

Outcome assessment in Ankylosing Spondylitis in focus

Citation for published version (APA):

Wanders, A. (2005). *Outcome assessment in Ankylosing Spondylitis in focus*. Datawyse / Universitaire Pers Maastricht.

Document status and date:

Published: 01/01/2005

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Summary in perspective

The studies described in this thesis cover important aspects of outcome assessment in Ankylosing Spondylitis (AS). For decades, research in AS was focused on eliciting the pathophysiological mechanisms operative in this heterogeneous disease, rather than on outcome assessment. The latter was considered unnecessary, since effective drugs were lacking and clinical trials were rarely performed in AS.

Two consecutive developments have changed this picture dramatically. The first development was the foundation of the international ASsessment in Ankylosing Spondylitis (ASAS) working group in 1995 in Amsterdam, the subsequent expert-based establishment of domains of outcome, and core sets of measurements to assess efficacy within these domains, and the endorsement of these core sets by the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiative in 1998. The second development was the discovery of biological drugs that have the potential to inhibit (the activity of) tumor necrosis factor (TNF)-alpha, and their introduction in chronic inflammatory diseases such as RA. Since anecdotal reports suggested efficacy of TNF-blocking drugs in AS and in Crohn's disease (a disease associated with AS) as well, and pharmaceutical industry expressed their willingness to further develop TNF-blocking drugs in this disease, an immediate need for appropriate outcome measures to be used in clinical trials arose. The application of TNF-blocking drugs in AS clinical trials has boosted the recent development of outcome measures in AS, has been a justification for the scientific work summarized in this thesis, and has set the research agenda for the forthcoming years.

In chapter 2, responsiveness and discriminatory capacity of measurement instruments included in the core set for disease controlling antirheumatic therapy (DC-ART) in AS was investigated, making use of the data of one of the first clinical trials with TNF-blocking therapy in AS. Responsiveness and discriminatory capacity are important characteristics of measurement instruments with respect to their use in clinical trials. It was shown that the majority of instruments in the DCART core set indeed fulfilled these prerequisites, which makes them useful as endpoints in clinical trials.

Chapter 3 focuses on measuring radiographic damage in AS. Radiographic damage was considered an important domain already by ASAS, but up to recently, an appropriately validated instrument to measure radiographic damage and progression of damage was lacking. In this chapter, three scoring methods to measure radiographic damage of the spine were compared with respect to the validation criteria proposed by OMERACT: 1) Truth, (*or*: does the instrument really measure what it is supposed to measure?); 2) Discrimination, (*or*: does the measure discriminate between groups of patients with different levels of severity?); and 3) Feasibility, (*or*: is it feasible to use this measure in the context of clinical trials?).

One scoring method, the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS), appeared to outreach the other two tested methods in terms of reliability (between readers) and sensitivity-to-change. Based on this study, the mSASSS was proposed as the method of choice to assess radiographic progression in AS clinical trials, and as such endorsed by OMERACT in 2004.

Chapter 4 focuses on a methodological issue pertaining to scoring radiographic progression in clinical trials: should radiographs be read “chronologically” (with known time order) or “paired” (with concealed time order)? The theoretical advantage of chronological reading (more signal picked up) can be a source of bias in clinical trials, because it may give the reader a direction to expected deterioration and the scores may be dependent on the magnitude of progression, which can be different in both treatment groups. At the other hand, paired reading may be at the cost of signal intensity. Signal intensity has always been a source of concern, since available data showed that measurable radiographic progression in patients with AS is slow. This study, however, showed that even with paired reading there was sufficient progression in a 2-year period to be picked up by the mSASSS. This finding is reassuring, since the methodology of measuring radiographic progression requested by registration authorities includes scoring with paired time order.

