



Universiteit
Leiden

The Netherlands

Raising the bar for classification and outcome assessment for clinical studies in axial spondyloarthritis

Boel, A.

Citation

Boel, A. (2022, October 18). *Raising the bar for classification and outcome assessment for clinical studies in axial spondyloarthritis*. Retrieved from <https://hdl.handle.net/1887/3483568>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3483568>

Note: To cite this publication please use the final published version (if applicable).



CHAPTER 11

NEDERLANDSE SAMENVATTING

Axiale spondyloartritis (axSpA) is een chronische reumatische ziekte die meestal ontstaat tussen het 20^e en 30e levensjaar. AxSpA wordt veroorzaakt door chronische ontsteking in de wervelkolom en de belangrijkste symptomen zijn chronische rugpijn (aanwezig voor ten minste 3 maanden) en stijfheid van de rug.

In tegenstelling tot andere reumatische aandoeningen, zoals reumatoïde artritis waarbij ontsteking kan leiden tot botafbraak, wordt axSpA gekenmerkt door de vorming van nieuw bot op de plek van ontsteking. Dit kan resulteren in botvergroeiingen in de sacro-iliacale gewrichten (het bekken) en de wervelkolom. Deze vergroeiingen dragen bij aan beperkingen in mobiliteit en fysiek functioneren, waardoor veel activiteiten van het dagelijks leven negatief worden beïnvloed. Aangezien axSpA meestal op relatief jonge leeftijd optreedt, zullen patiënten het grootste deel van hun leven moeten omgaan met deze ziekte. Naast pijn en stijfheid ervaren veel patiënten vermoeidheid en slaapproblemen, die allemaal een grote invloed hebben op de kwaliteit van leven en hun vermogen om deel te nemen aan dagelijkse activiteiten, zoals (huishoudelijk) werk.

Er zijn twee subtypes van axSpA: 1) radiografische axSpA (r-axSpA, ook bekend als ankyloserende spondylitis (AS)), gekenmerkt door aanzienlijke structurele schade aan de sacro-iliacale gewrichten (het bekken) zichtbaar op röntgenfoto's; en 2) niet-radiografische axSpA (nr-axSpA), gekenmerkt door klinische symptomen van axSpA in afwezigheid van duidelijke schade zichtbaar op röntgenfoto's. Nr-axSpA wordt vaak beschouwd als een vroeg stadium van de ziekte, wat impliceert dat patiënten gaandeweg r-axSpA kunnen ontwikkelen. Er zijn bepaalde risicofactoren voor het ontwikkelen van r-axSpA, zoals mannelijk geslacht, aanwezigheid van HLA-B27 gen (een erfelijke factor, welke leidt tot een verhoogde kans op ontstekingsziekten zoals axSpA), hoge ontstekingsactiviteit (d.w.z. verhoogde ontstekingswaarden in het bloed of ontsteking zichtbaar op MRI), en roken. Progressie van nr-axSpA naar r-axSpA wordt gezien bij ongeveer 5-20% van de patiënten in een tijdsperiode van 2-5 jaar, terwijl een deel van de patiënten nooit r-axSpA ontwikkelt. Nr-axSpA is dus meer dan alleen een vroeg stadium van ziekte, het is ook een ziekte-expressie. De klachten en ziekteactiviteit van patiënten met nr-axSpA zijn even ernstig en beperkend als die van patiënten met r-axSpA.

Er is geen eenduidig patroon waar alle patiënten met axSpA aan te herkennen zijn. Ziektekenmerken kunnen zelfs zeer verschillen tussen patiënten, wat axSpA een zogenaamd heterogene ziekte maakt. Naast de kenmerkende rugklachten zijn er verschillende andere klinische kenmerken die veel voorkomen bij patiënten met axSpA, de zogenaamde spondyloartritis (SpA) kenmerken. De meest voorkomende is inflammatoire rugpijn. Van inflammatoire rugpijn is volgens huidige criteria sprake als ten minste vier van de volgende vijf factoren aanwezig zijn: 1) ontstaan van de rugklachten vóór de leeftijd van 40; 2) geleidelijk ontstaan; 3) verbetering met beweging; 4) geen verbetering met

rust; en 5) nachtelijke pijn met verbetering bij het opstaan. Andere SpA kenmerken zijn een goede reactie op pijnstillers met een ontstekingsremmende werking (zogenaamd NSAIDs zoals naproxen of ibuprofen) en het voorkomen van SpA bij een familielid. Indien een eerstegraads (ouder, broer/zus, kind) of tweedegraads (grootouder, oom/tante, neef/nicht) familielid de diagnose SpA heeft spreken we van een positieve familieanamnese.

Ook klachten zoals ontstekingen in gewrichten buiten bekken en wervelkolom (artritis), ontstekingen van peesaanhechtingen (enthesitis) en zwellingen die leiden tot worstvormige vingers en tenen (dactylitis) zijn SpA kenmerken. Ontstekingen buiten de gewrichten en rug, zoals de ogen (uveitis), huid (psoriasis) en darmen (ziekte van Crohn en colitis ulcerosa, IBD) worden aangeduid als extra-musculoskeletale aandoeningen (niet tot het bewegingsapparaat behorend), en zijn ook SpA kenmerken. Verder zijn verhoogde ontstekingswaarden in het bloed en de aanwezigheid van de genetische factor HLA-B27 SpA kenmerken.

Aangezien axSpA bij vrijwel alle patiënten in ieder geval het bekken aantast, speelt beeldvorming van de sacro-iliacale gewrichten een cruciale rol bij de diagnose en classificatie van axSpA. Röntgenfoto's en MRI zijn de meest gebruikte beeldvormende technieken in de klinische praktijk. Er zijn echter beperkingen aan het gebruik van röntgenfoto's van de sacro-iliacale gewrichten bij patiënten met een vroege ziekte, omdat structurele veranderingen over het algemeen jaren duren. Daarom kan MRI van de sacro-iliacale gewrichten waardevolle informatie opleveren, omdat het de identificatie van actieve ontsteking mogelijk maakt, evenals de aanwezigheid van structurele veranderingen die het gevolg zijn van ontstekingen. Alle SpA-kenmerken zijn zeer nuttig bij de diagnose van axSpA en bij de classificatie van patiënten voor klinische onderzoeken. Bovendien bieden deze kenmerken belangrijke informatie over de ziekteprognose.

