



London, 20 January 2010
Doc. Ref. EMA/818496/2009

**OVERVIEW OF COMMENTS RECEIVED ON
DRAFT GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL
PRODUCTS USED IN THE TREATMENT OF OSTEOARTHRITIS**

Interested party (Organisations or individuals) that commented on the draft Guideline as released for consultation

Stakeholder No.	Name of Organisation or individual
1	OARSI, Osteoarthritis Research Society International
2	USZ, Department of Rheumatology and Institute of Physical Medicine, University Hospital of Zurich, Switzerland
3	AESGP, Association of the European Self-Medication Industry
4	EFPIA
5	GREES, Osteoarthritis section
6	EULAR

1. GENERAL COMMENTS – OVERVIEW:

Stakeholder No.	General Comment (if any)	Outcome (if applicable)
1	This draft has been reviewed by the executive committee and the chairs of the working groups involved in the OARSI initiative who leads the coordination of a critical appraisal of certain fundamentals of the science related to the design of clinical development programs for human drugs, biological products, and medical devices for the treatment and prevention of OA to assist the FDA as they work to finalize the draft guidance originally issued in July 1999. Since both agencies are working at the same time on the review of their own draft guidance document, OARSI sees an opportunity to share with EMEA their personal view.	
2	The overall draft text of this upgraded EMEA Guideline is very good and mostly reflects the actual state of knowledge in the field of OA therapy.	
3	No general comment.	
4	EFPIA welcomes the revision of the existing osteoarthritis (OA) Points to Consider document and the CHMP-EWP initiative to provide guidance on the clinical evaluation and potential label claims for the development of medicinal products for the treatment of OA. EFPIA wishes to raise the following key comments, regarding some of the concepts presented in the draft guideline. These key points are followed by other important comments presented according to the different sections of the draft guideline. In order to streamline the document, no editorial or typographical comments are provided. EFPIA major comments: We believe that the guideline should be more specific regarding the requirements for a claim for Symptom modifying drugs versus Structure modifying drugs. It should also address the acceptability of using MRI, which is becoming the standard for efficacy assessment for structure modifying	

Stakeholder No.	General Comment (if any)	Outcome (if applicable)
	drugs. 1. Harmonization of guidance across regions Harmonization of guidance across regions is highly desirable. There is a potential opportunity as the FDA is currently working with OARSI to come up with recommendations to update the current US FDA draft guidance for OA drug development. A public meeting is scheduled on 15 December 2009 to provide an update on work accomplished by the FDA OARSI Initiative. 2. Symptom modifying drugs: For symptoms modifying drugs that act on pain, the primary endpoint should be a symptomatic pain-related assessment and functional disability should be a key secondary endpoint. The CHMP/EWP should also consider that pain and functional improvement could be recognized separately. The guideline as currently written would exclude the possibility of studying new medicines as adjunctive therapy in combination with existing standard of care (analgesics and anti-inflammatory medications) as it currently refers to strict monotherapy setting. The CHMP/EWP should include wording to accommodate add-on therapy. The CHMP/EWP should clarify which level of evidence is expected to show that the drug has not detrimental effects (i.e. no clinically significant adverse effects on the structural changes) on structure. 3. Structure modifying drugs with no direct effect on symptoms While the guideline seems to include a path for the development of structure modifying drugs with no direct effect on symptoms, it states that symptom modifying effects have to be confirmed for the structure modifying drugs and specifically that the “benefit of radiological improvement” in the absence of any clinical improvement at the time of assessment (i.e. at least 2 years) is insufficient for marketing authorisation”. Consequently, this impedes such development options. It should be considered that it would be difficult to show statistical significance for pain considering the length of time that may be required for that a slow-down of structural disease progression would	

Stakeholder No.	General Comment (if any)	Outcome (if applicable)
	translate into a significant pain reduction. The acceptability to show a trend in improvement of symptoms when a correlation has been shown between structural outcome and pain and function should be considered and discussed. 4. Use of MRI Although the guideline recognises that there is an increasing body of evidence that MRI has the potential to evaluate joint OA it states that larger clinical studies are still needed to show its clinical relevance. Given the growing body of evidence on the potential of MRI to complement or replace X-ray, the CHMP/EWP should provide clear recommendation on data required to accept MRI taking into consideration that the guideline once adopted will not be updated for a few years. This is seen as an opportunity for synergy between the regulators and their stakeholders including industry.	
5	The draft guideline is considered to be an extremely valuable contribution, on line, with the current scientific view.	
6	After careful reviewing of this draft document, the EULAR steering committee members would like to thank and congratulate the EMEA persons in charge of the writing of such a guideline document - it perfectly summarizes current scientific knowledge in this area. We do feel that such document will strongly facilitate the design, the conduct, the report and the interpretation of clinical trials in this area. As such it enjoys our general support and commendation	

2. SPECIFIC COMMENTS ON TEXT (Stakeholder No. 1)

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		The separation of symptoms from structure, the recommendation to study drug effects for hip/knee separately from hand and from spine, the discussion of important factors to obtain on baseline history that may modify the drug	

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		effect have been strongly appreciated.	
		The inclusion of spine OA is interesting but the definition of spine OA is not clearly proposed.	No changes are considered necessary. This indication is an option that has to be justified by the applicant as a clear definition with respect to symptoms is considered problematic.
		The focus on pain intensity only as the primary symptom outcome maybe too restricted. The quality of pain is also a concern, for example, as is the impact on mood and sleep.	Partly accepted. Information for the selection of secondary endpoints is added: ‘Endpoints should be chosen in line with the pharmacological characteristic of the respective drug and the claimed indication.’ As primary endpoints pain and function are recommended. However, pain is generally accepted as main symptom because it is identified as most affecting disorder. Specific qualities of pain may be assessed with different respective questionnaires for example. However, these have to be validated. As secondary endpoints quality of life is listed. Insofar aspects as mood and sleep are indirectly sufficiently included and may be assessed. Anyhow mood and sleep have been included in brackets as example for quality of life measure in order to directly mention these important qualities.
		While there is mention of impact of pain on physical functioning as a co-primary outcome, none of these other effects/outcomes are discussed...	Not accepted. The impact of pain on function is not discussed as function is recommended as independent symptom whatever the cause of functional impairment might be. Other symptoms than pain and function are

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
			considered more appropriately listed as secondary endpoints.
Section 5.1		“numerical pain rating scales” should be mentioned -- for example, 0-10 scales -- which are all very well-validated in a large number of different pain conditions.	Accepted.
		The characterization of OA pain could be more sophisticated than “pain intensity” only.	Not accepted as the guideline is not intended to be used as a textbook. For a differentiation of pain quality assessment specific questionnaires can be used. The interpretation of the recommendations in this guideline should be adapted to the specific pharmacological characteristic of the respective drug and the claimed indication as mentioned above.
		The recommendation of a flare design as the gold standard for the assessment of pure efficacy of a drug versus placebo is reasonable, but once that is established in small numbers, larger real-world trials could be of interest.	We agree but to give advice to general research is not the intention of this guideline.
End of section 4		it is stated that all trials of symptomatic treatments require discontinuation of prior analgesic and anti-inflammatory medications, but shouldn't it also be possible to develop a medication to be used as an “add-on” treatment too, for example, NSAIDs? Discontinuing all meds may be a problem for many patients, and assessing add-on efficacy has merit in itself	Accepted. ‘In some cases when combined medication is assessed an add-on design may be used and patients should be stable on the respective medication prior to initiating treatment with the test drug.’
Lines 143 and 158		The limitation of enrolment to persons with symptomatic OA precludes enrolling persons at high risk of disease (e.g. post injury) and disease prevention. Line 158 appears to contradict this by suggesting enrolment of high risk individuals	Not accepted. The mentioned limitations of enrolment (line 143) are not clear as no limitations are described in that paragraph. In line 158 an option is mentioned to select specific patients with a more rapid progression in order to better detect positive effects on structure damage.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 150-1		It is unclear how this benefits clarity of the recommendation.	Not accepted. To mention MRI and its current status is considered adequate because this instrument may be used in future if sufficiently standardised. This has to be justified by the applicant.
Lines 283-290		Based on the recent literature, the paragraph on MRI could be modified since some clarifications on these questions have been provided.	See comment below on MRI.
Lines 358-259		Does this apply to all symptom modifying trials? How is it recommended structure be assessed? To adequately assess for no effect will require large sample sizes over long periods using current metrics.	Partly accepted. The subjunctive is knowingly used for the recommendation of the long term follow up because it is considered no need to assess e.g. topical NSAIDs with a long term follow up. The absence of negative effects on structure should be warranted within the safety long term data. It might be demonstrated by radiographic measure. A sentence in 7.3 is added.