Chapter 5 addresses a clinical problem: “Is it really necessary to take radiographs of the spine to be informed about the extent of damage, or is it sufficient to clinically determine spinal mobility by one or more of the available spinal mobility instruments?” This study shows that spinal mobility and radiographic damage are expectedly correlated on the group level, but that concordance between spinal mobility measurements and radiographic damage are insufficient to assure an appropriate classification at the individual patient level. The conclusion of this study therefore was that spinal mobility measures could not replace radiography of the spine in order to get an impression about structural damage in patients with AS, which gives justice to the inclusion of the domain radiographic damage in the ASAS core set for DCART in addition to the domain spinal mobility.

Chapter 6 of this thesis treads upon a methodological by-path: Selection of patients for a clinical trial at baseline based on prediction of the outcome of interest. Usually, patients selected for a clinical trial have a relatively bad prognosis (*or*: a high propensity on an unfavorable outcome) because it is widely accepted that such a selection ameliorates the statistical demonstration of a treatment difference. This study shows that the extent to which treatment contrast is determined by baseline propensity is dependent on how a treatment reduces the risk of an unfavorable outcome: by absolute risk reduction or by relative risk reduction. Though relative risk reduction models (*eg.* 50% reduction of the risk irrespective of the baseline risk) are most prevalent in pharmacology, a few examples from rheumatoid arthritis clinical trials suggest an

absolute risk reduction model for two treatment strategies, which may have implications for the selection of patients for such trials.

Chapter 7 describes the results of a randomized controlled trial in which all methodological principles of scoring radiographic progression as mentioned before were applied. The trial compared two strategies of treatment with non-steroidal anti-inflammatory drugs (NSAIDs) in patients with AS: one strategy with the continuous use of NSAIDs, one strategy with on- demand-only use of NSAIDs. The primary outcome measure was radiographic progression after a 2-year follow up period. Radiographic damage was scored by mSASSS, as outlined before, with radiographs presented with concealed time order (paired reading). The study not only showed that progression was measurable in a considerable proportion of patients (confirming sensitivity to change of the method), but also that continuous NSAID use was associated with significantly less radiographic progression as compared to on-demand-only NSAID use (discriminatory capacity). These results showed that 2-year radiographic progression is a feasible endpoint in AS clinical trials, confirmed that the mSASSS is an appropriate instrument to assess radiographic progression, and shed new light on the potential of NSAIDs as drugs that may influence structural damage in AS.

How will the studies described in this thesis contribute to clinical science in the field of AS in the near future?

It is beyond any doubt that clinical outcome research in AS in the year 2005 stands at a different level as compared to a decade ago. As stated above, evaluation of TNF-blocking drugs and outcome research in AS have mutually catalyzed each other's development, reaching a state that approximates the level of development of rheumatoid arthritis. Putatively as a consequence of the speed of development, a few areas of outcome assessment in AS are relatively untouched. First, unlike the situation in RA, and despite many efforts in the past, it is still not known how to predict a favorable or an unfavorable long-term outcome of the disease at presentation. Second, whilst the efficacy of TNF-blocking drugs on signs and symptoms of the disease is undisputed, it is still unknown whether these drugs truly influence long-term outcome, including structural damage. Third, and related to some extent, many experts in the field consider the success of TNF-blocking drugs in AS, which is based on subjective improvements, insufficiently substantiated by improvements on objective instruments that are beyond the influence of patient's perceptions.

Prediction of long-term outcome in AS, drug effects on structural damage, and objective improvements by drugs will be major fields of research in the forthcoming years, which all require a relevant (set of) outcome measure(s). There are good methodological arguments to postulate that imaging - plain radiography of the spine to assess structural damage and magnetic resonance imaging (MRI) of the spine to assess inflammation

and to some extent structural damage - will play an increasingly important role in the assessment of various types of outcome in AS.

