



London, 23 June 2005  
CPMP/EWP/4891/03

**COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE  
(CHMP)**

**DRAFT**

**GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL  
PRODUCTS FOR THE TREATMENT OF ANKYLOSING  
SPONDYLITIS**

<b>DISCUSSION IN THE EFFICACY WORKING PARTY</b>	January 2004 – April 2005
<b>TRANSMISSION TO CHMP</b>	June 2005
<b>RELEASE FOR CONSULTATION</b>	June 2005
<b>DEADLINE FOR COMMENTS</b>	December 2005

**Note:**

Any comments to this Guideline should be sent to the EMEA EWP Secretariat by e-mail: [line.jensen@emea.eu.int](mailto:line.jensen@emea.eu.int) or by fax: +44 20 74 18 86 13 by the end of December 2005.

## GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS FOR THE TREATMENT OF ANKYLOSING SPONDYLITIS

This Guideline is intended to provide guidance for the evaluation of new medicinal products in ankylosing spondylitis. This Guideline should be read in conjunction with Directive 2001/83/EC and all other pertinent elements outlined in current and future EU and ICH guidelines and regulations, especially those on:

- Dose-Response Information to Support Drug Registration (ICH E4),
- Statistical Principles for Clinical Trials (ICH E9),
- Choice of Control Group in Clinical Trials (ICH E10),
- The Extent of Population Exposure to Assess Clinical Safety for Drugs (ICH E1A).

This Guideline is intended to assist applicants during the development of medicinal products. It is only guidance; any deviation from guidelines should be explained and discussed in the Expert reports/Clinical Overview.

### INTRODUCTION

Ankylosing spondylitis (AS) belongs to the group of the spondyloarthritides (SpA). AS and psoriatic arthritis are the most important subtypes of SpA. AS is a chronic inflammatory disease that involves primarily the sacroiliac joints and the axial skeleton. Further, the entheses, peripheral joints and specific organ sites such as the anterior uvea may be involved. It is a largely genetically determined disease which has a strong association with the HLA-B27. The prevalence has been estimated to be between 0.1 % - 1.1 % of the population. AS is more common in males (male to female ratio is estimated to be 2-3:1). The disease tends to be more severe in men, in whom the spine is more frequently involved. Clinical manifestations usually begin in late adolescence or early adulthood (mean age of onset 26 years) and onset after age 45 is rare.

Clinical manifestations include lower back pain with predominant nocturnal pain and morning stiffness. Also chest pain, pain and swelling of peripheral joints and extra-articular tenderness may occur as well as several extraskeletal manifestations such as acute anterior uveitis, cardiac conduction defects and aortic valve disease or renal disease, mainly in the form of secondary renal amyloidosis.

AS is a chronic disease that causes a substantial amount of pain and disability. Functional limitations relate to inflammation in the early phases of disease but also increase with duration of disease due to new bone formation. Although most patients are able to maintain functional capacity there are also some patients with progressing disease who rapidly develop ankylosis at a young age. There are no solid prognostic parameters besides early radiographic progression.

The therapies which have been available until recently are of value but this is clearly limited. Physical therapy has a positive effect on stiffness and on spinal mobility and even on pain. NSAIDs are used to control pain with good response in most patients. Although it has been suggested that NSAIDs given on continuous basis may provide some benefit in terms of radiographic progression there is no evidence that NSAIDs may alter the course of the disease or prevent progression of disability. Systemic corticosteroids are in general of little benefit but they are used in selected patients, mainly for peripheral arthritis or extraskeletal manifestations. Intraarticular corticosteroids may also be used similar to other rheumatic diseases. Sulfasalazine is used and thought to have an effect on peripheral disease and extra-articular manifestations, but there is no evidence of effect in severe disease or in patients with substantial spinal involvement. Methotrexate is also sometimes tried based on limited data. Recently, some TNF- $\alpha$  antagonists have shown efficacy in the active disease and have been incorporated into the treatment of active moderate to severe AS patients.