2. SPECIFIC COMMENTS ON TEXT (Stakeholder No. 2)

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 369-371		Comments: Since years, both the scientific societies (Grees,OARSI, Omeract ...) and the investigators tried to significantly improve the design, the methodology and the outcomes of the clinical trials aiming at testing the effects of various symptomatic and/or chondroprotective drugs used in the treatment of OA at various locations (knee,hip, fingers...) .After the release of each new Guideline, all serious inverstigators did adapt their new trials accordingly and this is exactly what was done in the more recent multicentre, international large trials, for instance testing the effects of oral chondroitin sulphate in knee OA patients.	Partly accepted. Paragraph is accordingly revised.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		In addition, the guidelines did also clearly mention the requirements for testing a symptomatic drug versus a chondroprotective drug. In the domain of slow acting drugs for OA treatment, some of the drugs tested in the last years did provide enough evidences to be widely accepted as efficacious in terms of symptoms and for some of them also in terms of structure or disease progression. One of these Slow-acting drug for OA is oral chondroitin sulphate and this substance did show 1°) significant evidences to support the positive algo-functional effect of the drug in knee OA patients ; 2°) significant evidences to support a structure-modifying effect of the drug in knee OA patients. The requirements listed in the 369-371 sentence are basically impossible to fulfil based on the very demanding and clearly specified methodology applied to all the RCT's up to now ! As the design of symptomatic oriented RCT's and structure-modifying RCT's is absolutely not the same, there is definitively no way to fully compare and integrate the evolution of the pain and/or function data and the evolution of the Joint Space Narrowing at a specific target joint ! As OA trials are generally of long duration and do last a couple of years until completion, it would not be fair to fully change the rules---guidelines—when RCT's did already begin or did just finish. In addition, there is until now to the best of our knowledge absolutely no proof that there would be any significant correlation (negative or positive) between the radiological progression of joint space narrowing in OA patients and the level of pain and/or limitation of the function in the patients ! As a consequence, The sentence written under 6.2.1. Modification of structure, lines 369-371, should be fully suppressed from the final text of this upgraded CPMP Guideline. The rest of the document is fully acceptable and should be maintained as such. Proposed change (if any): Suppress the whole sentence located line 369-371	

2. SPECIFIC COMMENTS ON TEXT (Stakeholder No. 3)

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
79- 80		Comment: Please provide an example of a symptom modifying slow acting drug for clarity	Not accepted. To give this additional information is considered not necessary and in addition it would imply that the mentioned SYSADOA is more appropriate for a treatment than others.
88-90		Comments: It is said that symptom modifying treatments should demonstrate absence of adverse structural effects. It is unclear with what means this should be demonstrated, no specific endpoint is listed in 5.1, and section 5.2 primarily addresses limitations of the methods to assess structural damage, and may rather be intended to describe endpoints for disease modifying treatments. Would the design of studies targeting modification of structure (6.2.1) also apply for demonstration of the absence of potentially negative effects by symptomatic treatments? Proposed change: Specific guidance is warranted to what means are accepted to demonstrate absence of negative effects on structure.	Accepted. A sentence in 7.3 is added. The absence of negative effects on structure should be warranted within the safety long term data. It might be demonstrated by radiographic measure. Therefore this point could be added in section 7.3. For structure modifying drugs an earlier proof of symptom benefit and no deterioration is necessary. This should be mentioned
92-94		Comments: The common separation of pain into mild, moderate, severe is referred to and it is said that the severity of pain is to be “addressed in the claimed indication”. This may suggest that if a certain number of patients of each category were included in trials, a claim for mild to severe pain may be granted. With this it is implied that in order to get a claim “severe pain” no additional indications such as cancer pain need to be studied, and that studying patients of different pain severity has no implications for the choice of comparator. Proposed change: Please be more specific on how to adequately “address” the different pain intensities and what the potential label implications are.	Accepted. A specification is added: ‘according to different pharmacological profiles and potency of a drug’ Different pharmacological treatment or application may have different extent of pain reduction. Insofar the pain intensity which is treatable should be mentioned.
164-165		Comment: Please provide examples of treatments that may have a distinct impact on the outcome of OA studies.	Not accepted. It was agreed by the CHMP to delete the examples.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
251-252		Comment: This is the only place in this guideline the term “clinically relevant difference” is mentioned. It is unclear if this is related to only AUC of pain intensity. To what extent clinical relevance of treatment effects versus statistically significant differences are of importance for approval. It is said in 248-249 that a responder criterion should support the relevance and the “robustness” of the treatment effect. Overall, the guideline remains vague on the meaning and the criteria for clinical relevance to be used for approval. Proposed change: As this may be a major source of controversy, more detail and clarification of the agencies' view of what a “clinical relevant” treatment effect in this indication is, and what level of “robustness” of data is expected, would be very important for the planning of development programs.	Accepted. A more detailed explanation is added. In the paragraph ‘Primary endpoints’ the term clinically relevant effect is used as a main value for adequate sample size calculation. The clinically relevant difference versus placebo might vary between different drugs, study populations and also with reference to study design and measurement time point. The delta has to be predefined and sufficiently justified on the base of valid data from appropriate published clinical trials. A recently published analysis for example calculated the best mean difference in pain reduction versus placebo (between group comparisons) for several drugs in patients with knee OA. This analysis is listed in the references. Most of the clinical studies with OA patients have measured symptom improvement at a defined time point and not over the course of treatment (AUC). In addition, the duration of AUC measurement differs between the studies. Therefore at present it is considered not possible to predefine for the AUC an evidence based consistent value for a clinically relevant difference versus placebo. However, as secondary endpoint this measure can give helpful data about the course of improvement. In addition, responder rates can give different aspects of efficacy. Thus these data can support the results obtained from the primary endpoint and additionally demonstrate the robustness of the data.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
296-298		Comment: Typographical error Proposed change: "However, further work is still needed on how changes measured in some of these biochemical markers of cartilage turnover correlate with OA disease progression."	Cleared.
299-301		Comment: This summary paragraph addresses only secondary endpoints. Proposed change (if any): "Overall, JSN measured by plain X ray, with appropriately standardized methodology, is an acceptable primary endpoint for assessment of structural damage. The alternative technologies for the evaluation of the severity of OA, e.g. chondroscopy, MRI, scintigraphy, ultrasonography or biochemical measurements (serum, urine, joint fluids) may be considered as secondary endpoints"	See below.
326-330		Comments: Can an acceptable minimum duration for a clinical trial to show meaningful changes in structure be added here?	Not accepted. I discourage to add an acceptable minimum duration for a clinical trial to show meaningful changes in structure.
362-363		Comments: Please provide guidance on how one might prove the surrogacy value of "structural changes" for "modification of structure"	Not accepted. I do not believe that we have to provide guidelines, for the future, to define what validation of imaging techniques should do.
362		Comments: Reference to incorrect section number Proposed change (if any): As stated in section 5.2	Accepted.
363-365		Comments: Necessity of joint replacement and time to need for surgery are likely to depend on local practice and resource availability Proposed change: "Clinical endpoints, such as long-term clinical evolution (pain and disability), the necessity of joint replacement, and time to the need for surgery (with stratification to address variations in local practice and resource availability) are preferable in the assessment of efficacy of such drugs".	Not accepted.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
369 - 371		Comments: Please clarify what is meant by “clinical improvement” in the context of modification of structure alone. Minimum criteria for clinical improvement should be specified.	Not accepted. I discourage to define minimum criteria for clinical improvement.
381-383		Comments: Sentence structure needs to be modified. Proposed change: "For all trials, concomitant treatments (drugs or interventions) that are likely to affect joint structure or symptoms should be excluded and rescue treatment (including physical therapy) ..."	Accepted.
391-392		Comments: The rationale for requesting 2 months follow-up for adverse events after discontinuing study medication is being questioned. This does not appear to be a common standard applicable to other indications. This requirement might be advisable for some substance classes but should not be a general requirement.	Accepted. The requirement is listed only for the respective substance classes and not as a general requirement.