In de reumatologie zijn classificatiecriteria bedoeld om goed gedefinieerde, relatief homogene groepen patiënten te creëren voor klinisch onderzoek. Gevalideerde classificatiecriteria zijn van groot belang voor de interpretatie van onderzoeksbevindingen en vergelijkingen van resultaten tussen onderzoeken. De classificatiecriteria omvatten niet het hele spectrum van mogelijke uitingen van een ziekte, maar moeten zeer specifiek zijn om het onjuist labelen als het hebben van een ziekte tot een minimum te beperken.

Aangezien reumatische aandoeningen heterogeen van aard zijn, kunnen classificatiecriteria er niet in slagen om alle patiënten met axSpA te identificeren. Dit komt doordat classificatiecriteria gericht zijn op een meer homogene populatie dan in de dagelijkse klinische praktijk wordt gezien. Classificatiecriteria mogen dus niet worden gebruikt om patiënten te diagnosticeren, maar uitsluitend om patiënten te includeren in klinische onderzoeken.

Classificatiecriteria zouden van toepassing moeten zijn op alle patiënten met axSpA waar dan ook ter wereld. Dit is nodig om te garanderen dat de patiënten die worden geselecteerd voor deelname aan klinische onderzoeken wereldwijd hetzelfde zijn. In dit proefschrift willen we een internationaal perspectief bieden op de karakterisering van patiënten met axSpA-specifiek met betrekking tot de leeftijd bij aanvang van de symptomen en positieve familieanamnese van axSpA-, om te onderzoeken of classificatiecriteria inderdaad wereldwijd toepasbaar zijn.

Verder is het belangrijk dat alle onderzoeken die in verschillende delen van de wereld worden uitgevoerd dezelfde uitkomstmaten meten en deze ook op een vergelijkbare manier rapporteren. Dit om te zorgen dat bijvoorbeeld gegevens uit Amerikaanse onderzoeken kunnen worden vergeleken met onderzoeken die in Azië zijn uitgevoerd. Uitkomstmaten omschrijven alle meetinstrumenten die worden gebruikt om data te verzamelen in klinische studies, zoals een vragenlijst over pijn, of een score voor gewrichtspijn. Door het gebruik van dezelfde uitkomstmaten voorafgaand aan de start van de behandeling en na afloop van de behandeling, kunnen we iets zeggen over het effect dat de behandeling heeft gehad, bijvoorbeeld op pijn. Hierin is een belangrijke rol weggelegd voor een standaard set uitkomstmaten, ook wel een 'core outcome set' genoemd. De standaard set beschrijft de gegevens die ten minste gemeten/verzameld en gerapporteerd moeten worden wanneer er een studie wordt uitgevoerd. Door standaardisatie van metingen en rapportage wordt het mogelijk om directe vergelijkingen te maken tussen klinische onderzoeken naar de effectiviteit en veiligheid van behandelingen en wordt voorkomen dat alleen de gunstige uitkomsten gerapporteerd worden.

De eerste stap die moet worden gezet bij het ontwikkelen of vernieuwen van een standaard set is bepalen wat er gemeten moet worden, dit wordt gedefinieerd in zogenaamde domeinen. Daarna moet worden gedefinieerd hoe elk van de gekozen domeinen moet worden gemeten door de selectie van instrumenten (zoals vragenlijsten of gewricht-scores). De uiteindelijke standaard set zal zowel de geselecteerde domeinen als instrumenten bevatten.

Voor veel uitkomstmaten -zoals pijn of kwaliteit van leven- maken reumatologen en onderzoekers gebruik van de subjectieve informatie die door de patiënt wordt verstrekt, omdat er geen objectieve metingen beschikbaar zijn. Vandaar dat een groot deel van de uitkomstmaten die vaak worden gebruikt in axSpA zogenaamde patiënt-gerapporteerde uitkomstmaten zijn. Daarnaast heeft eerder onderzoek aangetoond dat artsen en patiënten verschillende opvattingen hebben van ziekteactiviteit en fysiek functioneren, wat het belang van de patiënt-gerapporteerde uitkomstmaten verder benadrukt.

Ook hebben de patiënt-gerapporteerde uitkomstmaten een grote rol gespeeld bij de erkenning dat de ziektelast vergelijkbaar is tussen patiënten met en zonder schade aan het bekken (het sacro-iliacaal gewricht) op de röntgenfoto's.

Het is bekend dat axSpA een nadelige invloed kan hebben op de kwaliteit van leven. Om die reden is het optimaliseren van de kwaliteit van leven op de lange termijn benoemd als het belangrijkste behandeldoel bij axSpA. Eerder onderzoek heeft aangetoond dat de kwaliteit van leven al verminderd is bij patiënten in de vroege fase van axSpA en dat deze kan worden verbeterd door de ziekteactiviteit te verminderen met een effectieve behandeling.

Bovendien gaat axSpA gepaard met een groot risico op beperking van de arbeidsproductiviteit gedurende het leven van de patiënt, wat bijdraagt aan substantiële maatschappelijke kosten van axSpA. De werkloosheidscijfers en arbeidsongeschiktheidscijfers zijn aanzienlijk hoger in vergelijking met de algemene bevolking. Ook komt overstappen naar een fysiek minder veeleisende baan of vervroegde uittreding vaker voor bij patiënten met axSpA. Verminderd vermogen om het werk adequaat uit te voeren (presenteïsme) en een toename van het aantal gemiste werkuren door ziekte (absenteïsme) leiden tot verminderde arbeidsproductiviteit bij patiënten met axSpA. Omdat axSpA vaak in de meest productieve periode van iemands leven ontstaat (rond het 20^e-30^e levensjaar), is onderzoek naar de impact van de ziekte op het vermogen om te werken erg belangrijk.