## 1. PATIENTS CHARACTERISTICS AND SELECTION OF PATIENTS

Ankylosing spondylitis is the most severe form and the prototype of the spondyloarthritides. This concept classifies patients with related diseases on the basis of the 1991 proposal of the European Spondyloarthropathy Study Group (ESSG). The other clinical entities included are psoriatic arthritis, reactive arthritis, inflammatory bowel disease-related arthritis and undifferentiated spondyloarthritides.

AS shares several features with these other subtypes of SpA but from a therapeutic point of view it needs to be considered separately. Therefore, it is important that AS is diagnosed using highly specific criteria to exclude patients with other spondyloarthritides where other therapeutic strategies may be considered.

The 1984 modified New York classification criteria are accepted criteria to diagnose AS with a high degree of specificity. Using these criteria the diagnosis relies strongly on the existence of changes in the sacroiliac joints exceeding grade II unilateral. The inter and intra-observer variability in the interpretation of radiographs is a known problem which may affect the specificity of the classification criteria. This should be addressed in the planning of the study in order to accurately document the existence of a grade II degree in sacroiliac changes.

The modified New York criteria, although specific, are not suitable to diagnose early patients with AS. Magnetic resonance imaging (MRI) may provide better visualisation of early stages of sacroiliitis but it is not yet sufficiently standardised to be used for classification.

There are three main clinical manifestations: 1. axial symptoms, 2. peripheral arthritis and 3. enthesitis. In addition, specific organ involvement may occur, the most common one is the anterior uveitis

Patients characteristics should be well documented: demographics, duration of the disease, previous and concomitant therapies, concomitant diseases including those specific diseases related to AS such as anterior uveitis or cardiac disease (conduction, valve problems). All specific diagnostic actions taken by physicians before including patients (e.g. screening for latent tuberculosis) should be described in the selection criteria of the protocol. In addition, there are some characteristics that may be considered in order to identify subpopulations where the benefit risk ratio of the new product might be different. These characteristics are the severity and extent of the disease, the disease activity, the spinal and peripheral joint involvement and the lack of response to previous treatments. Relevant identified subpopulations should be limited, justified and defined a priori in the study protocol. Statistical analysis should be adequate (see PtC concerning multiplicity in Clinical Trials). The consideration of a patient as non-responder to NSAIDs requires documentation of the lack of response with appropriate doses and treatment durations. Special attention should be paid to other patient characteristics that might be relevant depending on the mode of action of the product. In this sense, the levels of serologic markers of inflammation such as C reactive protein (CRP) or the erythrocyte sedimentation rate (ESR) should be considered.

The absence of HLA-B27 should not be an exclusion criterion. Patients between 16 and 18 years and those whose symptoms started prior to this should be not excluded from clinical trials.

Disease activity at the moment of the enrolment in the trials should be distinguished of the level of damage and functional disability reached by the patients due to the evolution of the disease until that moment. Activity of disease should be assessed by means of validated scales and considering several aspects of the disease such as pain or stiffness,.. The use of composite simple scales such as the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) is considered appropriate and may be complemented with the measurement of individual symptoms (e.g. pain) or global patient assessments of the disease with visual analogue scales. The levels of serologic markers of inflammation such as C reactive protein (CRP) or the erythrocyte sedimentation rate (ESR) should additionally be considered

In addition to the disease activity at a given time, the severity of the disease is determined by other characteristics such as the level of irreversible structural damage, the involvement of peripheral joints and other organs, the concomitant diseases and the unresponsiveness to previous treatments.

For including patients a moderate to severe disease activity should be required in order to show a sufficient treatment response (e.g. BASDAI >4 and nocturnal pain as measured by VAS >4). A

minimum duration of active disease should be established before entering into the trial. A duration of 4 weeks is sufficient to introduce products like NSAIDs, however a minimum duration of 3 months of active disease would be needed when assessing products to be used in patients not controlled with NSAIDs.

## **2. METHODS TO ASSESS EFFICACY**

In recent years, several experts have worked on defining the different domains to assess efficacy of products for AS as well as on the development and validation of methods to assess changes in those domains. Different domains may be assessed separately or using composite indexes that bring together the assessment of several domains.