2. SPECIFIC COMMENTS ON TEXT (Stakeholder No. 4)

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Executive Summary Line 42-56		Comments: EFPIA welcome the addition of an executive summary. However, to make it more useful it would be advisable to add the key points discussed further in the guideline document. Proposed change (if any): We suggest extending the scope of the summary to include key points of the guideline.	Not accepted. It was agreed in the CHMP to shorten the executive summary.
Line 43-44		Comments: The revised draft guideline defines OA as follows: “OA is a flaring degenerative arthropathy and a disorder which can potentially affect all synovial joints”. However, the definition provided doesn’t indicate that OA is a chronic and progressive disease, as is indicated later in the guideline. Additionally, it should be clarified if the disease should be called primary osteoarthritis, and if so, define this disease (e.g. osteoarthritis resulting from misalignment with or without	Partly accepted. The classification in primary and secondary is considered no longer usual as the outcome is the same. ‘Chronic’ is included. Detailed descriptions are mentioned in general information.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		obesity) and please specify the osteoarthritis sub segments you consider adequate. Proposed change (if any): We suggest re-phrasing as proposed “ OA is a flaring degenerative arthropathy and a <u>chronic progressive disease disorder</u> which can potentially affect all synovial joints.	
Lines 48-50 and 80-82		Comments: It is not clear from the sentence “Medicinal products with beneficial impact on structural progression of OA or disease modifying properties in erosive hand OA may be assessed in the future” if disease-modifying drugs are only being considered for erosive hand OA. Rewording the sentence is suggested to include any OA joint indication. The burden is on the sponsor to demonstrate efficacy. Proposed change (if any): “Medicinal products with beneficial impact on structural progression of OA or disease modifying properties in erosive hand OA <u>(including erosive hand OA)</u> may be assessed in the future.	Accepted.
Line 65		Comments: The guidance states: “OA is characterised by degeneration and regeneration of articular cartilage and bone.” We would prefer to argue on a more holistic description of the disease and are thus, proposing a new wording. Proposed change (if any): “OA is characterised by <u>unbalanced</u> degeneration and regeneration of articular cartilage and bone, <u>where the intrinsic repair mechanisms are insufficient.</u> ”	Accepted.
Line 67-68		Comments: The sentence “It has been suggested that radiologically diagnosed but asymptomatic OA is a precursor of symptomatic disease” is in our view, too restrictive and excludes other technique, such as MRI, which may be equally valid to determine this correlation. It is thus proposed a rewording of the	Not accepted as no publication is presented for the relation between MRI changes and future outcome. However, for a radiologically diagnosed population the correlation is assessed and proven.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		sentence. Proposed change (if any): “It has been suggested that radiologically diagnosed OA <u>diagnosed by structural changes (eg by X-rays or MRI)</u>, but asymptomatic OA is a precursor of symptomatic disease.”	
Line 73		Comments: It is well known clinically that these patients have a higher risk of OA pain vs. non-menisectomy patients. Proposed change (if any): “...degeneration <u>as well as a symptomatic disease..</u>”	Not accepted. The link between radiological progression and symptomatic disease is already mentioned above (line 68).
Line 76-79		Comments: This section provides an overview of the current pharmacological treatments available for the treatment of OA but does not mention opioid drugs, which are important pharmacological therapies used to treat OA pain. Proposed change (if any): We recommend including a reference to opioid drugs.	Partly accepted. Instead of paracetamol the more general term <u>analgesics</u> is listed.
Line 77-82		Comments: The treatment option described does not seem to reflect the European treatment guidelines (EULAR Recommendations 2003, Jordan et al, Ann Rheum Dis 2003; 62:1145-1155) for OA in which NSAID-therapies are recommended for chronic treatment of OA as well as acute treatment of flare. In this section the discussion on fast and slow onset of a drug seems to be mixed with that of acute and chronic treatment, i.e. fast onset only for acute flares and not for non-acute settings. SYSADOA is not a well-known terminology - please clarify what is meant by “symptom modifying slow acting drugs”. Proposed change (if any): Add that paracetamol, NSAIDs as well as symptom modifying slow acting drugs are also used in the chronic setting. Delete “slow acting drugs” or clarify the terminology.	Partly accepted. In lines 77-79 two different pharmacological effects are described: Drugs with the potency of quick symptom relief for acute conditions and symptomatic slow acting drugs with an onset of effect after 3-4 weeks which not appropriate to treat acute symptomatic conditions of OA. Both groups may be applied for long term treatment as OA is a chronic disease. NSAIDs as long as therapeutically necessary as short as possible. Also the EULAR in their recommendations uses SYSADOA: ‘symptomatic slow acting drugs for OA’ therefore it is considered a usual term. However the sentence can be shortened: ‘For non-acute treatment symptomatic slow acting drugs for OA (SYSADOA) are available.’

