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

Boel, A.

CHAPTER 10

SUMMARY AND GENERAL DISCUSSION

In the first part of this thesis, we provided an international perspective on the characterisation of patients with axial spondyloarthritis (axSpA). We investigated the similarities and differences between the modified New York (mNY) criteria for ankylosing spondylitis and the ASAS classification criteria for radiographic axSpA, and studied the distribution in age at onset and prevalence of a positive family history of axSpA outside of Europe. In the second part of this thesis, we described the process of the development of the core set for axSpA (i.e. the minimum and mandatory set of outcomes to be assessed in every trial) by updating the domains of the ASAS-OMERACT core set for ankylosing spondylitis. In the third part of this thesis, we increased knowledge on work and activity outcomes and health-related quality of life over time in chronic back pain patients with a diagnosis of axSpA or a suspicion thereof.

In this final chapter I will summarize the main findings of the studies presented here within, elaborate on the impact of these findings, and highlight remaining knowledge gaps and future perspectives.

INTERNATIONAL CHARACTERISATION OF AXIAL SPONDYLOARTHRITIS

In the first section of this thesis, we aimed to provide an international perspective on the characterisation of patients with axSpA. Several aspects of the classification criteria for axial spondyloarthritis were closely inspected for this purpose, starting with the nomenclature used to describe patients with axSpA with radiographic damage to the sacroiliac joints in **chapter 2**, for which we included data from several European cohorts[1-6], an American cohort[7] and a cohort including data from Asian, American and European patients[8,9]. Traditionally, patients with axSpA with definite structural changes on conventional radiographs were classified according to the mNY criteria as ankylosing spondylitis. However, an alternative was provided by the more recent ASAS axSpA criteria, wherein these patients could be classified as radiographic axSpA. Critics doubted whether both would classify the same patients[10], but this had never been assessed. In **chapter 2** we concluded that almost all patients with axSpA with radiographic sacroiliitis who met the mNY criteria also met the ASAS criteria for radiographic axSpA and vice-versa, which supported the interchangeable use of the terms ankylosing spondylitis and radiographic axSpA. Thus far, no consensus has been reached on whether one definition should be preferred over the other. Nonetheless, as we move towards one diagnosis (i.e. axial spondyloarthritis) with two subgroups for classification purposes only (i.e. radiographic and non-radiographic), in my opinion it would be desirable to use radiographic axSpA to describe these patients.

Furthermore, it is acknowledged nowadays that non-radiographic and radiographic axSpA have an equal disease burden and treatment with biological disease-modifying antirheumatic drugs (bDMARDs) is effective in both subgroups. Therefore, using axSpA to describe all patients with a diagnosis and radiographic vs non-radiographic to provide additional information on the expression of disease seems most appropriate.

The main cause for disagreement between patients classified by the axSpA criteria and mNY criteria was found to be the age at onset of back pain, which was introduced with the implementation of the ASAS criteria in 2009 and mainly based on data from Feldtkeller et al[11], which showed that onset after the age of 45 was seen in only 5% of the patients. Other studies performed since then have shown similar distributions regarding the age at onset[1,12-14], yet were also based on mainly European data. Therefore, **chapter 3** aimed to provide a worldwide perspective on the age at first symptom onset, to confirm whether the distribution of age at onset of axial symptoms was similar across the globe. Using data from the ASAS-PerSpA study we were able to confirm a similar distribution in age at first symptom onset in various geographical regions, and confirmed the vast majority of patient with axSpA indeed experienced their first symptoms before the age of 45 years. Compared to Feldtkeller et al.[11] and van der Linden et al.[15], we found a slightly lower percentage of patients with age at onset <45 years in this study, which might be explained by the fact that-contrary to these studies- patients with non-radiographic axSpA were also included in the current study. Another important finding of this study was the fact that HLA-B27 carriership was consistently associated with a younger age of symptom onset across the globe, as was male gender.