Radiography of the spine reveals structural damage, which is associated with spinal mobility and with function, and therefore is considered relevant to patients. What is still unknown is whether radiographic progression is determined by disease activity and/or inflammation, as in RA, but the analysis of databases of clinical trials and long-term observational studies will shed light on this relationship. A very important field of research will be the substantiation of inhibition of radiographic progression by TNF-blocking drugs. Most probably, the evaluation of a "structural damage claim" for such drugs will be unprecedented because normal trial designs do not suffice. It is widely considered unethical to perform placebo-controlled clinical trials in which patients are refrained from effective therapy for a period of at least 2 years, required to obtain a sufficiently large treatment contrast in radiographic progression. Technical developments in the processing of radiographs (*eg.* digitization) make it possible to match radiographs of actively treated patients with radiographs of patients from untreated cohorts in an indistinguishable manner, and offer these radiographs to readers for scoring in a concealed time order. Obviously, such an experimental set-up lacks the advantage of prognostic similarity of treatment groups, which is inherent to randomization. But differences between artificial treatment groups are only relevant in so far they determine radiographic progression. Unlike the situation in RA, up to now, thorough analysis with multiple demographic, clinical, laboratory and genetic variables has not revealed a single variable that contributes to explaining 2-year radiographic progression. It is therefore to be expected that – given the impossibility of a 2-year randomized controlled trial - a controlled cohort design, in which radiographic progression of actively treated AS patients is compared with radiographic progression of not concurrently treated control patients, will reveal useful information with regard to the potential of TNF-blocking drugs to inhibit progression of structural damage, and to fill in lacking information pertaining to objective outcome domains.

MRI of the spine is an imaging modality with promising prospects. MRI techniques visualize inflammation of the vertebrae and adjacent structures, as well as structural damage such as the formation of syndesmophytes and vertebral erosions. Preliminary data have shown that "MRI activity" of the spine can be suppressed by TNF-blocking drugs with a time interval of 6 weeks, whilst being unchanged in placebo treated patients. The actual place of MRI as a measurement instrument in AS clinical trials, and later on in clinical practice, is dependent on how accurately MRI inflammation reflects signs and symptoms of AS, function and quality of life, as expressed by patients, and how accurately MRI inflammation predicts impairment of spinal mobility and structural damage later on. An international working group under the umbrella of ASAS and OMERACT, and consisting of experts in the field of imaging, is now pursuing a thorough validation program in order to establish the value of MRI of the spine as an outcome instrument for AS. It is to be expected that "MRI-activity" will become important in the decision whether a patient with AS has active disease, whether this patient has an

unfavorable prognosis, and whether treatment with TNF-blocking drugs should be started. Other fields in development are biomarker research, genomics and proteomics which are also applied to AS. MRI in combination with plain radiography, together with the achievements of these new research fields, may help elucidating pathophysiological mechanisms underlying AS.

In summary, AS is a popular research field nowadays. A better understanding of the disease mechanisms, of its course and impact, and of the various effects of drugs is all within the scope of currently planned research. A high standard of outcome assessment is crucial in this process, and the studies described in this thesis may have added to this.

Samenvatting in perspectief

De in dit proefschrift beschreven studies bevatten belangrijke aspecten van *outcome assessment* (het beoordelen van de lange termijn uitkomst) van de ziekte van Bechterew. Het onderzoek op het gebied van de ziekte van Bechterew heeft zich tientallen jaren voornamelijk gericht op de pathofysiologische mechanismen van deze aandoening, en niet zozeer op *outcome assessment*. Dit laatste werd als onnodig beschouwd omdat er immers geen medicijnen beschikbaar waren die het ziektebeloop positief konden beïnvloeden. Daarnaast werd ook bijna geen onderzoek gedaan met medicamenten.

Twee opeenvolgende ontwikkelingen hebben dit beeld drastisch veranderd. Allereerst vond in 1995 de oprichting van de internationale werkgroep "ASsessment in Ankylosing Spondylitis (ASAS)" plaats te Amsterdam. Deze internationale groep van deskundigen definieerde de aspecten van de ziekte van Bechterew die gemeten dienden te worden voor verschillende situaties, bijvoorbeeld een klinische trial, of de dagelijkse praktijk. Ook selecteerde zij de bijbehorende meetinstrumenten. Deze ASAS activiteiten werden in 1998 bekrachtigd door OMERACT, het wereldwijde initiatief op het gebied van het ontwikkelen van uitkomstmaten bij reumatische ziekten.