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 88-89		Comments: For registration of symptom modifying drugs, the guideline requires “demonstration of a favorable effect on symptoms with no clinically significant adverse effects on the structural changes of the disease.” There is much room for interpretation of “no clinically significant adverse effects on the structural changes of the disease”: what period of treatment time is required, and how to measure structural changes and their clinical significance (e.g., radiographs, MRI, pain and function scales, etc), particularly since OA is a progressive disease with great heterogeneity. No specific endpoint is listed in section 5.1, and section 5.2 primarily addresses limitations of the methods to assess structural damage, and may rather be intended to describe endpoints for disease modifying treatments. Would the design of studies targeting modification of structure (section 6.2.1) also apply for demonstration of the absence of potentially negative effects by symptomatic treatments? Proposed change (if any): We would appreciate more clarity addressing the above issues and specific guidance to what means are accepted to demonstrate absence of negative effects on structure. We recommend the analysis of structural changes be relative change on a per patient basis, in order to reflect individual patient variability in this dimension.	Accepted. The absence of negative effects on structure should be warranted within the safety long term data. It might be demonstrated by radiographic measure. A sentence in 7.3 is added. Whether this will be demonstrated by means of an analysis of structural changes showing the relative change on a per patient basis belongs to the applicants decision.
Line 92-94		Comments: The common separation of pain into mild, moderate, severe is referred to and it is said that the severity of pain is to be “addressed in the claimed indication”. This may suggest that if a certain number of patients of each category were included in trials, a claim for mild to severe pain may be granted. With this it is implied that in order to get a claim “severe pain” no additional indications such as cancer pain need to be studied, and that studying patients of different pain severity has no implications for the choice of comparator. Proposed change (if any):	Accepted. A specification is added: ‘according to different pharmacological profiles and potency of a drug’ Different pharmacological treatment or application may have different extent of pain reduction. Insofar the pain intensity which is treatable should be mentioned.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Please be more specific on how to adequately “address” the different pain intensities and what the potential label implications are. Please include a definition of the pain scale range for mild, moderate and severe	
Line 104		Comments: It is not clear if a structure modification only indication is possible for a Structure Modifying Osteoarthritis Drug. From Line 104 and on it appears possible. However, from Line 369 it appears that clinical benefit needs to be demonstrated for registration. Is this a discrepancy or does this mean at least an indirect effect on symptoms must be demonstrated? In the later case, then this should be clearly stated along with a goal for magnitude of benefit (vs. placebo) and recommendations of tools to measure the clinical benefit. Proposed change (if any): We would appreciate more clarity addressing the above issues and the related comments to Lines 369-371, and specific guidance in line with the regulators expectations for registration of this group of drugs.	Paragraph is revised.
Line 106-107		Comments: The sentence “This, however, has to be confirmed.” should be clarified. As the guideline acknowledges that there are structure-modifying drugs with <u>no direct</u> effect on symptoms, some of which may show their effect on symptoms only after a very long period; it may not be feasible to “confirm” their effect on symptoms within the time frame of a confirmatory Phase III study. If confirmation in a clinical trial setting were required, then it would imply the drug would fall in the category “structure modifying, symptom relieving drugs”. There should be a difference between disease modification and symptoms modification. Time to joint replacement (or time to virtual joint replacement) may be an adequate assessment to demonstrate the clinical relevance of structural modification, but it would not be a feasible endpoint in the earlier stages of OA.	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Proposed change (if any): Replace “confirmed” by “documented”	
Section 4 Line 129-130		The section 4 “patients characteristics and selection of patients” does not specify the list of secondary causes of OA that should be excluded. Some of these probably still should be and it would be of benefit to know the view of the CHMP/EWP on this. Clarification would be helpful. Comments: The guideline states «Compounds having demonstrated efficacy either at the hip or at the knee level will be registered for 'treatment of osteoarthritis of the knee and the hip'. This statement should be further clarified e.g. if studies are performed in knees only the claim could still be treatment of osteoarthritis of the knee and the hip? Proposed change (if any):	Not accepted. As agreed in the CHMP the list of exclusion criteria was deleted.
Line 134-136		Comments: “In order to obtain indication 'treatment of osteoarthritis' in general a compound should demonstrate efficacy at the level of the hands and at the level of the knee or the hip.” The guideline doesn't provide any indication on the number of confirmatory trials that would be needed to achieve a 'treatment of osteoarthritis' claim. For example, would one confirmatory trial in the knee or hip and one trial in the hands be sufficient? We would appreciate more clarity addressing these issues. Proposed change (if any):	Not accepted. The guideline referred to the respective ICH E9.
Line 134-136		Comments: It is not clear whether the target joint in one study can include both hip and knee OA patients. It is also not clear if “limit the target joint to a single site” means limiting to one joint type per study (i.e. hip only or knee only in a study), or whether it means limiting to one joint location e.g. one study can enroll both hip and knee patients but the index joint (joint to be evaluated throughout the study for a given patient) should be identified (e.g. right or the left hip or knee).	Partly accepted. The following sentence is added: ‘Both knee and hip OA patients may be enrolled in the same study however the subgroups have to be stratified.’ The target joint should be measured at a single site. If the patient has two painful knees the most painful of both should be assessed.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Additionally, please clarify if 'lower limbs' would be classed as a 'single site', i.e. would a study with patients of either OA of the knee or hip be viewed as a homogeneous patient group? Proposed change (if any): Addition of the sentence: <u>"Both knee and hip OA patients may be enrolled in the same study as long as the index joint is defined for each patient upon enrollment (e.g. hip patient: target joint = Right hip)."</u>	
Line 137-139		Comments: In more than one instance there is mention that study inclusion criteria should limit the target joint to a single site. It is not clear whether this is intended to be at the patient level or the study level ie is it necessary to have <u>separate studies</u> of knee OA and hip OA as the target joint to support generalisation of the results or is it possible to study both knee and hip OA <u>in the same study</u>? Proposed change (if any): "To improve the homogeneity of the patient groups studied, inclusion criteria should limit the target joint to a single sites. However, s Simultaneous assessment of other joints <u>in the same study with subgroup analysis of each joint</u> is always possible and such results might generate supportive evidence for «general» OA efficacy."	See above.
Line 141-142		Comments: The guideline states: "A narrower age range will increase group homogeneity, possibly at the expense of the generalisability of the data obtained." It is unclear if the CHMP/EWP will accept increased homogeneity at the expense of generalizability and vice versa and how this could translate in the label. Additionally, please provide clarification if this could be reflected in the label resulting in a more limited labelling and provide adequate guidance on the appropriate populations to be studied for the indication (of primary OA). Proposed change (if any): "A narrower age range will <u>may</u> increase group homogeneity,	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		possibly at the expense of the generalisability of the data obtained.”	
Line 143-144		Comments: The guideline mentions, “To be enrolled in a study, patients should have both symptomatic and structural changes of OA in the target joint”. This statement seems to be unnecessarily exclusionary. What about a drug that blocked progression (structural and symptomatic) in a population that was initially defined as having structural OA (as demonstrated via X-ray, or even some other methodology that clearly demonstrated structural joint abnormalities) but was essentially asymptomatic at baseline in the signal knee? For example, one could potentially define a patient population with unilateral symptomatic/structural OA, but which also has an increased probability of progression in the asymptomatic target joint. A drug that blocked progression in the initially asymptomatic target joint would surely be a useful drug; however, that population would be excluded by this statement. Similarly, what about a population that was symptomatic at baseline in their signal knee but had no structural OA (at least as defined via X-ray)? These patients probably have OA but it is just not yet apparent on an X-ray. A drug that improved symptoms in this population would also be useful but they would be excluded by this statement. Proposed change (if any): It is suggested to remove this statement.	Not accepted. In order to diagnose OA structural changes should be assessed. In order to justify the therapeutic benefit both symptoms and structure should be documented. Therefore the paragraph should not be deleted.
Line 144-146		Comments: The current focus on “pain related to use” does not acknowledge that pain at rest is an important symptom too. Proposed change (if any): “This will mean pain related to use <u>or pain at rest</u> with radiological evidence of OA, e.g. according to the radiographic criteria of Kellgren and Lawrence and to the diagnostic criteria of the American College of Rheumatology (ACR) or the European League against Rheumatism (EULAR).”	Not accepted. As pain at rest is usually minor than pain on movement and pain on movement is usually selected as primary endpoint this pain quality should be the key symptom to be enrolled.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 150		Comments: The CHMP/EWP considers an indirect visualization of cartilage by radiography as better standardized compared to direct visualization by MRI, even though MRI may be superior. Proposed change (if any): Given that this guideline will likely be in effect for a few years, there should be an acceptable approach described for validating MRI for use in the course of a specific study. This could lead eventually to general acceptance of the use of this imaging approach.	Not accepted. I do not believe that we have to provide guidelines, for the future, to define what validation of imaging techniques should do.
Line 160-161		Comments: The guideline mentions, “inclusion of a specific risk group in studies decreases the potential for generalisation of the results” but doesn’t mention the impact this might have on labeling/indication statements. Please clarify what percentage of high-risk patients included in studies will decrease the potential for generalisation of results and how this will impact the label. Proposed change (if any): We recommend including in this statement the impact of including specific risk groups in studies on labeling/indication statements.	Accepted. The following sentence should be added: ‘A patient population should be selected with respect to the characteristic and the potency of the drug. This patient population appropriate to be treated will then be reflected in the indication.’ For some medicinal products it might be advantageous to mainly assess patients after meniscectomy known for rapid structure damage in order to better show a positive influence on joint structure. If mainly those patients are assessed the indication is considered limited.
Line 164-165		Comments: No examples of treatments that may have a distinct impact on the outcome of OA studies are provided. Proposed change (if any): Please provide examples of treatments that may have a distinct impact on the outcome of OA studies.	Not accepted. It was agreed by the CHMP to include a more general wording and to delete the list of exclusion criteria.
Line 166-168		Comments: The guideline states “However, the required minimum severity of symptoms related to disease in the target joint at entry will depend on the primary outcome measure being assessed, the potential mode and kinetics of action of the drug, and the joint sites involved.” This sentence is quite complex and would	Not accepted. The examples follow within the next sentences of this paragraph. ‘For example, a higher baseline level...’

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		benefit from clarification and being more explicit. Proposed change (if any): Please provide clarification on this sentence and make it more explicit. Examples may be useful.	
Line 170-171		Comments: The guideline states for “studies of <u>acute</u> symptom modifying drugs, it is recommended to include patients with at least a moderate to severe extent of symptoms”. However, the same criteria should be applied to medicines studied <u>chronically</u> . Additionally, it should be specified whether the example provided in the bracket “e.g. pain at least 40 mm as measured on a 100 mm visual analogue scale (VAS)” relates to “resting pain” or “pain-on-movement” (POM). The reference here should be “resting pain” since POM gives usually higher values. Proposed change (if any): “For studies of acute -symptom modifying drugs, it is recommended to include patients with at least a moderate to severe extent of symptoms (e.g. <u>resting</u> pain of at least...”.	Not accepted. The example for acute symptom modifying drugs does not exclude a long-term treatment but describes the characteristic of drugs with a quick onset of symptom relieving effects as NSAIDs have. To detect sufficient reduction of pain a baseline pain the lower limit level should be 40mm/100mm scale. This has been shown to be appropriate. However, the extent of mean baseline pain that is measured will be probably above that value. At this the value for pain on movement will be higher than for pain at rest.
Line 175-178		Comments: In the sentence “For structure modifying drugs, it is recommended to include patients with Kellgren and Lawrence radiographic entry criteria of grades 2 or 3 (i.e., sufficient remaining interbone distance to permit detection of worsening/progression) or a certain pre-defined amount of joint space width (in mm).” no range of amount of joint space width is provided. Proposed change (if any): We recommend providing an example of joint space width that would be acceptable for studies in this patient population.	Not accepted. I discourage to enter into a discussion pertaining to an appropriate Joint Space Width, at inclusion.
Line 179		Comments: CHMP/EWP should clarify whether a high BMI stratum is necessary (sort of a forced distribution of entry criteria). Proposed change (if any): We recommend providing adequate guidance in this section and	Partly accepted. This paragraph is worded in subjunctive to show specific aspects which might have an impact on the outcome in order to have the opportunity to avoid bias. To do so belongs to the decision of the