From a regulatory point of view, we distinguish the following claims of a therapy:

- 1) improvement of symptoms and signs such as pain and stiffness or enthesopathy,
- 2) improvement of physical function,
- 3) slowing or prevention of structural damage,
- 4) prevention of disability.

### **2.1 Main domains to be assessed in AS and instruments to be used in each domain**

#### Pain

Pain is adequately measured by means of patient self answered VAS. Patient should be asked for both specific pain at night due to AS and pain (without time restraints) due to AS. The question should refer to a recent past period (e.g. the past week or the past 48 hours).

Additional measures of pain may be provided by three out of the six items of the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) that ask for pain and discomfort during the last week. The three items assess 1) overall pain in neck, back or hip; 2) overall level of pain/swelling in joints other than neck, back or hip and 3) overall discomfort from any areas tender to touch or pressure.

#### Physical function

There are several acceptable instruments to measure physical function and its changes in patients suffering from AS. The most widely accepted and known instruments are two patient administered questionnaires: the Bath Ankylosing Spondylitis Functional Index (BASFI) and the Dougados Functional Index (DFI). The BASFI consists of 10 questions regarding ability to perform specific tasks as measured by visual analogue scales (VAS) whereas the DFI addresses similar aspects in 20 questions that should be answered by means of a three or five categories Likert scale. It appears that the BASFI is more sensitive to changes and easier to use than the DFI.

#### Spinal stiffness

Spinal morning stiffness that improves with movement is a relevant symptom related to inflammation in AS. The BASDAI index provides a good assessment of stiffness through the combination of two of its questions. One item measures the intensity of morning stiffness by means of a 100 mm VAS and the last item asks for the duration of stiffness from time of awakening. Both questions are referred to the situation during the previous week.

#### Patient global assessment

Patient subjective perception is an important complementary variable that may be measured by means of a visual analogue scale, asking the patients to inform on his/her global status during a recent past period (for example last week).

## Spinal Mobility

Spinal mobility is of great importance in AS and constitutes the most specific domain because other domains are common with many other rheumatic diseases. Although it may be difficult to detect changes in spinal mobility on the short term, spinal mobility is considered an important measure to assess efficacy.

Several instruments have been developed and a combination of them may be used in clinical trials. Chest expansion, modified Schober test, lateral spinal flexion and occiput to wall distance are amongst the most known methods to measure spinal mobility.

The Bath Ankylosing Spondylitis Metrology Index (BASMI) is a combined measure of spinal mobility and hip function that has been used in several clinical trials and is also an acceptable method. However, this index does not include the well established measures previously mentioned and therefore, if the BASMI index is used, it is recommended to add some of the other relevant single measures (e.g. the occiput to wall distance , chest expansion,..).

## Structural damage

The assessment of structural damage and progression in AS is mainly based on radiography. Osteodestructive and osteoproliferative changes are detectable and acute and chronic changes need to be differentiated. Structural damage may be detected by conventional radiographs, by MRI/T1 weighted sequences and by DEXA or quantitative CT measurements. Several radiographic scoring systems have been developed such as the Bath Ankylosing Spondylitis Radiology Index (BASRI), the Stoke Ankylosing Spondylitis Spinal Score (SASSS) and the more recent modified SASSS which appears to be the most sensitive method with a good reliability.

There is still limited knowledge about the relation of these radiographic scoring systems with the underlying pathophysiological mechanisms and the clinical course of the disease. In addition, its sensitivity to change appears modest in AS patients (i.e. a considerable number of patients with no changes in these scores in one year follow-up) and there is limited knowledge on its use in long-term clinical trials.

Radiographic changes might be used as a measure of structural damage, but at present, the clinical relevance of observed differences is unclear. Therefore, clinically relevant differences should be predefined and justified.

## **2.2. Other domains and instruments to be assessed**

### Peripheral joints and entheses

The affection of peripheral joints is a complementary assessment with modest specificity in measuring efficacy of products to treat AS. It may be measured as the count of the swollen or tender joints based on the 44 joints counts without grading or weighting.

The assessment of affected entheses is time consuming if it is done in an extensive manner (e.g. Mander entheses index based on 66 entheses)). On the opposite end, the simple question included in the BASDAI about “discomfort because areas tender to touch or pressure” may not be sensitive enough to capture changes in entheses inflammation status. Several indices have been developed in order to provide a feasible method to assess enthesopathy. The MASES index based on 13 entheses may be an acceptable instrument but also other instruments if validated and reliable might be appropriate.

### Acute phase reactants

Although levels of C reactive protein (CRP) or the erythrocyte sedimentation rate (ESR) may be related to the activity of the disease and its prognosis, there are no data to support them as useful surrogate variables to assess efficacy in AS.

### Quality of life (QoL)

It may be assessed either using some specific scales (e.g. ASQoL) or general instruments (e.g. SF-36). The use of accepted multidimensional scales assessing QoL may provide complementary information to

the efficacy demonstrated by the main variables. These multidimensional scales are preferred over specific physical QoL scales that are closely related to the improvement of symptoms and physical function.

### 2.3 Main efficacy end points

The use of a composite measure based on the previous domains assessments is an appropriate way to assess the efficacy of a product. For this purpose only validated composite endpoints are acceptable as primary or secondary endpoints, provided that consistency is shown between different measures of the composite as well as with other single efficacy measures. It is very important that response criteria are adequately justified, chosen before the study is started and thresholds are predefined.

The main efficacy end point will depend on the type of product and the intended therapeutic claim

#### 2.3.1 Medicinal products intended to improve symptoms/physical function

The ASAS working group, based on previous trials with NSAIDs, has proposed a measure that has been now used in several trials to measure the efficacy of new products on the symptoms of AS. The ASAS Response Criteria (ASAS 20) is defined as an improvement of at least 20% and absolute improvement of at least 10 units on a 0-100mm scale in at least 3 of the following domains:

- Patient global assessment measured on a VAS scale with extremes labelled “none” and “severe.”
- Pain assessment represented by the average of total and nocturnal pain scores, both measured on a VAS scale with extremes labelled “no pain” and “most severe pain.”
- Function represented by BASFI average of 10 questions regarding ability to perform specific tasks as measured by VAS with extremes labelled “easy” and “impossible.”
- Morning stiffness, represented by the average of the last 2 questions on the 6-question BASDAI regarding morning stiffness as measured by VAS: one with extremes labelled “none” and “very severe”; the other marking duration of morning stiffness between “0” and “2 or more hours.”

Additionally, absence of deterioration (of at least 20% and absolute change of at least 10 units on a 0–100 mm scale) in the remaining domain should be documented.

The percentage of patients reaching an ASAS 20 response may be an acceptable primary efficacy end point for some products (e.g. NSAIDs). However, its use as a primary efficacy endpoint may not be appropriate to evaluate a new product from a new therapeutic class where major improvement is intended to be assessed. Furthermore the ASAS 20 composite does not include the assessment of the spine mobility, which is a relevant efficacy parameter in AS

Two new indexes have been considered to improve the assessment:

With regard to the magnitude of the clinical response, it has been proposed to use the ASAS 40 response criteria. It is defined, similarly to ASAS 20, as an improvement of at least 40% and absolute improvement of at least 20 units on a 0-100mm scale in at least 3 of the 4 domains, without worsening in the remaining domain. This index has been already used in several trials and may be considered an appropriate primary efficacy end point to assess major clinical response

With regard to spinal mobility, there is a proposal to supplement the ASAS index with two additional domains: spinal mobility and acute phase reactants. However, this index (ASAS 5/6) defines the response as at least 20% improvement in any 5 of the 6 domains, with the possibility of leaving out the improvement in spinal mobility and assigning an important efficacy consideration to the change in biochemical markers.

In summary, the percentage of patients with ASAS 20 or ASAS 40 is considered an appropriate primary end point to assess the efficacy of a new medicinal product intended to improve symptoms of AS. Spinal mobility must be considered either a co-primary end point or an important secondary end-point.

### 2.3.2 Additional claim to prevent structural damage

There are several radiological scores and recently, based on consensus the modified SASSS has been chosen as the preferred validated and sensitive method to assess radiological changes. However, to address possible claims of prevention of structural damage, imaging methods should be accompanied by a demonstration of an effect on the clinical consequences of the slowing or prevention of the structural damage, i.e. an effect on spinal mobility .

Radiographs should be taken on fixed and predefined time points and be assessed by at least two assessors blinded for the allocation of the patient to type of treatment, chronological sequence of the radiographs and initial assessment(s) of the other assessor(s). The method for obtaining the final score should be described in detail (e.g. consensus) and be predefined.. Handling of missing information should be described and justified

## 2.4 Secondary end points

### 2.4.1 Spinal mobility

Spinal mobility should be considered as an important secondary end-point, even for symptomatic short-term treatments

### 2.4.2 Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)

BASDAI is a composite index that includes the assessment by the patients of their symptoms of pain, discomfort, stiffness and fatigue. It is a widely used measure of disease activity and its changes with treatment should be assessed. The percentage of patients with clinical response as measured by an improvement of at least a 50% from the baseline score in BASDAI is considered useful to judge the clinical benefit of a treatment.

### 2.4.3 Global patient assessment

2.4.4 Other secondary end points may be the individual components of the ASAS instrument as well as individual assessments of the main domains of the disease. Additional endpoints may be the ASAS 40 or 70 or the ASAS 5/6 as well as the peripheral tender joints and swollen joint count (change and percent change from baseline).

## 3. STRATEGY AND DESIGN OF CLINICAL TRIALS

### 3.1 Early Studies in Man

Specific dose response studies should be performed in patients with AS. There are several antecedents of different response to medicinal products in patients with AS as compared to the same product in other rheumatic diseases (i.e. rheumatoid arthritis) or other AS- related non-articular disorders (i.e. inflammatory intestinal disease). Therefore, dose guidance provided by previous studies in other related disorders is of limited value. An appropriate dose finding should be performed in patients with AS in order to find the posology regimen with the most favourable benefit-risk ratio in this particular disease.

Whenever appropriate depending of the mechanism of action, efforts should be done to find different doses or intervals according to the respective patient characteristics (i.e. severity, inflammation,).

Placebo controlled parallel group studies are recommended.

The ASAS 20 composite index is an appropriate measure to the exploratory trials and short duration trials (e.g. 12-24 weeks) may be enough to demonstrate efficacy on symptoms of the disease.

### 3.2 Therapeutic Confirmatory Studies

Medicinal Products with a claim of improvement of symptoms and physical function:

Conventional treatment of AS consists of NSAID combined with physical therapy which are enough to control pain in most patients as well as to improve physical function. Therefore, new products belonging to therapeutic classes other than NSAIDs are expected to be combined with this standard therapy in

patients with insufficient control of their symptoms. The improvement of symptoms include pain, stiffness and physical function

#### Study design

Studies should have a randomised, double blind, parallel group design. Efficacy of products claiming improvement in symptoms and disease activity or function are generally established by means of placebo controlled trials that may well be add-on trials where all patients receive physical therapy and NSAIDs.

Products belonging to new therapeutic classes may need also comparison against anti TNF treatments in order to properly assess the benefit risk ratio of the new product. A three-arm trial is recommended.

The concomitant standard therapy should be carefully documented and its impact on results analysed based on a pre-established plan. Also the previous use and response to standard therapy should be documented.

#### Main end points and study duration

The primary end point depends on the expected extent of response induced by the product. With products like NSAIDs, the percentage of patients with an ASAS 20 improvement at 6-12 weeks may be an appropriate end point. For products other than NSAIDs (e.g. TNF inhibitors), percentage of patients with an ASAS 20 or preferably ASAS 40 at 12 or 24 weeks is a more appropriate end point. It is expected that a concomitant improvement in spinal mobility is also demonstrated.