Een tweede belangrijke ontwikkeling was de ontdekking van biologische medicijnen die in staat waren om tumor necrosis factor alpha te remmen, de zogeheten anti-TNF middelen. Tumor necrosis factor alpha is een stof die deel uitmaakt van een ontstekingscascade bij chronische ontstekingsziekten. De introductie van deze middelen bij de behandeling van patiënten met reumatoïde artritis liet spectaculaire resultaten zien. Op kleinere schaal werd er ook geëxperimenteerd met patiënten met de ziekte van Crohn (een ziekte die geassocieerd is met de ziekte van Bechterew) en bij patiënten met de ziekte van Bechterew. Deze eerste resultaten waren positief. De farmaceutische industrie wilde dan ook de anti-TNF middelen laten registreren voor de ziekte van Bechterew, waarvoor grote klinische trials noodzakelijk zijn, met geschikte meetinstrumenten. Het kan rustig worden gesteld dat de toepassing van anti-TNF middelen bij de ziekte van Bechterew heeft gezorgd voor een enorme stimulans voor het onderzoek op het gebied van *outcome assessment* bij de ziekte van Bechterew. Deze ontwikkelingen zijn een rechtvaardiging voor het wetenschappelijke werk zoals samengevat in dit proefschrift, en hebben tevens de onderzoeksagenda voor het Bechterew onderzoek voor de komende jaren bepaald.

Met de komst van de anti-TNF middelen lijkt het erop dat er medicatie beschikbaar is gekomen waarmee het ziektebeloop fundamenteel kan worden beïnvloed. Of dit ook daadwerkelijk het geval is zal door onderzoek in klinische trials moeten worden uitgezocht. Voor dit doel heeft de ASAS een speciale set meetinstrumenten ontwikkeld, de DCART core set. In hoofdstuk 2 wordt een studie beschreven waarin de gevoeligheid en het onderscheidend vermogen van de instrumenten behorend tot deze set onderzocht wordt. Hierbij wordt gebruik gemaakt van de resultaten van een van de

eerste trials waarin anti-TNF medicatie bij Bechterew patiënten werd getest. Het bleek dat het merendeel van de geselecteerde instrumenten over een voldoende gevoeligheid en onderscheidend vermogen beschikten, hetgeen ze tot bruikbare instrumenten maakt voor het beoordelen van klinische trials.

Hoofdstuk 3 richt zich op het meten van radiologische schade van de ziekte van Bechterew. Door de ASAS werkgroep werd radiologische schade als een belangrijke uitkomst gezien, echter een gevalideerd en geschikt instrument om radiologische schade en het voortschrijden van deze schade (progressie) te meten, ontbrak. In dit hoofdstuk worden drie scoringsmethodes, die de radiologische schade van de wervelkolom meten, vergeleken op basis van validatie-criteria zoals voorgesteld door de OMERACT. Deze criteria zijn als volgt: 1) Meet het instrument daadwerkelijk wat het veronderstelt te meten? 2) Kan het instrument onderscheid maken tussen groepen patiënten met een verschillende ernst van de aandoening? . 3) Is het praktisch haalbaar om het instrument toe te passen in de context van een klinische trial? Deze drie criteria worden in het OMERTACT filter aangeduid met de volgende Engelse termen; truth, discrimination and feasibility.

Eén scoringsmethode, de gemodificeerde Stoke Ankylosing Spondylitis Spinal Score (mSASSS), bleek beter dan de andere twee met betrekking tot betrouwbaarheid (dat wil zeggen: de hoogste mate van overeenkomst tussen verschillende gebruikers van de methode) en de mSASSS bleek het meest gevoelig voor veranderingen. Gebaseerd op deze studie werd de mSASSS voorgesteld als de methode van eerste keus om radiologische progressie in klinische trials bij Bechterew patiënten te beoordelen. Dit voorstel werd in 2004 door OMERACT bekrachtigd.