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		clarification whether all parameters should be stratified at randomisation.	applicant. This paragraph is closely connected to long term studies and corresponding to structure modifying drugs. Insofar it should be connected with the paragraph before without a space.
Lines 183-189, and also 381-384		Comments: The guideline refers to the strict monotherapy setting with discontinuation of all prior analgesic. The guideline should also expand on criteria for add-on designs since this in many ways would mimic the clinical reality in OA. These sections, as currently written, would exclude the possibility of studying a new medicine as adjunctive therapy in combination with an existing medicine. Trials on top of SOC are not possible if this is maintained, would be an issue for trial which primarily effect structure with derived benefit on function and pain. If dual-mode of action compounds are used and the structural modification aspect is of primary interest, concomitant analgesics are permissible and beneficial effects on pain and function might be inferred via reduced analgesic consumption. The need for discontinuation of prior analgesic and anti-inflammatory medications including topical agents and steroid injections prior to initiation of the study is supported for monotherapy drugs. However, the sentence could be interpreted that these medications are not allowed during the entire study, which would be unethical, especially in a placebo-controlled setting. Therefore, it should be specified that these medications are allowed during the study but should be discontinued sufficiently long before assessment of symptomatic endpoints. Proposed change (if any): We recommend re-wording this section to make a distinction between the requirements for a new monotherapy medicine vs. an adjunctive medicine. Additionally, specific suggestions are: - should the paragraph starting “All symptom-oriented studies require ...” be kept, it is suggested to replace “All” by	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		“Most”. - the following sentence “When such medication is required during the study, they should be discontinued before each assessment of symptomatic endpoints” should be added after “..prior to initiating treatment, with the test drug in order to permit an evaluation of unmodified pain severity”.	
Section 5: Methods to assess efficacy Line 199		Comments: “radiographic features” would imply that X-ray only is accepted. Other methods of determining structural features e.g. MRI should also be considered. Proposed change (if any): Replace “radiographic features” by “structural changes”	Not accepted. The most important OA characteristics are listed which are radiographically detectable. If MRI is used this should to be separately described and justified.
Lines 199-200, 268-290, 299-301		Comments: Hip OA inclusion based solely on radiograph may be problematic in research (and treatment). The guideline states that MRI isn’t good for evaluation, and to concentrate on the plain radiograph. Hip X-rays are not always particularly clear cut, or taken or read well, many times secondary to patient placement. Often “hip OA” patients in both practice and studies are not, in fact, pure hip OA. Another problem with hip pain (suspected OA) is a concurrence of diagnoses, ie, a radiograph that appears positive, and may be for early (mild) or moderate OA, but there is also a secondary diagnosis, which is not looked at if the X-ray is possibly positive (piriformis syndrome, myofascial injury, joint effusions, etc). On lines 199 and 200 it explains what to look for in a hip radiograph- but sometimes more investigations with MRI etc is needed for full evaluation to diagnose OA (the radiograph looks for narrowed or loss of joint space, presence of osteophytes, subchondral sclerosis and subchondral bone cysts). Other tools that may be useful include Ultra-Sound and tomography as well as arteriography/venography.	Not accepted. There is already an exemplary listing of alternative technologies for an evaluation. lines 299-301. To refer to further more unusual methods of evaluation is considered not necessary. If such an additional technology for a correct diagnosis is necessary this belongs to the decision of the physician. To do so is not excluded by the paragraph.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Proposed change (if any): in line 299 “Overall, the alternative technologies for the evaluation of the severity of OA, e.g. chondroscopy, MRI, scintigraphy, ultrasonography, <u>tomography, arteriography/venography</u> or biochemical measurements (serum, urine, joint fluids) are helpful in diagnosing OA, particularly OA of the hip , and may be considered as secondary endpoints.	
Line 200		Comments: Reference to subchondral bone lucencies is unclear. Does this mean bone marrow oedema or that subchondral bone lucencies are a separate (and worthy) manifestation. Please clarify.	Accepted. This point should be deleted as only the most important characteristics should be listed.
Line 201-202		Comments: “In OA of the knee outcome measures for both symptoms and structures are better validated for medial tibiofemoral disease than for lateral or patellofemoral disease”. The purpose of this sentence is unclear and could have potential impact on enrollment criteria. For example, what is the impact of including patients with lateral and patellofemoral disease? Patellofemoral OA (chondromalacia) represents a distinct medical need and therefore, investigations in this setting should be encouraged. The sentence should be re-written to accommodate for future validation of measures for lateral and patellofemoral disease. Proposed change (if any): Provide details addressing the above issues and include a reference for this sentence.	Not accepted. In most clinical trials assessing structure-modifying OA drugs patients with medial tibiofemoral OA were selected in whom joint space narrowing (JSN) was measured only in the medial compartment. Moreover, Boegard, showed that in contrast to the patellofemoral joint the tibiofemoral joint allows a detection of radiographic progression. Insofar although this is a statement that was already included in the 1997 version of the NfG OA and despite Altman (1986) recommended the patellofemoral OA (chondromalacia) the information that the medial tibiofemoral disease is better validated is still valid. Boegard TL , (Rudling O , Petersson IF, Jonsson K . Joint space width of the tibiofemoral and the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2 year follow-up. <i>Osteoarthritis Cartilage</i> 2003; 11 :370–6)
Line 203		Comments: It is recommended to mention that reduction or slowing of joint space narrowing (JSN) at spine or reduction in disk prolapses or “bulging disk” would be clinical meaningful endpoints to be considered.	Not accepted. These endpoints are not validated.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 215-221		Comments: The section on primary endpoint could benefit from a clearer message around the choice of primary and secondary endpoints, e.g. consequences for the indication wording choosing functional disability as a primary endpoint or not.	Not accepted. As function is a primary endpoint and clinically relevant and statistically significant improvement is shown this can be included into the wording of the indication. No additional explanation is considered necessary.
Line 217-240		Comments: We accept that functional disability is important but prefer to see this as a key secondary endpoint. The requirement for co-primary endpoints makes the hurdle for demonstrating efficacy higher for bringing forward new analgesia products. Proposed change (if any): Remove functional disability as co-primary endpoint but replace as important secondary endpoint for a label claim of signs and symptoms, or allow for efficacy assessment of only pain or only function which would result in a restricted label claim.	This issue should be discussed. Proposal: “Functional disability <u>can be</u> an important additional primary endpoint for symptom modifying drugs and should preferably be included as co-primary endpoint. Studies should be powered to demonstrate a clinically relevant and statistically significant effect on pain and <u>optional</u> on functional disability.”
Lines 223-226		Comments: As in the past industry was confronted with refusal of studies that were statistically significant however not judged clinically relevant by regulators, it would be useful to indicate what would be the minimum accepted clinical effect on pain (e.g., 7 mm for resting pain or 10 mm for pain-on-movement, for 100 mm VAS). Proposed change (if any): Add at the end of the paragraph “ An effect size such as 7 mm for resting pain or 10 mm for POM (100 mm VAS or equivalent on Likert scale) could be acceptable as clinical relevant. It is however recommended that sponsors seek scientific advice.	Not accepted. A respective sentence is included. However to justify the value for the delta for the individual clinical study belongs to the applicant on the base of data from according published clinical studies.
Line 224		Comments: The present wording is not scientifically correct and up to date. Likert scale stands for a categorical scale and is normally a 5, 7 or 11-point scale, but the number of scale point could also be other numbers. The VAS allows for continuous variable. The NRS 0-10 Likert scale is a categorical scale. These have a high correlation in results, and are equally preferred. Proposed change (if any): The Likert scale is a Verbal Rating Scale (<u>5 or 7-point scale</u>) <u>based on verbal pain intensity</u>	Partly accepted. NRS and the more applicable description for the Likert scale are included. However, a 5 point scale is considered to small.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		<u>descriptors.</u>	
Lines 229-234 and 240-241		Comments: For both these sections, we recommend providing examples of the tools considered to be validated today and also indicating that sponsors can recommend other tools they consider validated, which would reflect changes occurring in the future.	Not accepted. It was agreed by the CHMP not to include specific questionnaires. Main point is that the questionnaire is validated.
Line 251-252		Comments: In line 251 the parentheses after ‘course of pain intensity’ indicates that ‘area under the curve (AUC)’ is an appropriate method of assessing time course of the secondary endpoints. There is confounding of AUC with duration of time in the study, and also AUC may be hard to (clinically) interpret. Since there are other ways to define the course of pain intensity, we recommend placing an “e.g.” in front of the area under the curve example. An alternative may be to assess efficacy endpoints using a range of landmark analyses, with appropriate consideration for discontinued patients. In this way the time-course may be evaluated in a consistent way (compared to the primary time-point), and aspects such as onset and maintenance of benefit can be assessed. This is the only place in this guideline the term “clinically relevant difference” is mentioned. It is unclear if this is related to only AUC of pain intensity. To what extent clinical relevance of treatment effects versus statistically significant differences are of importance for approval. It is said in 248-249 that a responder criterion should support the relevance and the “robustness” of the treatment effect. Overall, the guideline remains vague on the meaning and the criteria for clinical relevance to be used for approval. As this may be a major source of controversy, more detail and clarification of the agencies' view of what a “clinical relevant” treatment effect in this indication is, and what level of “robustness” of data is expected, would be very important for the planning of development programs. Proposed change (if any):	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		“As secondary endpoint, the course of pain intensity (<u>e.g.</u> area under the curve (AUC)) might be used. However, the clinically relevant difference should be justified. <u>An alternative may be to assess efficacy endpoints using a range of land mark analyses, with appropriate consideration for discontinued patients. In this way the time-course may be evaluated in a consistent way (compared to the primary time-point), and aspects such as onset and maintenance of benefit can be assessed.</u> ”	
Line 253		Comments: For ease of reading, an introductory sentence before the bullet points on other secondary endpoints should be introduced. It is understood that these are examples only and the sponsor may choose from the list. Proposed change (if any): Add before line 253 “<u>Other secondary endpoints may include:</u>”	Accepted.
Lines 253-266		Comments: Functional disability (if not assessed as co-primary endpoint) – in this list it gives the option of being a secondary endpoint. Is this a discrepancy with the earlier section which indicates functional disability should be a co-primary? It would be preferable to clarify if functional disability can be a key secondary endpoint or to be able to assess only pain or only function with a restricted label claim for registration. This should be aligned with lines 217-221.	See proposal line 217-221 (Function as optional co-primary endpoint.)
Line 263		Comments: It would be helpful to clarify the following bullet: “Total score of a questionnaire”. Proposed change (if any): Provide some examples of relevant questionnaires.	Not accepted. Not to list examples was agreed by the CHMP. Explanation: Total scale, not only subscales as e.g. ‘subscale pain’.
Line 266 and following		Comments: Sections 5.1 and 5.2 refer either to endpoints or to assessment of structural damage, but inflammation is not mentioned. However	<u>Biochemical markers as secondary endpoints</u> There are relatively few comments on this particular issue. There