AS is a chronic disease and therefore, symptomatic treatment is expected to be maintained on the long term. Therefore, although efficacy may be demonstrated in 12-24 weeks trial, maintenance of the effect in longer trials (e.g. 1 year) should be demonstrated. To establish that a symptom-modifying drug does not have deleterious effects, structural changes should be monitored for at least one year. In addition, the adequate duration of treatment should be addressed and data after stopping therapy should be provided.

#### Medicinal Products with a claim of slowing or prevention of structural damage

Confirmatory trials to demonstrate an effect on prevention of structural damage and subsequent function, spinal mobility and disability should be parallel group controlled trials of long duration (e.g. at least 2years).

Trials should be ideally double blind placebo controlled trials. However, it is acknowledged that such a long duration of a placebo controlled trial may raise feasibility and ethical concerns.

From a therapeutic point of view, patients with mild disease activity may probably be enrolled in such a long trial as an add-on trial over standard therapy with NSAIDs, physical therapy or corticosteroids if needed. However, depending on the type of product (i.e. frequent/painful parenteral administrations) such a long placebo controlled trial may also have feasibility problems. On the other hand, this population may not be suitable to demonstrate prevention of structural damage due to their slight progression.

Patients with severe disease activity cannot be maintained in a placebo-controlled trial for a long period because of the recent availability of some therapies other than NSAIDs (i.e. anti-TNF). Therefore, unless an add-on therapy over anti-TNFs was the aim of the therapy, alternative designs should be explored. A possible alternative may be a trial with a randomised start of the active treatment. Differences between groups may be sustained at the end of the 2 or 3 years period reflecting the difference in the start of treatment.

Slowing of radiographic progression may itself not constitute a definite patient benefit and it is still not accepted surrogate for long term clinical benefit. Although there is indirect evidence that, by favourably modifying the natural history of AS in terms of structural changes, long-term clinical benefit will occur in a large proportion of patients, it would be expected that an applicant will provide additional evidence to support this surrogacy.



Differences in radiological scores alone are not considered clinically relevant in AS if they are not accompanied by differences in spinal mobility, function and an adequate clinical response. Although the radiological score may be the primary end point it should be accompanied by a clinical co-primary end point, preferably related to spinal mobility. Clinical response (e.g. percentage of patients with an ASAS 20 or 40 response) should be an important secondary end point.

#### **4. CLINICAL SAFETY EVALUATION**

##### **4.1 Specific adverse events to be monitored**

In addition to those that may be related to the pharmacological actions of the product (infections, malignancies,...) it is necessary to give some reassurance on the lack of deleterious effect on the affected joints. Monitoring of structural changes during at least one year may provide reasonable information

##### **4.2 Extent of population exposure to assess clinical safety**

Complete safety database should be submitted for assessing a new product. When AS is an additional indication for an already approved product, safety data obtained in trials in other indications can be considered, provided that the dosage regimen is the same and population is expected to behave similarly (e.g. rheumatoid arthritis or psoriatic arthritis).

Considering the need for confirmatory trials for efficacy as well as the rest of studies specific in AS, it is possible that efficacy trials may provide also controlled safety data.

##### **4.3 Long term safety**

Ankylosing spondylitis is a prevalent chronic disease and treatments will need to be approved for long-term treatment. The need for specific long term trials to demonstrate efficacy and the effect on structural changes will provide an adequate safety data base, likely to go beyond the minimum requirement of the CHMP/ICH guidance requesting 300-600 patients treated for 6 months or 100 patients treated for 12 months. Although this may depend on the characteristics of the product, in general, safety data from periods longer than one year are recommended.

##### **4.4 Development of antibodies**

Whenever the development of antibodies (neutralising antibodies and others) (NAB) may be expected, the rate and therapeutic consequences of this fact should be studied. Factors that influence the appearance of NAB such as duration and dose of the treatment or the concurrent use of other medicinal products should be analysed.