Hoofdstuk 4 richt zich op de methodologie van de scorings volgorde van röntgenfoto's om radiologische progressie te beoordelen. De vraag hierbij is of de foto's van een patiënt gescoord moeten worden in een chronologische, dus bekende volgorde, dan wel in een willekeurige volgorde. Het theoretische voordeel van het scoren in chronologische volgorde is dat er een grotere verandering wordt gezien. Maar dit voordeel is ook gelijk een nadeel. Het bekend zijn van de volgorde kan er immers toe leiden dat men geneigd is meer schade te zien dan er daadwerkelijk is, omdat men meer schade verwacht. Het scoren van de foto's in willekeurige volgorde heeft dit nadeel niet. Maar men is geneigd behoudender te scoren, en het gevaar bestaat dat een daadwerkelijk bestaand verschil tussen twee verschillende groepen niet wordt gezien. Op die manier gaat het scoren met een willekeurige volgorde ten koste van signaalintensiteit. De signaalintensiteit is altijd al een bron van zorg geweest omdat de progressie bij de ziekte van Bechterew zeer langzaam verloopt. De studie beschreven in dit hoofdstuk laat echter zien dat zelfs met het scoren van de foto's in willekeurige volgorde middels de mSASSS, er meetbare progressie optreedt gedurende een follow up duur van 2 jaar.

In hoofdstuk 5 komt de volgende klinische vraag aan de orde: " Is het werkelijk nodig om röntgenfoto's van de wervelkolom te nemen om geïnformeerd te worden over de mate van schade, of is het voldoende om klinisch de beweeglijkheid van de wervelkolom te bepalen met behulp van één of meer meetinstrumenten? Deze studie laat zien dat er zoals verwacht een correlatie bestaat tussen spinale mobiliteit en radiologische schade, maar dat de overeenkomst tussen deze spinale mobiliteit en radiologische schade onvoldoende is om de individuele patiënt correct te classificeren. De conclusie van deze studie is dan ook dat het meten van de beweeglijkheid van de wervelkolom niet als vervanging kan dienen voor röntgenfoto's om een indruk te krijgen van de structurele schade bij Bechterew patiënten. Deze conclusie rechtvaardigt de inclusie van het beoordelen van radiologische schade naast het beoordelen van de spinale mobiliteit in de DC-ART core set van de ASAS.

Hoofdstuk 6 van dit proefschrift beschrijft een uitstapje op een methodologisch zijpad: selectie van patiënten voor een klinische trial. Normaal gesproken worden patiënten geselecteerd die een slechte prognose hebben of te wel een hoge kans op een ongunstig resultaat. Dit omdat een dergelijke selectie de kans het grootst maakt dat statistisch kan worden aangetoond dat een behandeling daadwerkelijk effect heeft. De in dit hoofdstuk beschreven studie laat zien dat de mate waarin selectie het behandelingscontrast beïnvloedt afhangt van de manier waarop de behandeling het risico van de ongunstige uitkomst reduceert; door absolute dan wel relatieve risico reductie. Hoewel relatieve risico reductie modellen (bijvoorbeeld 50% minder kans op de ongunstige uitkomst aan het eind van de behandeling ten opzichte van het begin risico) het meest voorkomen binnen de farmacologie, suggereren een aantal voorbeelden afkomstig uit trials met reumatoïde arthritis patiënten dat een absoluut risico reductie model werkzaam is. Dit zou implicaties kunnen hebben voor de selectie van patiënten voor klinische trials.