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		joint inflammatory mediators play an important role in mediating pain in OA. A decrease of inflammation markers could support the improvement of symptoms and this type of markers would fall first in the symptoms modifying drugs. This type of marker could be tertiary endpoints or exploratory. Proposed change (if any): Add a header “Tertiary endpoints” and below add: Joint inflammatory mediators play an important role in mediating pain in OA. A decrease of inflammation markers could support the improvement of symptoms and this type of markers would fall first in the symptoms modifying drugs. This type of marker could be additional endpoints.	are mostly semantic. Action: I have no objection in including, on line 320 “... <u>measurement of progression of OA as well as for the prediction or evaluation of response to pharmacological intervention with compound...</u>”. I am not in favour of including inflammation markers in the structure-modifying section.
Line 267-301		Comments: Please make it clear what should be the primary endpoint for assessing structural changes. The section only deals with various surrogate endpoints, but it is not clear if any of those are acceptable as primary endpoints. Section 6.2.1 recommends the use of clinical endpoints; however that is not completely clear from the discussion in Section 5.2. Proposed change (if any): Clearly state what is currently acceptable as primary endpoint and what could be acceptable in the future given further validation of imaging techniques.	See below comments to MRI
Lines 286, 287, 289		Comments: Volume is only one of several quantitative morphological outcomes, other being cartilage thickness, denuded bone area, etc. Proposed change (if any): “Volume” to be replaced by “MRI-based quantitative cartilage morphology”; “structural changes quantified by MRI”	See below comments to MRI
Lines 289 – 290		Comments: It is expected that if the correlation between clinical relevance and MRI measurements were confirmed in large longitudinal studies, these studies would support the use of MRI measures as primary endpoints in Phase III.	<u>Use of MRI as primary endpoint to assess the structure-modifying effect</u> Several Stakeholders mention this possibility. However, in my opinion, no demonstration of the surrogacy value of MRI has been, unequivocally obtained. I do not believe that it is our role,

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Proposed change (if any): “As with other structural changes, Hence, larger longitudinal studies are still needed to further confirm the clinical relevance of MRI measured cartilage volume in OA disease progression. <u>Such studies might support the use of these measures as primary endpoint in studies of structure-modifying OA drugs</u>”	in such a guidelines document, to clarify what the requirements for accepting the surrogacy of MRI are. This has been previously documented in EMEA publications. There are, however, some interesting suggestions in the Stakeholders’ comments. Action: On line 314, I suggest to delete “measured cartilage volume”. The sentence should read “... <u>relevant MRI-based quantitative cartilage morphology in OA disease progression</u>”. Similarly, on line 311, I suggest to delete “volume and thickness” and on line 312, to delete “volume”. If needed, we could add, at the end of line 315 “<u>MRI measures may be an acceptable primary endpoint when there clinical relevant has been shown</u>”.
Line 291-203		Comments: The word “enzyme” is redundant and considered restrictive if “biological markers” are referred to. It is proposed to refer to enzymes as an example of possible biological markers. Additionally, this paragraph should also cover the fact that cartilage regeneration, not only degradation, can occur during OA or following treatment with a drug. Proposed Changes (if any): During progression of OA many biological markers will be released in synovial fluid, blood and urine, reflecting either degradation or synthesis of cartilage, bone or synovium (enzymes, matrix fragments, growth factors, etc).	I agree to replace line 317-318 by “... <u>tissue and bone. During progression of OA, many biological markers will be released in synovial fluid, blood and urine, reflecting either degradation or synthesis of cartilage, bone or synovium (enzymes, matrix fragments, growth factors, etc...).</u> <u>There has been progress...</u>”
Line 294-295		Comments: Markers could act as predictors of progression but also of response Proposed change (if any): “There has been progress into the use of some of these markers for the prediction or measurement of progression of OA as well as for the <u>prediction or</u> evaluation of response to pharmacological intervention with compounds of potential	See above comments to biomarkers.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		structure modifying activity.”	
Line 297		Comments: The word « now » should be exchanged with « how ». Biomarkers reflect not only changes in cartilage turnover therefore “cartilage turnover” should be deleted. Proposed change (if any): “However, further work is still needed on how changes measured in some of these biochemical markers of cartilage turnover correlate with OA disease progression or are modulated by the compound...”	Accepted.
Line 299-301		Comment: This summary paragraph addresses only secondary endpoints. Proposed change (if any): “Overall, <u>JSN measured by plain X-ray, with appropriately standardized methodology, is an acceptable primary endpoint for assessment of structural damage. MRI measures may be an acceptable primary endpoint when their clinical relevance has been shown.</u> The alternative technologies for the evaluation of the severity of OA, e.g. chondroscopy, MRI, scintigraphy, ultrasonography or biochemical measurements (serum, urine, joint fluids) may be considered as secondary endpoints”.	Accepted. We agree to change line 324-326 into “<u>Overall JSN measured by plain x-ray, with appropriately standardised methodology, is an acceptable primary endpoint for assessment of structural damage. The alternative technologies for the evaluation of the severity of OA, e.g. chondroscopy, MRI, scintigraphy, ultrasonography or biochemical measurements (serum, urine, joint fluids) may be considered as secondary endpoints</u>”.
Section 6: Strategy and design of clinical trials Line 312-330		Comment: It should be explained how study designs should look like addressing both symptoms and structure. Maybe alternatively separate trials are necessary? Comments: No mention of biomarkers such as CTX II as endpoints. Given the ongoing IMI initiative, the role of biomarkers should be encouraged and clarified.	Not accepted. C-terminal crosslinking telopeptide of type II collagen may be assessed as additional value. There is no exclusion. The heading of the secondary endpoint is: ‘Other secondary endpoints may include:’
Line 310-311		Comments: The guideline states, “If the medicinal product is for intra-articular administration the residence time in the joint and the systemic availability of the active substance should be	Partly accepted. As this procedure may burden the patients these data may be obtained only for some medicinal products as e.g.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		investigated”. PK requirement for intra-articular compounds requiring profile of retention in joint should be further elucidated, especially as repeated arthrocenteses are considered burdensome to patients. The modelling of data from a minimal number of arthrocenteses (clinical development) and the applicability of pre-clinical data to aid in such modelling could be considered or other exploratory methods. In addition, please clarify which metrics to use for the residence time in the joint; e.g. plasma synovial fluid	Corticosteroids. In that way respective information is added.
Line 317-318		Comments: The guideline states that “Medicinal products to be used in (fixed) combinations need appropriate studies to find the best dose regimen for the intended combination” without defining the word “appropriate”. Proposed change (if any): Clarify acceptable designs for combination or adjunctive therapies. Ideally, PK/PD can play a role to elucidate this.	Not accepted. Proof of positive contribution qualitatively and quantitatively and appropriateness of pharmacokinetic profiles together are needed. Add-on design is already included.
Lines 337-341		Comments: In more than one instance there is mention that study inclusion criteria should limit the target joint to a single site (see also comment to Lines 137-138). It is not clear whether this is intended to be at the patient level or the study level ie is it necessary to have <u>separate studies</u> of knee OA and hip OA as the target joint to support generalization of the results or is it possible to study both knee and hip OA <u>in the same study</u> ? Proposed change (if any): “Because of the heterogeneity of OA, limiting the number of different joints investigated can limit the potential for generalisation of the results. In each trial one joint, preferably the hip or the knee, should be selected as the target joint, although simultaneous assessment of further joints <u>in the same study with subgroup analysis of each joint</u> is possible. The primary analysis population should be defined according to the intention to treat principle. The design and the duration of these studies may differ according to the properties of the drug.”	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 343		Comments: The guidance states, “Modification of symptoms: Studies should have a randomised, double blind, parallel group design.” The possibility of using a crossover design should also be considered in view of the increased statistical power (e.g. this could reduce the number of patients). Proposed change: “Studies should have <u>ideally</u> a randomised, double blind, parallel group design”	Not accepted. Studies may have an add-on design, a dose escalation phase or a dose free interval but a crossover design is considered not appropriate because of differences in the course of the disease and an impact of the phase I medication on the outcome of phase II.
Line 344-345		Comment: “Efficacy of products claiming improvement in symptoms is generally established by means of placebo controlled trials with an established active comparator as a relative control.”An active control may not always be appropriate. For example, in situations where a new medicine is studied in a treatment resistant population. Proposed change: “Efficacy of products claiming improvement in symptoms is generally established by means of placebo controlled trials, with an established active comparator as a relative control <u>unless the treatment is for patients that have no other treatment options</u>”	Partly accepted. The wording ‘<u>normally</u>’ and ‘<u>appropriate</u>’ is included. The situation where a treatment resistant population is assessed is considered rather applicable in RA patients. However, it might be the situation that for a new therapeutic concept an appropriate active comparator is not available.
Line 346-350		Comments: Please clarify what is meant by “claiming acute symptom relief” in relation to claiming “symptom relief” (e.g. refer to Cox2s in OA)? Terminology of acute symptom relief and symptom modifying slow acting drugs is confusing and consequently what actual duration of studies are being suggested is unclear. The interpretation from AZ is the product for acute symptom relief are also used for chronic treatment, e.g. NSAIDs, weak opioids and that the majority of products used in OA pain would then fall into the acute category. Proposed change (if any): Propose to change the wording to “Maintenance of improvement should be evaluated after at least 3 months for medicinal	Accepted.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		products claiming symptom relief. For symptom modifying slow acting drugs for which the effect is building up slowly , an evaluation after at least 6 to 12 months may be considered appropriate”	
Line 351-352		Comments: The guideline states that “Although efficacy [...] may be demonstrated in trials with shorter duration, long-term maintenance of the effect should be shown” but the duration is not specified. In lines 347 to 350 an indication is given of duration of efficacy trials depending on type of drug. It is assumed that this covers the ‘long-term maintenance’ in line 352, but this is not clear. Please clarify. Proposed change (if any): Please specify what long term maintenance corresponds to (at least 3 months for compounds claiming acute symptom relief; 6 to 12 months for symptom modifying slow acting drugs)	Accepted. For clarifying the sentence a description more in detail with an example is added: <u>In some cases depending on the claimed treatment duration and the type of drug (e.g. topical NSAIDs)</u> efficacy (change in symptom intensity) may be demonstrated in trials with shorter duration, however <u>generally</u> long-term maintenance of the effect should be shown.
Line 353-354		Comments: The guideline states “a double-blind placebo-controlled phase may be followed by a long-term double-blind phase with the remaining active comparator only [...]”. However, the guidance doesn’t define the outcome that would be considered acceptable for a long-term trial with an active comparator. For example, would non-inferiority be the expected standard? Proposed change (if any): Clearly define the outcome that would be considered acceptable for a long-term trial with an active comparator.	Accepted. The sentence is reworded and describes the long-term double-blind phase more in detail.
Lines 355-357		Comments: The guideline proposes a randomised withdrawal period for symptom modifying slow acting drugs. However it is implied that this “period” be part of the main study design: in a long-term follow-on study. The guideline notes it might be possible to use a randomized withdrawal period but is such a sub-trial in the longer-term trials <u>required</u> or is this a nice to have? Proposed change (if any): “For symptom modifying slow acting drugs it might be possible	Partly accepted. A randomized withdrawal period is an option of long term treatment design. If maintenance of efficacy within a drug free period or an interval dose recommendation is claimed this is considered an option for a study design. For a better understanding, the sentence is reworded.