Dit proefschrift is opgebouwd rondom drie thema's: 1) een internationaal perspectief bieden op de kenmerken van patiënten met axSpA; 2) beschrijven van het proces van de ontwikkeling van een standaard set uitkomstmaten voor axSpA; en 3) meer kennis vergaren over werkproductiviteit en kwaliteit van leven bij chronische rugpijnpatiënten met de diagnose axSpA of een vermoeden daarvan. In dit laatste hoofdstuk worden de bevindingen van dit proefschrift samengevat en in een breder perspectief geplaatst. Ten slotte zullen er aanbevelingen worden gedaan voor toekomstig onderzoek.

INTERNATIONALE CLASSIFICATIE VAN AXIALE SPONDYLOARTRITIS

In het eerste deel van dit proefschrift wilden we een internationaal perspectief bieden op de classificatie van patiënten met axSpA. Hiervoor zijn verschillende aspecten van de classificatiecriteria voor axSpA nauwkeurig onderzocht. Te beginnen met de naamgeving die wordt gebruikt om patiënten met axSpA met radiografische schade aan de sacro-iliacale gewrichten (het bekken) te beschrijven in **hoofdstuk 2**. Van oudsher werden patiënten met axSpA met onomkeerbare structurele veranderingen op röntgenfoto's geclassificeerd

volgens de modified New York (mNY) criteria als ankyloserende spondylitis (AS). Echter, met het ontstaan van de recentere ASAS axSpA-criteria is een alternatief ontstaan, waarmee deze patiënten konden worden geclassificeerd als radiografische axSpA. Critici betwijfelden of beide classificatiesets dezelfde patiënten zouden classificeren, maar dit was nooit onderzocht. In **hoofdstuk 2** concludeerden we dat bijna alle patiënten met axSpA met radiografische sacroiliitis die aan de mNY-criteria voldeden, ook voldeden aan de ASAS-criteria voor radiografische axSpA en vice versa. Dit betekent dat de termen AS en radiografische axSpA uitwisselbaar zijn. Tot nu toe is er geen overeenstemming bereikt of de ene definitie de voorkeur verdient boven de andere. Desalniettemin, aangezien we op weg zijn naar één diagnose (d.w.z. axiale spondyloartritis) met twee subgroepen enkel voor classificatiedoeleinden (d.w.z. radiografische en niet-radiografische axSpA), zou het naar mijn mening wenselijk zijn om radiografische axSpA te gebruiken om deze patiënten te beschrijven. Bovendien wordt tegenwoordig erkend dat niet-radiografische en radiografische axSpA een gelijke ziektelast hebben en behandeling met anti-reumatische geneesmiddelen (bDMARDs) is effectief gebleken in beide groepen patiënten. Daarom lijkt het gebruik van axSpA het meest geschikt om alle patiënten met een diagnose te beschrijven en de toevoeging radiografisch versus niet-radiografisch om aanvullende informatie te geven over de expressie van ziekte.

Erkennen dat de termen AS en radiografische axSpA uitwisselbaar zijn vergroot de vergelijkbaarheid tussen onderzoeken, aangezien beide termen dezelfde patiënten beschrijven. Dit zorgt er ook voor dat onderzoek uitgevoerd in cohorten van AS kan worden vergeleken met meer recent gepubliceerde artikelen over radiografische axSpA-cohorten. Dit is van enorm belang voor medicatieonderzoeken. Als de effectiviteit van een bepaald medicijn in het verleden is bewezen, worden ze niet onderworpen aan verdere gerandomiseerde klinische onderzoeken om de effectiviteit ervan te beoordelen, omdat het onethisch zou zijn om patiënten effectieve medicatie te onthouden. Om het effect van medicatie grootschalig te kunnen onderzoeken worden vaak zogenaamde meta-analyses uitgevoerd, hierin worden resultaten van een aantal vergelijkbare klinische studies gebundeld. Door gegevens van radiografische axSpA in meta-analyses op te nemen, kunnen vergelijkingen worden gemaakt waarbij rekening wordt gehouden met alle soorten behandelingen, inclusief diegene welke zijn onderzocht bij AS. Dit betekent dat de behandeling kan worden gestart zonder dat er eerst een nieuw onderzoek moet worden uitgevoerd.

De belangrijkste reden voor discrepanties tussen patiënten geclassificeerd volgens de axSpA-criteria en mNY-criteria bleek de leeftijd waarop rugpijn ontstaan is. De leeftijd waarop de rugpijn ontstaat werd geïntroduceerd in de ASAS-criteria in 2009 en is voornamelijk gebaseerd op gegevens van Feldtkeller en collega's gepubliceerd in 2003. Zij toonden aan dat het ontstaan van rugpijn na de leeftijd van 45 bij slechts 5%

van de patiënten voorkwam. Vergelijkbare verdelingen in het ontstaan van de rugpijn zijn gevonden in andere onderzoeken die sindsdien zijn uitgevoerd, maar ook deze waren gebaseerd op voornamelijk Europese gegevens. **Hoofdstuk 3** was erop gericht een wereldwijd perspectief te bieden op de leeftijd waarop de eerste symptomen zich voordeden. Dit om te bevestigen of de leeftijd bij het ontstaan van de rugklachten over de hele wereld gelijk was. In **hoofdstuk 3** hebben we bevestigd dat de leeftijd bij het ontstaan van de rugklachten vergelijkbaar was in verschillende geografische regio's. Ook konden we bevestigen dat de overgrote meerderheid van de patiënten met axSpA inderdaad hun eerste symptomen ervoeren vóór de leeftijd van 45 jaar. Verder toonde onze studie aan dat over de hele wereld de rugpijn op jongere leeftijd ontstaat in patiënten die het HLA-B27 gen dragen. Ook vonden we dat bij mannen de rugpijn op iets jongere leeftijd ontstaat dan bij vrouwen.

Deze gegevens zijn erg belangrijk, omdat het bevestigt dat het criterium leeftijd bij aanvang van rugpijn kan worden toegepast op patiënten overal ter wereld. Verder impliceren deze gegevens dat axSpA zich op jongere leeftijd manifesteert bij HLA-B27-positieve en mannelijke patiënten. Het lijkt er dus op dat de leeftijd bij het ontstaan van de rugklachten een heel nuttig hulpmiddel is bij het identificeren van personen met een risico op axSpA. Aangezien slechts een zeer klein deel van de patiënten symptomen ontwikkelt na de leeftijd van 45 jaar, is het zeer onwaarschijnlijk dat een patiënt die rugklachten krijgt op het moment dat hij of zij een stuk ouder is dan 45 jaar axSpA heeft, wat belangrijke kennis is voor de klinische praktijk omdat het onnodige diagnostiek kan voorkomen.