Hoofdstuk 7 beschrijft de resultaten van een gerandomiseerde gecontroleerde trial waarin alle eerder beschreven methodologische principes worden toegepast. In de trial worden twee behandelingsstrategieën met non-steroidale anti-inflammatoire medicijnen (NSAIDs) bij Bechterew patiënten vergeleken. Bij de eerste behandelstrategie gebruikt de patiënt voortdurend een NSAID en bij de tweede strategie gebruikt de patiënt alleen een NSAID als hij/zij klachten heeft, dan wel symptomen bemerkt ("naar behoefte"). Als primaire uitkomstmaat werd gekeken naar de radiologische progressie over 2 jaar. De radiologische progressie werd gescoord middels de mSASSS waarbij de foto's in willekeurige volgorde werden bekeken. Deze studie liet niet alleen zien dat er bij een aanzienlijk deel van de patiënten inderdaad progressie zichtbaar was maar ook dat het continue gebruik van NSAIDs geassocieerd was met een significant lagere radiologische progressie als werd vergeleken met de patiënten die een NSAID gebruikten naar behoefte. Deze resultaten bevestigen dat radiologische progressie over 2 jaar een praktisch haalbare uitkomstmaat is in trials met Bechterew patiënten. Tevens

bevestigen deze resultaten dat de mSASSS een geschikt instrument is voor het meten van radiologische progressie en werpen zij nieuw licht op de rol die NSAIDs spelen in de behandeling van Bechterew patiënten.

Hoe kunnen de studies die zijn beschreven in dit proefschrift bijdragen aan het klinisch wetenschappelijk onderzoek naar de ziekte van Bechterew in de nabije toekomst?

Het is zonder enige twijfel een feit dat het klinisch onderzoek met betrekking tot *outcome assessment* van de ziekte van Bechterew van een geheel ander nivo is dan een decennium eerder. Zoals reeds eerder vermeld hebben de komst van anti-TNF middelen en het onderzoek van *outcome assessment* bij Bechterew elkaars ontwikkeling gekatalyseerd. Dit heeft voor *outcome assessment* geleid tot een status die ongeveer gelijk is aan die van *outcome assessment* in het onderzoeksveld van reumatoïde artritis, een veld dat tot voor kort een grote voorsprong had. Door de enorme snelheid waarmee de ontwikkelingen zich binnen het Bechterew onderzoek hebben voltrokken zijn er een aantal gebieden die relatief weinig aandacht hebben gehad. Ten eerste is het nog steeds niet bekend, ondanks diverse pogingen, hoe bij vroege ziekte het langere termijn beloop voorspeld kan worden, dit in tegenstelling tot de situatie bij reumatoïde artritis. Ten tweede: het effect van anti-TNF middelen op klachten en symptomen van Bechterew is onomstreden, maar het is nog steeds onbekend of deze middelen de lange termijn uitkomst beïnvloeden, inclusief structurele schade aan de wervelkolom. Als derde punt sluit hierbij aan dat een aantal experts het succes van anti-TNF blokkers bij de ziekte van Bechterew beschouwen als een succes dat gebaseerd is op subjectieve symptomen, en niet op objectieve instrumenten (instrumenten die niet kunnen worden beïnvloed door de perceptie van de patiënt).

Deze drie punten, het voorspellen van de lange termijn uitkomst van Bechterew, de effecten van medicatie op structurele schade, en objectieve verbeteringen door medicatie, zullen de komende jaren belangrijke onderwerpen zijn in het Bechterew onderzoek. Al deze zaken vereisen relevante uitkomstmaten. Er zijn goede methodologische argumenten om te stellen dat de conventionele röntgenfoto van de wervelkolom om structurele schade te beoordelen en een MRI van de wervelkolom om de ontsteking te beoordelen (en tot op bepaalde hoogte ook de structurele schade), een steeds belangrijker rol zullen gaan spelen bij het beoordelen van de uitkomst van de ziekte van Bechterew.

Radiologie van de wervelkolom laat structurele schade zien, die geassocieerd is met spinale beweeglijkheid en met functioneren. Door deze associatie wordt radiologie beschouwd als relevant voor de patiënt. Het is nog steeds onbekend of radiologische progressie bepaald wordt door ziekte activiteit en/of ontsteking, zoals bij reumatoïde artritis het geval is. Analyses van gegevens van klinische trials en langdurig observationeel onderzoek zullen hopelijk meer licht werpen op deze relatie.

Een zeer belangrijk onderwerp van onderzoek zal zijn het bevestigen van het remmende effect van de anti-TNF middelen op radiologische progressie bij Bechterew

patiënten. Hierbij doet zich echter een ethisch probleem voor. Het is onverantwoord om een placebo gecontroleerde trial te doen gedurende 2 jaar, de tijd die nodig is om een verschil in radiologische progressie aan te tonen, en daarmee de patiënten behorend tot de placebo groep een therapie die effectief is om de klinische verschijnselen te bestrijden, te onthouden.

Technische ontwikkelingen, zoals het digitaliseren van oude röntgenfoto's, maken het mogelijk om röntgenfoto's van patiënten die met anti-TNF middelen zijn behandeld te matchen met reeds bestaande röntgenfoto's van patiënten die niet behandeld zijn en deze aan te bieden om te scoren en te vergelijken. Omdat in een dergelijke experimentele opzet geen sprake is van randomisatie (toevalsverdeling) zijn de behandelgroepen niet per definitie gelijk qua prognose van de ziekte. Echter een verschil in prognostische gelijkheid is alleen relevant als het gaat om factoren die de uitkomst, in dit geval radiologische progressie, beïnvloeden. En voor Bechterew, in tegenstelling tot reumatoïde artritis, is ondanks grondige analyse van meerdere demografische, klinische, laboratorium en genetische variabelen, er geen enkele variabele aan het licht gekomen die als prognostische factor aangemerkt kan worden. Vandaar dat kan worden verwacht dat een experiment als hierboven beschreven bruikbare informatie zal opleveren met betrekking tot het effect van anti-TNF middelen op radiologische schade.

MRI van de wervelkolom is een beeldvormende techniek met veelbelovende verwachtingen. Middels MRI kan ontsteking van de wervels en van de aanliggende structuren zichtbaar gemaakt worden, evenals structurele schade, zoals syndesmofyten en vertebrale erosies. Voorlopige gegevens hebben laten zien dat ontstekingsactiviteit zoals door MRI wordt weergegeven, kan worden onderdrukt door het gedurende 6 weken gebruiken van anti-TNF middelen, terwijl deze ontstekingsactiviteit in met placebo behandelde patiënten onveranderd was gebleven. De plaats van MRI als een meetinstrument voor Bechterew in klinische trials, en later in de klinische praktijk, is afhankelijk van de vraag hoe accuraat de ontstekingsactiviteit zoals gezien middels MRI, de klachten en symptomen, functie en kwaliteit van leven van Bechterew patiënten weergeeft, en hoe accuraat de ontstekingsactiviteit zoals waargenomen op MRI de spinale mobiliteit en structurele schade voorspelt. Een internationale werkgroep bestaand uit experts op het gebied van beeldvorming is nu onder de paraplu van ASAS en OMERACT bezig om de waarde van MRI van de wervelkolom als meetinstrument vast te stellen. Naar verwachting zal ontstekingsactiviteit zoals door MRI waargenomen een belangrijk item worden in de besluitvorming of een Bechterew patiënt een actief ziektebeloop heeft, of deze patiënt een ongunstige prognose heeft en of behandeling met anti-TNF gestart moet worden. Op het gebied van onderzoek naar biomarkers, genen en eiwitten zijn er ontwikkelingen gaande die ook hun toepassing in het Bechterew onderzoek zullen vinden. MRI in combinatie met conventionele radiologie en de resultaten van de hier bovengenoemde onderzoeksvelden kunnen behulpzaam zijn bij het ontrafelen van het pathofysiologisch mechanisme dat ten grondslag ligt aan de ziekte van Bechterew.

Samengevat kan worden gesteld dat het onderzoek op het gebied van Bechterew momenteel volop in de belangstelling staat. Een beter begrip van de ziektemechanismen, van het ziektebeloop en de gevolgen, en van de vele effecten van behandeling met medicijnen zijn allen binnen het bereik van het huidige lopende en geplande onderzoek. In dit proces is een hoge standaard met betrekking tot *outcome assessment* cruciaal, en de studies beschreven in dit proefschrift hebben daar zo mogelijk aan bijgedragen.