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		to show that the beneficial effect is sustained long-term by means of a withdrawal period in which actively treated patients and responders to treatment are randomised to discontinue or continue treatment. <u>Generation of this type of data is not a requirement for approval but would be helpful information for the prescriber and patient</u> ”.	
Line 358		Comment: The sentence “To establish that a symptom modifying drug does not have deleterious effects on the joint, structural changes should be monitored for at least one year.” is contradictory to the earlier statement that for medicinal products claiming acute symptom relief three months should be sufficient. Proposed change (if any): Please distinguish between the different drug types.	Accepted. Respective sentences are included. Also in paragraph 7.3 an addition is included.
Line 358-359		Comments: It is assumed that this sentence means that structural changes should be monitored for a total period of 1 year (i.e. for 1 year after the start of the treatment period), but this is not totally clear. Please clarify? We also suggest that for drugs that have not generated prior signals of a potential deleterious effect, this information may be gathered post-treatment. Proposed change (if any): “To establish that a symptom-modifying drug does not have deleterious effects on the joint, structural changes should be monitored for at least one year <u>after the first start of treatment</u>.”	Accepted.
Line 362-363		Comments: Please provide guidance on how one might prove the surrogacy value of “structural changes” for “modification of structure” Any example/proposal to provide?? Comments: Reference to incorrect section number Proposed change (if any): As stated in section H.2–<u>5.2</u>	see below comments line 369-371

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Line 363-365		Comments: Necessity of joint replacement and time to need for surgery are likely to depend on local practice and resource availability. Proposed change (if any): “Clinical endpoints, such as <u>long-term clinical evolution (pain and disability)</u>, the necessity of joint replacement, <u>and</u> time to the need for surgery and long-term clinical evolution (pain and disability) are preferable in the assessment of efficacy of such drugs”. <u>Influence of local practice should also be addressed.</u>	see below comments line 369-371
Lines 369-371		Comments: The guideline distinguishes structure-modifying, symptom-relieving drugs from structure-modifying drugs with no direct effect on symptoms. Only for the former, a symptom-modifying effect should be required for registration purposes. However, it is not clear if a structure modification only indication is possible: from Line 104 section it appears possible but, from lines 369-371 it appears clinical benefit needs to be demonstrated. A paragraph pointing out the differences for the two classes should be added. With a structure-modifying, only indirectly symptom-modifying drug, the time to symptomatic improvement may only become evident after 10-20 years. As the clinical relevance of structural change such as JSN is acknowledged, there should be no need to provide these data at time of registration. Please also discuss the acceptability of a conditional MAA approval in case this requirement is kept. Otherwise the distinction made by the guideline is irrelevant. We suggest the following rewording:"in the absence of any clinical improvement-benefit vs. placebo at the time...." It would also be helpful if there was some additional clarity regarding the endpoints (pain/function...etc) and the magnitude of change that would define a clinical benefit vs. placebo, i.e. is statistically different from placebo sufficient or is there some absolute % difference that we will need to shoot for with respect to the clinical benefit of a disease modifying drug? Is there any	<u>Requirement to concomitantly show an effect on symptoms and structure</u> This point has been raised by various Stakeholders and was already extensively discussed in the past. Whereas I still do believe that it will be hard to obtain registration for a drug having only shown its ability to prevent or reduce Joint Space Narrowing, without any effect on clinical symptom, I think that the suggestion made by EULAR is reasonable. Action: Line 404-403 “the benefit...authorisation” should be deleted and replaced by “<u>a trend in improvement of symptoms and/or a correlation between structural outcome and pain and function evolution will support the surrogacy value of x-ray changes</u>”. This would be also in accordance with the requirements from Stakeholder 2 and Stakeholder 4. I do not believe that we should accept the proposal to define a threshold of efficacy for clinical changes. Mentioning, as we did that it should be “<u>of a clinically relevant magnitude</u>” seems sufficient to me. However, as requested by Stakeholder 4, we could include, on line 396 “<u>time to the need for virtual or actual surgery</u>”. Similarly, in accordance with the suggestion of Stakeholders 4 and 5, this paragraph (line 403) could end by a sentence: “<u>The</u>

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		opportunity to get approval for a compound that produced robust evidence of structural improvement in 2-year phase III trials, but only a trend for (or statistically significant) symptomatic benefit vs. placebo, if the expectation was that additional clinical benefit would be defined in phase IV trials (e.g. time to actual or virtual joint replacement)? Proposed change (if any): The sentence “The benefit of a radiographic improvement... authorisation” should either be deleted or changed to read “The benefit of a radiographic improvement in the absence of any clinical <u>benefit improvement</u> at the time of assessment (i.e. at least 2 years) is difficult to predict and <u>should be confirmed post-approval</u> is insufficient for marketing authorisation.	<u>clinical benefit of a radiographic improvement, if not demonstrated during the trial, should be confirmed post-approval</u>”. I suggest also to approve the comments suggesting to change, on line 115, the word “confirmed” by the word “<u>documented</u>”.
Lines 381-383		Comments: Sentence structure needs to be modified. Proposed change: "For all trials, concomitant treatments (drugs or interventions) that are likely to affect joint structure or symptoms changes assessment should be excluded prior to assessment of symptomatic endpoints and rescue treatment (including physical therapy) ..."	Accepted.
Section 7: Clinical safety evaluation Lines 391-392		Comments: The guideline proposes that adverse events that occur after drug discontinuation be evaluated and documented for at least 2 months post-study. The general standard for follow up is 5 half-lives or 30 days, whichever is longer. We would like to understand the rationale for 60 days follow-up, especially with typical 18-24 hour half-life medications? Is there a scientific rationale for anticipating delayed toxicities that appear during wash out? This may be relevant on a case-by-case basis, but we do not understand this requirement as a routine matter. Proposed change (if any):	Partly accepted. This sentence refers to potentially occurring of CV effects after stopping treatment with coxibs. A more generalised wording is used (see above comments by stakeholder 1)

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		Adverse effects that occur after drug discontinuation should be evaluated and documented for at least <u>5 half-lives or 30 days, 2 month</u> post study.	