Hoewel de symptomen bij HLA-B27-positieve en mannelijke patiënten op wat jongere leeftijd optreden, betekent uiteraard niet dat de diagnose axSpA bij HLA-B27-negatieve en vrouwelijke patiënten niet moet worden overwogen. Het besef dat de aanvang van de symptomen gemiddeld iets later kan zijn bij HLA-B27-negatieve en vrouwelijke patiënten, is dus iets wat een arts mee moet nemen in zijn of haar overwegingen. Wat de classificatiecriteria betreft, lijkt het criterium van een ontstaan van klachten <45 jaar dus geldig, aangezien de overgrote meerderheid van de patiënten met axSpA symptomen ontwikkelden vóór deze leeftijd over de hele wereld. Aangezien classificatiecriteria zijn gericht op het creëren van een homogene groep patiënten, lijkt het leeftijd-criterium een nuttig hulpmiddel bij het uitsluiten van de minder typische patiënten.

Het derde en laatste aspect van de classificatiecriteria dat in dit proefschrift wordt besproken, is de positieve familieanamnese van spondyloarthritis, oftewel het voorkomen van SpA bij een familielid. De waarde van een positieve familieanamnese in zijn huidige vorm is eerder in twijfel getrokken, omdat de definitie niet is getest of gevalideerd voordat deze in de ASAS-classificatie werd opgenomen. **Hoofdstuk 4** beschreef de prevalentie van een positieve familieanamnese van spondyloarthritis (d.w.z. hoe vaak een positieve

familieanamnese voorkomt bij patiënten met axSpA) in verschillende geografische regio's en de relatie met het HLA-B27 gen. We ontdekten dat axSpA de meest voorkomende vorm van spondyloartritis is in een positieve familieanamnese. Verder bleek de relatie tussen een positieve familieanamnese van axSpA en HLA-B27-dragerschap onafhankelijk van een positieve familieanamnese voor andere vormen van spondyloartritis. Deze bevindingen bevestigen dat het verband tussen een positieve familieanamnese en HLA-B27-dragerschap grotendeels wordt veroorzaakt door een positieve familieanamnese voor axSpA in een wereldwijd cohort. Dit werd eerder al aangetoond in cohorten met voornamelijk Europese en enkele Aziatische patiënten.

De bevindingen gepresenteerd in **hoofdstuk 4** van dit proefschrift in combinatie met eerder onderzoek naar familieanamnese (in de ASAS-, DESIR- en SPACE-cohorten) suggereren dat het tijd is om dit criterium kritisch te evalueren. Alle beschikbare gegevens tonen aan dat axSpA de meest voorkomende vorm van spondyloartritis is in een positieve familieanamnese. Verder wordt het verband tussen een positieve familieanamnese en HLA-B27-dragerschap grotendeels bepaald door een positieve familieanamnese voor axSpA. Gezien de consistente bevindingen over de hele wereld, moet worden onderzocht of de huidige definitie van een positieve familieanamnese in de classificatiecriteria kan worden versmald naar een positieve familieanamnese van axSpA.

Verdere discussie en toekomstig onderzoek

Uniformiteit in classificatie en rapportage helpt de wereldwijde communicatie van wetenschappelijke en klinische bevindingen, wat het begrip van het ontstaan en beloop van axSpA en diens behandeling vergroot. Het feit dat het niet duidelijk was of twee belangrijke elementen in de ASAS-classificatiecriteria (d.w.z. leeftijd bij ontstaan klachten en positieve familieanamnese) van toepassing waren op alle axSpA-patiënten wereldwijd, wijst op een kwetsbaarheid in het wetenschappelijke proces. Aangezien patiënten uit verschillende continenten kunnen verschillen in hun ziektepresentatie, is het belangrijk dat er voldoende patiënten van over de hele wereld worden geïncludeerd om ervoor te zorgen dat de classificatiecriteria representatief zijn voor alle patiënten wereldwijd. Het ASAS PerSpA-cohort leverde het bewijs voor de haalbaarheid van een dergelijk onderzoek: door internationale samenwerking en slim gebruik te maken van een elektronisch dataverzamelingssysteem was het mogelijk om data te verzamelen in 24 landen over de hele wereld. Door de opstartkosten van een onderzoek te verlagen (bijvoorbeeld door een elektronisch gegevensverzamelingssysteem aan te bieden) en het aanbieden van een (bescheiden) vergoeding, wordt de drempel om deel te nemen verlaagd voor landen met minder financiële middelen.

De validiteit van de ASAS-classificatiecriteria voor axSpA zijn in het verleden in twijfel getrokken. Hierbij werd gezegd dat de classificatiecriteria te breed zijn, wat zou kunnen leiden

tot verschillen in de samenstelling van patiënten populaties in verschillende onderzoeken. Anderen benoemen dat de verhoogde sensitiviteit van de huidige ASAS criteria ten koste van specificiteit een reden tot bezorgdheid is, aangezien classificatiecriteria gericht zijn op het creëren van homogene onderzoekspopulaties en daarom moeten streven naar de hoogst mogelijke specificiteit. Critici benadrukken dus het belang van her-evaluatie van de ASAS axSpA-classificatiecriteria om de specificiteit te verbeteren en daarmee heterogeniteit te verminderen binnen de groep van axSpA-patiënten die volgens deze criteria zijn geïnclassificeerd.

In 2019 startten ASAS en SPARTAN de CLASSIC-studie (Classification of Axial Spondyloarthritis Inception Cohort) met als doel de ASAS-classificatiecriteria voor axSpA opnieuw te evalueren. Voor dit doel worden patiënten met chronische rugpijn met verdenking op axSpA geïnccludeerd, en een diagnose van axSpA of geen axSpA wordt gesteld na zorgvuldige evaluatie van klinische, laboratorium- en beeldvormingsresultaten. Doordat patiënten met en zonder een diagnose van axSpA worden geïnccludeerd, is het niet alleen mogelijk om de sensitiviteit en specificiteit van de classificatiecriteria als geheel te onderzoeken, maar ook de afzonderlijke componenten ervan.

In principe blijven de classificatiecriteria ongewijzigd als een sensitiviteit $\geq 75\%$ en specificiteit $\geq 90\%$ wordt behaald in de CLASSIC studie. Desalniettemin kan het noodzakelijk zijn voor sommige van de afzonderlijke componenten te beoordelen of deze ongewijzigd moeten blijven, ongeacht of de sensitiviteit en specificiteit van de classificatiecriteria als geheel wordt behaald.

Een van de afzonderlijke componenten die onderzocht moet worden, is de definitie van een positieve familieanamnese. Hierbij moet er rekening mee worden gehouden dat een positieve familieanamnese voor verschillende doeleinden kan worden gebruikt. De eerste toepassing van de familieanamnese is als een van de klinische criteria in de ASAS-classificatiecriteria. Om de waarde voor dit doel te bepalen, moet de oorspronkelijke definitie worden vergeleken met een herziene definitie (d.w.z. een positieve familieanamnese die alleen axSpA in een eerste- of tweedegraads familielid omvat). Daarnaast moet worden onderzocht of het gewicht geschikt is en of het een onafhankelijk SpA-kenmerk moet blijven naast HLA-B27 dragerschap.

Ten tweede kan familieanamnese worden gebruikt als een indicatie voor HLA-B27-positiviteit. Dit is met name relevant in situaties waarin HLA-B27-testen niet nuttig zijn (bijv. in de huisartsenpraktijk waar de axSpA-prevalentie laag is) of niet mogelijk (bijv. hoge kosten in landen met minder financiële middelen). Zoals getoond in **hoofdstuk 4** van dit proefschrift was er in het Midden-Oosten en Noord-Afrika een hoger percentage patiënten voor wie HLA-B27 niet beschikbaar was in vergelijking met de andere geografische regio's.

Bovendien komt HLA-B27 minder vaak voor in het Midden-Oosten en Noord-Afrika, wat aangeeft dat in deze regio informatie over familieanamnese bijzonder waardevol kan zijn. Ten slotte kan de familieanamnese worden gebruikt als een risicofactor voor de ontwikkeling van axSpA. Eerder onderzoek heeft aangetoond dat zodra de HLA-B27-status bekend is, een positieve familieanamnese niet bijdraagt aan de kans op een axSpA-diagnose. Daarom lijkt het gebruik ervan als een risicofactor voor de ontwikkeling van axSpA beperkt tot situaties waarin HLA-B27 niet beschikbaar is. De voorspellende waarde van de verschillende SpA-entiteiten in een familieanamnese voor het risico op het ontwikkelen van axSpA is in het verleden niet beoordeeld. Echter, gebaseerd op het feit dat een relatie met HLA-B27 dragerschap alleen werd gevonden voor axSpA in de PerSpA- en ASAS-cohorten en voor axSpA en uveitis in de DESIR- en SPACE-cohorten, is het zeer waarschijnlijk dat de definitie van een positieve familieanamnese ook voor dit doel opnieuw gedefinieerd dient te worden. Opmerkelijk is het feit dat een verband tussen HLA-B27 en uveitis slechts in twee Europese cohorten werd gevonden, wat nogmaals benadrukt hoe belangrijk het is om patiënten van over de hele wereld te includeren wanneer een beslissing wordt genomen over een gewijzigde definitie, die van toepassing zou moeten zijn op alle patiënten wereldwijd.

Aangezien classificatiecriteria worden gebruikt bij de selectie van patiënten voor klinische onderzoeken, moet hun belangrijkste doel blijven om een homogene groep patiënten te creëren. Echter, er moet rekening worden gehouden met de aard van de ziekte waarvoor ze worden gebruikt.

AxSpA is een ziekte met een grote variatie in symptoomexpressie, daarom kunnen de classificatiecriteria niet te restrictief zijn, omdat de patiënten met 'minder typische' symptomen dan niet geïnclassificeerd zullen worden. De CLASSIC studie biedt een unieke kans om verschillende combinaties van criteria te beoordelen.

Concluderend, in het eerste deel van dit proefschrift is het belang benadrukt van het opnemen van patiënten uit verschillende landen met verschillende etnische achtergronden. Toekomstig onderzoek moet zich richten op het verder verbeteren van de ASAS-classificatiecriteria- en de componenten ervan- door patiënten van over de hele wereld te includeren.

ASAS/OMERACT STANDAARD SET UITKOMSTMATEN VOOR AXIALE SPONDYLOARTRITIS

In het eerste deel van dit proefschrift werd het belang van wereldwijde toepasbaarheid van classificatiecriteria besproken. Deze zorgen ervoor dat over de hele wereld dezelfde patiënten worden geselecteerd voor deelname aan klinische onderzoeken, waardoor directe vergelijkingen mogelijk zijn tussen onderzoeken die zijn uitgevoerd in verschillende landen. Op eenzelfde manier maken gestandaardiseerde beoordeling en rapportage van resultaten directe vergelijkingen mogelijk tussen onderzoeken die verschillende therapieën onderzoeken, waarover in het tweede deel van dit proefschrift wordt gediscussieerd.

Een standaard set uitkomstmaten beschrijft de minimale en verplichte set uitkomstmaten die moeten worden beoordeeld en gerapporteerd in alle klinische onderzoeken van een specifiek ziektebeeld en/of patiëntpopulatie. Een standaard set uitkomstmaten bestaat uit domeinen (“wat te meten”) en instrumenten (“hoe te meten”). De standaard set die momenteel in axSpA wordt gebruikt, is de ASAS-OMERACT standaard set uitkomstmaten voor ankyloserende spondylitis (AS). Sinds de ontwikkeling van de oorspronkelijke standaard set meer dan twee decennia geleden, is het duidelijk geworden dat axSpA in feite een ziektespectrum is dat uit twee subtypes bestaat: radiografische axSpA en niet-radiografische axSpA. Daarnaast hebben er grote ontwikkelingen plaatsgevonden in de meetinstrumenten die gebruikt worden in axSpA, zoals het gebruik van MRI, de ontwikkeling van de Ankylosing Spondylitis Disease Activity Score (ASDAS)-voor het meten van ziekteactiviteit-, en de ASAS Health Index-voor het meten van kwaliteit van leven en algehele gezondheid-. Ten slotte is er vooruitgang geboekt in de methodologie rond de ontwikkeling van standaard sets uitkomstmaten, waardoor ASAS besloot dat het tijd was om de standaard set te herzien.

De eerste stap naar een vernieuwde standaard set uitkomstmaten was om te beoordelen of de domeinen die in de oorspronkelijke standaard set vertegenwoordigd waren nog steeds relevant zijn. Om informatie te verzamelen over het belang van de uitkomstmaten van patiënten met axSpA en experts op het gebied van axSpA, werd een Delphi-enquête uitgevoerd. Uit de resultaten van deze Delphi-enquête-beschreven in **hoofdstuk 5** van dit proefschrift- leerden we dat patiënten met axSpA een andere mening hadden over de uitkomsten die in alle onderzoeken naar axSpA beoordeeld moeten worden dan de experts die betrokken zijn bij hun behandeling. Patiënten gaven de voorkeur aan een allesomvattende benadering, terwijl experts een onderscheid bleken te maken tussen uitkomsten die van cruciaal belang zijn voor verschillende typen behandelingen. Volgens de experts waren meer objectief meetbare domeinen, zoals structurele schade en mobiliteit, het meest cruciaal om te meten in onderzoeken waarin anti-reumatische behandelingen worden onderzocht. Dit zijn behandelingen waarbij men verwacht dat het beloop van de

ziekte fundamenteel beïnvloed wordt, d.w.z. meer dan alleen bestrijden van de klachten. De beoordeling van de meer subjectieve domeinen zoals pijn, stijfheid en algemeen functioneren en gezondheid werd vooral belangrijk bevonden voor onderzoeken die zich richten op symptoom bestrijding. Deze gegevens stelden ons in staat om een korte lijst samen te stellen met de belangrijkste domeinen om te beoordelen in alle studies naar axSpA. Deze gegevens vormden de basis voor de ontwikkeling van de vernieuwde standaard set uitkomstmaten voor axSpA.

De Delphi-enquête is een veelgebruikte methode die wordt ingezet om meningen van een grote groep deelnemers te verzamelen, ofwel om een prioriteitenlijst te maken omtrent onderzoeksonderwerpen, om itemlijsten in te korten, of voor het verzamelen van feedback. Voor een methode die zo vaak wordt gebruikt, is informatie over de methodologie opvallend schaars. In **hoofdstuk 6** wilden we inzicht geven in het effect van het kiezen van een bepaalde uitnodigingstechniek op de uitkomst van de Delphi, door twee veelgebruikte uitnodigingsbenaderingen te vergelijken: 1) Alle deelnemers uitnodigen voor volgende rondes, ongeacht of zij hebben geantwoord op de vorige ronde; of 2) Alleen de deelnemers uitnodigen voor volgende rondes die de vorige ronde hebben voltooid. We ontdekten dat er geen effect is op de uiteindelijke uitkomst van de Delphi. Wel stelden we dat het de voorkeur kan hebben om deelnemers die een ronde hebben gemist toch uit te nodigen voor volgende rondes, omdat deze benadering minder gevoelig is voor het niet-willekeurige verlies van meningen dat zou kunnen leiden tot valse overeenstemming. Bovendien zorgt deze aanpak ervoor dat het eindresultaat de mening weergeeft van iedereen die was uitgenodigd om deel te nemen.

De Delphi-enquête was een klein onderdeel van een veel grotere project om de standaard set uitkomstmaten voor AS te vernieuwen, zodat deze toepasbaar is op alle patiënten met axSpA. In **hoofdstuk 7** beschreven we het proces dat leidde tot de ASAS-OMERACT core outcome set voor axSpA. De domeinen van de vernieuwde standaard set uitkomstmaten voor axSpA zijn vergelijkbaar met de oorspronkelijke standaard set voor AS. Beiden bevatten de domeinen fysiek functioneren, ochtendstijfheid, pijn, vermoeidheid en ziekteactiviteit in hun kern. Daarnaast is het meten van structurele schade alleen verplicht voor studies die anti-reumatische therapieën onderzoeken, waarbij men verwacht dat het beloop van de ziekte fundamenteel beïnvloed wordt (meer dan alleen klachten). Het eerste opvallende verschil is de toevoeging van algemeen functioneren en gezondheid in de kern van de vernieuwde standaard set uitkomstmaten voor axSpA. Hiermee wordt de impact van axSpA op andere aspecten van het leven gemeten, wat de afgelopen jaren meer aandacht heeft gekregen. Ten tweede is het opmerkelijk dat mobiliteit van de wervelkolom niet langer deel uitmaakt van de kern in de standaard set voor axSpA. Dit is te verklaren door het gebrek aan standaardisatie, onvermogen verandering in mobiliteit van de wervelkolom goed weer te geven en de slechte test-hertest betrouwbaarheid. Test-

hertest betrouwbaarheid betekent het vermogen van een test om hetzelfde resultaat te geven als de test wordt herhaald op een ander moment, en de overige omstandigheden gelijk zijn gebleven. Nu de domeinen zijn vastgesteld, moeten geschikte meetinstrumenten worden geselecteerd om deze domeinen te meten.

De veranderingen ten opzichte van de oorspronkelijke standaard set voor AS zullen implicaties hebben voor toekomstig onderzoek, omdat er net andere domeinen beoordeeld moeten worden dan eerder is gedaan. Helaas vermindert dit voor sommige aspecten van ziekte de vergelijkbaarheid met oudere onderzoeken (bijv. mobiliteit van de wervelkolom), maar tegelijkertijd zorgt het voor een betere vergelijkbaarheid in toekomstige onderzoeken omdat er meer duidelijkheid is over welke uitkomstmaten moeten worden gemeten en gerapporteerd. Bovendien zijn alle belanghebbenden die baat hebben bij een vernieuwde standaard set uitkomstmaten betrokken bij de ontwikkeling ervan, wat de acceptatie zal vergroten. De volgende stap is om voor elk van de geselecteerde domeinen de beste meetinstrumenten te selecteren. Hierbij is een zorgvuldige afweging van de meeteigenschappen, haalbaarheid en bruikbaarheid van de meetinstrumenten van belang.

Een van de meeteigenschappen van de meetinstrumenten is de test-hertest betrouwbaarheid. In **hoofdstuk 8** wordt de test-hertest betrouwbaarheid van de meetinstrumenten die zijn gebruikt in drie recente klinische studies in axSpA onderzocht. Uit deze studie hebben we geconcludeerd dat hoewel de meeste meetinstrumenten zijn ontwikkeld voor radiografische axSpA, ze ook betrouwbaar zijn bevonden voor niet-radiografische axSpA. Bovendien liet deze studie zien dat meetinstrumenten met meerdere items robuuster bleken te zijn tegen meetfouten vergeleken met meetinstrumenten die slechts uit één item bestaan. Verder onderzoek zal de andere meeteigenschappen moeten onderzoeken voordat een definitief besluit kan worden genomen over welke instrumenten het beste geschikt zijn om de geselecteerde domeinen te meten.

Verdere discussie en toekomstig onderzoek

De standaard set uitkomstmaten voor AS die momenteel wordt gebruikt om te bepalen welke data minimaal moeten worden verzameld in elke klinische studie in axSpA, bevatte niet direct na ontwikkeling een instrument voor elk domein. Er waren niet direct specifieke instrumenten gedefinieerd voor de beoordeling van vermoedheid of enthesitis (ontstekingen van peesaanhechtingen), omdat er op het moment van ontwikkeling van de standaard set uitkomstmaten geen gevalideerd instrument beschikbaar was. Als gevolg hiervan zijn verschillende instrumenten gebruikt om deze domeinen te beoordelen, waardoor vergelijkingen tussen onderzoeken worden belemmerd, terwijl die erg belangrijk zijn voor het beoordelen van de effectiviteit van de behandeling. Daarom is het van groot belang dat de vernieuwde standaard set één specifiek instrument voor elk

domein adviseert (met de mogelijkheid om er meer toe te voegen). Deze instrumenten dienen niet alleen te worden gemeten, maar, belangrijker nog, in elke studie moet de resultaten op gestandaardiseerde wijze worden gerapporteerd, wat één-op-één vergelijking van onderzoeken en de ontwikkeling van behandelaanbevelingen mogelijk maakt. De oorspronkelijke standaard set wordt veelvuldig gebruikt, wat aangeeft dat de beschikbaarheid van een standaard set leidt tot een gestructureerde verzameling van informatie in klinische onderzoeken. Er was echter nogal wat variatie in de instrumenten die werden gebruikt om de informatie te verzamelen, wat het belang benadrukt om voor elk domein één specifiek instrument aan te bevelen. Bovendien bleek dat niet alle verzamelde informatie ook werd gerapporteerd. Zo bevat BASDAI een maat voor vermoeidheid, maar deze werd vaak niet specifiek gerapporteerd en daarom kon op basis van de gepresenteerde gegevens geen conclusie worden getrokken over het effect van de onderzochte therapie op vermoeidheid.

Het belangrijkste doel van een standaard set uitkomstmaten is het beschrijven van de domeinen en instrumenten die ten minste in elke studie moeten worden beoordeeld. Naarmate ons begrip van een ziekte toeneemt, of het ziektebeloop verandert door eerdere herkenning en effectieve therapie, kan dit leiden tot de ontwikkeling van nieuwe instrumenten (bijvoorbeeld de ASAS Health Index) of validatie van bestaande instrumenten. Er is een redelijke kans dat deze nieuwe instrumenten beter presteren dan bestaande instrumenten, en daarmee de voorkeur krijgen om een bepaald domein te meten, wat het vernieuwen van een standaard set uitkomstmaten vereist. Helaas is het vernieuwen van een standaard set een langdurig en tijdrovend proces, en je kunt je afvragen of het proces zijn doel overstijgt en of er meer standaard sets (d.w.z. meer gestandaardiseerde metingen) zouden zijn als het proces gebruiksvriendelijker zou zijn. Een mogelijkheid om het proces te vereenvoudigen is het regelmatig herzien van een standaard set (bijv. om de 10 jaar), waarbij kan worden besloten een instrument te vervangen als is aangetoond dat het nieuwe instrument beter presteert dan het bestaande, zonder dat daarvoor alle stappen die nodig zijn voor de ontwikkeling van een nieuwe standaard set hoeven worden doorlopen. Om dit te doen, zou een volledige vergelijking van alle meeteigenschappen tussen het nieuwe en bestaande instrument een eerste vereiste zijn. Echter, het regelmatig veranderen van een standaard uitkomst set heeft het nadeel dat nieuwer en ouder onderzoek niet langer één-op-één te vergelijken is.

Ondanks de tekortkomingen zijn standaard sets waardevol in onderzoek, omdat ze zorgen voor meer transparantie in het geneesmiddelenregistratieproces door directe vergelijkingen met eerder geregistreerde geneesmiddelen. Ook kunnen zij bijdragen aan een betere acceptatie van nieuwe behandelingen in het veld, omdat een directe vergelijking van uitkomstmaten de prestaties van het nieuwe geneesmiddel laten zien in relatie tot eerder geaccepteerde/meer bekende medicijnen. Het verleden heeft ons

geleerd dat het belangrijk is om per domein concreet te definiëren welke instrumenten moeten worden gebruikt en welke optioneel kunnen worden toegevoegd. Bovendien garandeert het gebruik van een instrument niet dat alle verzamelde gegevens ook worden gerapporteerd. Daarom moet de standaard set uitkomstmaten voor axSpA niet alleen de domeinen en instrumenten bieden die in elke studie moeten worden beoordeeld, maar ook specifieke instructies voor het rapporteren van verzamelde data.

Concluderend, standaard sets uitkomstmaten zijn een waardevol hulpmiddel bij het beoordelen van resultaten, maar toekomstig onderzoek zou moeten bekijken of hun ontwikkelingsproces kan worden verbeterd.

PATIËNT GERAPPORTEERDE UITKOMSTMATEN IN VROEGE AXIALE SPONDYLOARTRITIS

In het laatste deel van dit proefschrift bespreken we gezondheids-gerelateerde kwaliteit van leven en beperkingen in werkproductiviteit in axSpA. In **hoofdstuk 9** lieten we zien dat werkproductiviteit evenals gezondheids-gerelateerde kwaliteit van leven verbeterden gedurende twee jaar follow-up bij patiënten met chronische rugpijn verdacht van axSpA. Deze verbetering werd aangetoond voor patiënten met en zonder een diagnose van axSpA, maar patiënten met axSpA vertoonden een grotere verbetering in vergelijking met rugpijnpatiënten zonder axSpA.

De resultaten uit **hoofdstuk 9** hebben ons geleerd dat we in de loop van de tijd enige verbetering kunnen verwachten bij alle chronische rugpijnpatiënten die verdacht worden van axSpA, ongeacht de diagnose. Dit kan naast de behandeling te maken hebben met het feit dat patiënten bij hun eerste bezoek aan de reumatoloog de meest ernstige klachten ervaren. De klachten kunnen in de loop van de tijd op natuurlijke wijze verbeteren (regressie naar het gemiddelde). Bovendien konden we concluderen dat een diagnose van axSpA een onafhankelijke voorspeller is van verbetering in gezondheids-gerelateerde kwaliteit van leven en verbetering in werkproductiviteit bij patiënten met chronische rugpijnklachten. De verschillen in uitkomsten tussen degenen met en zonder een diagnose van axSpA bleven bestaan wanneer de analyses beperkt werden tot patiënten die geen anti-reumatische medicatie kregen, wat aangeeft dat behandeling met deze medicatie de verschillen tussen de groepen niet verklaarde. Belangrijk is dat, ondanks de verbeteringen in de tijd, de kwaliteit van leven en werkproductiviteit nog steeds slechter waren in vergelijking met de algemene bevolking. Deze resultaten benadrukken het belang van het optimaliseren van de gezondheids-gerelateerde kwaliteit van leven en de sociale participatie van patiënten met axSpA op de lange termijn. Dit wordt ook beschreven in de huidige axSpA-behandelingsrichtlijnen als een van de belangrijkste behandeldoelen.

Verdere discussie en toekomstig onderzoek

In **hoofdstuk 9** hebben we gevonden dat patiënten met de diagnose axSpA een grotere verbetering in werkproductiviteit en gezondheids-gerelateerde kwaliteit van leven vertoonden vergeleken met diegenen zonder een diagnose van axSpA. We vermeldden dat dit mogelijk te verklaren is door een verschil in beschikbare behandelingsopties, maar dit kan ook worden verklaard door een verschil in ziektepercepties en daaropvolgende coping-mechanismen (d.w.z. de manier waarop men met de ziekte omgaat). Een van de vragen die we ons stelden is of de diagnose een invloed kan hebben op hoe een patiënt zijn of haar ziekte ervaart en dit vervolgens de coping-mechanismen beïnvloedt. Toekomstig onderzoek zou zich moeten richten op het verkrijgen van meer inzicht in de psychologische effecten van het stellen van een diagnose, of alleen weten 'wat er met je aan de hand is' invloed heeft op hoe klachten worden ervaren. Verder zou moeten worden onderzocht of de kwaliteit van leven nog verder kan verbeteren als patiënten geïnformeerd worden over mogelijke coping-mechanismen. Tegelijkertijd kunnen adequate coping-mechanismen een invloed hebben op de werkproductiviteit. Aangezien axSpA het leven beïnvloedt van jonge mensen, is er een enorme waarde in het uitbreiden van ons begrip van effectieve therapieën -hetzij medicamenteus, educatief of psychologisch- die een positief effect hebben op de kwaliteit van leven en werkproductiviteit.

Een ander aspect dat aandacht verdient, is het gebruik van algemene versus ziekte-specifieke vragenlijsten om kwaliteit van leven te beoordelen. Het belangrijkste voordeel van het gebruik van een algemene vragenlijst is dat de scores kunnen worden vergeleken met scores van patiënten met andere (chronische) ziekten of de algehele bevolking, wat de maatschappelijke waarde vergroot. Daarentegen hebben ziekte-specifieke vragenlijsten (zoals de ASAS Health Index) betrekking op meer ziekte-specifieke vragen, die een duidelijker inzicht geven in het effect van ziekte op de kwaliteit van leven van patiënten met axSpA, waardoor ze wellicht van hogere wetenschappelijke waarde zijn. In **hoofdstuk 9** maakte het gebruik van de SF-36 (een vragenlijst die wordt gebruikt om de kwaliteit van leven in kaart te brengen) een directe vergelijking mogelijk van patiënten met en zonder een diagnose van axSpA en er kon ook een vergelijking met de algemene bevolking worden gemaakt. Directe vergelijking van degenen die wel en geen diagnose van axSpA kregen, toonde ons dat er een verschil in verbetering in de tijd is tussen deze groepen, een feit dat onopgemerkt zou zijn gebleven als we een ziekte-specifieke vragenlijst hadden gebruikt. Desalniettemin had het gebruik van een ziekte-specifieke vragenlijst kunnen leiden tot inzicht in welke aspecten (indien aanwezig) hebben bijgedragen aan de verbetering van de kwaliteit van leven over tijd. Daarom kan een ziekte-specifieke vragenlijst de voorkeur hebben in het onderzoeken van patiënten met een chronische ziekte over de tijd. Aangezien het zeer waarschijnlijk is dat de gezondheids-gerelateerde kwaliteit van leven bij deze patiënten verminderd zal blijven in vergelijking met de algemene bevolking, zou inzicht in aspecten die bijdragen aan verbetering of verslechtering van de kwaliteit van leven in

de loop van de tijd waardevoller kunnen zijn, omdat dit tot nieuwe behandeldoelen zou kunnen leiden. Uiteindelijk zal onderzoek altijd vragen om het maken van keuzes en het vinden van de optimale balans tussen kosten en beloning, dus wat 'de juiste keuze' is, hangt af van de onderzoeksvraag, de beschikbare gegevens en de beschikbare middelen.