The third and final aspect of the classification criteria under review in this thesis was the positive family history of spondyloarthritis. The value of a positive family history in its current form has been questioned previously, as the definition was not tested nor validated prior to inclusion in the ASAS classification criteria. **Chapter 4** described the prevalence of a positive family history of spondyloarthritis in various geographical regions and its relationship with HLA-B27 carriership. We found that axSpA is the most common entity of spondyloarthritis in a positive family history, and the association between a positive family history of axSpA and HLA-B27 carriership was independent of a positive family history of other SpA entities. These findings confirmed that the association between a positive family history and HLA-B27 status is largely driven by a positive family history for axSpA in a worldwide cohort, which was previously shown in cohorts that included mostly European and some Asian patients[16-18]. However, we were not able to show an association between a positive family history of axSpA and HLA-B27 carriership in the Middle East & North Africa. This may be explained by the fact that the prevalence of HLA-B27 positive disease was much lower compared to the other regions, which is in line with other research performed in the Middle East[19,20]. Additionally, the Middle East & North Africa showed a high prevalence

of a positive family history of axSpA in HLA-B27 negative patients compared to the other regions, where there was a very low prevalence of a positive family history of axSpA in HLA-B27 negative patients. Nonetheless, a positive family history of axSpA was the most common amongst patients who reported a positive family history in this region, identical to the other geographical regions.

Implications from these findings

Acknowledging that ankylosing spondylitis and radiographic axSpA are interchangeable -as confirmed in **chapter 2-** increases comparability between studies, since both terms describe the same patients. This also ensures that research performed in ankylosing spondylitis cohorts can be compared to more recently published articles on radiographic axSpA cohorts. This is of tremendous importance for medication trials. If the effectiveness of a given medication has been proven in the past, they are not subjected to further randomised clinical trials assessing its effectiveness, as it would be unethical to withhold effective medication from patients. Including data from r-axSpA in meta-analyses allow for comparisons taking into account all treatment types, including those that have been investigated in ankylosing spondylitis. This means treatment can be initiated without the need for conducting a trial first.

The data described in **chapter 3** showed that the age at symptom onset was similar in all investigated geographical regions and the age at symptom onset was consistently lower in HLA-B27 positive patients compared to their HLA-B27 negative counterparts, and also consistently lower in male compared to female patients. This data was long overdue, as it confirms the age at onset criterion can be applied to patients anywhere in the world, rather than just the European, North American and Asian patients in which it was developed. Further, these data imply axSpA manifests at an earlier age in HLA-B27 positive and male patients. Thus, it appears that age at onset is a helpful tool in identifying those at risk of axSpA in the group of patients who present to the rheumatologists with chronic back pain complaints. Given that only a very small proportion of patients develop symptoms after the age of 45 years, it is very unlikely that a patient above this age will be diagnosed with axSpA, which is important knowledge for clinical practice. Even though symptoms occur at a somewhat earlier age in HLA-B27 positive patients and male patients, this does not imply a diagnosis of axSpA in HLA-B27 negative patients and female patients should not be considered. Being aware that symptom onset may be somewhat later in HLA-B27 negative and female patients might result in earlier consideration of an axSpA diagnosis in these 'less typical' patients, which might subsequently reduce diagnostic delay. As for the classification criteria, the entry criterion of an age at onset <45 years seems valid, as the vast majority of patients with axSpA developed symptoms before this age across the globe. As classification criteria are aimed at creating a homogenous group of patients, the age at onset criterion seems a useful tool in excluding the less typical patients.

The findings presented in **chapter 4** of this thesis combined with previous research on family history (in the ASAS, DESIR and SPACE cohorts[16-18]) suggest it is time to critically re-evaluate this criterion. All available data show that axSpA is the most common entity as part of a positive family history and the association between a positive family history and HLA-B27 status is largely driven by a positive family history for axSpA. Given the consistent findings across studies and across the globe, it should be investigated whether the current expert definition of a positive family history in the classification criteria may be redefined to only include the presence of a positive family history of axSpA. The criterion for a positive family history is not only present in the classification criteria for axSpA, but also in the classification criteria for peripheral SpA. As these are used to classify a different subset of patients, it is very likely the definition of a positive family history will have different implications in this patient population. In the event of a redefinition of the family history criterion for axSpA, it would be expected the performance of the family history criterion should be assessed for peripheral SpA too.

Further discussion and future perspectives

Uniformity in classification and reporting aids global communication of scientific, clinical, and epidemiological findings which enhances understanding of the pathogenesis and treatment of axSpA. The fact that it was not clear if two major elements in the ASAS classification criteria (i.e. age at onset and positive family history) applied to all axSpA patients worldwide, points to a flaw in the scientific process. As patients from different continents may vary in their disease presentation[21], sufficient patients from all over the world should be included from the get-go, to ensure the classification criteria are representative for all patients across the globe. The ASAS-PerSpA cohort[21] provided proof for the feasibility of such a study: through international collaboration and smart use of an electronic data collection system it was possible to collect data in 24 countries across the globe. By decreasing the start-up costs of a study (e.g. by providing an electronic data collection system), the threshold to partake will be lowered for countries with less financial funds.

In the past, the validity of the ASAS classification criteria for axSpA has been questioned, as it has been argued that its complex two-arm selection design and its broad spectrum may lead to differences among the composition of patients in different studies[10]. Others have stated that the increased sensitivity of the two-arm design compared to a classification set that included only the imaging arm (82.9% compared to 66.2%) at the cost of specificity (84.4% compared to 97.3%)is a cause for concern, as classification criteria are aimed at creating homogeneous study populations and should therefore aim for the highest possible specificity[22]. Thus, critics emphasize the importance of revision of the ASAS axSpA classification criteria in order to improve specificity and reduce heterogeneity within the group of axSpA patients classified using these criteria[10,22]. Nonetheless, the imaging arm in its current form is not perfect either, as it is well-known

that there is no perfect agreement between readers, and inflammatory lesions on MRI can also be found in healthy controls[23,24].

In 2019 ASAS and SPARTAN initiated the CLASSIC study (CLassification of Axial Spondyloarthritis Inception Cohort) with the aim to re-evaluate the sensitivity and specificity of the ASAS classification criteria for axSpA and provide training in assessment of imaging and diagnosis of patients with axSpA worldwide. For this purpose, patients with chronic back pain suspected of axSpA are included, and a diagnosis of axSpA or no axSpA is made after careful evaluation of clinical, laboratory and imaging results. By design, the CLASSIC study includes patients with and without a diagnosis of axSpA, allowing for assessment of the sensitivity and specificity not only of the classification set as a whole, but also its individual components. In principle, the classification criteria will remain unchanged if they show ≥75% sensitivity and ≥90% specificity in CLASSIC. Nonetheless, for some of the individual components it may be advisable to assess whether they should remain unchanged, regardless of whether the sensitivity and specificity of the classification set as a whole is reached.

One of the individual components that should be investigated is the value of the definition of a positive family history. Herein, it should be taken into account that a positive family history can be used for different purposes. The first application of the family history is as one of the clinical criteria in the ASAS classification criteria. To determine its value for this purpose, the original definition should be compared with a redefined definition (i.e. a positive family history that solely includes axSpA in a first- or second-degree family member). Additionally, it should be investigated whether its weight is appropriate, and whether it should remain an independent SpA feature in addition to HLA-B27. This should be assessed separately for the ASAS criteria for axSpA and peripheral SpA.

Secondly, family history can be used as a proxy for HLA-B27 positivity. This is particularly relevant in situations where HLA-B27 testing is not useful (e.g. in general practice where the axSpA prevalence is low) or not possible (e.g. high costs in countries with lower funds). As shown in **chapter 4** of this thesis, in the Middle East & North Africa there was a higher percentage of patients for whom HLA-B27 was not available compared to the other geographical regions. Furthermore, HLA-B27 is less prevalent in the Middle East & North Africa[19,20], indicating that in this region information on family history may be especially valuable.

Finally, family history can be used as a risk factor for the development of axSpA. Notably, it has been shown that as soon as HLA-B27 status is known, a positive family history does not contribute to the likelihood of an axSpA diagnosis[17]. Therefore, its use as a risk factor for the development of axSpA seems limited to situations in which HLA-B27

is unavailable. The predictive value of the individual SpA entities in a family history for the risk of developing axSpA have not been assessed in the past. However, based on the fact that an association with HLA-B27 was only found for axSpA in the PerSpA and ASAS cohorts[16] and for axSpA and uveitis in the DESIR and SPACE cohorts[18], it is very likely that the definition of a positive family history should be redefined for this purpose too. Of note, the fact that an association between HLA-B27 and uveitis was only found in two European cohorts, emphasises yet again the importance of including patients from around the world when a decision is made regarding a changed definition, which should be applicable to all patients worldwide.

As classification criteria are used in the selection of patients for clinical studies, their most important aim should remain to create a homogeneous group of patients, yet the nature of the disease for which they are employed should be taken into account. AxSpA is a disease with a great variation in symptom expression, hence the classification criteria cannot be too restrictive, as it will cut out a chunk of patients with 'less typical' symptoms. The CLASSIC study provides a unique opportunity to assess different combinations of criteria, or more likely, different weights for the various symptoms.

In conclusion, the first part of this thesis has emphasised the importance of including patients from various countries with various ethnic backgrounds. Future research should focus on further enhancing the ASAS classification criteria -and its components- by including patients from across the globe.

ASAS/OMERACT CORE SET FOR AXIAL SPONDYLOARTHRITIS

The first section of this thesis reviewed the importance of global applicability of classification criteria, as they ensure the same patients are selected for participation in clinical trials worldwide, allowing for direct comparisons between studies executed in different geographical regions. Likewise, standardised assessment and reporting of results allows for direct comparisons between studies investigating different treatments, or identical treatments in populations from a different ethnicity or background, which is debated in the second part of this thesis.

A core outcome set describes the minimum and mandatory set of instruments that should be assessed and reported all clinical studies of a specific health condition, population and setting[25,26]. A core outcome set consists of domains (*what to measure*) and instruments (*how to measure*). The core outcome set currently used in axSpA is the ASAS-OMERACT core outcome set for ankylosing spondylitis[27,28]. Since the development of the original core outcome set over two decades ago, it has become apparent that axSpA is in fact a

disease spectrum that consists of two subtypes: radiographic axSpA and non-radiographic axSpA[29]. Additionally, major advances have occurred in the outcome assessments in the field of axSpA, such as the use of magnetic resonance imaging[30], the development of the Ankylosing Spondylitis Disease Activity Score (ASDAS)[31], validated enthesitis scores[32], and the ASAS-health index[33] . Finally, progress regarding the methodology surrounding the development of core sets has occurred, all of which made ASAS decide it was time to update the core outcome set.

The first step towards an updated core outcome set was to assess whether the domains represented in the original core set are still relevant. In order to gather information from patients with and experts in axSpA regarding the importance of outcomes to be assessed in all trials in axSpA, a Delphi survey was conducted. From the results of this Delphi survey -described in **chapter 5** of this thesis- we learned patients with axSpA had a different opinion regarding the outcomes that have to be assessed in all trials investigating therapies in axSpA than the experts involved in their treatment. Patients preferred an all-inclusive approach, whereas experts turned out to make a distinction between outcomes of critical importance for different treatment settings. According to the experts in the Delphi exercise, more objectively measurable domains such as structural damage and mobility were most critical to be measured in settings investigating disease modifying therapies, whereas the importance of assessing the more subjective domains such as pain, stiffness, and overall functioning & health was considered limited to the settings investigating symptom modifying therapies.

The Delphi survey is a common method used to gather opinions from a large group of participants, either to prioritise research topics, for importance ratings or to reduce item lists. For something used so often, guidance on the fundamentals of the methodology is scarce. In **chapter 6** we aimed to provide insight in the effect of choosing a certain invitation technique on the outcome of the Delphi, by comparing two often used invitation approaches: 1) Invite all participants to subsequent rounds, irrespective of response to the previous round; or 2) Invite only those participants who completed the previous round to subsequent rounds.

We found there is no effect on the final outcome of the Delphi, but argued it may be preferential to invite participants who missed a round to subsequent rounds, as this approach is less sensitive to the non-random loss of opinions that could lead to false consensus. Additionally, this approach ensures the end-result displays the opinion of all those invited.

The Delphi survey was a small part of a much larger effort to update the core outcome set for ankylosing spondylitis, to ensure applicability to all patients with axSpA. In **chapter 7** we described the process that led to the ASAS-OMERACT core domain set for axSpA. At

its core, the resulting core domain set is similar to the original core set for ankylosing spondylitis[34,35], in that both include the domains physical function, morning stiffness (called spinal stiffness in the core set for ankylosing spondylitis), pain, fatigue and disease activity (represented by patient global assessment in the core set for ankylosing spondylitis) in their core and structural damage (referred to as spine radiograph in the core set for ankylosing spondylitis) for DMARD settings only. The first noticeable difference is the addition of overall functioning and health in the core of the updated core domain set for axSpA, representing the impact of axSpA on other aspects of life and how this has received increased attention in the past years. Secondly, the removal of spinal mobility from the core is remarkable, which was caused by the lack of standardisation and poor reliability and sensitivity to change of spinal mobility outcomes[36,37]. Now the domains have been established, appropriate instruments need to be selected to measure these domains.

As a preparatory step towards the selection of instruments for the core set for axSpA, **chapter 8** described the test-retest reliability of the outcomes assessed in three recent randomised controlled trials in axSpA. From this study we concluded that even though most instruments were developed for radiographic axSpA they were also found reliable for non-radiographic axSpA. Furthermore, this study provided evidence in favour of multi-item instruments, as they were found to be more robust against measurement error.

Implications from these findings

Using the data collected in the Delphi survey-described in **chapter 5**-, we were able to compose a concise list of domains of which all stakeholders agree these are the most important domains to assess in all trials of axSpA. These data provided the basis for the development of the updated core domain set for axSpA. **Chapter 7** revealed the process underlaying the development of and emanating endorsement of the core domain set for axSpA. The changes made compared to the original core set for ankylosing spondylitis will have implications for future research, as it requires the assessment and reporting of slightly different outcomes than have been done previously. Unfortunately, for some aspects of disease (e.g. spinal mobility) this reduces comparability with older studies, yet at the same time ensures more comparability in future trials due to increased clarity on which outcomes should be measured and reported. Furthermore, all stakeholders who will benefit from an updated core set were involved in its development, which will increase the uptake. Next steps include the selection of the best instruments to assess the selected domains for which careful consideration includes assessment of measurement properties, feasibility and usability of the candidate instruments. The assessment of test-retest reliability and measurement error was described in **chapter 8** of this thesis, which provided proof that the assessed instruments are reliable for all patients with axSpA. Further research will have to investigate the other measurement properties (such as construct validity, and discrimination) before a final decision can be made on which

instruments are best suitable to measure the endorsed domains. Finally, **chapter 6** filled a knowledge gap regarding who to invite when conducting a Delphi survey. With this paper we provided a first piece of methodological guidance regarding the Delphi survey.

Further discussion and future perspectives

The core set for ankylosing spondylitis[34,35] which is currently used to determine which outcomes should be assessed as minimum in each clinical trial conducted in axSpA does not contain an instrument for each domain. No specific instruments were defined for the assessment of fatigue or enthesitis, because at the time of development of the core set there was no validated instrument available[28]. As a result, various instruments have been used to assess these domains, impeding comparisons between trials that are ever so important for assessing treatment efficacy. Therefore, it is vital the updated core set will advise one specific instrument for each domain (with the potential of adding more) that will not only have to be measured, but more importantly will have to be reported in each trial, enabling one-on-one comparison of trials and the development and update of treatment recommendations. The original core outcome set was well-implemented[38], indicating the implementation of a core set leads to structured collection of information for the endorsed domains in clinical trials. However, there was quite some variation in the instruments used to collect the information, emphasising the importance of recommending one specific instrument for each domain. Moreover, the review found that not all information that was collected was also reported[38]. For example, BASDAI includes a measure of fatigue, but frequently this was not reported separately and therefore no conclusion could be drawn on the effect of the investigated therapy on fatigue based on the presented data.

The main aim of a core outcome set is to provide the minimum and mandatory set of domains and instruments that has to be assessed in every trial. As our understanding of a disease increases, or the course of disease changes as a result of earlier recognition and effective therapy, this may lead to the development of new instruments (e.g. the ASAS-HI) or validation of existing instruments. There is a fair chance that these new instruments outperform existing instruments, and become the preferred instrument to measure a given domain, which would require an update of the core outcome set. Unfortunately, updating a core set is a lengthy and time-consuming process, and one might wonder whether the process surpasses its goal and whether there would be more core sets (i.e. more standardised measurements) if the process was more user-friendly. One option to simplify the process might be the regular review of a core set (e.g. every 10 years), in which it can be decided to replace an instrument if the new instrument has been shown to outperform the existing one, without having to go through all the steps required for the development of a new core set. In order to do so, a complete comparison on all psychometric properties between the new and existing instrument would be a pre-requisite.

For all its flaws, core outcome sets are valuable tools in research, as they allow for more transparency in the drug registration process (due to direct comparisons with previously registered drugs) and to a better acceptance of new treatments in the field (because a direct comparison of outcome measures shows its performance in relation to previously accepted/more familiar drugs). The past has taught us it is important to specifically define which instruments should be used to assess each domain, and which can be optionally added. Additionally, the use of an instrument does not guarantee all collected data is reported too. Hence, in addition to providing the domains and instruments to be assessed in each trial, the core set for axSpA should provide specific instructions for the reporting of data as well.

A Delphi survey was used to gain consensus among experts in and patients with axSpA. One can argue this is not the best way to collect information regarding the importance of domains in axSpA, as its main aim is to strive for consensus among participants. Though consensus tends to dilute the strength of less favoured opinions, and replaces individuality by group opinion, it ensures all the involved can accept the final outcome. An additional benefit of the Delphi survey is that it can be completed online (i.e. no space- and time constraints) and anonymously, thereby providing a safe environment to express what could be considered a less favourable opinion. One major drawback with the taken approach is the lack of standardisation in the methodology to execute Delphi surveys, which decreases its validity as a tool for such important aspects of research. Within this thesis we made an effort to critically assess one aspect in the methodology of the Delphi survey. Providing clear methodological instructions for the use of Delphi surveys will improve its validity and more importantly, the validity of its results. Future research should focus on other elements of the methodology, such as how feedback should be provided between rounds, or how many panellists and stakeholder groups should be invited. Within these methodological recommendations, separate instructions should be provided for the various applications of a Delphi survey (e.g. a preference of care evaluation requires a different approach than the selection of domains for a core set), such that clear guidance will become available for researchers who wish to perform a Delphi survey.

In conclusion, core sets are a valuable tool in outcome assessment, but future research should investigate whether their development process can be smoothened. As for Delphi surveys, there is no doubt they have proven their worth as a means of gathering opinions and reaching consensus, yet the lack of methodological guidance should be addressed in the future.

PATIENT REPORTED OUTCOMES IN EARLY AXIAL SPONDYLOARTHRITIS

In the final part of this thesis, we discussed health-related quality of life and work and activity impairments in axSpA. In **chapter 9** we showed work and activity outcomes as well as health-related quality of life improved over two years of protocolized follow-up in patients with chronic back pain suspected of axSpA. This improvement was shown for patients with and without a diagnosis of axSpA, yet patients diagnosed with axSpA showed a larger improvement compared to those without a diagnosis of axSpA.

Implications from these findings

The results from **chapter 9** have taught us we can expect some improvement with time in all chronic back pain patients suspected of axSpA-regardless of diagnosis-, which apart from treatment may relate to the fact that at their first visit to the rheumatologist a patient's complaints are at their most severe and thus might improve naturally with time (i.e. regression to the mean). Furthermore, as there was a control group available in this study (i.e. those without a diagnosis) we were able to conclude a diagnosis of axSpA is an independent predictor of improvement in health-related quality of life and work- and activity outcomes in patients with chronic back pain complaints. Importantly, sensitivity analyses showed that the differences in outcomes between those with and without a diagnosis of axSpA remained when analyses were restricted to patients not on biological therapy, indicating that treatment with biologicals did not explain the differences between the groups. Importantly, despite the improvements over time, outcomes were still impaired compared to the general population. These results emphasise the importance of optimising long-term health-related quality of life and social participation of patients with axSpA, which is also described in the current axSpA treatment guidelines as one of the primary treatment goals[39].

Further discussion and future perspectives

Chapter 9 showed that patients who were diagnosed with axSpA after two years of protocolised follow-up showed a larger improvement in their health-related quality of life, work productivity loss and activity impairment than chronic back pain patients who did not get diagnosed with axSpA. In **chapter 9** we mentioned that this might be due to a difference in available treatment options, but this might also be explained by a difference in illness perceptions and subsequent coping mechanisms. One of the questions we asked ourselves is whether being diagnosed could have an impact on how a patient perceives his/her illness and subsequently influence coping mechanisms. Future research should focus on getting more insight in the psychological effects of getting a diagnosis, whether simply knowing 'what is wrong with you' has an impact on how complaints are perceived, and whether quality of life might improve even further if patients are taught adequate

coping mechanisms. Simultaneously, adequate coping mechanisms might have an impact on work-outcomes. Work productivity loss is based on presenteeism (reduced ability to perform one's job adequately) and absenteeism (the hours missed from work due to disease), if effective treatment combined with adequate coping mechanisms can reduce presenteeism, subsequently work productivity loss will improve. As axSpA affects the lives of people that are in the prime of their life, there is an immense value in increasing our understanding of effective therapies -either medicinal, educational or psychological- that have a positive effect on quality of life and work- and activity outcomes.

Another aspect that deserves attention is the use of generic versus disease-specific questionnaires to assess health-related quality of life. The main advantage of using a generic questionnaire is that the scores can be compared to scores from patients diagnosed with other (chronic) diseases or healthy controls, augmenting societal value. Contrary, disease specific questionnaires (such as the ASAS-HI and Ankylosing Spondylitis Quality of Life survey) pertain more disease specific questions, providing a clearer insight in the effect of disease on quality of life of patients with axSpA and could be considered of higher scientific value. In **chapter 9** the use of the SF-36 allowed for a direct comparison of patients with and without a diagnosis of axSpA and also a comparison with the general population. Directly comparing those who did and did not receive a diagnosis of axSpA showed us that in fact there is a difference in improvement over time between these groups, a fact that would have gone unnoticed had we used a disease-specific questionnaire. Nonetheless, using a disease specific questionnaire could have led to insight in which aspects (if any) contributed to the improvement in quality of life over time. Therefore, a disease-specific questionnaire may be preferred in a longitudinal cohort of patients with a chronic disease. As it is very likely health-related quality of life will remain impaired in these patients compared to the general population, insight in aspects that contribute to improvement or worsening of quality of life over time might be more valuable, as this could bring about new treatment goals. In the end, research will always demand making choices and finding the optimal balance between cost and reward, hence, what constitutes as 'the right choice' will depend on the question at hand, the available data and many other variables that are beyond our control.

REFERENCES

1. Rudwaleit M, Haibel H, Baraliakos X, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. *Arthritis Rheum* 2009;60(3):717-27.
2. Muñoz-Fernández S, Carmona L, Collantes E, et al. A model for the development and implementation of a national plan for the optimal management of early spondyloarthritis: the Esperanza Program. *Ann Rheum Dis* 2011;70(5):827-30.
3. Spoorenberg A, van Tubergen A, Landewé R, et al. Measuring disease activity in ankylosing spondylitis: patient and physician have different perspectives. *Rheumatology* 2005;44(6):789-95.
4. Canhão H, Faustino A, Martins F, et al. Reuma.pt - the rheumatic diseases portuguese register. *Acta Reumatol Port* 2011;36(1):45-56.
5. Ciurea A, Scherer A, Exer P, et al. Tumor necrosis factor α inhibition in radiographic and nonradiographic axial spondyloarthritis: results from a large observational cohort. *Arthritis Rheum* 2013;65(12):3096-106.
6. van den Berg R, de Hooge M, van Gaalen F, et al. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the Spondyloarthritis Caught Early (SPACE) cohort. *Rheumatology* 2013;52(8):1492-9.
7. Lee W, Reveille JD, Davis JC, et al. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. *Ann Rheum Dis* 2007;66(5):633-38.
8. Rudwaleit M, Landewé R, van der Heijde D, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. *Ann Rheum Dis* 2009;68(6):770-6.
9. Rudwaleit M, van der Heijde D, Landewé R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68(6):777-83.
10. van der Linden S, Akkoc N, Brown MA, et al. The ASAS Criteria for Axial Spondyloarthritis: Strengths, Weaknesses, and Proposals for a Way Forward. *Curr Rheumatol Rep* 2015;17(9):62.
11. Feldtkeller E, Khan MA, van der Heijde D, et al. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatol Int* 2003;23(2):61-6.
12. Ciurea A, Scherer A, Weber U, et al. Age at symptom onset in ankylosing spondylitis: is there a gender difference? *Ann Rheum Dis* 2014;73(10):1908-10.
13. Chung HY, Machado P, van der Heijde D, et al. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. *Ann Rheum Dis* 2011;70(11):1930-6.
14. Skare TL, Leite N, Bortoluzzo AB, et al. Effect of age at disease onset in the clinical profile of spondyloarthritis: a study of 1424 Brazilian patients. *Clin Exp Rheumatol* 2012;30(3):351-7.
15. van der Linden SM, Valkenburg HA, de Jongh BM, et al. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27(3):241-9.
16. van Lunteren M, Sepriano A, Landewé R, et al. Do ethnicity, degree of family relationship, and the spondyloarthritis subtype in affected relatives influence the association between a positive family history for spondyloarthritis and HLA-B27 carriership? Results from the worldwide ASAS cohort. *Arthritis Res Ther* 2018;20(1):166.
17. van Lunteren M, van der Heijde D, Sepriano A, et al. Is a positive family history of spondyloarthritis relevant for diagnosing axial spondyloarthritis once HLA-B27 status is known? *Rheumatology* 2019;58(9):1649-54.
18. Ez-Zaitouni Z, Hilkens A, Gossec L, et al. Is the current ASAS expert definition of a positive family history useful in identifying axial spondyloarthritis? Results from the SPACE and DESIR cohorts. *Arthritis Res Ther* 2017;19(1):118.
19. Mustafa KN, Hammoudeh M, Khan MA. HLA-B27 Prevalence in Arab Populations and Among Patients with Ankylosing Spondylitis. *J Rheumatol* 2012;39(8):1675-7.
20. Ziade N, Abi Karam G, Merheb G, et al. HLA-B27 prevalence in axial spondyloarthritis patients and in blood donors in a Lebanese population: Results from a nationwide study. *Int J Rheum Dis* 2019;22(4):708-14.
21. López-Medina C, Molto A, Sieper J, et al. Prevalence and distribution of peripheral musculoskeletal manifestations in spondyloarthritis including psoriatic arthritis: results of the worldwide, cross-sectional ASAS-PerSpA study. *RMD Open* 2021;7(1)
22. Akkoc N, Khan MA. ASAS classification criteria for axial spondyloarthritis: time to modify. *Clin Rheumatol* 2016;35(6):1415-23.
23. de Winter J, de Hooge M, van de Sande M, et al. Magnetic Resonance Imaging of the Sacroiliac

Joints Indicating Sacroiliitis According to the Assessment of SpondyloArthritis international Society Definition in Healthy Individuals, Runners, and Women With Postpartum Back Pain. *Arthritis Rheumatol* 2018;70(7):1042-48.
24. Weber U, Jurik AG, Zejden A, et al. Frequency and Anatomic Distribution of Magnetic Resonance Imaging Features in the Sacroiliac Joints of Young Athletes: Exploring "Background Noise" Toward a Data-Driven Definition of Sacroiliitis in Early Spondyloarthritis. *Arthritis Rheumatol* 2018;70(5):736-45.
25. Boers M, Kirwan JR, Tugwell P. OMERACT Handbook, 2018.
26. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials* 2017;18(3):280.
27. van der Heijde D, van der Linden S, Bellamy N, et al. Which domains should be included in a core set for endpoints in ankylosing spondylitis? Introduction to the ankylosing spondylitis module of OMERACT IV. *J Rheumatol* 1999;26(4):945-47.
28. van der Heijde D, van der Linden S, Dougados M, et al. Ankylosing spondylitis: plenary discussion and results of voting on selection of domains and some specific instruments. *J Rheumatol* 1999;26(4):1003-5.
29. Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis Rheum* 2005;52(4):1000-8.
30. Rudwaleit M, Jurik AG, Hermann K-GA, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68(10):1520-27.
31. Lukas C, Landewé R, Sieper J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;68(1):18-24.
32. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, et al. Assessment of enthesitis in ankylosing spondylitis. *Ann Rheum Dis* 2003;62(2):127-32.
33. Kiltz U, van der Heijde D, Boonen A, et al. Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. *Ann Rheum Dis* 2015;74(5):830-35.
34. van der Heijde D, Bellamy N, Calin A, et al. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. *J Rheumatol* 1997;24(11):2225-9.
35. van der Heijde D, Calin A, Dougados M, et al. Selection of instruments in the core set for DC-ART, SMARD, physical therapy, and clinical record keeping in ankylosing spondylitis. Progress report of the ASAS Working Group. Assessments in Ankylosing Spondylitis. *J Rheumatol* 1999;26(4):951-4.
36. Marques ML, Ramiro S, Goupille P, et al. Measuring spinal mobility in early axial spondyloarthritis: does it matter? *Rheumatology* 2019;58(9):1597-606.
37. Ogdie A, Duarte-García A, Hwang M, et al. Measuring Outcomes in Axial Spondyloarthritis. *Arthritis Care & Research* 2020;72(S10):47-71.
38. Bautista-Molano W, Navarro-Compán V, Landewé RB, et al. How well are the ASAS/OMERACT Core Outcome Sets for Ankylosing Spondylitis implemented in randomized clinical trials? A systematic literature review. *Clin Rheumatol* 2014;33(9):1313-22.
39. van der Heijde D, Ramiro S, Landewé R, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 2017;76(6):978-91