2. SPECIFIC COMMENTS ON TEXT (Stakeholder No. 6)

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
lines 45 to 49 and lines 77 to 80		Comments: We are not in total agreement with the differences in the indication of the different symptomatic drugs. We would be more in favor to state that, until now, concerning the use of pharmacological systemic treatments, two categories have been proposed mainly based on their onset of action e.g. quick acting (e.g. within 1 to 2 weeks) and slow acting (e.g. within 2 to 3 months). The latter have been termed SYSADOA for Symptomatic slow acting drugs for OA. Obviously, both categories are used in context of either an acute flare of the disease or in the case of a chronic painful condition.	Accepted.
Patients characteristics and selection of patients, lines 127 to 136		Comments: C1 Labeling with regard to the evaluated joints We do agree with the current proposals. Proposed change: However, could you please confirm the following sentence (e.g. line 135-136): <i>"In order to obtain indication 'treatment of osteoarthritis' in general a compound should demonstrate efficacy at the level of the hands and at the level of the knee or the hip"</i> ? Could you please confirm that the applicants will not have to evaluate spinal OA to get this labeling? Moreover, concerning spinal OA, we consider that such an indication should be on the "research agenda" and at least, should differentiate lumbar and cervical spinal OA.	Not accepted. We admit that spinal OA is a problematic indication. Therefore the applicant is requested to demonstrate the validity of the endpoints chosen and their clinical relevance for that specific indication (see section 5 Methods to assess efficacy) to claim that indication.
line 143-146		Comments:	Not accepted. Inclusion criteria e.g. patient characteristics and

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		C2 Characteristics of the patients: In order to facilitate the design of trials, we would propose the following sub-sections concerning inclusion criteria (which are different from the outcome measures). Proposed change: C.2.1. Definition of the disease yes/no. This sub-section should be focused on the classification criteria.	their documentation as well as outcome measure and their measurement instruments at entry are considered adequately listed.
lines 166-174 and lines 175-178		C.2.2. Definition of an active disease. This sub-section should describe the following: - the minimum level of symptoms (<i>e.g.</i> pain function) at entry with the description of the use or not of a flare design. - the presence/requirement or not of objective signs of inflammation such as hyarthrodial effusion evaluated at physical examination/ultra-sonography, MRI sub-chondral bone edema. C.2.3. Definition of a severe disease. This sub-section should refer to the structural damage. For this purpose, we recommend to refer to the Kellgren and Lawrence scoring system for the evaluation of symptomatic trials and to a threshold of plain X-rays JSN for structural modifying trials.	Partly accepted. The definition of the severity of disease is of importance when the characteristics of the patients are documented. Therefore we included 'Objective signs of inflammation such as hyarthrodial effusion may be additionally documented.' into according section (line 152). The proposed structure is considered more to generally determine OA definitions. At present the order serves to give information to all issues which are needed when planning the study. Here a detailed description of the primary outcome measure is of importance. The use of the Kellgren and Lawrence scoring system for the evaluation of symptomatic trials is mentioned in line 143-149) A threshold of plain X-rays JSN for structural modifying trials is mentioned in lines 175-178.
lines 157-161 and 152-156		C.2.4. Definition of a potential severe disease. This section should focus on the predisposing factors of subsequent structural progression (<i>e.g.</i> lines 157-161). C.2.5. Other characteristics. We do agree that a minimum level of information should be given concerning the demographics and the concomitant and/or previous therapies (<i>e.g.</i> lines 152-156).	Predisposing factors are important issues in research. Specific patient populations with altered or severe disease are considered only an option for selection.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
Section Methods to assess efficacy, lines 209-212		Comments: You are proposing 3 categories of claims: “From a regulatory point of view the following potential claims should be distinguished: 1. improvement of symptoms such as pain, 2. improvement of functional disability, 3. slowing or prevention of structural damage.” We have the following comments: 1. We would like to combine 1 and 2 in the following: “Improvement of symptoms such as pain and function (impairment in daily activities)”. 2. We would like to propose as the second one your current #3 (<i>e.g.</i> slowing or preventing structural damage). 3. We would like to propose as the third one: “prevention of functional disability”. This latter has never been evaluated but should be very clinically relevant. 4. We are wondering whether you could consider the following: “preventing requirement for total articular replacement” or you could consider our third proposal “prevention of functional disability” in case a trial is able to show a difference in such outcome (<i>e.g.</i> rate of total articular replacement). From a general point of view, we would be strongly in favor of adding a specific section concerning this potential outcome measure (<i>e.g.</i> requirement to total articular replacement or indication to total articular replacement).	To merge pain and function is for discussion. As <u>prevention</u> of functional disability has never been evaluated it is considered until now not appropriate as a claim which can be measured in licensing trials. Thus, at present this claim should not be included. However, it is agreed that in general this issue is very clinically relevant. 4. I am not particularly in favour of including a specific indication “preventing requirement for total articular replacement”, since this reflect only one of the potential clinical endpoints to be considered and since no studies ever used this as a primary endpoint.
Section medicinal products to improve symptoms, line 226		Comments: We have a detailed comment concerning the reference period (<i>e.g.</i> line 226). We propose more flexibility (<i>e.g.</i> the past 24 or 48 hours).	Accepted.
Section Assessment of		Comments: We do agree with this section. However, as an introduction to this section (<i>e.g.</i> line 268), we suggest that one recalls the main radiological features of	Not accepted.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
structural damage, line 268		OA (<i>e.g.</i> joint space narrowing, osteophytes, cysts, geodes) and to recall that JSN should still be the radiological variable of interest.	This issue is considered adequately listed.
Section G.2.1. study design, line 353		Comments: G.1. Duration of the placebo period of a symptomatic trial We would be in favor of stating matters as follows: “time point appropriate to show the maximum effect over placebo” and by providing some examples such as 12 weeks in case of a quick acting symptomatic drug, 12 to 24 weeks in case of a slow acting symptomatic drug, more than 24 weeks in case of a potential structural modifying drug. G.2. Duration of the long-term extension of the placebo symptomatic trial We do agree that such extension is very useful. However, we feel that the line 353 referring to the long-term double blind period is unclear. We would prefer to state that the evaluation of the sustainability of the symptomatic effect should be done in a study (which can be the extension of the placebo trial) of at least 6 to 12 months duration in which an active conventional comparator is included.	Partly accepted. A better description of the quick acting symptomatic drugs is included. The time point for maximum effect over placebo should be applied from the literature for the according substances as there may be distinct differences in their pharmacodynamic mode of action. Thus, general examples cannot be given. However, the duration of clinical study to show maintenance of effect is listed. Accepted. One option could be to re-randomise and switch the placebo patients to the active comparator or the test product in order to get controlled efficacy long term data. The other possibility is an open label extension phase. The following sentence in brackets is included: (<i>e.g.</i> re-randomise and switch the patients from the placebo group to the active comparator or the test group)
line 356 to 357		Comments: Moreover, we have some difficulties to understand the interest in a discontinuation trial to demonstrate such a sustainability effect (<i>e.g.</i> line 356 to 357). We can understand a discontinuation trial in order to demonstrate a symptomatic effect of a potential structural modifying drug which failed to demonstrate a relevant symptomatic effect in the “conventional” designs. However, we do not understand such discontinuating design to demonstrate the sustainability of the symptomatic effect.	Accepted, the sentence is reworded. The evaluation of the number of symptomatic flares is additionally included. Meant is that a SYSADOA might show a lasting effect over <i>e.g.</i> 3 or 6 month although the treatment is finished. This can be shown by randomisation to discontinue or continue treatment.

Line No	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
		We would like to emphasize that the concept of “sustainability” is different from that of “remanent” or “carry-over” for which a discontinuating trial might be also of interest. We would be in favor to evaluate such long-term (<i>e.g.</i> few months) comparative trial using a conventional approach (<i>e.g.</i> AUC) but also by evaluating the number of symptomatic flares.	
Section G.2.3. Concomitant intervention		Comments: Could we consider such concomitant interventions as potential outcome measures? For example, the amount of analgesics intake in case of NSAID/Coxibs trials; the amount of analgesics/NSAIDs/Coxibs in case of SYSADOA and/or potential structural modifying drugs; the rate of total articular replacement whatever the study drugs.	Accepted. Heading is included. Potentially all outcome measure can be used as secondary endpoints, however, they have to be justified, validated and/or should support the claimed indication in a meaningful way. In section 5.1 a heading is included for the secondary endpoints as follows: Other secondary endpoints may include: