que la 4^e ligne de traitement biologique devient moins importante avec un horizon temporel plus court (15 ans).

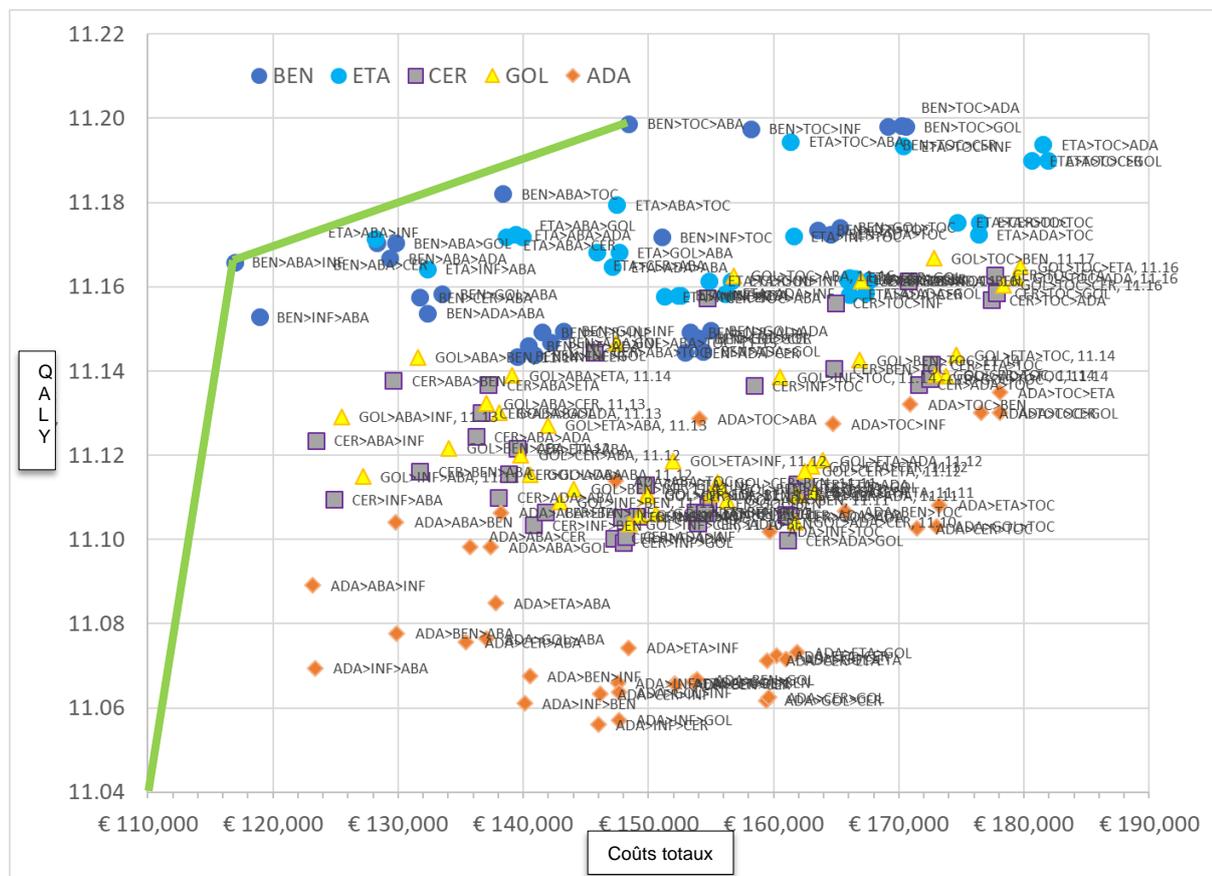
1.3. L'analyse de l'efficacité

Étant donné que les séquences commençant par le biosimilaire de l'étanercept étaient les moins coûteuses et tendaient également à être parmi les plus « efficaces », elles dominaient généralement la plupart des autres séquences. En effet, dans l'analyse de référence, **seules deux séquences, toutes deux commençant par le biosimilaire de l'étanercept (ETAB-->ABA-->INF et ETAB-->TOC-->ABA), étaient sur la frontière d'efficacité** (figure 1). Le RDCR pour la séquence avec tocilizumab en 2^e ligne par rapport à celle avec abatacept en 2^e ligne était extrêmement élevé et supérieur aux seuils de disposition à payer retenus dans certains pays étrangers. La valeur élevée de ce RDCR s'expliquait principalement par le faible différentiel de gain de QALY (**tableau 3**).

Tableau 3. Résultats de l'analyse de référence (résultats actualisés)

Séquence	Total coûts (€)	Total QALY	RDCR (€/QALY)
ETAB>ABA>INF	116 912	11,1658	référence
ETAB>TOC>ABA	148 363	11,1987	955 778

Figure 1. Frontière d'efficacité



2. Exploration de l'incertitude sur l'évaluation économique des séquences de traitements de fond biologiques

Les catégories potentielles d'incertitude identifiées étaient l'incertitude sur les paramètres d'entrée du modèle (ou incertitude paramétrique) et l'incertitude sur la structure du modèle (incertitude structurelle) dont les sources pouvaient être liées à ses hypothèses.

Dans le cadre de la présente évaluation, l'accent a été notamment mis sur l'exploration de l'incertitude paramétrique sur les variables d'intérêt du modèle PR France. L'incertitude structurelle a en partie été explorée par des analyses de scénarios déterministes relatives à des hypothèses structurelles, et non par des approches formelles à l'instar des techniques de « *model averaging method* » (Price *et al.*, 2011) (173) ou « *discrepancy method* » (Strong *et al.*, 2012) (174). L'application de ces dernières approches exige à la fois une argumentation clinique/exhaustive de tous les scénarios plausibles définissant l'incertitude structurelle (Ghabri *et al.*, 2018) (175) et des « révélations » des poids/probabilités associés aux scénarios structurels définis, par exemple, par un groupe d'experts (Soares, 2015) (176).

2.1. Les analyses de scénarios déterministes

2.1.1. Choix des analyses de scénarios déterministes

L'intérêt des analyses de sensibilité déterministes univariées en réalisant des diagrammes de Tornado a été discuté lors de la 3^e réunion du groupe de travail du 17 juillet 2018. La pertinence de ce type d'analyses s'est avérée très limitée dans le cadre de notre évaluation économique réalisée sur un nombre très important de séquences de traitements (180 séquences). En effet, ces analyses multiplient le nombre de comparaisons par paires. Dans certains cas particuliers, elles pourraient être réalisées en choisissant deux séquences à comparer.

Il a été convenu de réaliser des analyses de scénarios déterministes qui ont été jugées cliniquement et/ou économiquement pertinentes. Vingt-quatre analyses de scénarios déterministes (cf. tableau 4) ont été discutées et validées par le groupe de travail PR.

Tableau 4. Analyses de sensibilité du modèle PR France : analyses de scénarios déterministes

Paramètres/choix structurants du modèle PR France	Analyse de référence	Analyse de scénarios déterministes	Rationnel
Horizon temporel	Vie entière (40 ans)	1) 15 ans ; 2) 60 ans	15 ans : horizon temporel court 40 ans : tous les patients ne sont pas décédés 60 ans : permet de vérifier que tous les patients décèdent
Population d'analyse	Activité modérée à sévère	3) Activité sévère ; 4) Érosions ; 5) Tous	3- Seulement les patients avec une activité sévère à l'entrée du modèle 4- Seulement les patients présentant une érosion à l'entrée du modèle 5- Scénario incluant les patients avec une faible activité
Instrument permettant le calcul du score d'utilité	EQ-5D	6) SF-6D	Le SF-6D est un instrument alternatif à l'EQ-5D et disponible dans la cohorte ESPOIR
Temps de vérification de la réponse EULAR	6 mois	7) 9 mois ; 8) 12 mois	Hypothèse validée par les experts du groupe de travail PR
Taux d'actualisation	4% (et 2 après 30 ans)	9) 2,5%	Une des analyses de sensibilité recommandée par le guide HAS
Hazard ratio mortalité associée à la PR	1,54	10) 1,00	Scénario supposant une mortalité associée à la PR (littérature)
Le changement de traitement augmente avec la prise en compte de la TT	Non	11) Oui	Études cliniques contenant quelques éléments soulignant une moins bonne adhésion à la trithérapie
Réduction taux de changement L2, L3	0,3 ; 0,5	12) 0 13) 0,8 ; 0,9	Hypothèse validée par les experts du groupe de travail PR
Perte réponse L2, L3 (%)	52,2; 35,3; 5,6 ; 7,7%	14) 0%	Scénario qui élimine la perte d'efficacité qui n'est pas suffisamment démontrée dans la littérature. Ce scénario permet de pallier les limites de l'essai clinique ROC utilisé dans l'analyse de référence
Traitement après L3	Rituximab	15) TOC	Hypothèse fondée sur les avis CT et explorée par les experts cliniques du groupe de travail PR
Prob de diminution et d'augmentation de l'activité	24,8%, 21,3%	16) 0%	Scénario qui élimine la variabilité qui n'est pas clairement quantifiée dans la littérature
Séquences à analyser	180 (sans TT)	17) 234 (avec TT)	Hypothèse fondée en partie sur les avis CT et validée par les experts du groupe de travail PR
Retour DAS à la valeur initiale si échec	Non	18) Oui	Hypothèse explorée par les experts cliniques du groupe de travail PR
Coûts examens et consultations	Divers	19) -10% ; 20) +10%	Hypothèse explorée par les experts cliniques du groupe de travail PR
Variation prix des médicaments	0%	21) -25%	Hypothèse explorée par les experts cliniques du groupe de

Paramètres/choix structurants du modèle PR France	Analyse de référence	Analyse de scénarios déterministes	Rationnel
Règles décision « switch ¹⁶ »	Règles 1, 2, 3	22) Règles 1, 2	travail PR Scénario alternatif aux règles de décision « switch » du modèle PR France
Prob Switch si réponse ≥ 1.2	50%	23) 0% ; 24) 100%	Hypothèse explorée par les experts cliniques du groupe de travail PR

2.1.2. Les résultats des analyses de scénarios déterministes

Tous les résultats des analyses de scénarios déterministes sont présentés dans le tableau 5.

Tableau 5. Résultats des analyses de scénarios déterministes pour les séquences situées sur la frontière d'efficacité.

Séquences	Coût total (€)	Total QALY	RDCR (€/QALY)
Horizon temporel			
15 ans			
ETAB>ABA>INF	80 434	7,4394	
ETAB>TOC>ABA	104 940	7,4726	737 897
100 ans			
ETAB>ABA>INF	119 528	11,4822	
ETAB>TOC>ABA	151 477	11,5134	1 026 453
ETAB>TOC>CER	172 771	11,5134	1 774 522 447
Pas de réduction du taux de changement de traitements L2 et L3 « no switch reduction »			
ETAB>ABA>INF	116 752	11,1510	
ETAB>TOC>ABA	151 641	11,1813	1 150 618
ETAB>TOC>CER	173 877	11,1814	292 572 578
Absence de variabilité de l'activité (DAS 28)			
ETAB>ABA>INF	116 912	12,5122	
ETA>ABA>INF	128 131	12,5802	164 975
ETA>ABA>TOC	147 472	12,6276	407 946
ETA>TOC>ABA	161 325	12,6562	484 751
ETA>TOC>CER	180 582	12,6631	2 804 264
Pas de perte d'efficacité Lignes 2 et 3			
ETAB>ABA>INF	116,743	11,1651	
ETAB>TOC>ABA	149 582	11,1999	945 219
ETA>TOC>ABA	161 361	11,2028	3 990 123
ETA>TOC>CER	180 949	11,2036	23 656 950
Considérer le TOC en 4^e ligne et plus			

¹⁶ Légende :

Règle 2 : Activité >3.

Règle 3 : Activité 3 et réponse <1,2, ou selon une probabilité si la réponse >1,2.

Règle 4 : Activité 2, aucune réponse et L2+.

ETAB>ABA>TOC	138 356	11,1822	
ETAB>TOC>ABA	148 363	11,1987	602 960
Retour au DAS initial après un échec			
ETAB>INF>ABA	117 673	10,7864	
ETAB>ABA>INF	118 196	10,7935	74 014
ETAB>ABA>TOC	140 977	10,8494	407 217
ETA>ABA>TOC	151 525	10,8554	1 761 852
ETA>TOC>GOL	171 819	10,8594	5 034 449

Tableau 5. (suite) Analyses de scénarios déterministes

Pas d'échecs (règle 2)			
Séquences	Coût total actualisé (€)	Total QALY	RDCR (€/QALY)
ETAB>ABA>INF	17 049	11,1570	
ETAB>TOC>ABA	149 593	11,1831	1 246 971
ETAB>TOC>CER	169 569	11,1846	13 748 302
100 % échecs (règle 3)			
ETAB>ABA>INF	116 017	11,1817	
ETAB>ABA>TOC	137 810	11,2033	1 010 734
ETAB>TOC>ABA	147 603	11,2113	1 224 254
ETA>TOC>ABA	159 542	11,2171	2 045 879
2^e source sur les données d'utilité : SF-6D			
ETAB>ABA>INF	116 912	11,2286	
ETAB>TOC>ABA	148 363	11,2467	1 744 941
Repousser la décision de changement de traitement à 9 mois			
ETAB>ABA>INF	116 998	11,1809	
ETA>ABA>INF	131 974	11,2088	535 495
ETA>ABA>TOC	148 571	11,2183	1 760 044
ETA>TOC>CER	183 400	11,2211	12 169 427
Repousser la décision de changement de traitement à 12 mois			
ETAB>ABA>INF	118 015	11,1835	
ETA>ABA>INF	135 228	11,2167	518 221
ETA>ABA>TOC	151 073	11,2204	4,309 176
ETA>TOC>ABA	168 063	11,2213	18 941 309
Taux d'actualisation égal à 2,5 %			
ETAB>ABA>INF	139 631	13,3883	
ETAB>TOC>ABA	176 304	13,4232	1 049 755
Pas d'ajustement de la mortalité associée à la PR (HR=1)			
ETAB>ABA>INF	122 673	11,7562	
ETAB>TOC>ABA	155 221	11,7861	1 088 417
Diminution plus importante du taux de changement de traitement « switch »			
ETAB>ABA>INF	115 873	11,1514	
ETAB>TOC>ABA	154 563	11,1836	1 200 794
Population incluant toutes les catégories du DAS			
ETAB>ABA>INF	116 185	11,1948	
ETAB>ABA>TOC	137 863	11,2215	811 638
ETAB>TOC>ABA	149 152	11,2269	2 071 232
ETAB>TOC>CER	170 890	11,2291	9 976 383
Analyse de référence intégrant la trithérapie (TT)			
TT>ETAB>INF	110 410	11,07	
ETAB>ABA>INF	116 912	11,1658	65 623
ETAB>TOC>ABA	148 363	11,1987	955 778

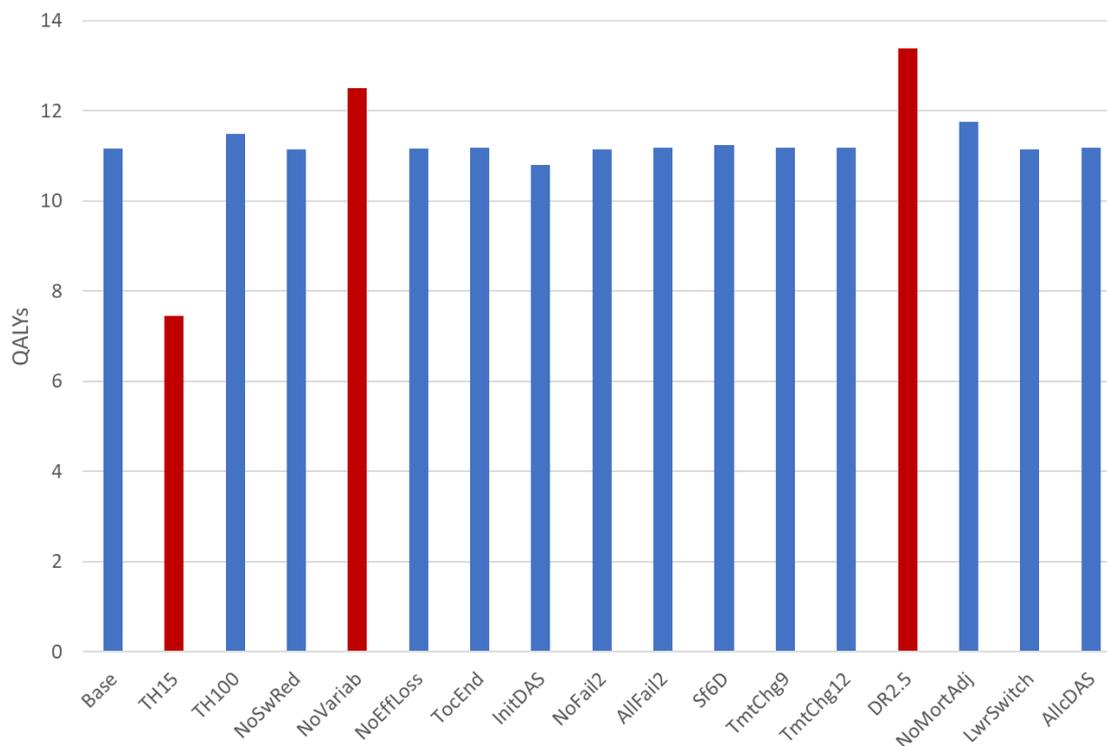
Tableau 5 (suite). Analyses de scénarios déterministes

Séquences	Coût total (€)	Total QALY	RDCR (€/QALY)
Activité sévère à l'entrée du modèle PR France (DAS 28 > 5,1)			
ETAB>ABA>INF	116 509	11,0385	
ETAB>ABA>CER	127 888	11,0566	628 284
ETAB>ABA>TOC	138 393	11,0699	791 174
ETA>ABA>TOC	148 757	11,0750	2 029 488
ETAB>TOC>CER	167 011	11,0750	231 062 700
Présence d'érosions (modifications structurales) à l'entrée du modèle			
ETAB>ABA>INF	122 321	11,2699	
ETA>ABA>TOC	155 118	11,3116	787 439
ETA>INF>TOC	165 694	11,3162	2 257 400
ETA>CER>TOC	177 173	11,3170	14 511 702
Baisse des prix des médicaments			
Baisse de 25 %			
ETAB>ABA>INF	97 691	11,1658	
ETAB>TOC>ABA	119 366	11,1987	658 713
Baisse de 50 %			
ETAB>ABA>INF	76 326	11,175	
ETAB>TOC>ABA	88 664	11,206	397 110
ETAB>TOC>ADA	99 463	11,207	17 878 619
Baisse des coûts (autres que les coûts d'acquisition) de 10 %			
ETAB>ABA>INF	114 300	11,1658	
ETAB>TOC>ABA	145 755	11,1987	955 904
Augmentation des coûts (autres que les coûts d'acquisition) de 10 %			
ETAB>ABA>INF	119 524	11,1658	
ETAB>TOC>ABA	150 971	11,1987	955 652
Prise en compte seulement des pneumonies			
ETAB>ABA>INF	114 755	11,1755	
ETAB>TOC>ABA	146 657	11,2065	1 026 778
ETAB>TOC>ADA	168 096	11,2071	35 494 492
Utiliser le paramètre de changement de traitement en gardant la courbe csDMARD <i>TTSwitchParm</i>			
TT>INF>ABA	91 358	11,0699	
TT>ABA>ETAB	99 816	11,0871	492 523
TT>ETA>ABA	109 577	11,1031	608 677
TT>TOC>ABA	121 361	11,1182	784 549
TT>TOC>ETA	148 605	11,1278	2 823 794
Règle 3 non appliquée			
ETAB>ABA>INF	117 293	11,1622	
ETAB>TOC>ABA	148 299	11,1950	945 664

➤ Les résultats de santé

Les analyses déterministes de scénarios ont mis en évidence une certaine variabilité dans les estimations des QALYs, mais le positionnement d'une séquence par rapport à une autre avait tendance à ne pas changer. La figure 2 montre l'impact des différents scénarios sur le total des QALYs pour la séquence **ETAB-->ABA-->INF**. Il augmentait dans les scénarios relatifs à la réduction de l'horizon temporel à 15 ans et à l'absence de variabilité dans le DAS 28.

Figure 2. QALYs obtenus avec la séquence **ETAB-->ABA-->INF** à travers tous les scénarios d'analyse



Légende :

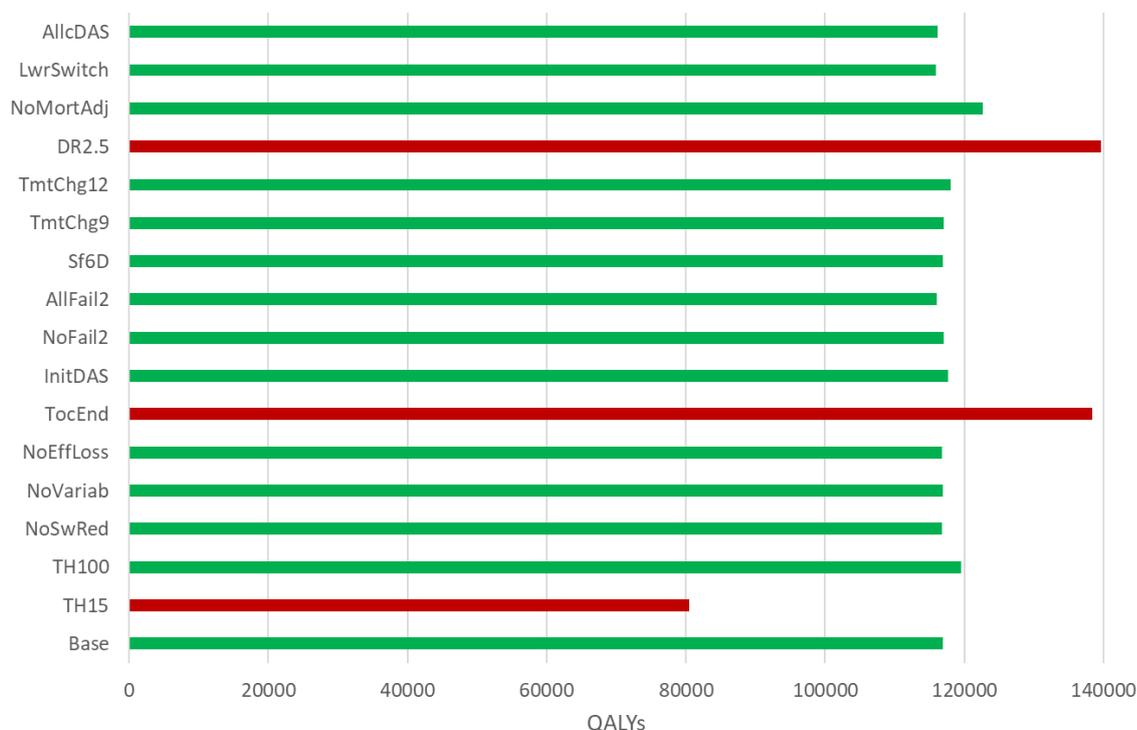
Base : analyse de référence ; **TH15** : horizon temporel de 15 ans ; **NoSwRed** : absence de réduction du taux de changement de traitement ; **NoVariab** : absence de variabilité de l'activité (DAS 28) ; **NoEffLoss** : absence de perte d'efficacité lors de l'utilisation d'un deuxième Anti-TNF ; **TocEnd** : tocilizumab pris en 4^e ligne biologique ; **NoFail2** : absence d'échecs par la règle 2 ; **AllFail2** : 100 % d'échecs par la règle 2 ; **SF-6D** : utilités calculées à partir du SF-6D ; **TmtChg9** : repousser la décision de changement de traitement à 9 mois ; **TmtChg12** : repousser la décision de changement de traitement à 12 mois ; **DR2.5** : taux d'actualisation égal à 2,5 % ; **NoMortAdj** : pas d'ajustement de la mortalité associée à la PR (HR=1) ; **LwrSwitch** : diminution moins importante du taux de changement de traitements ; **AllcDAS** : population incluant toutes les catégories du DAS.

➤ Les coûts

La variation des coûts totaux est également relativement limitée et les positions relatives des séquences changeaient très peu (figure 3). À l'instar des QALYs, un taux d'actualisation plus faible et un horizon temporel plus court avaient un impact majeur sur les coûts. En revanche, l'hypothèse d'absence de variabilité de l'activité de la PR n'affectait pas les coûts. L'utilisation du tocilizumab après la troisième ligne augmentait les coûts.

La prise en compte des seuls coûts associés à la pneumonie dans les événements indésirables avait un impact faible sur les coûts totaux actualisés des séquences de traitements. Par exemple, le coût de la séquence de référence ETAB-->ABA-->INF baissait simplement de 1,8 %.

Figure 3. Coûts totaux de la séquence ETAB-->ABA-->INF selon les scénarios d'analyse



➤ La prise en compte de la trithérapie comme stratégie de référence

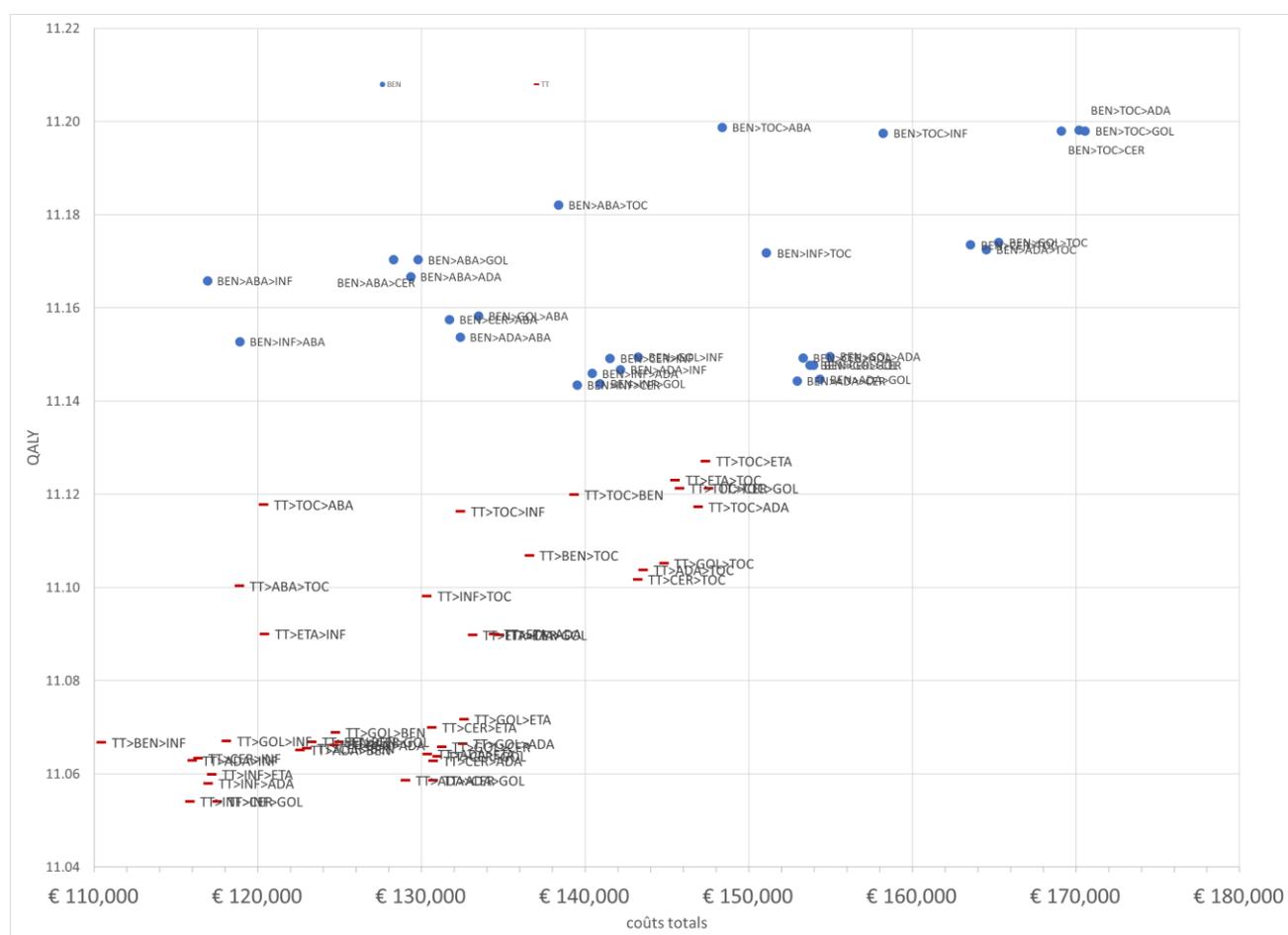
La prise en compte de la trithérapie en début de séquence (au lieu d'un traitement biologique) ajoute **54 séquences dans les stratégies de séquences à comparer**. La plupart d'entre elles entraînent des coûts similaires à ceux des séquences composées entièrement par des bDMARDs, mais une efficacité moindre et elles étaient donc dominées.

Toutefois, **une seule séquence, TT-->ETAB-->INF**, n'était pas dominée par la séquence débutant par le biosimilaire de l'étanercept la plus efficace **ETAB-->ABA-->INF**. Si on considère la séquence débutant par la trithérapie **TT-->ETAB-->INF** comme stratégie de référence, **le RDCR de la séquence débutant par le biosimilaire de l'étanercept devient 65 623 par QALY** (tableau 5).

Tableau 6. Résultats de l'analyse de référence – Scénario de la prise en compte de la trithérapie

Séquence	Total coût (€)	Total QALY	RDCR (€/QALY)
TT>ETAB>INF	110 410	11,0668	
ETAB>ABA>INF	116 912	11,1658	65 623
ETAB>TOC>ABA	148 363	11,1987	955 778

La figure 4 a été seulement construite dans le but d'avoir une idée claire sur la TT. En effet, dans ce cas, l'analyse des séquences a été faite sur 54 séquences.

Figure 4. Analyse des séquences de trithérapie avec celles débutant par le biosimilaire de l'éтанercept


➤ L'analyse de l'efficacité

L'analyse de l'efficacité des séquences de traitements restait inchangée dans la plupart des scénarios analysés et les RDCRs restaient supérieurs à 100 000 €/QALY (sauf le scénario incluant la trithérapie). Les séquences débutant avec le biosimilaire de l'éтанercept en 1^{re} ligne avaient tendance à être plus efficaces, en particulier avec l'abatacept en 2^e ligne. Comme dans l'analyse de référence, la séquence reposant sur l'utilisation de l'infliximab en 3^e ligne était la séquence la

moins coûteuse et la plus efficace. D'autres séquences pouvaient « rejoindre » la frontière d'efficacité selon le scénario en question. Par exemple, dans le scénario d'absence de variabilité de l'activité de la PR dans le DAS 28, quatre séquences débutant par l'étanercept rejoignaient la frontière d'efficacité (tableau 7). D'autres séquences débutant par ETAB étaient « très proches » de la frontière d'efficacité.

Tableau 7. Séquences situées sur la frontière d'efficacité : scénario d'absence de variabilité du DAS 28

Séquences	Total coût actualisé (€)	Total QALY	RDCR (€/QALY)
ETAB>ABA>INF	116 912	11,1658	
ETA>ABA>INF	128 131	12,5802	164 975
ETA>ABA>TOC	147 472	12,6276	407 946
ETA>TOC>ABA	161 325	12,6562	484 751
ETA>TOC>CER	180 582	12,6631	2 804 264

Pour une baisse simultanée des prix des bDMARDs de 50 %, une troisième séquence s'ajoute sur la frontière d'efficacité (ETAB>TOC>ADA). Même si le RDCR de la séquence ETA>ABA>INF vs ETAB>ABA>INF baisse environ de 40 %, les niveaux des RDCRs entre les séquences restaient supérieurs à 100 000 €/QALY.

2.2. Les analyses de sensibilité probabilistes

Les analyses probabilistes ont été réalisées à l'aide de l'interface DICE. L'approche d'interprétation des analyses probabilistes a été validée par le coordinateur projet, un économiste-statisticien et deux rhumatologues faisant partie du groupe de travail PR.

2.2.1. Choix des séquences dans l'analyse probabiliste

Le groupe de travail PR a émis un avis favorable à un choix pragmatique de séquences intégrées dans les analyses de sensibilité probabilistes :

1. il était techniquement difficile de réaliser des simulations probabilistes incluant 180 séquences, de gérer et d'interpréter un nombre très important de sorties des analyses probabilistes ;
2. à notre connaissance, et abstraction faite de la faisabilité des simulations sur toutes les séquences de l'analyse de référence, la littérature scientifique actuelle ne propose pas de méthodologie appropriée à ce cas de figure. **Dans le cadre de notre évaluation, l'analyse probabiliste de séquences de traitements est une problématique originale.**

L'analyse probabiliste s'est ainsi limitée à 10 séquences représentatives de la prise en charge de la PR parmi toutes les séquences incluses dans l'analyse de référence. Ces 10 stratégies (séquences) retenues incluaient les séquences déjà situées sur la frontière d'efficacité : **1) ETAB->ABA->INF ; 2) ETAB->TOC->ABA**. Les deux séquences **CERT-->-ABA-->INF** et **GOL-->-ABA-->INF** n'ont pas été retenues parce qu'elles devaient aboutir à des résultats équivalents. Le tableau 8 présente les 10 séquences retenues dans l'analyse probabiliste.

Tableau 8. Les dix séquences retenues dans l'analyse probabiliste

	Séquences retenues
1	ADA-->ABA-->INF
2	ETAB-->TOC-->INF
3	ETAB-->ADA-->INF
4	ETAB-->ABA-->ADA
5	ETAB-->ABA-->TOC
6	ADA-->ABA-->TOC
7	ADA-->ETAB-->TOC
8	ADA-->TOC-->ABA
9	ETAB-->ABA-->INF
10	ETAB-->TOC-->ABA

Il est à souligner qu'en présence d'un nombre très élevé de séquences, il n'a pas été possible d'établir le nombre optimal de simulations aléatoires requis pour la réalisation de l'analyse de sensibilité probabiliste. Toutefois, pour chaque séquence, 1 000 répliques ont été réalisées pour couvrir l'étendue de l'incertitude.

2.2.2. Spécifications des distributions des paramètres intégrés dans les analyses probabilistes

Le tableau 9 décrit les distributions des paramètres intégrés dans les analyses probabilistes. Les distributions des principaux paramètres ont été spécifiées en considérant notamment :

- les valeurs d'efficacité comparative (les probabilités de réponses EULAR) générées par le processus CODA ;
- des décompositions de Cholesky réalisées tenant compte des corrélations des paramètres d'intérêt, après avoir effectué :
 - les estimations de modèles mixtes à effets aléatoires régressant les scores d'utilité sur l'activité de la PR, la catégorie du HAQ et leurs interactions,

- des estimations de modèles linéaires généralisés à effets aléatoires régressant les catégories du HAQ sur l'activité de la PR,
- les probabilités du changement de l'activité d'une visite à la suivante générées par des distributions bêta non corrélées.

Les autres paramètres n'ont pas été intégrés dans les analyses probabilistes pour les raisons suivantes :

- les entrées non associées à des distributions statistiques telles que les prix et les taux d'actualisation ;
- les distributions statistiques pouvant être spécifiées d'une manière arbitraire mais discutables telles que les distributions des coûts et des événements indésirables entraînant des hospitalisations ;
- les paramètres dont les distributions statistiques ne pouvaient pas être spécifiées tels que la diminution des probabilités de réponses EULAR associées à la 2^e ligne de traitement après une réponse inadéquate à un anti-TNF α . Les données publiées dans Gutemberg *et al.* (2016) ne permettent pas d'estimer les paramètres de la distribution Dirichlet qui devait être la distribution statistique appropriée à ce cas.

Tableau 9. Spécifications des distributions relatives aux paramètres d'intérêt intégrés dans l'analyse probabiliste

Paramètre	Cas de référence	Type de distribution	a	b
Trithérapie <i>pGood</i>	0,19	CODA	NA	NA
Trithérapie <i>pModOrGood</i>	0,57	CODA	NA	NA
Infliximab <i>pGood</i>	0,28	CODA	NA	NA
Infliximab <i>pModOrGood</i>	0,69	CODA	NA	NA
Adalimumab <i>pGood</i>	0,3	CODA	NA	NA
Adalimumab <i>pModOrGood</i>	0,7	CODA	NA	NA
Étanercept <i>pGood</i>	0,56	CODA	NA	NA
Étanercept <i>pModOrGood</i>	0,89	CODA	NA	NA
Golimumab <i>pGood</i>	0,39	CODA	NA	NA
Golimumab <i>pModOrGood</i>	0,78	CODA	NA	NA
Certolizumab <i>pGood</i>	0,38	CODA	NA	NA
Certolizumab <i>pModOrGood</i>	0,77	CODA	NA	NA
InfliximabRem <i>pGood</i>	0,28	CODA	NA	NA
InfliximabRem <i>pModOrGood</i>	0,69	CODA	NA	NA
InfliximabInf <i>pGood</i>	0,28	CODA	NA	NA
InfliximabInf <i>pModOrGood</i>	0,69	CODA	NA	NA
InfliximabFli <i>pGood</i>	0,28	CODA	NA	NA
InfliximabFli <i>pModOrGood</i>	0,69	CODA	NA	NA
ÉtanerceptETAB <i>pGood</i>	0,52	CODA	NA	NA
ÉtanerceptETAB <i>pModOrGood</i>	0,69	CODA	NA	NA
Abatacept <i>pGood</i>	0,28	CODA	NA	NA
Abatacept <i>pModOrGood</i>	0,69	CODA	NA	NA
Tocilizumab <i>pGood</i>	0,63	CODA	NA	NA
Tocilizumab <i>pModOrGood</i>	0,92	CODA	NA	NA

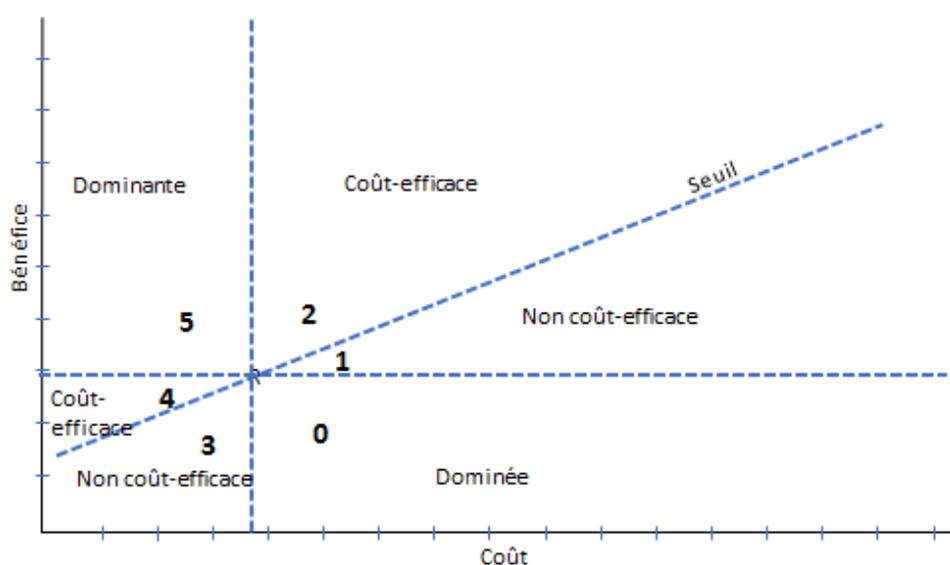
Tableau 9 (suite). Spécifications des distributions relatives aux paramètres d'intérêt intégrés dans l'analyse probabiliste

Paramètre	Cas de référence	Type de Distribution	a	b
Rituximab <i>pGood</i>	0,15	CODA	NA	NA
Rituximab <i>pModOrGood</i>	0,65	CODA	NA	NA
Probabilité de diminuer DAS	0,248	Bêta	145	440
Probabilité d'augmenter DAS	0,213	Bêta	125	461
HAQ GEE Intercept1	-1,7424	Cholesky/Normale	2,2548	0,1641
HAQ GEE Intercept2	-0,5853	Cholesky/Normale	-0,4403	0,1389
HAQ GEE Intercept3	0,7856	Cholesky/Normale	-0,3756	0,1133
HAQ GEE Intercept4	2,2128	Cholesky/Normale	0,0120	0,1008
HAQ GEE activité	3,1986	Cholesky/Normale	0,0349	-0,1619
HAQ GEE activité	2,3975	Cholesky/Normale	0,5995	-0,1590
HAQ GEE activité	1,391	Cholesky/Normale	0,0561	-0,1505
HAQ GEE activité	0	Cholesky/Normale	0,0000	0,0000
EQ5D Intercept	-0,01202	Cholesky/Normale	-0,4034	-0,0067
EQ-5D utilité0	0,1761	Cholesky/Normale	-0,8583	0,0017
EQ-5D activité	0,1705	Cholesky/Normale	-0,3980	0,0002
EQ-5D activité	0,1267	Cholesky/Normale	2,7783	0,0002
EQ-5D activité	0,08656	Cholesky/Normale	0,3013	0,0002
EQ-5D cHAQ	0,5902	Cholesky/Normale	-0,1092	0,0053
EQ-5D cHAQ	0,4661	Cholesky/Normale	-0,7010	0,0053
EQ-5D cHAQ	0,3189	Cholesky/Normale	-0,4042	0,0053
EQ-5D cHAQ	0,1535	Cholesky/Normale	1,4293	0,0051

2.2.3. Traitement et interprétation des résultats des ASP

Habituellement, les sorties des analyses de sensibilité probabilistes (ASP) sont présentées sous forme de nuage de dispersion situé sur un plan coût/efficacité et/ou une courbe d'acceptabilité coût-efficacité. Le calcul se fait en comptant la proportion de simulations dans lesquelles une intervention se situe dans le **secteur 2** « coût-efficace » selon un seuil (disposition à payer) donné par rapport au **secteur 1** « non coût-efficace » (figure 5). Dans l'analyse probabiliste, le seuil varie et les proportions correspondant à chaque seuil sont tracées. Ce processus est raisonnablement faisable lorsqu'une intervention de référence est clairement définie (c'est-à-dire la moins coûteuse) car seuls les **secteurs 1 et 2** sont en jeu, même si une ou plusieurs interventions peuvent occasionnellement tomber dans le **quadrant dominé 0**.

Figure 5. Les secteurs permettant de construire la courbe d'acceptabilité coût-efficacité



Comme notre évaluation mobilise un nombre très important de séquences de traitements (180 séquences) et comme nous ne disposons pas **d'une séquence standard** pouvant être facilement assimilée à une stratégie de référence, il devient difficile d'interpréter les sorties de l'ASP. Un nuage de points sera difficile à lire car il y aura dix mille points. De plus, il n'est pas également évident d'imaginer comment construire une courbe d'acceptabilité : en effet, pour n'importe quelle simulation, la séquence de référence peut être différente des autres séquences, et l'interprétation de la position de chaque séquence dépend du fait qu'il s'agit ou non de la séquence de référence. Si la séquence analysée n'est pas la séquence de référence, les mêmes règles de l'approche classique des ASP (c'est-à-dire : elles doivent se situer dans l'un des secteurs 0-2) s'appliquent ;

mais si la séquence analysée est la séquence de référence, alors la question est de savoir où sont situées les autres séquences.

Pour proposer une solution à cette situation, nous pouvons considérer les six secteurs, y compris ceux qui impliquent un coût moindre [3-5]. Lorsqu'une séquence n'est pas la séquence de référence, les secteurs 0-2 s'appliqueront (par rapport à la séquence de référence). Lorsqu'une séquence analysée est la séquence de référence, nous pouvons attribuer un score de 3 à 5 pour indiquer où se trouvent les autres séquences.

Les étapes de ce processus sont les suivantes :

1. pour chaque simulation, déterminer **quelle est la séquence de référence** (la moins coûteuse) ;
2. pour les séquences ne représentant pas des séquences de référence, déterminer dans quel secteur **elles se situent (0-2)** ;
3. pour **la séquence de référence**, attribuer **5** si toutes **les autres séquences sont dominées**, **4** s'il y en a au moins une qui **est coût-efficace**, sinon **3** ;
4. pour chaque séquence, dériver **la distribution de fréquence pour les scores associés aux différents secteurs** ;
5. **répéter pour d'autres valeurs du seuil** (ceci affectera seulement les scores 1-4).

2.2.4. Résultats

Les figures 6 à 8 présentent les distributions des dix séquences par scores associés aux différents secteurs pour trois dispositions à payer (30 000 €/QALY ; 50 000 €/QALY et 100 000 €/QALY). Les autres sorties (i.e. tableaux correspondant aux distributions des bénéfices monétaires nets) sont données dans l'annexe 6.3.

Les résultats des analyses probabilistes étaient en accord avec les résultats de l'analyse de référence et des analyses de scénarios déterministes.

Pour les trois dispositions à payer, les distributions des séquences selon les scores associés aux différents secteurs coût-efficacité étaient comparables. La séquence **ETAB-->ABA-->INF** dominait plusieurs fois la plupart des séquences mais jamais toutes en même temps. Dans 83 % des simulations, elle était la séquence de référence et il n'existait pas une autre séquence qui était coût-efficace. Dans 15 à 17 % des simulations, il y avait au moins une autre séquence qui était coût-efficace (ADA->ABA->INF). Parmi les neuf séquences systématiquement dominées et/ou non

coût-efficaces, les profils des « sous-groupes » de distributions de séquences suivants étaient assez similaires :

- **ADA-->ABA->INF** et **ETAB-->ADA-->INF** ;
- **ADA-->ABA-->TOC**, **ADA-->ETAB-->TOC** et **ADA-->TOC-->ABA** ;
- **ETAB-->TOC-->INF** et **ETAB-->TOC-->ABA**.

À titre informatif, si la disposition à payer était égale à 100 000 €/QALY, les séquences **ETAB-->ABA-->ADA** et **ETAB-->ABA-->TOC** devenaient coûts-efficaces par rapport à **ADA-->ABA-->INF** comme séquence de référence dans 13 % (respectivement 4 %) des simulations.

Figure 6. Analyse probabiliste : distribution des séquences selon les scores associés aux différents secteurs coût-efficacité – Disposition à payer : 30 000 €/QALY

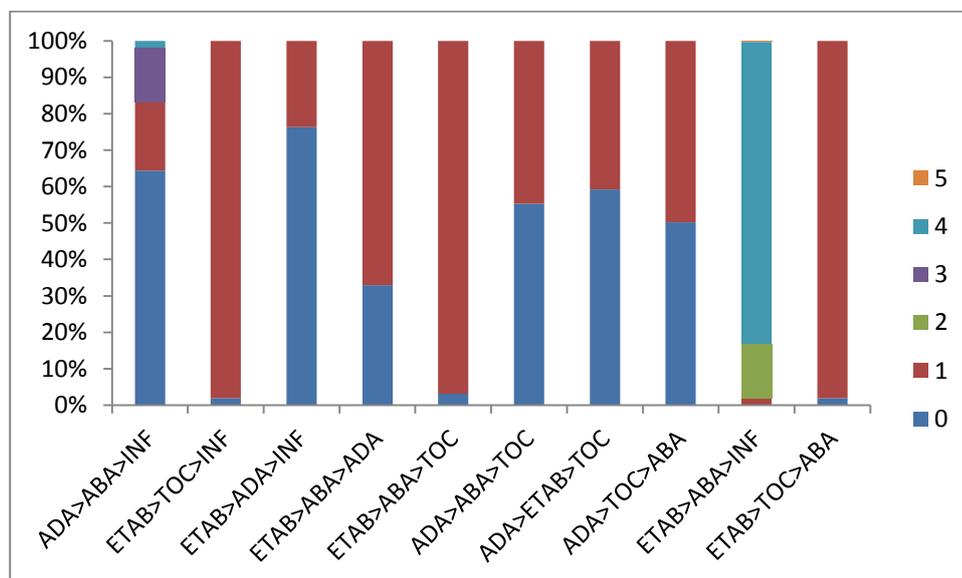


Figure 7. Analyse probabiliste : distribution selon les scores associés aux différents secteurs coût-efficacité – Disposition à payer 50 000 €/QALY.

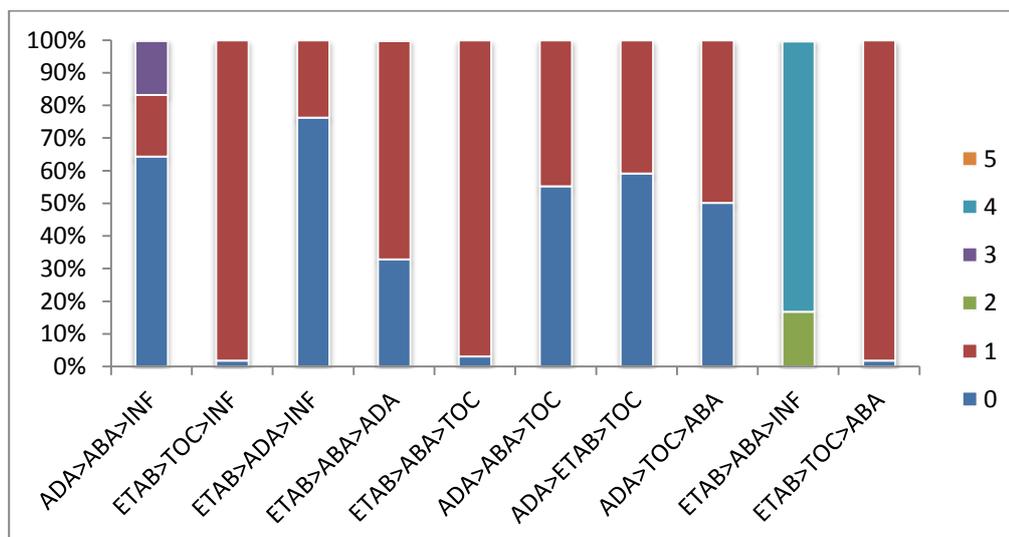
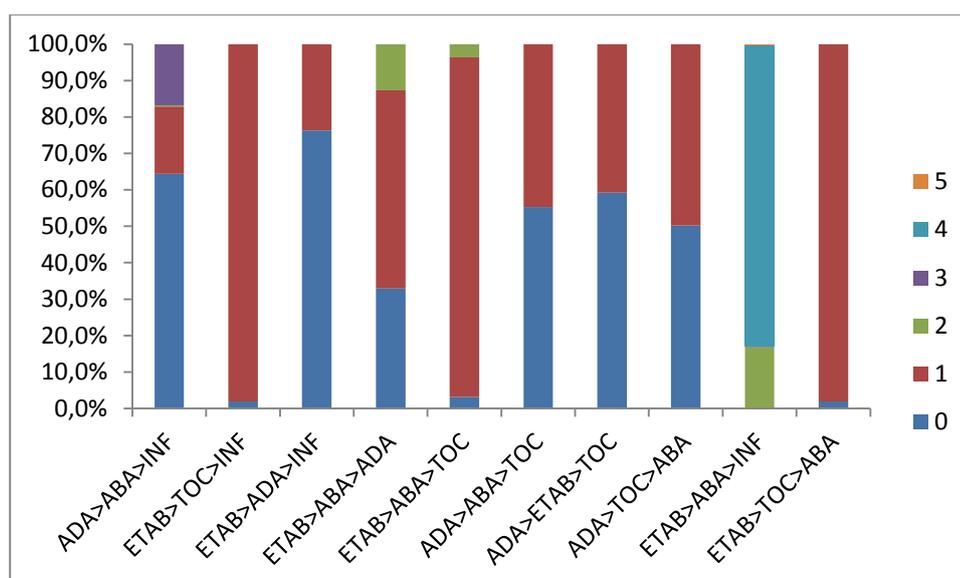


Figure 8. Analyse probabiliste : distribution selon les scores associés aux différents secteurs coût-efficacité – Disposition à payer 100 000 €/QALY.



Légende :

- **Quand la séquence n'est pas celle de référence :** 0=Dominée, 1=Non coût-efficace (CE), 2=Coût-efficace pour l'analyse référence.
- **Quand la séquence est celle de référence :** 3=Au moins une autre séquence est CE, 4=Aucune autre séquence est CE, 5=Toute autre séquence est dominée.

3. Analyses complémentaires

3.1. Analyses complémentaires n° 1

3.1.1. Rationnel

En réponse à une proposition des rhumatologues du groupe de travail PR, des analyses complémentaires ont été réalisées. L'objectif de ces analyses était d'intégrer certaines séquences de bDMARDs qui n'ont pas été incluses dans le rationnel de l'analyse de référence : par exemple, **des séquences débutant la première ligne biologique par un non-anti-TNF α (ex. : TOC ou ABA).**

3.1.2. Séquences sélectionnées

Le choix a porté sur les huit séquences suivantes qui ont été ajoutées aux séquences situées sur la frontière d'efficacité de l'analyse de référence (**ETAB-->ABA-->INF** et **ETAB-->TOC-->ABA**) :

- TOC-->ABA-->ETAB ;
- TOC-->ABA-->ADA ;
- TOC-->ADA-->ABA ;
- TOC-->ADA-->ETAB ;
- ABA-->TOC-->ETAB ;
- ABA-->TOC-->ADA ;
- ABA-->ETAB-->TOC ;
- ABA-->ADA-->TOC.

Légende: ABA : abatacept ; TOC : tocilizumab ; INF : infliximab ; ETAB : biosimilaire de l'étanercept ; ADA : adalimumab.

3.1.3. Données utilisées

Les données utilisées étaient les mêmes que celles utilisées dans l'analyse de référence. Les données de sources étrangères n'ont pas été incluses.

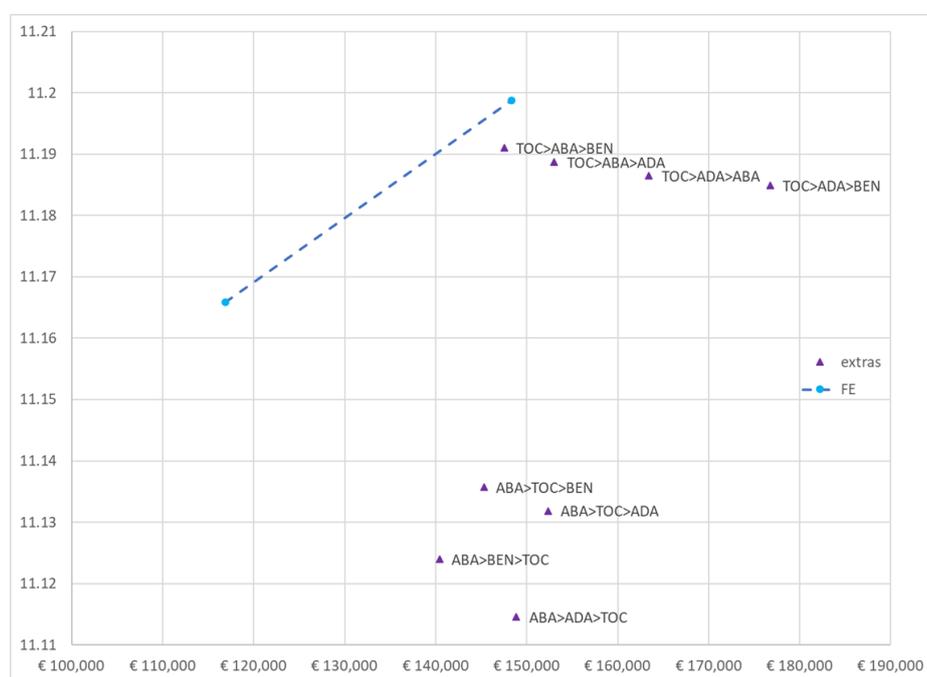
3.1.4. Résultats

Les résultats (coût actualisé, QALY total et RDCR) sont présentés dans le tableau 10 et la frontière d'efficacité est présentée dans la figure 9. Des résultats similaires à l'analyse de référence ont été obtenus. **Ceci montre que même si l'on modifie le choix de certaines séquences de traitements en proposant d'autres bDMARDs en première ligne biologique (ex. : TOC, ABA), les séquences ETAB->ABA->INF et ETAB->TOC->ABA restent toutes les deux sur la frontière d'efficacité.**

Tableau 10. Analyse complémentaire : résultats incluant des séquences ne faisant pas partie de l'analyse de référence

Séquences	Total coût actualisé (€)	Total QALY	RDCR (€/QALY)
ETAB-->ABA-->INF	116 912	11,1658	Référence
ETAB-->TO-->ABA	148 363	11,1987	955 778

Figure 9. Analyse complémentaire : frontière d'efficacité incluant des séquences ne faisant pas partie de l'analyse de référence



3.2. Analyses complémentaires n° 2

3.2.1. Rationnel

Les tsDMARDs (tofacitinib et baricitinib), le nouveau bDMARD (ex. : sarilumab) et biosimilaire (ex. : IMRALDI, biosimilaire de l'adalimumab) ont fait l'objet d'une analyse de scénario complémentaire. En effet, ces produits n'ont pas été inclus dans l'analyse de référence parce qu'ils ne faisaient pas partie des méta-analyses en réseau réalisées, étant donné qu'ils ont été évalués par la CT seulement fin 2017 et en 2018. **Le groupe de travail PR et certains experts du groupe de lecture ont proposé d'explorer l'évaluation économique de ces produits tout en reconnaissant notamment la disponibilité réduite des données concernant les tsDMARDs.**

À cet égard, une sélection de séquences intégrant ces produits a été approuvée par les rhumatologues du groupe de travail PR. Le but était d'explorer des exemples de séquences dans lesquels les tsDMARDs (tofacitinib, baricitinib) et le nouveau bDMARD (sarilumab) pourraient être prescrits conformément à leur AMM européenne (prescription après échec d'au moins un traitement de fond (synthétique ou biologique ou synthétique ciblé)).

3.2.2. Séquences sélectionnées

Trente séquences ont été sélectionnées (tableau 11) :

- des séquences incluant les tsDMARDs en 1^{re}, en 2^e et en 3^e lignes de traitement de fond ciblé (biologique et/ou synthétique) ;
- des séquences incluant le sarilumab en 1^{re} ou en 2^e ou en 3^e ligne de traitement de fond biologique ;
- des séquences incluant le biosimilaire de l'adalimumab (ADB) dans une sélection de lignes de traitements de fond biologiques ;
- des séquences conservant en première ligne le biosimilaire de l'étanercept et en faisant varier les 2^e et 3^e lignes de traitement entre bDMARD et/ou tsDMARDs ;
- des séquences intégrant le rituximab en 3^e ligne biologique.

Tableau 11. Séquences incluses dans l'analyse intégrant les tsDMARDs et sarilumab

Séquences
ABA-->SAR-->RTX
ADB-->ABA-->INF
ADB-->SAR-->ABA
ADB-->TOC-->ABA
BAR-->ABA-->INF
BAR-->ABA-->TOC

BAR-->TOC-->ABA
ETAB-->ABA-->ADB
ETAB-->ABA-->BAR
ETAB-->ABA-->SAR
ETAB-->ABA-->TOF
ETAB-->ADB-->ABA
ETAB-->ADB-->INF
ETAB-->BAR-->ABA
ETAB-->BAR-->INF
ETAB-->SAR-->ABA
ETAB-->SAR-->INF
ETAB-->TOC-->ADB
ETAB-->TOC-->BAR
ETAB-->TOC-->TOF
ETAB-->TOF-->ABA
ETAB-->TOF-->INF
SAR-->ABA-->INF
SAR-->ABA-->RTX
SAR-->ABA-->TOF
SAR-->ADB-->ABA
SAR-->ETAB-->ABA
TOF-->ABA-->INF
TOF-->ABA-->TOC
TOF>TOC>ABA

Légende: ABA : abatacept ; SAR : sarilumab ; RTX : rituximab ; TOC : tocilizumab ; BAR : baricitinib ; INF : infliximab ; ETAB : biosimilaire de l'étanercept ; TOF : tofacitinib.

3.2.3. Données utilisées

La source des données utilisée pour estimer les probabilités de réponses EULAR relatives aux baricitinib, tofacitinib et sarilumab provenait des données des méta-analyses en réseau réalisées par l'ICER en 2017 (13) portant sur des patients atteints de PR ayant été traités par des csDMARDs et/ou des bDMARDs (annexe A6.4). Les caractéristiques de cette population mixte étaient assez comparables à celles des AMM européennes de ces trois produits et à celles de la population de l'analyse de sensibilité de la méta-analyse HAS & SchARR n° 2 (cf. section 4.3.2.

Analyse de sensibilité, page 115 et annexe A6.4) qui incluait des réponses ACR provenant d'essais cliniques dans lesquels une proportion des patients avait été traitée par des bDMARDs.

L'évaluation économique de l'ICER était principalement fondée sur la simulation des probabilités de réponses ACR. Afin d'estimer les probabilités de réponses EULAR relatives aux baricitinib et sarilumab, les réponses définies en termes de probabilité ACR de ces produits ont été estimées en se fondant sur la distribution des probabilités de réponses ACR ($f(\text{ACR}20)$, $f(\text{ACR}50)$, $f(\text{ACR}70)$) et transformées en probabilités EULAR selon les formules suivantes :

- **Probabilité bonne réponse EULAR :**

$$p_{\text{Bonne}} = p_{\text{Bonne}|\text{MoinsdeACR}20} * f_{\text{MoinsdeACR}20} + p_{\text{Bonne}|\text{ACR}20} * f_{\text{ACR}20} + p_{\text{Bonne}|\text{ACR}50} * f_{\text{ACR}50} + p_{\text{Bonne}|\text{ACR}70} * f_{\text{ACR}70}$$

- **Probabilité réponse EULAR modérée :**

$$p_{\text{Modérée}} = p_{\text{Mod}|\text{MoinsdeACR}20} * f_{\text{MoinsdeACR}20} + p_{\text{Mod}|\text{ACR}20} * f_{\text{ACR}20} + p_{\text{Mod}|\text{ACR}50} * f_{\text{ACR}50} + p_{\text{Mod}|\text{ACR}70} * f_{\text{ACR}70}$$

où $p_{X|\text{niveau ACR}}$ est la probabilité du résultat EULAR conditionnel au niveau atteint ACR et $f_{\text{Niveau ACR}}$ est la fraction des patients avec le niveau ACR donné pour chaque médicament.

Des exemples de transformation des probabilités de réponses ACR en EULAR existent dans la littérature économique. Dans l'évaluation des bDMARDs du NICE en 2016, Stevenson *et al.* (2016) (12) ont utilisé et validé une formule de transformation « *mapping* » des probabilités de réponses ACR en probabilités EULAR. Jansen *et al.* (2017) (133) ont également appliqué la formule de Stevenson *et al.* (2016) dans le cadre d'une évaluation économique portant sur des bDMARDs et des tsDMARDs aux États-Unis. Dans notre cas, les formules de transformation des probabilités ACR en probabilités EULAR suivaient le même principe de Stevenson *et al.* (2016), mais présentaient quelques différences ; en l'occurrence, dans notre cas, les formules de transformation ont été utilisées sous forme agrégée (le modèle PR France ne simulait pas les probabilités de réponses ACR à l'échelle individuelle), alors que dans les modèles de Jansen *et al.* (2017) (133) et Stevenson *et al.* (2016) (12), l'effet du traitement a été obtenu à l'aide des probabilités de réponses ACR issues des méta-analyses en réseau, et lors de la simulation de chaque patient dans ces modèles, les auteurs assignaient une réponse EULAR en utilisant la distribution des probabilités de réponses ACR.

Le tableau 12 donne les probabilités de réponses EULAR estimées, en transformant les résultats ACR obtenus dans trois méta-analyses (HAS référence, HAS sensibilité, ICER). Pour le tofacitinib,

les estimations des probabilités EULAR ont été obtenues en transformant les résultats de l'analyse de sensibilité HAS & SchARR ; et pour le baricitinib et le sarilumab, les estimations des probabilités de réponses EULAR ont été obtenues en transformant les résultats ICER (en jaune).

Tableau 12. Probabilités des réponses EULAR utilisées dans l'analyse de scénario

	HAS & SchARR (Analyse de référence)			HAS & SchARR (Analyse de sensibilité)			ICER (Analyse de référence)		
	pNulle	pModérée	pBonne	pNulle	pModérée	pBonne	pNulle	pModérée	pBonne
ABT i.v.+	30%	28%	42%	32%	28%	40%	30%	27%	43%
ABT s.c.+	25%	26%	49%	29%	27%	44%	30%	28%	42%
ADA+	26%	27%	47%	30%	28%	42%	32%	28%	40%
CTZ+	30%	29%	42%	20%	23%	58%	22%	24%	55%
ETN+	22%	24%	54%	23%	25%	52%	21%	23%	56%
GOL+	24%	26%	49%	26%	27%	47%	30%	27%	43%
IFX+	29%	28%	43%	31%	29%	40%	32%	28%	40%
TCZ+	20%	24%	56%	25%	25%	50%	28%	26%	46%
TOF 10mg+				28%	27%	45%	32%	29%	39%
Baricitinib +							30%	28%	42%
Sarilumab +							28%	27%	45%

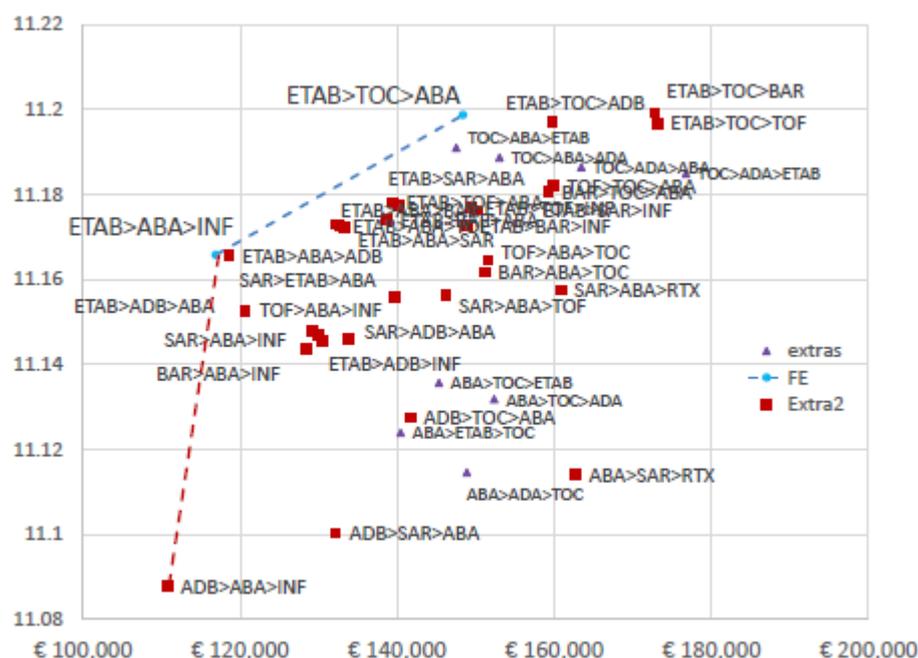
N.B : Le signe (+) indique une association du produit avec le méthotrexate.

Comme l'adalimumab et son biosimilaire sont des traitements équivalents, les estimations des réponses EULAR relatives au biosimilaire de l'adalimumab ont été approximées par les probabilités EULAR relatives à l'adalimumab qui provenaient des méta-analyses de SchARR & HAS (84). En revanche, le prix d'acquisition du biosimilaire de l'adalimumab a été utilisé dans les simulations du modèle PR France.

3.2.4. Résultats

Les résultats de l'analyse de scénario complémentaire sont résumés dans la figure 10 et le tableau 12. L'examen de la figure 10 permet de constater les éléments suivants :

- une séquence incluant l'ADAB se situait sur la frontière d'efficacité. Elle était la séquence la moins coûteuse ;
- les séquences contenant le tofacitinib ou le baricitinib ne se situaient pas sur la frontière d'efficacité.

Figure 10. Frontière d'efficacité – Analyse complémentaire 2

Légende : **extras** : sont les analyses incluant TOC ou ABA en première ligne biologique réalisées dans l'analyse complémentaire n° 1 ; **Extra2** : sont les analyses avec les nouvelles séquences (analyse complémentaire Tr 2) ; **FE** : frontière d'efficacité.

Le tableau 13 synthétise les résultats de l'estimation de cette analyse complémentaire pour les séquences qui n'ont pas été dominées.

Tableau 13. Résultats de l'estimation de l'analyse de scénario complémentaire

Séquence	Total coût actualisé (€)	Total QALY actualisé	RDCR
ADAB>ABA>INF	110 724	11,0879	
ETAB>ABA>INF	116 912	11,1658	79 407
ETAB>TOC>ABA	148 363	11,1987	17 878 619

Les résultats de cette analyse complémentaire complètent ceux de l'analyse de référence. Pratiquement toutes les analyses de scénarios réalisées ont montré que les séquences débutant par un biosimilaire d'un anti-TNF (étanercept ou adalimumab) étaient situées sur la frontière d'efficacité et pouvaient même être les séquences les moins coûteuses : la séquence débutant par le biosimilaire de l'étanercept (ETAB>ABA>INF) dans l'analyse de référence, et la séquence débutant par le biosimilaire de l'adalimumab (ADAB>ABA>INF) qui devenait dans cette analyse complémentaire la séquence la moins coûteuse.

Le RDCR de la séquence ETAB>ABA>INF vs ADAB>ABA>INF était égal à 79 407 €/QALY. En termes d'ordre de grandeur absolue, ce RDCR était inférieur aux RDCRs obtenus dans l'analyse de référence et les analyses de scénarios qui étaient supérieurs à 100 000 €/QALY à quelques exceptions près. Il demeurait largement supérieur au seuil de disposition à payer accepté au

Royaume-Uni. Cependant, force est de constater qu'une baisse différenciée de prix plus soutenue d'un biosimilaire (ex. : biosimilaire de l'adalimumab dans cette analyse complémentaire) pourrait faire diminuer le RDCR de manière plus importante.

Les séquences contenant le baricitinib et le tofacitinib ne se situaient pas sur la frontière d'efficience mais elles n'étaient pas trop éloignées de cette courbe. Ce résultat paraissait en partie compréhensible compte tenu de l'absence de comparaison de ces produits avec les bDMARDs en 3^e ligne de traitement de la PR et, d'une manière générale, du faible recul de ces traitements en termes d'efficacité et de tolérance, comparativement aux bDMARDs comme l'indiquaient les avis de la CT relatifs à ces produits (baricitinib [\(41\)](#), tofacitinib [\(80\)](#)).

4. Discussion

Les résultats de l'évaluation économique ont été estimés en accord avec les choix structurants du guide de la HAS pour l'évaluation économiques (horizon vie entière, taux d'actualisation en vigueur, perspective collective, analyse coût-utilité adaptée aux spécificités de la pathologie) et les recommandations de l'ISPOR sur les bonnes pratiques de modélisation. Comme l'approche standard de l'exploration de l'incertitude n'aborde pas les analyses probabilistes de séquences de traitements, une approche méthodologique exploratoire permettant une interprétation des résultats a été proposée.

- Les résultats de santé

Les QALYs ont été estimés à partir du questionnaire EQ-5D conformément aux recommandations du guide de la HAS (en analyse de référence) et du questionnaire SF-6D (en analyse de scénarios déterministes). La comparaison des distributions des scores d'utilité extraits de l'EQ-5D et du SF-6D montrait des résultats différents. Le SF-6D avait un plus grand niveau de sensibilité au changement en cas d'amélioration. L'EQ-5D était plus sensible à la détérioration. Ces constats étaient en accord avec les conclusions des travaux de Gaujoux-Viala *et al.* (2013) (26) sur les instruments de mesure de l'utilité dans la PR en France.

En termes de différentiel de QALY entre les séquences de traitements de fond biologiques, les très faibles différences de QALYs entre les séquences paraissaient cliniquement plausibles. Une certaine comparabilité de l'amplitude du différentiel de QALY a été observée dans les évaluations récentes de l'ICER (2017) et de Jansen *et al.* (2017). À titre indicatif, le différentiel entre les stratégies de bDMARDs présentant les valeurs de QALY maximale et minimale dans Jansen *et al.* (2017) était de 0,063. En revanche, le différentiel de QALY des séquences de bDMARDs vs séquences de csDMARDs était nettement supérieur aux différentiels de QALY inter-séquences. Il est à noter que ces deux études américaines reposaient sur des modèles de micro-simulation différents de PR France et approximaient l'évolution de l'activité de la maladie à long terme à l'aide du critère de l'incapacité fonctionnelle (HAQ), contrairement au modèle PR France qui est fondé sur l'activité de la maladie.

À défaut de données montrant l'efficacité radiologique des bDMARDs ainsi que leurs effets sur les années de vie gagnées, l'efficacité relative de chaque bDMARD a été seulement traduite par la réponse EULAR. Ainsi, une bonne réponse EULAR pourrait ne pas se traduire par un gain de long terme optimal, à la fois en termes d'amélioration de l'incapacité fonctionnelle, d'amélioration de l'état mental et de baisse de la douleur en raison de la forte variabilité de l'activité de la maladie observée en vie réelle. Ceci pourrait être illustré par une analyse de scénario extrême (en faveur des bDMARDs) où la variabilité de l'activité de la PR serait artificiellement éliminée : les QALYs (en termes de comparaisons intra-séquences) augmentaient de plus d'un an. Cependant, en

termes de différentiel de QALY, les séquences situées sur la frontière d'efficacité de cette analyse de scénario restent proches (cf. tableau 5).

Une discussion sur les limites du QALY et les perspectives complémentaires à son utilisation dans la PR sont intégrées dans la conclusion générale du projet.

- *Les coûts*

Les séquences de traitements se différencient plus facilement en termes de coûts qu'en termes de QALY. Cette différenciation a été plus marquée lorsque les séquences débutaient par le biosimilaire de l'étanercept (ETABepali®). À l'instar des conclusions des études françaises sur les coûts (Fautrel *et al.*, 2016 (8) ; Beck *et al.*, 2015 (15)), les coûts d'acquisition des bDMARDs représentaient la part la plus importante des coûts totaux.

Il est important de rappeler que notre évaluation économique s'est principalement focalisée sur les postes de coûts permettant d'analyser un différentiel de coûts entre les bDMARDs. Une analyse exhaustive des coûts, au-delà des postes clés présentant des différences entre les traitements de fond synthétiques, ne faisait pas l'objet de la problématique de ce projet.

- *L'analyse de l'efficacité*

Les valeurs de RDCR étaient élevées et supérieures aux seuils retenus dans certains pays tels que le Royaume-Uni, et ce, même dans le scénario de l'absence de variabilité du DAS 28. Dans une analyse de scénarios sur les utilités, l'utilisation des scores d'utilités estimés à partir du SF-6D, qui avait un plus grand niveau de sensibilité au changement en cas d'amélioration de l'état du patient, aboutissait à un RDCR encore plus élevé que celui obtenu en utilisant les utilités estimées à partir de l'EQ-5D.

Cette évaluation économique est, à notre connaissance, la première évaluation de technologies de santé qui aborde :

- la pertinence de l'intégration des biosimilaires dès la première ligne de traitement biologique dans la prise en charge de la PR en France. Deux séquences, toutes deux commençant par le biosimilaire de l'étanercept, étaient sur la frontière de l'efficacité de l'analyse de référence. Ce résultat est confirmé par l'analyse complémentaire intégrant le biosimilaire de l'adalimumab : deux séquences, l'une commençant par le biosimilaire de l'adalimumab (la moins coûteuse), l'autre commençant par le biosimilaire de l'étanercept, étaient sur la frontière d'efficacité ;

- l'importance de l'intégration de la trithérapie dans une stratégie de prise en charge de la PR. L'évaluation l'intègre en effet en 1^{re} ligne « biologique », avant un biosimilaire en 2^e ligne biologique. Aujourd'hui, l'utilisation de la trithérapie est très peu fréquente en France, comme démontré dans la cohorte ESPOIR. Les enjeux cliniques et économiques relatifs à l'utilisation de la trithérapie pour des patients présentant une réponse inadéquate au MTX ont été abordés dans des essais cliniques tels que l'essai RACAT (O'Dell *et al.*, 2013), les recommandations EULAR (2013)

et les évaluations économiques (Bansback *et al.*, 2017). Notre analyse de l'efficacité explorant l'utilisation de la trithérapie suggère que la séquence TT-->ETAB-->INF est la moins coûteuse parmi les 234 séquences analysées (les 180 séquences de l'analyse de référence plus 54 séquences débutant par la trithérapie). La prise en compte de la trithérapie pour des patients présentant une réponse inadéquate au MTX pourrait réduire davantage les coûts si les résultats des études françaises en vie réelle concordaient avec les résultats des essais cliniques réalisés aux USA et montraient une bonne adhérence et une acceptation des patients de son mode d'administration par rapport aux anti-TNF α et aux autres traitements de fond ciblés. À ce propos, un programme hospitalier de recherche clinique (PHRC) comparant les bDMARDs à la trithérapie est en cours (BIO-3 trial, 2014) ([177](#)). La publication des résultats de ce PHRC est prévue en février 2020.

Comparaison des résultats du modèle PR France aux autres modèles

La comparaison des résultats de cette évaluation à d'autres se heurte aux difficultés suivantes :

- la très forte variabilité multidimensionnelle des résultats identifiés dans la revue systématique de la littérature (ex. : choix structurants, indication analysée, variabilité inter-temporelle, variabilité intra-pays, variabilité inter-pays) ;
- les séquences de traitements de fond biologiques étaient systématiquement comparées à des séquences de traitements de fond synthétiques dans la littérature (ex. : ICER, 2017 ([13](#)) ; Stevenson *et al.*, 2016 ([12](#))) ;
- toutes les séquences de traitements de fond biologiques se limitaient à un nombre très réduit de séquences qui différaient souvent par leur début. De plus, à l'exception de quelques études (ICER, 2017 ([13](#)), Jansen *et al.*, 2017 ([133](#))) qui proposaient au moins deux lignes différentes dans chaque séquence, les autres évaluations se limitaient à des séquences qui se distinguaient seulement par la première ligne de traitement biologique (ex. : Stevenson *et al.*, 2016 ([12](#))) ;
- la modélisation selon la prise en charge de la PR dans le pays en question, les pratiques cliniques influençant le choix des variables décrivant l'évolution de l'activité de la PR (DAS 28 vs HAQ). Le modèle de Sheffield utilisé dans l'évaluation du NICE de 2015 (Stevenson *et al.*, 2016) se fondait sur la réponse EULAR mais ne reposait pas sur l'évolution de l'activité de la PR selon le DAS 28. **De ce fait, il n'était pas adapté à la modélisation de la prise en charge de la PR en France : ses simulations reposaient essentiellement sur l'évolution de l'indicateur d'incapacité fonctionnelle (le HAQ).** Au même titre que le modèle de Sheffield, les modèles utilisés dans les études américaines de l'ICER

(2017) et Jansen *et al.* (2017) reposaient également sur l'évolution du HAQ. L'annexe 6.4 propose une comparaison des principaux éléments de la structure de ces modèles par rapport à PR France.

À défaut de support méthodologique commun entre ces modélisations et à titre informatif, les résultats de santé et/ou de RDCRs identifiés dans les évaluations récentes retenues dans la grille d'analyse de la littérature (ICER, 2017 ; Jansen *et al.*, 2017 et Stevenson *et al.*, 2016) vont grossièrement dans le même sens que certains résultats du modèle PR France. Par exemple, dans les deux études américaines, le différentiel de QALY entre les séquences de traitements de fond biologiques était faible, à peu près égal à 0,063 dans l'étude de Jansen *et al.* (2017) (133). Cette dernière utilisait la réponse EULAR pour décrire l'efficacité comparative à court terme, approximait l'évolution de l'activité de la PR par l'incapacité fonctionnelle (i.e. le HAQ) et estimait les utilités à l'aide d'une fonction de transformation du HAQ.

De même, dans les deux études américaines, les RDCRs étaient supérieurs à 100 000 €/QALY et pourtant le comparateur commun était une séquence de traitements de fond synthétiques. Dans une moindre mesure, l'évaluation de Stevenson *et al.* (2016) (12) comparant des séquences de traitements biologiques (et non biologiques), qui se différençaient uniquement au niveau de la 1^{re} ligne de traitement biologique, aboutissait à des RDCRs supérieurs à 45 000 €/QALY.

Une mise à jour de cette évaluation devrait pallier notamment les limites suivantes :

- comme la plupart des évaluations économiques identifiées dans la revue systématique de littérature économique, PR France ne modélisait pas l'adhésion. Il est admis en pratique que l'adhésion aux anti-TNF α est assez similaire. Néanmoins, la collecte des données comparatives sur l'adhésion de la trithérapie est plus que jamais nécessaire pour confirmer/infirmar les gains économiques résultant de son introduction dans la prise en charge des patients présentant une réponse inadéquate à un traitement de fond synthétique (i.e. MTX). Une mauvaise observance peut impacter de façon plus ou moins grande l'efficacité définie dans l'étude contrôlée vs biothérapies et donc diminuer l'adhésion à la trithérapie ;
- le choix de la cohorte ESPOIR était certainement optimal par rapport aux données françaises disponibles en vie réelle. Cependant, la combinaison d'autres sources de données médico-administratives avec la cohorte ESPOIR permettrait non seulement d'intégrer des séquences fondées sur d'autres stratégies de traitements (ex. : la diminution des doses en cas de rémission établie), mais aussi d'améliorer l'estimation des hospitalisations liées aux événements indésirables sévères/graves en termes d'incidence pour toutes les classes de médicaments. De ce fait, l'aboutissement du projet de l'appariement des données de la cohorte ESPOIR avec les données médico-administratives (SNIIRAM) est très attendu ;

- la prise en compte d'une perspective plus large tenant compte des restes à charge et de certains types de coûts tels que des coûts directs non médicaux et des coûts indirects est conditionnée par la disponibilité de données pertinentes sur les différentes classes de traitements de fond synthétiques ;
- les traitements de fond ciblés synthétiques (i.e. le baricitinib, tofacitinib) et certains biosimilaires (i.e. les biosimilaires de l'adalimumab et du rituximab) évalués par la CT, soit après la validation des résultats des méta-analyses en réseau présentés dans le chapitre 3, soit au moment de la rédaction de ce rapport, ont été intégrés dans les analyses économiques complémentaires. Les résultats de ces analyses **doivent être lus avec prudence** en raison de l'importante incertitude générée par l'utilisation des probabilités de réponses EULAR obtenues à l'aide d'une transformation de probabilités ACR et par l'utilisation d'une source étrangère de données issues de méta-analyse en réseau (ICER, 2017) et qui ne sont pas complètement conformes aux caractéristiques de la population de l'analyse de référence des méta-analyses en réseau réalisées dans ce projet. À titre indicatif, contrairement aux méta-analyses SchHARR & HAS (Simpson *et al.*, 2019) (94), le design de celles de l'ICER (13) incluait des études observationnelles, et les critères d'analyse clinique privilégiaient principalement les probabilités de réponse ACR20. La prochaine mise à jour de la méta-analyse en réseau du modèle PR France pourra inclure ces traitements après avoir collecté suffisamment de données d'efficacité et de tolérance en vie réelle sur les nouveaux traitements de fond ciblés tsDMARDs, la trithérapie (association de trois traitements de fond synthétiques) et les nouveaux biosimilaires ;
- le faible nombre d'essais cliniques comparant directement les traitements de fond en 2^e ligne biologique et par conséquent les contraintes de réaliser des méta-analyses en réseau (*a minima* des comparaisons indirectes) appropriées sont une source d'incertitude sur les paramètres d'efficacité du modèle PR France au-delà de la première ligne de traitement biologique (après échec d'un csDMARD). Le choix des données issues de l'essai clinique pragmatique ROC et l'ensemble des analyses de scénarios supposant par exemple des taux d'échecs secondaires comparables entre les traitements, ont permis d'explorer cette incertitude.

5. Conclusion

L'évaluation économique présentée dans ce projet a permis de quantifier les résultats de santé, les coûts et le ratio coût-efficacité des séquences de traitements biologiques décrivant le parcours médicamenteux des patients atteints de PR modérée à sévère et présentant une réponse inadéquate et/ou une tolérance insuffisante à un traitement de fond biologique (i.e. MTX) en France. Les résultats étaient principalement fondés sur une structure de modélisation transparente (PR France), reposant sur des règles décisionnelles correspondant aux pratiques françaises : les réponses EULAR issues de la méta-analyse HAS & Sheffield et les équations du modèle PR France fondées sur les données en vie réelle extraites de la cohorte ESPOIR.

Le différentiel de QALY entre les séquences de traitements de fond biologiques résultant de l'efficacité clinique induite par la réponse EULAR était très faible. Cet écart s'atténue dans le scénario où la variabilité de l'activité de la PR est complètement éliminée.

L'impact du choix du traitement de fond biologique (ou trithérapie) sur le différentiel de coûts était plus important que sur le différentiel de QALY : les séquences débutant par le biosimilaire de l'éta-nercept (ETABepali®) étaient moins coûteuses que les séquences correspondantes débutant par d'autres anti-TNF α . Les séquences les moins coûteuses étaient **ETAB-->ABA-->INF** (analyse de référence) et **ADAB-->ABA-->INF** (analyse de scénario complémentaire). La majeure partie du différentiel du coût total était expliquée par les coûts d'acquisition des traitements.

L'analyse de référence a identifié deux séquences se situant sur la frontière d'efficacité. Les valeurs des RDCRs pour la deuxième séquence sur la frontière (et pour celles qui ne sont pas dominées dans les scénarios) restent élevées, quel que soit le scénario analysé, et étaient quasiment toutes supérieures à 100 000 €/QALY. Concernant l'analyse de scénarios sur les utilités, l'utilisation des scores estimés à partir du SF-6D – qui avaient un plus grand niveau de sensibilité au changement en cas d'amélioration de l'état du patient – aboutissait à un RDCR encore plus élevé que celui obtenu en utilisant les utilités estimées à partir de l'EQ-5D. En dépit de son usage en analyse de scénario, le SF-6D aboutissait à une estimation du RDCR plutôt « conservatrice » par rapport à celle de l'EQ-5D (Gaujoux-Viala *et al.*, 2013) ([26](#)).

Cette évaluation économique a analysé deux aspects d'actualité dans la prise en charge de la PR en France : elle a permis non seulement de consolider la place économique des biosimilaires dans le début du parcours médicamenteux des patients atteints de PR et présentant une réponse inadéquate au MTX, mais suggère également qu'une séquence de traitements moins coûteuse peut être ajoutée à l'arsenal thérapeutique actuel si la trithérapie est donnée à la place de la première ligne de traitement biologique, et si un biosimilaire est utilisé en deuxième ligne biologique. Ceci a été illustré par l'identification de la séquence la moins coûteuse se situant sur la frontière

d'efficience dans l'analyse secondaire intégrant la trithérapie. Cependant, la pertinence « médico-économique » de la trithérapie, qui ne représente pas un traitement standard en France, est étroitement liée à son adhérence, au même titre que son acceptation par le patient. Ces deux aspects font actuellement l'objet d'un PHRC qui comparera la trithérapie aux traitements de fond biologiques utilisés dans la prise en charge actuelle de la PR (BIO-3 trial, 2014) ([177](#)).

Conclusion générale / Synthèse

La polyarthrite rhumatoïde (PR) est une maladie rhumatismale auto-immune, inflammatoire, chronique, évoluant par poussées provoquant des gonflements et des douleurs au niveau des poignets, des mains ou des pieds, déformant et détruisant progressivement les articulations. Elle touche également d'autres organes extra-articulaires tels que les poumons et altère la qualité de vie du patient, peut conduire à une incapacité fonctionnelle et à la difficulté d'exercer une activité professionnelle. En France, le nombre de personnes en affection de longue durée (ALD) pour PR « grave évolutive » était de 226 640 au 31 décembre 2015 ([178](#)). Selon la même source, ce nombre a augmenté de 3,4 % en 2013, de 4,4 % en 2014 et de 4,6 % en 2015.

Comparativement à la littérature anglo-saxonne, les évaluations économiques des traitements de fond biologiques réalisées en France sont peu nombreuses et la plupart d'entre elles portent sur des estimations agrégées des coûts. Ce rapport constitue la première évaluation française réalisée en parallèle des travaux menés par la CNAM dans le cadre de son plan de maîtrise médicalisée sur le bon usage des traitements de fond biologiques. Il répond à la saisine de la Direction de la Sécurité sociale concernant l'« Évaluation médico-économique des biothérapies dans la prise en charge de la PR » en réalisant, d'une part, une analyse coût-utilité de 180 stratégies de prise en charge de la PR en France définies en termes de séquences de traitements de fond biologiques (et leurs biosimilaires) et en discutant, d'autre part, les résultats de l'analyse de l'efficacité en tenant compte de l'incertitude entachant les données disponibles.

1. Quelles étaient les méthodes utilisées pour répondre à l'objectif du projet ?

- **une revue systématique de la littérature clinique et la réalisation de méta-analyses en réseau** qui portaient sur deux critères de résultats cliniques : la réponse selon les critères européens proposés par la Ligue européenne de lutte contre le rhumatisme (EULAR) et celle fondée sur les critères définis par la Société américaine de rhumatologie (ACR). L'indication de la population d'entrée du modèle PR France visait les patients atteints de PR active modérée à sévère, présentant une réponse inadéquate ou une tolérance insuffisante au méthotrexate (MTX) ;
- à la différence de certaines méta-analyses réalisées par la Cochrane et d'autres organismes d'évaluation des technologies de santé (ICER, 2017 et CADTH, 2018) et à l'instar du NICE (Stevenson *et al.*, 2016), le critère clinique principal retenu était la réponse EULAR. Ce choix a permis d'estimer des différences de scores de probabilités de réponses, robustes et issues des méta-analyses en réseau. Ces estimations présentent l'avantage de tenir directement compte de l'aspect ordinal des réponses EULAR alors que les estimations classiques reposent fréquemment sur un lien logistique permettant d'estimer des risques relatifs standards (i.e. odds ratio) ;

- **une revue systématique de la littérature économique adaptée** à l'analyse des stratégies en termes de séquences de traitement (1^{re} ligne, 2^e ligne, 3^e ligne et plus). Cette méthode s'est fondée sur des grilles d'évaluation standards validées (Ghabri *et al.*, 2017), de *reporting* (CHEERS, 2013) et adaptées aux analyses critiques (combinaison de l'OMERACT, 2003 ; Drummond, 2005 et CHEERS, 2013). Elle a principalement porté sur l'indication à l'entrée du modèle PR France, à savoir les patients atteints de PR active modérée à sévère présentant une réponse inadéquate ou une tolérance insuffisante au méthotrexate, et a été complétée par l'indication associée aux lignes suivantes (i.e. patients atteints de PR ayant une réponse inadéquate à au moins un anti-TNF α) ;
- **un modèle de décision économique de novo**, nommé PR France, fondé sur une simulation du parcours médicamenteux du patient atteint de PR active modérée à sévère présentant une réponse inadéquate ou une tolérance insuffisante au méthotrexate. Il repose sur :
 - un algorithme décisionnel de la prise en charge de la PR en pratique clinique, des probabilités de réponses EULAR issues de la méta-analyse en réseau HAS & SchARR fondée sur des essais cliniques randomisés,
 - des données en vie réelle issues de la cohorte française ESPOIR « Étude et suivi des polyarthrites indifférenciées récentes » permettant des estimations de la qualité de vie (des utilités) selon deux questionnaires (EQ-5D et SF-6D) et l'équation de changement de traitements de fond synthétiques et biologiques,
 - un algorithme sur les coûts de suivi (i.e. consultations médicales),
 - les changements dans l'activité de la PR et les ajustements de doses des traitements administrés selon les pratiques cliniques françaises.

La structure de ce modèle a été nourrie par deux protocoles d'étude sur la cohorte ESPOIR et sur les événements indésirables sévères et/ou graves susceptibles d'entraîner une hospitalisation. Parmi les cinq études françaises portant sur les patients atteints de PR en vie réelle, la cohorte ESPOIR était la base de données la plus appropriée pour les besoins du modèle PR France en raison de son suivi régulier d'un ensemble de patients atteints de PR depuis 2003. Quant à l'analyse portant sur les événements indésirables, elle représente un choix pragmatique fondé sur les évaluations et les réévaluations des traitements de fond biologiques réalisées par la CT, les données sur les événements indésirables déposées par les entreprises du médicament à l'Agence européenne du médicament avant et après les autorisations de mise sur le marché des médicaments (EudraVigilance) et une synthèse des méta-analyses récentes.

2. Quelles étaient les réponses aux questions principales du projet ?

- **L'efficacité comparative :**
 - pour les patients ayant une PR modérée à sévère en échec de traitement de fond synthétique conventionnel csDMARD (ex. : méthotrexate), les méta-analyses en réseau réalisées dans le cadre de ce projet ont montré que, si le comparateur commun était un csDMARD (i.e. méthotrexate), les associations du méthotrexate avec l'éтанercept (ou son biosimilaire) et avec le tocilizumab étaient caractérisées par des réponses ACR et EULAR élevées. En revanche, les résultats de méta-analyses en réseau ne détectaient pas de différences significatives entre les biothérapies (bDMARDs) ;
 - l'analyse complémentaire exploratoire des résultats de l'efficacité comparative des tsDMARDs récemment évalués par la CT complète les résultats des méta-analyses réalisées. Elle a montré que les associations des tsDMARDs avec le MTX étaient plus efficaces que le MTX en monothérapie et qu'on ne pouvait pas différencier en termes de bénéfice clinique un tsDMARD d'un bDMARD.
- **La revue systématique de la littérature économique :**
 - a confirmé les aspects économiques communs à la majorité des évaluations économiques portant sur la PR (ex. : typologie de coûts, horizon temporel, actualisation, présentation des résultats) ;
 - n'a pas apporté de réponse claire aux questions suivantes : la mobilisation des données de tolérance et de qualité de vie nécessaires à la modélisation, la prise en compte des utilités dans la PR, l'intégration d'au moins trois lignes de traitements biologiques différentes dans les séquences de traitements biologiques ;
 - a mis en évidence l'absence ou le peu de données portant sur le contexte français.
- **L'analyse de l'efficience :**
 - l'évaluation économique présentée dans ce projet a permis de quantifier les résultats de santé, les coûts et le ratio coût-efficacité **des séquences de traitements biologiques** à partir de la simulation du parcours médicamenteux des patients atteints de PR :
 - *le différentiel de QALY* entre les séquences étudiées, résultant de l'efficacité clinique mesurée par la réponse EULAR, était très faible. Ce différentiel augmentait dans le scénario où la variabilité de l'activité de la PR était complètement éliminée,
 - *l'impact du choix des séquences sur le différentiel de coûts* était plus important que pour les QALYs. Alors que la différence dans les QALYs **actualisés n'était que de 1 %**, les **coûts actualisés différaient de plus de 35 %**. Les séquences débutant par le biosimilaire de l'éтанercept (ETAB Benepali®) étaient moins coûteuses que les séquences correspon-

- dantes débutant par d'autres anti-TNF α . La séquence la moins coûteuse était le biosimilaire de l'étanercept suivi par l'abatcept et l'infliximab (ETAB \rightarrow ABA \rightarrow INF). La majeure partie du différentiel du coût total était expliquée par les coûts d'acquisition des traitements,
- les séquences commençant par le biosimilaire de l'étanercept étaient les moins coûteuses et tendaient également à être parmi les plus « efficaces », elles dominaient généralement la plupart des autres séquences. **Seules deux séquences, commençant chacune par le biosimilaire de l'étanercept (ETAB->ABA->INF et ETAB->TOC->ABA), étaient situées sur la frontière d'efficience dans l'analyse de référence.** Le RDCR de ETAB->TOC->ABA vs ETAB->ABA->INF s'élevait à 955 778 €/QALY. Ce constat sur les séquences commençant par les biosimilaires a été confirmé dans l'analyse complémentaire incluant le biosimilaire de l'adalimumab : **une séquence commençant par le biosimilaire de l'adalimumab (ADAB->ABA->INF) et une autre par le biosimilaire de l'adalimumab (ETAB->ABA->INF) étaient situées sur la frontière d'efficience,**
 - l'efficience des séquences de traitements restait globalement inchangée dans la plupart des scénarios analysés. Même dans le cas d'une baisse de 50 % des prix des traitements de fond biologiques, les RDCRs restaient supérieurs à 100 000 €/QALY,
 - *l'analyse de scénario* incluant la trithérapie en début de séquence (à la place d'un traitement de fond biologique) faisait apparaître une seule séquence, TT->ETAB->INF, non dominée par la séquence débutant par le biosimilaire de l'étanercept la plus « efficace » : ETAB->ABA->INF. Si l'on considère la séquence débutant par la trithérapie TT->ETAB->INF comme stratégie de référence, le RDCR de la séquence débutant par le biosimilaire de l'étanercept était de 65 623 € par QALY.

3. Quel est l'impact attendu des résultats en matière d'aide à la décision publique et pour les professionnels de santé ?

Un des objectifs de l'évaluation économique réalisée dans ce rapport était de fournir au décideur public, aux professionnels de santé et aux prescripteurs des informations sur les séquences de traitements efficaces dans la prise en charge de la PR, tant le poids financier actuel et futur de ces thérapies est important pour l'assurance maladie.

- Compte tenu des niveaux de preuve cliniques et économiques disponibles, cette évaluation économique a notamment abordé deux thèmes d'actualité dans la prise en charge de la PR en France : **l'intégration** des biosimilaires dans les stratégies de prise en charge de la PR **définies en termes de séquences de traitements** et la place de la trithérapie dans un parcours médicamenteux efficace du patient atteint de PR.
- Cette évaluation a permis de confirmer la place économique des biosimilaires dans le début du parcours médicamenteux des patients atteints de PR et présentant une réponse inadéquate au méthotrexate. Les séquences commençant par le biosimilaire de l'éta nercept dominaient généralement la plupart des autres séquences.
- Les coûts d'acquisition des traitements de fond biologiques représentent le poste qui contribue le plus au coût direct total d'une séquence de traitements sur l'horizon de vie du patient atteint de PR.
- La prise en compte de la trithérapie a permis d'identifier la séquence la moins coûteuse se situant sur la frontière d'efficacité dans l'analyse de scénarios. Cependant, le bénéfice médico-économique de la trithérapie (qui ne représente pas aujourd'hui un traitement standard dans la prise en charge de la PR en France) reste étroitement lié à la démonstration de la qualité de son adhésion thérapeutique (i.e. degré d'acceptation du patient vis-à-vis de la trithérapie).

4. Quelques actions à mener pour améliorer l'efficacité de la prise en charge de la PR

- à défaut d'une diminution différenciée importante dans les prix des molécules mères de bDMARDs,
 - le biosimilaire de l'éta nercept (ou le biosimilaire de l'adalimumab) pourrait être utilisé en première ligne de traitement biologique (après échec d'un csDMARD),
 - il ne semble pas « efficace » d'utiliser successivement deux anti-TNF α dans une perspective médico-économique ;
- l'analyse économique complémentaire ne positionne pas les séquences incluant les tsDMARDs comme stratégies « efficaces » ou « moins coûteuses », mais ce résultat est

entaché d'une forte incertitude liée aux données utilisées et notamment au faible recul en termes d'efficacité et de tolérance comparativement aux biothérapies disponibles ;

- compte tenu du différentiel de qualité de vie faible entre les biosimilaires et leur molécule mère et au vu des résultats sur les différentiels des coûts entre les séquences de traitements, la place des biosimilaires dans le parcours médicamenteux des patients atteints de PR doit être consolidée par des analyses d'impact budgétaire portant sur **les biosimilaires actuels et futurs** disponibles sur le marché. Dans le contexte international actuel marqué par l'arrivée de nouveaux traitements innovants et très onéreux, la soutenabilité financière du système de l'assurance maladie est un enjeu majeur pour les autorités publiques. **Dans ce contexte, le renforcement de la mise en concurrence des traitements de fond biologiques et de leurs biosimilaires est une opportunité pour générer des économies tout en garantissant un accès à des traitements de qualité à de nombreux patients.** Il est à remarquer que la part de marché des biosimilaires destinés à être utilisés en ambulatoire reste limitée en France ;
- **l'adhésion des patients à la trithérapie doit faire l'objet d'études en vie réelle dans lesquelles elle sera comparée aux traitements de fond biologiques au même titre que les traitements synthétiques ciblés utilisés dans la prise en charge actuelle de la PR active modérée à sévère en France.** Une acceptation de la trithérapie par le patient pourrait allonger la durée de traitement après l'échec d'un traitement de fond synthétique (i.e. le méthotrexate) et diminuer par la même occasion les coûts du traitement, compte tenu du moindre coût de la trithérapie. Le besoin d'études sur l'adhésion au traitement et son acceptabilité est également valable dans le cas des biosimilaires et des traitements de fond ciblés. En effet, au moment de la diffusion des résultats du PHRC en 2020, les rhumatologues auront utilisé plus largement les inhibiteurs de JAK, en monothérapie par voie orale. En termes d'observance, les patients seront donc face à un traitement quotidien par voie orale (les JAKi) *versus* une trithérapie par voie orale ;
- abstraction faite de l'analyse de l'efficacité des traitements de fond biologiques, la mise en place d'études complémentaires intégrant les préférences des patients (i.e. perception de la rémission par le patient) serait très utile pour optimiser l'analyse des dimensions non captées par le QALY. À titre indicatif, la perception de la rémission par le patient pourrait porter sur les trois dimensions suivantes : l'absence/la réduction des symptômes, l'impact relatif de la réduction de l'activité (ou de la modification structurale) de la maladie sur la vie quotidienne du patient, et le sentiment d'un retour à une « vie normale ». **Cette perspective contribuerait également à accroître l'implication du patient dans la prise en charge de sa maladie en identifiant des facteurs ayant un impact sur son adhésion**

thérapeutique et sa participation active dans le choix des traitements (ex. : choix de la trithérapie et des biosimilaires). Le groupe de travail PR a unanimement souligné l'importance de cette perspective d'exploration de l'analyse de la perception de la rémission par le patient. À ce propos, un rationnel de protocole portant sur ce thème a été proposé à la HAS en juin 2017 ;

- **la finalisation du projet de la SFR consistant à appairer la cohorte ESPOIR avec les données médico-administratives issues du Système national d'information inter-régimes de l'assurance maladie (SNIIRAM) permettrait d'explorer d'autres stratégies efficaces de prise en charge en cas de rémission établie,** et d'optimiser la prise en compte de l'incidence et des coûts des événements indésirables sévères et/ou graves entraînant des hospitalisations ;
- **l'impact des comorbidités sur les coûts, au même titre que sur la qualité de vie des patients atteints de PR (et des maladies chroniques d'une manière générale), est peu pris en compte dans les évaluations économiques actuelles.** Il doit être davantage exploré en utilisant, par exemple, d'autres types de données (registres ou enquêtes) renseignant sur leurs effets déjà captés par les instruments EQ-5D-5L ou le SF-6D. Si l'impact des événements aigus (infections graves) sur les pertes d'utilités (qualité de vie) des patients atteints de PR a été pris en compte dans cette évaluation ainsi que dans celles du NICE et de l'ICER, l'évaluation de l'impact d'autres événements indésirables liés à l'utilisation des traitements de fond biologiques à long terme (ex. : cancers) pose problème, tant sur le plan conceptuel que pratique. À ce jour, aucun surrisque de cancer n'est associé à l'usage des bDMARDs, en dehors d'une augmentation des cancers de la peau (non mélanome), mais à titre de précaution les recommandations proposent de ne pas prescrire un bDMARD chez les patients ayant un antécédent néoplasique de moins de 5 ans. Si avant ce délai de rémission, un traitement biologique est indispensable, alors le rituximab est choisi en raison de son mécanisme d'action particulier. Cette pratique clinique explique la surreprésentation des événements néoplasiques observés chez les patients traités par le rituximab, qui ne représente pas un surrisque lié au traitement mais un surrisque lié aux antécédents particuliers des patients éligibles à ce traitement (population à risque de rechute d'un cancer récent exclue pour les autres bDMARDs) ;
- au-delà des comparaisons des stratégies de traitements de fond synthétiques ou biologiques, une analyse en termes de coûts indirects de la PR serait très utile : elle permettrait non seulement de chiffrer ce coût pour le décideur public, mais également de nourrir la réflexion sur les modalités permettant une prise en charge plus précoce et individualisée des patients atteints de PR. À ce propos, **des études montrant l'impact de la PR sur les arrêts de travail des patients actifs sont très attendues ;**

- **une actualisation de cette évaluation économique** s'avérera nécessaire après avoir collecté des données d'efficacité et de tolérance en vie réelle sur les traitements de fond ciblés.

Références bibliographiques

- Canadian Agency for Drugs and Technologies in Health. Drugs for the management of rheumatoid arthritis: clinical evaluation. Project Protocol. Ottawa: CADTH; 2016.
<https://www.cadth.ca/dv/drugs-management-rheumatoid-arthritis-clinical-evaluation-project-protocol>
- Guillemin F, Saraux A, Guggenbuhl P, Roux CH, Fardellone P, Le Bihan E, *et al.* Prevalence of rheumatoid arthritis in France: 2001. *Ann Rheum Dis* 2005;64(10):1427-30.
- Caisse nationale d'assurance maladie. Prévalence des ALD en 2013 [En ligne] 2016.
<http://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/affection-de-longue-duree-ald/prevalence/prevalence-des-ald-en-2013.php>
- Société française de rhumatologie, Gaujoux-Viala C, Gossec L, Cantagrel A, Dougados M, Fautrel B, *et al.* Recommendations of the French Society for Rheumatology for managing rheumatoid arthritis. *Joint Bone Spine* 2014;81(4):287-97.
- European League Against Rheumatism, Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73(3):492-509.
- European League Against Rheumatism, Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76(6):960-77.
- Société française de rhumatologie, Daien C, Hua C, Gaujoux-Viala C, Cantagrel A, Dubremetz M, *et al.* Actualisation des Recommandations de la Société française de rhumatologie pour la prise en charge de la polyarthrite rhumatoïde. *Revue du Rhumatisme* 2019;86(1):8-24.
- Fautrel B, Cukierman G, Joubert JM, Laurendeau C, Gourmelen J, Fagnani F. Healthcare service utilisation costs attributable to rheumatoid arthritis in France: Analysis of a representative national claims database. *Joint Bone Spine* 2016;83(1):53-6.
- Hallert E, Husberg M, Bernfort L. The incidence of permanent work disability in patients with rheumatoid arthritis in Sweden 1990-2010: before and after introduction of biologic agents. *Rheumatology* 2012;51(2):338-46.
- Haute Autorité de Santé. Polyarthrite rhumatoïde : aspects thérapeutiques hors médicaments et chirurgie - aspects médico-sociaux et organisationnels. Saint-Denis La Plaine: HAS; 2007.
http://www.has-sante.fr/portail/upload/docs/application/pdf/argumentaire_pr_non_med_071018.pdf
- Nystad TW, Fenstad AM, Furnes O, Fevang BT. Predictors for orthopaedic surgery in patients with rheumatoid arthritis: results from a retrospective cohort study of 1010 patients diagnosed from 1972 to 2009 and followed up until 2015. *Scand J Rheumatol* 2018;47(4):282-90.
- Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens J, *et al.* Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. *Health Technol Assess* 2016;20(35):1-610.
- Institute for Clinical and Economic Review. Targeted immune modulators for rheumatoid arthritis: effectiveness & value. Boston: ICER; 2017.
- Maravic M. Economic impact of rheumatoid arthritis (RA) biotherapies in France. *Joint Bone Spine* 2010;77(4):319-24.
- Beck M, Velten M, Rybarczyk-Vigouret C, Covassin J, Sordet C, Michel B. Analysis and breakdown of overall 1-year costs relative to inpatient and outpatient care among rheumatoid arthritis patients treated with biotherapies using health insurance claims database in Alsace. *Drugs Real World Outcome* 2015;2(3):205-15.
- Kobelt G, Woronoff AS, Richard B, Peeters P, Sany J. Disease status, costs and quality of life of patients with rheumatoid arthritis in France: the ECO-PR Study. *Joint Bone Spine* 2008;75(4):408-15.
- Fautrel B, Gaujoux-Viala C. Aspects médico-économiques de la polyarthrite rhumatoïde. *Bull Acad Natle Méd* 2012;196(7):1295-306.
- Sany J, Cohen JD, Combescure C, Bozonnet MC, Roch-Bras F, Lafon G, *et al.* Medico-economic evaluation of infliximab in rheumatoid arthritis--prospective French study of a cohort of 635 patients monitored for two years. *Rheumatology* 2009;48(10):1236-41.
- Saroux A, Gossec L, Goupille P, Bregman B, Bocard E, Dupont D, *et al.* Cost-effectiveness modelling of biological treatment sequences in moderate to severe rheumatoid arthritis in France. *Rheumatology* 2010;49(4):733-40.
- Vanier A, Mariette X, Tubach F, Fautrel B. Cost-effectiveness of TNF-blocker injection spacing for patients with established rheumatoid arthritis in remission: an economic evaluation from the spacing of

- TNF-blocker injections in rheumatoid arthritis trial. *Value Health* 2017;20(4):577-85.
21. Alemao E, Al MJ, Boonen AA, Stevenson MD, Verstappen SMM, Michaud K, *et al.* Conceptual model for the health technology assessment of current and novel interventions in rheumatoid arthritis. *PLoS ONE* 2018;13(10):e0205013.
22. Madan J, Ades AE, Welton NJ. An overview of models used in economic analyses of biologic therapies for arthritis--from current diversity to future consensus. *Rheumatology* 2011;50 Suppl 4:iv10-8.
23. Barton P. Development of the Birmingham Rheumatoid Arthritis Model: past, present and future plans. *Rheumatology* 2011;50 Suppl 4:iv32-iv8.
24. Tosh J, Archer R, Davis S, Stevenson M, Stevens JW. Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying antirheumatic drugs: a NICE single technology appraisal. *Pharmacoeconomics* 2013;31(8):653-61.
25. European Medicines Agency. Guideline on clinical investigation of medicinal products 5 other than NSAIDs for treatment of rheumatoid arthritis. London: EMA; 2015.
https://www.ema.europa.eu/documents/scientific-guideline/draft-guideline-clinical-investigation-medicinal-products-other-non-steroidal-anti-inflammatory_en-0.pdf
26. Gaujoux-Viala C, Fautrel B, Guillemin F, Flipo RM, Bourgeois P, Rat AC. Who are the patients with early arthritis with worse than death scores on the EQ-5D? Results from the ESPOIR cohort. *Rheumatology* 2013;52(5):832-8.
27. Hernandez-Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology* 2013;52:944-50.
28. Gaujoux-Viala C, Rat AC, Guillemin F, Flipo RM, Fardellone P, Bourgeois P, *et al.* Responsiveness of EQ-5D and SF-6D in patients with early arthritis: results from the ESPOIR cohort. *Ann Rheum Dis* 2012;71(9):1478-83.
29. Haute Autorité de Santé. Évaluation médico-économique des biothérapies dans la prise en charge de la polyarthrite rhumatoïde. Feuille de route. Saint-Denis La Plaine: HAS; 2015.
http://www.has-sante.fr/portail/upload/docs/application/pdf/2015-12/feuille_de_route_-_prise_en_charge_de_la_polyarthrite_rhumatoide.pdf
30. Haute Autorité de Santé. Choix méthodologiques pour l'évaluation économique à la HAS. Saint-Denis La Plaine: HAS; 2011.
http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-11/guide_methodo_vf.pdf
31. Verstappen SM. Rheumatoid arthritis and work: The impact of rheumatoid arthritis on absenteeism and presenteeism. *Best Pract Res Clin Rheumatol* 2015;29(3):495-511.
32. Canadian Agency for Drugs and Technologies in Health. Drugs for the management of rheumatoid arthritis: clinical evaluation. Ottawa: CADTH; 2018.
https://www.cadth.ca/sites/default/files/pdf/HT0010_RA_Report.pdf
33. Caro JJ. Discretely Integrated Condition Event (DICE) simulation for pharmacoeconomics. *Pharmacoeconomics* 2016;34(7):665-72.
34. Caro JJ, Moller J. Adding events to a Markov model using DICE simulation. *Med Decis Making* 2018;38(2):235-45.
35. Combe B. ESPOIR and the French database management: what have we learned from the first years of follow-up? *Clin Exp Rheumatol* 2014;32(5 Suppl 85):S-153-7.
36. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097.
37. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33(5):607-17.
38. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013;346:f1049.
39. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Standard methods for the economic evaluation of health care programme. Oxford: Oxford University Press; 2015.
40. Gabriel SE, Drummond M, Maetzel A, Boers M, Coyle D, Welch V, *et al.* OMERACT 6 economics working group report: a proposal for a reference case for economic evaluation in rheumatoid arthritis. *J Rheumatol* 2002;30(4):886-90.
41. Haute Autorité de Santé. Olumiant 2 mg, comprimé pelliculé. Avis de la Commission de la Transparence 21 juin 2017. Saint-Denis La Plaine: HAS; 2017.
https://www.has-sante.fr/portail/upload/docs/evamed/CT-16116_OLUMIANT_PIC_INS_Avis%20CT16116.pdf
42. Alamanos Y, Voulgari PV, Drosos AA. Incidence and Prevalence of Rheumatoid Arthritis, Based on the 1987 American College of Rheumatology Criteria: A Systematic Review. *Semin Arthritis Rheum* 2006;36:182-8.
43. Doran MF, Pond GR, Crowson CS, O'Fallon WM, Gabriel SE. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. *Arthritis Rheum* 2002;46(3):625-31.

44. Myasoedova E, Davis JM, 3rd, Crowson CS, Gabriel SE. Epidemiology of rheumatoid arthritis: rheumatoid arthritis and mortality. *Curr Rheumatol Rep* 2010;12(5):379-85.
45. Institut national de santé publique Québec. Surveillance de la polyarthrite rhumatoïde au Québec : prévalence, incidence et mortalité. Montréal: INSPQ; 2017.
46. Widdifield J, Bernatsky S, Paterson JM, Tomlinson G, Tu K, Kuriya B, *et al.* Trends in excess mortality among patients with rheumatoid arthritis in Ontario, Canada. *Arthritis Care Res* 2015;67(8):1047-53.
47. Coste J, Jouglu E. Mortality from rheumatoid arthritis in France, 1970-1990. *Int J Epidemiol* 1994;23(3):545-52.
48. Ziadé N, Jouglu E, Coste J. Population-level influence of rheumatoid arthritis on mortality and recent trends: a multiple cause-of-death analysis in France, 1970-2002. *J Rheumatol* 2008;35(10):1950-7.
49. Avouac J, Amrouche F, Meune C, Rey G, Kahan A, Allanore Y. Mortality profile of patients with rheumatoid arthritis in France and its change in 10 years. *Semin Arthritis Rheum* 2017;46(5):537-43.
50. Minaur NJ, Jacoby RK, Cosh JA, Taylor G, Rasker JJ. Outcome after 40 years with rheumatoid arthritis: a prospective study of function, disease activity, and mortality. *J Rheumatol Suppl* 2004;69:3-8.
51. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, *et al.* Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. *Joint Bone Spine* 2013;80(1):29-33.
52. England BR, Sayles H, Mikuls T, Michaud K. Mortality trends in rheumatoid arthritis during the biologic era, 1998 to 2011. Dans: 2014 ACR/ARHP Annual Meeting 2014.
53. Agca R, Heslinga SC, Rollefstad S, Heslinga M, McInnes IB, Peters MJ, *et al.* EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. *Ann Rheum Dis* 2017;76(1):17-28.
54. Lindhardsen J, Gislason GH, Jacobsen S, Ahlehoff O, Olsen AM, Madsen OR, *et al.* Non-steroidal anti-inflammatory drugs and risk of cardiovascular disease in patients with rheumatoid arthritis: a nationwide cohort study. *Ann Rheum Dis* 2014;73(8):1515-21.
55. van den Hoek J, Boshuizen HC, Roorda LD, Tijhuis GJ, Nurmohamed MT, van den Bos GA, *et al.* Mortality in patients with rheumatoid arthritis: a 15-year prospective cohort study. *Rheumatol Int* 2017;37(4):487-93.
56. Lacaille D, Avina-Zubieta JA, Sayre EC, Abrahamowicz M. Improvement in 5-year mortality in incident rheumatoid arthritis compared with the general population-closing the mortality gap. *Ann Rheum Dis* 2017;76(6):1057-63.
57. Zhang Y, Lu N, Peloquin C, Dubreuil M, Neogi T, Avina-Zubieta JA, *et al.* Improved survival in rheumatoid arthritis: a general population-based cohort study. *Ann Rheum Dis* 2017;76(2):408-13.
58. Myasoedova E, Gabriel SE, Matteson EL, Davis JM, 3rd, Therneau TM, Crowson CS. Decreased cardiovascular mortality in patients with incident rheumatoid arthritis in recent years: dawn of a new era in cardiovascular disease in RA? *J Rheumatol* 2017;44(6):732-9.
59. Roubille C, Richer V, Starnino T, McCourt C, McFarlane A, Fleming P, *et al.* The effects of tumour necrosis factor inhibitors, methotrexate, non-steroidal anti-inflammatory drugs and corticosteroids on cardiovascular events in rheumatoid arthritis, psoriasis and psoriatic arthritis: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74(3):480-9.
60. Ajeganova S, Andersson ML, Frostegard J, Hafstrom I. Disease factors in early rheumatoid arthritis are associated with differential risks for cardiovascular events and mortality depending on age at onset: a 10-year observational cohort study. *J Rheumatol* 2013;40(12):1958-66.
61. Barnabe C, Martin BJ, Ghali WA. Systematic review and meta-analysis: anti-tumor necrosis factor alpha therapy and cardiovascular events in rheumatoid arthritis. *Arthritis Care Res* 2011;63(4):522-9.
62. Rao VU, Pavlov A, Klearman M, Musselman D, Giles JT, Bathon JM, *et al.* An evaluation of risk factors for major adverse cardiovascular events during tocilizumab therapy. *Arthritis & rheumatology* 2015;67(2):372-80.
63. Bili A, Tang X, Pranesh S, Bozait R, Morris SJ, Antohe JL, *et al.* Tumor necrosis factor alpha inhibitor use and decreased risk for incident coronary events in rheumatoid arthritis. *Arthritis Care Res* 2014;66(3):355-63.
64. Innala L, Moller B, Ljung L, Magnusson S, Smedby T, Sodergren A, *et al.* Cardiovascular events in early RA are a result of inflammatory burden and traditional risk factors: a five year prospective study. *Arthritis Res Ther* 2011;13(4):R131.
65. Listing J, Kekow J, Manger B, Burmester GR, Pattloch D, Zink A, *et al.* Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNF alpha inhibitors and rituximab. *Ann Rheum Dis* 2015;74(2):415-21.
66. Puolakka K, Kautiainen H, Pohjolainen T, Virta L. No increased mortality in incident cases of rheumatoid arthritis during the new millennium. *Ann Rheum Dis* 2010;69(11):2057-8.
67. Sparks JA, Chang SC, Liao KP, Lu B, Fine AR, Solomon DH, *et al.* Rheumatoid arthritis and mortality among women during 36 years of prospective follow-up:

- results from the nurses' health study. *Arthritis Care Res* 2016;68(6):753-62.
68. Juge PA, Dieudé P. Pneumopathies interstitielles diffuses au cours de la polyarthrite rhumatoïde. *Rev Rhum* 2017.
69. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38(1):44-8.
70. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, *et al.* A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* 2003;42(2):244-57.
71. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, *et al.* Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7(4):R796-806.
72. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyll LH, Funovits J, *et al.* American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;70(3):404-13.
73. Haute Autorité de Santé. Polyarthrite rhumatoïde évolutive grave. Guide affection de longue durée. Guide médecin. Saint-Denis La Plaine: HAS; 2008. http://www.has-sante.fr/portail/upload/docs/application/pdf/gm_polyarthrite_web.pdf
74. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23(2):137-45.
75. Guillemin F, Braincon S, Pourel J. Mesure de la capacité fonctionnelle dans la polyarthrite rhumatoïde: adaptation française du Health Assessment Questionnaire (HAQ). *Rev Rhum Mal Osteoartic* 1991;58(6):459-65.
76. Pollard LC, Choy EH, Gonzalez J, Khoshaba B, Scott DL. Fatigue in rheumatoid arthritis reflects pain, not disease activity. *Rheumatology* 2006;45(7):885-9.
77. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. The arthritis impact measurement scales. *Arthritis Rheum* 1980;23(2):146-52.
78. Leplège A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation and preliminary psychometric evaluation. *J Clin Epidemiol* 1998;51(11):1013-23.
79. Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). *Arthritis Care Res* 2011;63 Suppl 11:S4-13.
80. Haute Autorité de Santé. Xeljanz 5 mg, comprimé pelliculé. Avis de la Commission de la Transparence 27 septembre 2017. Saint-Denis La Plaine: HAS; 2017. https://www.has-sante.fr/portail/upload/docs/evamed/CT-16198_XELJANZ_PIC_INS_Avis2_CT16198.pdf
81. Haute Autorité de Santé. La prise en charge de votre polyarthrite rhumatoïde. Vivre avec une polyarthrite rhumatoïde. Guide affection de longue durée. Guide patient. Saint-Denis La Plaine: HAS; 2008. http://www.has-sante.fr/portail/upload/docs/application/pdf/2009-07/gp_ald_polyarthrite.pdf
82. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, *et al.* 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Care Res* 2016;68(1):1-25.
83. Gaujoux-Viala C, Gossec L, Cantagrel A, Dougados M, Fautrel B, Mariette X, *et al.* Recommandations de la société française de rhumatologie pour la prise en charge de la polyarthrite rhumatoïde. *Revue du Rhumatisme* 2014;81(4):303-12.
84. University of Sheffield, Haute Autorité de Santé. Biological treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs and after the failure of conventional disease-modifying anti-rheumatic drugs only: systematic review update and network meta-analyses update. Birmingham: ScHARR; 2017.
85. Ministère des solidarités et de la santé, Ministère de l'action et des comptes publics. Arrêté du 6 novembre 2017 modifiant la liste des médicaments agréés à l'usage des collectivités et divers services publics mentionnée à l'article L. 5123-2 du code de la santé publique. *Journal Officiel* 2017;(0263).
86. Haute Autorité de Santé. Kineret 100mg/0,67ml, solution injectable en seringue préremplie. Avis de la Commission de la Transparence 17 septembre 2014. Saint-Denis La Plaine: HAS; 2014.
87. Higgins J, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ Open* 2011;343(d5928).
88. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7(4):434-55.
89. Singh JA, Hossain A, Tanjong Ghogomu E, Kotb A, Christensen R, Mudano AS, *et al.* Biologics or tofacitinib for rheumatoid arthritis in incomplete responders to methotrexate or other traditional disease-modifying anti-rheumatic drugs: a systematic review and network

- meta-analysis. *Cochrane Database Syst Rev* 2016;(5):CD012183.
90. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29(7-8):932-44.
91. Singh JA, Hossain A, Tanjong Ghogomu E, Mudano AS, Tugwell P, Wells GA. Biologic or tofacitinib monotherapy for rheumatoid arthritis in people with traditional disease-modifying anti-rheumatic drug (DMARD) failure: a Cochrane Systematic Review and network meta-analysis (NMA). *Cochrane Database Syst Rev* 2016;11:CD012437.
92. Singh JA, Hossain A, Tanjong Ghogomu E, Mudano AS, Maxwell LJ, Buchbinder R, *et al.* Biologics or tofacitinib for people with rheumatoid arthritis unsuccessfully treated with biologics: a systematic review and network meta-analysis. *Cochrane Database Syst Rev* 2017;3:CD012591.
93. Hazlewood GS, Barnabe C, Tomlinson G, Marshall D, Devoe DJ, Bombardier C. Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic disease modifying anti-rheumatic drugs for rheumatoid arthritis: A network meta-analysis. *Cochrane Database Syst Rev* 2016;(8):CD010227.
94. Simpson EL, Ren S, Hock ES, Stevens JW, Binard A, Pers YM, *et al.* Rheumatoid arthritis treated with 6-months of first-line biologic or biosimilar therapy: an updated systematic review and network meta-analysis. *Int J Technol Assess Health Care* 2019:1-9.
95. Kaneko Y, Atsumi T, Tanaka Y, Inoo M, Kobayashi-Haraoka H, Amano K, *et al.* Comparison of adding tocilizumab to methotrexate with switching to tocilizumab in patients with rheumatoid arthritis with inadequate response to methotrexate: 52-week results from a prospective, randomised, controlled study (SURPRISE study). *Ann Rheum Dis* 2016;75(11):1917-23.
96. Kivitz A, Olech E, Borofsky M, Zazueta BM, Navarro-Sarabia F, Radominski SC, *et al.* Subcutaneous tocilizumab versus placebo in combination with disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. *Arthritis Care Res* 2014;66(11):1653-61.
97. Ogata A, Tanimura K, Sugimoto T, Inoue H, Urata Y, Matsubara T, *et al.* Phase III study of the efficacy and safety of subcutaneous versus intravenous tocilizumab monotherapy in patients with rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2014;66(3):344-54.
98. Burmester GR, Rubbert-Roth A, Cantagrel A, Hall S, Leszczynski P, Feldman D, *et al.* A randomised, double-blind, parallel-group study of the safety and efficacy of subcutaneous tocilizumab versus intravenous tocilizumab in combination with traditional disease-modifying antirheumatic drugs in patients with moderate to severe rheumatoid arthritis (SUMMACTA study). *Ann Rheum Dis* 2014;73(1):69-74.
99. Amano K, Matsubara T, Tanaka T, Inoue H, Iwahashi M, Kanamono T, *et al.* Long-term safety and efficacy of treatment with subcutaneous abatacept in Japanese patients with rheumatoid arthritis who are methotrexate inadequate responders. *Mod Rheumatol* 2015;25(5):665-71.
100. Takeuchi T, Miyasaka N, Zang C, Alvarez D, Fletcher T, Wajdula J, *et al.* A phase 3 randomized, double-blind, multicenter comparative study evaluating the effect of etanercept versus methotrexate on radiographic outcomes, disease activity, and safety in Japanese subjects with active rheumatoid arthritis. *Mod Rheumatol* 2013;23(4):623-33.
101. Emery P, Vencovsky J, Sylwestrzak A, Leszczynski P, Porawska W, Baranaskaite A, *et al.* A phase III randomised, double-blind, parallel-group study comparing SB4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis* 2017;76(1):51-7.
102. Yoo DH, Racewicz A, Brzezicki J, Yatsyshyn R, Arteaga ET, Baranaskaite A, *et al.* A phase III randomized study to evaluate the efficacy and safety of CT-P13 compared with reference infliximab in patients with active rheumatoid arthritis: 54-week results from the PLANETRA study. *Arthritis Res Ther* 2016;18:82.
103. Choe JY, Prodanovic N, Niebrzydowski J, Staykov I, Dokoupilova E, Baranaskaite A, *et al.* A randomised, double-blind, phase III study comparing SB2, an infliximab biosimilar, to the infliximab reference product Remicade in patients with moderate to severe rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis* 2017;76(1):58-64.
104. Matsubara T, Inoue H, Iwahashi M, Yamazaki A, Takeuchi T. THU0125 A multi-center, double-dummy, double-blind study of subcutaneous (SC) abatacept (ABA) compared with intravenous (IV) ABA in Japanese rheumatoid arthritis patients with inadequate response to methotrexate. *Ann Rheum Dis* 2014;71(Suppl 3):197.
105. Iwahashi M, Inoue H, Matsubara T, Tanaka T, Amano K, Kanamono T, *et al.* Efficacy, safety, pharmacokinetics and immunogenicity of abatacept administered subcutaneously or intravenously in Japanese patients with rheumatoid arthritis and inadequate response to methotrexate: a Phase II/III, randomized study. *Mod Rheumatol* 2014;24(6):885-91.
106. Takeuchi T, Yamanaka H, Ishiguro N, Miyasaka N, Mukai M, Matsubara T, *et al.* Adalimumab, a human anti-TNF monoclonal antibody, outcome study for the prevention of joint damage in Japanese patients with early rheumatoid arthritis: the HOPEFUL 1 study. *Ann Rheum Dis* 2014;73(3):536-43.
107. Takeuchi T, Kaneko Y, Atsumi T, Tanaka Y, Inoh M, Kobayashi H. Adding tocilizumab or switching to tocilizumab monotherapy in RA patients with inadequate response to methotrexate: 24 week results from a randomized controlled study (SURPRISE study). *Ann Rheum Dis* 2013;72(Suppl3):62.

108. Emery P, Vencovsky J, Sylwestrzak A, Leszczynski P, Porawska W, Stasiuk B. Additional efficacy results of SB4 (Etanercept biosimilar) up to week 100: comparison between continuing SB4 and switching from reference etanercept (Enbrel) to SB4. *Arthritis Rheumatology* 2016;68(S10).
109. Emery P, Vencovsky J, Sylwestrzak A, Leszczynski P, Porawska W, Stasiuk B, *et al.* Long-term safety and efficacy of SB4 (etanercept biosimilar) in patients with rheumatoid arthritis: Comparison between continuing SB4 and switching from etanercept reference product to SB4. *Ann Rheum Dis* 2016;75:236.
110. Vencovsky J, Sylwestrzak A, Leszczynski P, Porawska W, Baranauskaite A, Tseluyko V, *et al.* A phase III randomised, double-blind clinical study comparing SB4, an etanercept biosimilar, with etanercept reference product (Enbrel) in patients with moderate to severe rheumatoid arthritis despite methotrexate therapy (52-week results). *Ann Rheum Dis* 2015;74:467-8.
111. Hanyang D, Prodanovic N, Jaworski J, Miranda P, Ramiterre E, Lanzon A. Efficacy and safety of CT-P13 (infliximab biosimilar) over two years in patients with rheumatoid arthritis: Comparison between continued CT-P13 and switching from infliximab to CT-P13. *Arthritis Rheum* 2013;65(12):3319.
112. Lee SJ, Yoo D, Park W, Muller-Ladner U, Pyo TN. Statistical evaluation of joint damage progression in patients with rheumatoid arthritis treated with infliximab or biosimilar infliximab (CT-P13) anti-TNF therapy: A role of sensitivity analysis for missing data evaluating similarity. *Annals of the Rheumatic Diseases Conference* 2014;73.
113. Yoo D, Racewicz A, Brzezicki J, Yatsyshyn R, Tobias Arteaga E, Baranauskaite A. A phase 3 randomised controlled trial to compare CT-p13 with infliximab in patients with active rheumatoid arthritis: 54 week results from the Planetra study. *Ann Rheum Dis* 2013;72(Suppl 3):A73-A.
114. Yoo DH, Hrycaj P, Miranda P, Ramiterre E, Piotrowski M, Shevchuk S, *et al.* A randomised, double-blind, parallel-group study to demonstrate equivalence in efficacy and safety of CT-P13 compared with innovator infliximab when coadministered with methotrexate in patients with active rheumatoid arthritis: the PLANETRA study. *Ann Rheum Dis* 2013;72(10):1613-20.
115. Yoo D, Park W, Brzosko M, Geher P, Andersone D, Jaworski J. Disease activity assessment using the DAS28, CDAI and SDAI and effect of anti-drug antibody on clinical response in a randomized, double-blind, comparative trial of CT-P13 and innovator infliximab: Planetra study. *Annals of the Rheumatic Diseases Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR.* 2014.
116. Choe JY, Prodanovic N, Niebrzydowski J, Staykov I, Dokoupilova E, Baranauskaite A. A randomized, double-blind, phase III study comparing SB2, an Infliximab biosimilar, to the Infliximab reference product (Remicade) in patients with moderate to severe rheumatoid arthritis despite methotrexate therapy: 54 week results. abstract number 2056. 2015 ACR/ARHP Annual Meeting. 2015.
117. Choe JY, Smolen JS, Keystone E, Genovese MC, Choi J, Rho YH. Efficacy and safety analysis by overall anti-drug antibody results up to week 30 in patients with rheumatoid arthritis treated with sb2 (an infliximab biosimilar) or infliximab reference product in phase III study. *Ann Rheum Dis* 2016;75(Suppl 2):232.1-.
118. Durez P, Mysler E, Smolen JS, Choe JY, Choi J, Rho YH. The 54-week results of interferon- γ release assay in a phase III study comparing SB2, an infliximab biosimilar, to infliximab reference product in patients with rheumatoid arthritis: table 1. *Ann Rheum Dis* 2016;75(Suppl 2):235.3-6.
119. Smolen JS, Choe JY, Prodanovic N, Niebrzydowski J, Staykov I, Dokoupilova E, *et al.* Comparable safety and immunogenicity and sustained efficacy after transition to SB2 (an infliximab biosimilar) vs ongoing infliximab reference product in patients with rheumatoid arthritis: results of phase III transition study. *Ann Rheum Dis* 2016;75(Suppl 2):488.1-.
120. Kremer JM, Russell AS, Emery P, Abud-Mendoza C, Szechinski J, Westhovens R, *et al.* Long-term safety, efficacy and inhibition of radiographic progression with abatacept treatment in patients with rheumatoid arthritis and an inadequate response to methotrexate: 3-year results from the AIM trial. *Ann Rheum Dis* 2011;70(10):1826-30.
121. van der Heijde D, Tanaka Y, Fleischmann R, Keystone E, Kremer J, Zerbini C, *et al.* Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. *Arthritis Rheum* 2013;65(3):559-70.
122. Yamamoto K, Takeuchi T, Yamanaka H, Ishiguro N, Tanaka Y, Eguchi K, *et al.* Efficacy and safety of certolizumab pegol without methotrexate co-administration in Japanese patients with active rheumatoid arthritis: the HIKARI randomized, placebo-controlled trial. *Mod Rheumatol* 2014;24(4):552-60.
123. Singh JA, Hossain A, Mudano AS, Tanjong Ghogomu E, Suarez-Almazor ME, Buchbinder R, *et al.* Biologics or tofacitinib for people with rheumatoid arthritis naive to methotrexate: a systematic review and network meta-analysis. *Cochrane Database Syst Rev* 2017;5:CD012657.
124. Alfonso-Cristancho R, Armstrong N, Arjunji R, Riemsma R, Worthy G, Ganguly R, *et al.* Comparative effectiveness of biologics for the management of rheumatoid arthritis: systematic review and network meta-analysis. *Clin Rheumatol* 2017;36(1):25-34.
125. Smolen JS, Aletaha D. Interleukin-6 receptor inhibition with tocilizumab and attainment of disease

- remission in rheumatoid arthritis: the role of acute-phase reactants. *Arthritis Rheum* 2011;63(1):43-52.
126. Uttley L, Bermejo I, Ren S, Martyn-St James M, Wong R, Scott DL, *et al.* Tofacitinib for treating rheumatoid arthritis after the failure of disease-modifying anti-rheumatic drugs: an evidence review group perspective of a NICE single technology appraisal. *Pharmacoeconomics* 2018;36(9):1063-72.
127. Ren S, Bermejo I, Simpson E, Wong R, Scott DL, Young A, *et al.* Baricitinib for previously treated moderate or severe rheumatoid arthritis: an evidence review group perspective of a NICE single technology appraisal. *Pharmacoeconomics* 2018;36(7):769-78.
128. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed. London: NICE; 2016.
129. Fautrel B, Foltz V, Bourgeois P. Mortalité dans les rhumatismes inflammatoires chroniques. *Revue du Rhumatisme* 2006;73(5):530-1.
130. Mauskopf J, Samuel M, McBride D, Mallya UG, Feldman SR. Treatment sequencing after failure of the first biologic in cost-effectiveness models of psoriasis: a systematic review of published models and clinical practice guidelines. *Pharmacoeconomics* 2014;32(4):395-409.
131. Ghabri S, Autin E, Poullie AI, Josselin JM. The French National Authority for Health (HAS) guidelines for conducting budget impact analyses (BIA). *Pharmacoeconomics* 2018;36(4):407-17.
132. Michaud K, Vera-Llonch M, Oster G. Mortality risk by functional status and health-related quality of life in patients with rheumatoid arthritis. *J Rheumatol* 2012;39(1):54-9.
133. Jansen JP, Incerti D, Mutebi A, Peneva D, MacEwan JP, Stolshek B, *et al.* Cost-effectiveness of sequenced treatment of rheumatoid arthritis with targeted immune modulators. *J Med Econ* 2017;20(7):703-14.
134. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. *Rheumatology* 2004;43(1):62-72.
135. Diamantopoulos A, Benucci M, Capri S, Berger W, Wintfeld N, Giuliani G, *et al.* Economic evaluation of tocilizumab combination in the treatment of moderate-to-severe rheumatoid arthritis in Italy. *J Med Econ* 2012;15(3):576-85.
136. Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. *Arthritis Rheum* 2003;48(6):1530-42.
137. Athanasakis K, Tarantilis F, Tsalapati K, Konstantopoulou T, Vritzali E, Kyriopoulos J. Cost-utility analysis of tocilizumab monotherapy in first line versus standard of care for the treatment of rheumatoid arthritis in Greece. *Rheumatol Int* 2015;35(9):1489-95.
138. Beresniak A, Baerwald C, Zeidler H, Kruger K, Neubauer AS, Dupont D, *et al.* Cost-effectiveness simulation model of biologic strategies for treating to target rheumatoid arthritis in Germany. *Clin Exp Rheumatol* 2013;31(3):400-8.
139. Heather EM, Payne K, Harrison M, Symmons DP. Including adverse drug events in economic evaluations of anti-tumour necrosis factor-alpha drugs for adult rheumatoid arthritis: a systematic review of economic decision analytic models. *Pharmacoeconomics* 2014;32(2):109-34.
140. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, *et al.* A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess* 2006;10(42):iii-iv, xi-xiii, 1-229.
141. Merkesdal S, Kirchhoff T, Wolka D, Ladinek G, Kielhorn A, Rubbert-Roth A. Cost-effectiveness analysis of rituximab treatment in patients in Germany with rheumatoid arthritis after etanercept-failure. *Eur J Health Econ* 2010;11(1):95-104.
142. Malottki K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, *et al.* Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. *Health Technol Assess* 2011;15(14):1-278.
143. Claxton L, Taylor M, Gerber RA, Gruben D, Moynagh D, Singh A, *et al.* Modelling the cost-effectiveness of tofacitinib for the treatment of rheumatoid arthritis in the United States. *Curr Med Res Opin* 2018;34(11):1991-2000.
144. Russell A, Beresniak A, Bessette L, Haraoui B, Rahman P, Thorne C, *et al.* Cost-effectiveness modeling of abatacept versus other biologic agents in DMARDs and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis. *Clin Rheumatol* 2009;28(4):403-12.
145. Scholz S, Mittendorf T. Modeling rheumatoid arthritis using different techniques: a review of model construction and results. *Health Econ Rev* 2014;4(1):18.
146. Joensuu JT, Huoponen S, Aaltonen KJ, Kontinen YT, Nordstrom D, Blom M. The cost-effectiveness of biologics for the treatment of rheumatoid arthritis: a systematic review. *PLoS ONE* 2015;10(3):e0119683.
147. Chevreul K, Haour G, Lucier S, Harvard S, Laroche ML, Mariette X, *et al.* Evolution of direct costs in the first years of rheumatoid arthritis: impact of early versus late biologic initiation--an economic analysis based on the ESPOIR cohort. *PLoS ONE* 2014;9(5):e97077.

148. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care* 2005;21(2):240-5.
149. Adarkwah CC, van Gils PF, Hiligsmann M, Evers SM. Risk of bias in model-based economic evaluations: the ECOBIAS checklist. *Expert Rev Pharmacoecon Outcomes Res* 2016;16(4):513-23.
150. Ghabri S, Stevenson M, Moller J, Caro JJ. Trusting the results of model-based economic analyses: is there a pragmatic validation solution? *Pharmacoeconomics* 2018.
151. Bermejo I, Ren S, Simpson E, Clowes M, Scott DL, Young A, *et al.* Sarilumab for previously-treated moderate or severe rheumatoid arthritis: an evidence review group perspective of a NICE single technology appraisal. *Pharmacoeconomics* 2018;36(12):1427-37.
152. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, *et al.* Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Med Decis Making* 2012;32(5):678-89.
153. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, *et al.* Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making* 2012;32(5):733-43.
154. Caro JJ, Briggs AH, Siebert U, Kuntz KM, Force I-SMGRPT. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making* 2012;32(5):667-77.
155. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, *et al.* Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Making* 2012;32(5):722-32.
156. Chevalier J, de Pouvourville G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* 2013;14(1):57-66.
157. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92.
158. Brazier J, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851-9.
159. Gottenberg JE, Brocq O, Perdriger A, Lassoued S, Berthelot JM, Wendling D, *et al.* Non-TNF-targeted biologic vs a second anti-tnf drug to treat rheumatoid arthritis in patients with insufficient response to a first anti-TNF drug: a randomized clinical trial. *JAMA* 2016;316(11):1172-80.
160. Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, *et al.* Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum* 2006;54(9):2793-806.
161. Mariette X, Gottenberg JE, Ravaud P, Combe B. Registries in rheumatoid arthritis and autoimmune diseases: data from the French registries. *Rheumatology* 2011;50(1):222-9.
162. de La Forest Divonne M, Gottenberg JE, Salliot C. Revue systématique des registres de polyarthrites rhumatoïdes sous biothérapie dans le monde et méta-analyse sur les données de tolérance. *Revue du Rhumatisme* 2017;84(3):199-207.
163. de La Forest Divonne M, Gottenberg JE, Salliot C. Safety of biologic DMARDs in RA patients in real life: A systematic literature review and meta-analyses of biologic registers. *Joint Bone Spine* 2017;84(2):133-40.
164. Kobelt G. Treating to target with etanercept in rheumatoid arthritis: cost-effectiveness of dose reductions when remission is achieved. *Value Health* 2014;17(5):537-44.
165. Lekander I, Borgstrom F, Lysholm J, van Vollenhoven RF, Lindblad S, Geborek P, *et al.* The cost-effectiveness of TNF-inhibitors for the treatment of rheumatoid arthritis in Swedish clinical practice. *Eur J Health Econ* 2013;14(6):863-73.
166. Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. *Arthritis Rheum* 2000;43(12):2751-61.
167. Favalli EG, Pregnotato F, Biggioggero M, Becciolini A, Penatti AE, Marchesoni A, *et al.* Twelve-year retention rate of first-line tumor necrosis factor inhibitors in rheumatoid arthritis: real-life data from a local registry. *Arthritis Care Res* 2016;68(4):432-9.
168. Bertin P, Fagnani F, Duburcq A, Woronoff AS, Chauvin P, Cukierman G, *et al.* Conséquences de la polyarthrite rhumatoïde sur la trajectoire professionnelle, la productivité et l'employabilité : étude PRET. *Rev Rhum* 2015;82:306-11.
169. Fautrel B, Woronoff-Lemsi MC, Ethgen M, Fein E, Monnet P, Sibilia J, *et al.* Impact of medical practices on the costs of management of rheumatoid arthritis by anti-TNFalpha biological therapy in France. *Joint Bone Spine* 2005;72(6):550-6.
170. Rat AC, Boissier MC. Rheumatoid arthritis: direct and indirect costs. *Joint Bone Spine* 2004;71(6):518-24.
171. Juillard-Condât B, Constantin A, Cambon-Thomsen A, Bourrel R, Taboulet F. Coût de la polyarthrite rhumatoïde en France: comparaison leflunomide/etanercept. *Thérapie* 2007;62(2):137-42.
172. Juillard-Condât B, Constantin A, Cambon-Thomsen A, Bourrel R, Taboulet F. Impact of

etanercept on the costs of rheumatoid arthritis (RA): results from a French observational study. *Joint Bone Spine* 2008;75(1):25-8.

173. Price MJ, Welton NJ, Briggs AH, Ades AE. Model averaging in the presence of structural uncertainty about treatment effects: influence on treatment decision and expected value of information. *Value Health* 2011;14(2):205-18.

174. Strong M, Oakley JE, Chilcott J. Managing structural uncertainty in health economic decision models: a discrepancy approach. *Appl Statist* 2012;61(Part 1):25-45.

175. Ghabri S, Cleemput I, Josselin JM. Towards a new framework for addressing structural uncertainty in health technology assessment guidelines. *Pharmacoeconomics* 2018;36(2):127-30.

176. Soares M. Eliciting beliefs in health technology assessment to characterize uncertainties. *Value & Outcomes Spotlight* 2015:10-2.

177. Hôpitaux universitaires de Strasbourg. Randomized controlled clinical trial evaluating methotrexate + biologic versus methotrexate, salazopyrine and hydroxychloroquine in patients with rheumatoid arthritis and insufficient response to methotrexate. The BIO-3 trial. PHRC National 2014. Version 1.1. Strasbourg; 2014.

178. Caisse nationale d'assurance maladie. Statistiques et publications. Données statistiques. Affections de longue durée [En ligne] 2018.

<https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/affection-de-longue-duree-ald/index.php>

ÉQUIPE PROJET & REMERCIEMENTS

Équipe projet

- Ce rapport a été coordonné et rédigé par Salah Ghabri au sein du Service évaluation économique et santé publique (SEESP).
- La revue systématique et les méta-analyses en réseau sur les traitements de fond biologiques ont été réalisées en collaboration avec le Pr Matthew D. Stevenson (Département d'économie de la santé et de sciences de la décision, Université de Sheffield et Comité scientifique d'évaluation des technologies de santé, *National Institute for Health and Care Excellence*, NICE) et son équipe (Dr Emma Simpson, Dr Kate Ren, Dr Rachel Archer, Dr Emma Hock) qui ont participé à la réalisation et la validation de la méta-analyse en réseau (conception, méthodes et interprétation) et la rédaction de l'article portant sur ce thème.
- Le Pr Jaime Caro (Université de McGill et *London School of Economics*) a collaboré à :
 - la réalisation du modèle PR France dans le cadre d'une convention de partenariat avec la HAS. Il a mis en place une modélisation fondée sur la simulation DICE (*Discretely Integrated Condition Event*) ;
 - l'élaboration du protocole d'analyse de la cohorte ESPOIR, la rédaction et la présentation des documents techniques destinés au groupe de travail (GT PR).
- L'organisation logistique et le travail de secrétariat ont été réalisés par Samantha Fernandes (assistante projet polyarthrite rhumatoïde, SEESP).
- La recherche documentaire a été effectuée par Gaëlle Fanelli (documentaliste) et Yasmine Lombry (assistante documentaliste).

Remerciements :

- les membres du groupe de travail PR de leur contribution à l'avancement de ce projet et de leur relecture du rapport final : Dr Aymeric Binard, Dr Yves-Marie Pers, Franck Maunoury, Hans Martin Spath, Sandrine Rollot, Morgane Beck et François Bocquet et particulièrement Dr Aymeric Binard et Dr Yves-Marie Pers de leur participation continue à la conceptualisation du modèle économique PR France ;
- les membres de la CEESP de leurs remarques lors de l'examen du rapport final PR et particulièrement les deux rapporteurs, Lionel Perrier et Sophie Coté-Mesnier ;
- les membres du groupe de lecture PR : les Pr Bruno Fautrel, Gérard de Pouverville, Isabelle Durand-Zaleski, Irina Cleemput, Bernard Combe, Alain Saraux et Cecile Goujoux-Viala ; Dr Gwenaël le Teuff, Dr Mathilde Perron et Mme Nathalie Rincheval ;

- le comité scientifique de la cohorte ESPOIR qui a autorisé la HAS à faire des analyses statistiques sur les données de la cohorte ESPOIR. En particulier, les Pr Francis Berenbaum, Olivier Vittecoq, Bruno Fautrel et Mme Nathalie Rincheval (coordinatrice nationale de la cohorte ESPOIR), qui a réalisé un contrôle qualité indépendant sur les données utilisées, les estimations des utilités et les critères cliniques (ex. : DAS 28, HAQ) permettant le calcul des équations du modèle ;
- Catherine Rumeau-Pichon (adjointe à la directrice, Direction de l'évaluation médicale, économique et de santé publique (DEMESP) et chef de service du Service évaluation économique et santé publique (SEESP)) de ses remarques sur les messages clés du rapport ;
- Dr Olivier Scemama (adjoint au chef de service, SEESP), Anne-Isabelle Poullié et Erwan Autin (chefs de projet, SEEP) pour leurs relectures continues du rapport ;
- Dr Anne d'Andon (chef de service, Service évaluation des médicaments à la HAS, SEM) et Mathilde Grande (adjointe au chef de service, SEM) de leur aide pour consulter tous les dossiers des industriels portant sur les traitements de fond biologiques pour la période 2000-2018 ;
- Sarah Koné et Carine Busin (chefs de projet, SEM) de leur relecture du rapport (ou d'une partie du rapport) et leurs réponses aux questions cliniques relatives aux avis de la CT ;
- Frédérique Fleish (analyste, Service archivage – HAS) de son aide pour identifier et fournir les dossiers des industriels portant sur les traitements de fond biologiques au cours de la période 2000-2016 ;
- le Pr Andreas Gerber-Grote (IQWIG, Allemagne et ZHAW *School of Health Professions*, Suisse), le Pr Jean-Michel Josselin (ex-membre de la CEESP, Université de Rennes 1), le Pr Robert Launois (REES France), et le Pr Jean-Luc Harousseau (ancien président de la HAS, 2011-2015) de leurs conseils et leurs encouragements à la réalisation de ce projet ;
- le Pr Pelham Barton (Université de Birmingham) de sa vérification de la première version du modèle PR France et le Dr Jörgen Möller (Université de Lund, Suède et Evidera, Londres) pour sa réalisation de la validation interne et indépendante du modèle PR France ;
- le Pr David Scott et le Dr Fowzia Ibrahim (*King's College Hospital NHS Foundation Trust*, Londres, Royaume-Uni) d'avoir fourni les données portant sur l'essai clinique TACIT ;
- les internes de santé publique et de pharmacie :
 - o particulièrement, le Dr Chloé Herpin de sa collaboration dans la synthèse des données sur les coûts, la conception de l'algorithme des coûts et du changement des traitements biologiques et de son aide pour la préparation de la deuxième réunion du groupe de travail PR,

- Laurent Lam pour sa collaboration dans la préparation et l'analyse du protocole sur les événements indésirables,
- Jacques Li, Laurent Lam et Tristan Le Magoarou pour leur participation à la validation de l'adaptation de la grille « *Consolidated Health Economic Evaluation Reporting Standards* » (CHEERS) à l'analyse de la polyarthrite rhumatoïde et l'extraction indépendante des références de la littérature économique.



Toutes les publications de l'HAS sont téléchargeables sur
www.has-sante.fr