London, 26 February 2015
EMA/37217/2015
Rheumatology Immunology Working Party (RIWP)

Overview of external comments received on the "Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus, and lupus nephritis" (EMA/CHMP/51230/2013)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	EFPIA
2	Technology Appraisals Programme, National Institute for Health and Care Excellenc
3	TEVA
4	Lupus Research Institute
5	Pfizer, Inc

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

1. Overarching comments

Stakeholder no.	General comment (if any)	Outcome (if applicable)
1	EFPIA welcome the release of the guideline on the clinical development of medicinal products for the treatment of SLE, cutaneous lupus and lupus nephritis. This is a sound, comprehensive document that collates current understanding of the taxonomy of SLE and, generally, recommends sound approaches to clinical development, based on current understanding and standard of care assessment instruments. The following major comments were identified: 1. **Specific claims:** it would be useful to highlight in a separate section the specific claims that the endpoints referenced in the document would support. Many terms are used to describe the endpoints such as reduction in disease activity, major clinical response or remission, and partial response, or prevention and reduction of flare, without a clear understanding of what criteria would define these endpoints. Consistency around these criteria and definitions throughout the document would add clarity. A Reference section listing the sources of the classifications mentioned in the document (e.g. SLICC SLE criteria, CLASI score, BILAG, ECLAM, etc) could be appropriate. 2. **Lupus manifestations beyond those identified in the draft guideline:** there seems to be a focus on two specific organ manifestations (cutaneous and nephritis). Guidance on studying some of the other organ specific manifestations such as polyarthritis, anti-phospholipid syndrome, or CNS would also be useful since limiting the discussion to cutaneous and nephritis	Overarching comment – GL to be reviewed to ensure consistency and cross-reference to terms and list of abbreviations and source references It is not possible to advise on endpoints for other organ manifestations as these are not validated. Furthermore the GL states that although guidance is not provided for these manifestations these patients need not be excluded. Should validated assessment scales be developed in other organ

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	appears to limit the potential populations to study. 3. **Patient population and stratification:** A great emphasis is placed on the need to clearly define the patient population at baseline and to control for disease heterogeneity and risk factors by balancing numerous factors on randomization at baseline. This priority is underscored numerous times in various sections. Although this is to be desired, stratification on more than a few criteria is impractical, if not, logistically impossible, given the recruitment limitations and requirements to meet other study-specific inclusion/exclusion criteria. It might be preferable to rank their importance. Finally, for prevention of long term damage, it is recommended that rather than requiring inclusion of only patients without preexisting damage, the draft guideline recommend that stratification for baseline damage is used to control for variable baseline damage as focusing on subjects with short duration of disease only would limit the population to be studied. 4. **Dose Response studies and use of extrapolation:** many biologicals are investigated for the use in 2 or more autoimmune disease conditions in parallel or sequentially. The draft guideline does not provide for extrapolation of dose finding information across indications and seems to request data from a dose response study in SLE LN patients specifically. In our view, extrapolation of dose finding information across indications should be considered as an acceptable alternative if a suitable	manifestations in SLE there is no bar to a sponsor conducting a trial and seeking scientific advice as new approaches cannot be foreseen in a guidance document. Due to the limited regulatory experience up to now, cutaneous lupus is not specifically covered in the guideline any longer. Partly accepted One priority is stratification on baseline damage scores. Accepted. Text revised accordingly.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	justification is provided. 5. **Baseline values:** As SLE has a waxing and waning course and the many parameters measured in the course of treatment in clinical studies are compared against baseline values, the importance of accuracy in the measurement of the baseline values should be stressed. Guidance would be useful on when baseline values might be measured relative to the start of treatment and situations where more than one measurement of the same parameter should be made (to obtain a more accurate estimate or indicate an underlying trend) In addition to "baseline", various terms have been used to describe early study events, for example, "at the beginning of the study" (line 185), "the start of the investigational therapy" (line 205), "randomization" (line 206), "the enrolment phase" (line 217) and "at study entry" (line 226). If more than one of these terms is used to describe the same event, it would be better to opt for one term only. Where the use of any term could lead to a misunderstanding or there is a need to distinguish between two similar terms, a brief definition would be useful. 6. **Lupus nephritis and assessment of partial response:** Consideration should be given to allowing partial response as the primary endpoint in a lupus nephritis trial (with complete response as a secondary endpoint) or analysis of renal response as an ordinal endpoint (no response=0 points, partial response= 1 point, or complete response = 2 points). 7. **Abbreviations:** it would be useful to update section 8, knowing the many abbreviations in this guideline, which would help the	Agreed – review GL Similar point made by LRI – for review throughout document One recommendation is that at baseline retrospective data be collected on disease activity for 12 months. Following baseline assessment a screening period of 2 months following trial entry should be encouraged in order to obtain repeated and accurate estimates and indicate any underlying trends in disease activity before starting treatment in the trial. Accepted. Text revised accordingly

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	reading. Having the section upfront could be also very useful. In addition, EFPIA have the following specific comments on text as detailed in section 2.	Accepted–abbreviation list to be extended and move to top of GL in line with RA GL.
2	It would be preferable not to use the term 'subjects' when referring to people/patients throughout the document	Accepted –GL to be modified
3	No clear differentiation of Induction and Maintenance in Clinical Practice: There is no clarity around the concept of induction and maintenance. In section 5.1, it states the aim of the study is to prevent severity of flares or reduce corticosteroid use while not being offset by worsening of overall condition. For individual patients, there is no hard and fast rule for when the induction period ends and when the maintenance period begins. Therefore it is proposed not to use terms that lack clarity like induction and maintenance but consider "treatment of lupus" as a continuum of disease therapy. While adjustments to the therapeutic regimen can and will occur for some agents this may not be true for all drugs. Therefore the idea of a mandatory regulatory distinction between "induction" and "maintenance" is not supported. Because the induction time frame may vary considerably, the study duration time frames for approvals related to "induction" or "maintenance" seems confusing. It may be more clinically relevant to demonstrate a meaningful clinical response at one year, perhaps with longer term follow-up demonstrating longer term maintenance of effect and continued safety. In addition to the above considerations, since lupus is a rare disease,	Accepted. The guideline has been completely revised to address this aspect. As it is written, the guideline still allows for potential medicinal products only targeting short-term induction, but a systematic distinction is not any longer required.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	separating between induction and maintenance will result in clinical development programmes which are not feasible in terms of sample size. This would be the case when separating induction and maintenance studies (recruiting different patients) or as a result of an integrated design in which one group of patients will be recruited, and only patients demonstrating an adequate response during induction, will be re-randomized to maintenance.	
3	1) In terms of Baseline characteristics of the patient population: With respect to (section 4.2.3) "...Patients should be stratified for randomization by relevant baseline characteristics pertinent for risk profiling e.g. histological class of lupus nephritis, level of proteinuria, and/or serum creatinine for ability to achieve remission; while other risk factors relevant for intended claim (e.g. ability to achieve remission, renal relapses or progression of renal failure) should be reported and the most important factors should be identified beforehand and taken into consideration by inclusion of these factors into the analysis model". There are quite a few issues with requesting pre-study baseline data: • Pre-study baseline laboratory data will not have been collected in a uniform way, by a central laboratory in consistent units or analyzed by consistent methods. How does the Agency suggest normalizing all of the laboratory assessments from different laboratories to a uniform baseline? • If all baseline laboratory characteristics couldn't be normalized to a single baseline norm, how does the Agency suggest using pre-study baseline laboratory data to stratify patients?	Not accepted. In terms of lab measurements historically – if these are from standard tests from labs that are accredited and subjects to External Quality assessment then there should be no difficulty in accepting these measurements for historical data.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	• Baseline standard of medical care: Medical practices differ from region to region and country to country. Anti-malarials, NSAIDS, statins and ACE/ARBs may not be used consistently even in Western EU and US/Canada. In large global studies, it is unlikely that all patients will have equal access to these medications. In particular statins and ACEI/ARBs would be expected to have an effect on long term outcomes like cardiovascular disease to which patients with SLE are predisposed. In some lupus manifestations, prior treatment with medications could also have a marked effect on study outcomes; for example, patients who have received anti-malarials long term may have a better outcome in lupus nephritis studies than patients who have not received anti-malarials prior to study entry. Can the Agency provide any guidance for how these common concomitant medications should be handled in large global multi-center studies to minimize variability and consequently difficulties in interpretation of the clinical data?	SOC will have to be justified Stratification by region is expected
	• Controlling for patient diversity across treatment arms can be difficult in large multi-center studies. It is difficult to study a broad patient population with excluding sub-populations that cannot reasonably be studied together. For lupus there are some subpopulations that have been identified as having particularly severe manifestations (for example, patients of African descent, Hispanic patient, Asian patients, and patients with certain antibody markers). Given concerns about power for each subpopulation and multiplicity in analysis planning, it is impossible to stratify for all of these subpopulations. Can the Agency provide some guidance what patient subgroups are of	Major prognostic factors such as ethnicity should be stratified. Need to prioritise stratification factors depending on included population and disease characteristics.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	particular importance for certain manifestations of lupus? If there are certain subpopulations which the Agency feels must be represented in studying certain manifestations of lupus, can the Agency provide guidance on this? Study of such sub-populations can demonstrate a trend but will not be powered. Can the Agency provide some guidance on this issue? • Pre-study biopsies: when patients are referred to study investigators from less experienced sites, there are often concerns with regards to the original biopsy interpretations. Furthermore, in patients with active glomerulonephritis, progression of renal damage may be rapid so that a biopsy obtained within 6 months of a study is not a good indicator of the patient's histologic status at the time the patient enrolls in a trial. For these reasons, can the Agency comment on their preference for obtaining biopsies at the time of study entry?	Histological blocks taken in the past can be re—reviewed centrally for the purposes of a clinical trial
3	Secondary EP: throughout the document many possible secondary EP are discussed. In light of the need to control for multiplicity, can the Agency address if these end-points should be considered as key/major secondary for which control of multiplicity is required or are they secondary end-points?	The number of and weight given to secondary endpoints should be tailored to the aim of the trial. Relevant endpoints are noted in the document.
4	Systemic lupus erythematosus (SLE) is known to have a waxing and waning course. Since the many parameters measured during the course of treatment in clinical trials are compared against baseline values, the importance of accuracy in the measurement of baseline values should be stressed. Guidance would be useful on when baseline values should be measured relative to the start of treatment, and situations where	Suggest - within 2 months of start of treatment (See above

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	more than one measurement of the same baseline parameter should be made. It should be noted that in addition to the term 'baseline, many other terms have been used throughout the draft guideline to describe early study events, including 'at the beginning of the study', 'the start of the investigational therapy, 'the enrolment phase', and 'at study entry'. If more than one of these terms is used to describe the same event, it would be helpful to confirm one term only and use it throughout the guideline. If these terms have different meanings or could lead to a misunderstanding, a brief definition of each would be helpful to avoid confusion. In general, it would be useful to include a reference section outlining the various classifications appearing through the document and their sources (e.g., SLICC SLE criteria, CLASI score, BILAG, etc.), in addition to addressing central validation of clinical endpoints to improve the accuracy of complex measures such as the BILAG and SLEDAI. Finally, currently employed composite measures of disease manifestations often do not use numerical assessments of signs and symptoms (i.e., arthritis). It would be useful to employ numerical measures of individual measure, such as the number of swollen and tender joints, whenever possible.	p5 of this document) List of abbreviations and list of references should suffice for this This is already incorporated as the component of the composite indices have to be reported
5	Overall, this draft guideline on interventional clinical trials designed to support evaluation of the treatment of systemic lupus erythematosus (SLE) and related disease is clear and logical. We agree with EMA's assessment that SLE is complex, there are no internationally validated diagnostic criteria for SLE, and possible confounders include background/concomitant medications and length	Pfizer have not formally requested a meeting. For any questions on major deviations from guideline recommendations, Companys should seek scientific advice.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	of time required to detect a treatment effect. We encourage the Agency to participate in additional dialogue in the design of specific studies, particularly regarding those aspects that relate to clinical safety evaluation. If relatively small numbers of subjects are enrolled, a pragmatic approach to long-term safety will be needed at authorisation to protect patient safety while providing access to a promising new treatment.	

2. Major Comments on each Section

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
4.2.3. Lupus nephritis (no specific lines)	3	Please clarify what the Agency means by ”ability to achieve remission”. If by this, the Agency means, reporting the baseline biopsy chronicity score, extent of tubulointerstitial fibrosis, or pre-flare eGFR, it would be helpful to include this information in the guidance. It may be helpful to provide examples of what the Agency feels constitute “ability to achieve remission”. Since the Agency is looking for baseline characteristics that assess ability to achieve remission, stratification of the population to balance for this across treatment arms will also have to be handled in the analytical plan. There is no universally accepted definition of “major/complete response”. A “major/complete response” for an LN patient who had a normal pre-flare eGFR and was non-nephrotic at baseline would probably not be the same as a major/complete response for a patient who enters a study with an eGFR of 45 and a urine protein of 15 grams/24 hours. This term does not account for inter-patient differences in baseline renal function parameters in large multi-center registration studies. If the Agency will only consider valuable improvement to include resumption of normal or near normal eGFR and protein excretion, Sponsors may look to eliminate these patients from	Accepted. Text revised accordingly. Experience with relevant conducted studies is now taken on board. Partial remission might be accepted as primary endpoint if cut-off well justified. Text revised also to provide more clarity on the treatment duration.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		registration studies. It is suggested to to include Clarification for how patients with renal damage at baseline should be considered in evaluating them for improvement. For example, does the Agency consider reduction in the slope of 1/Cr to be an acceptable surrogate for prevention of progression to ESRD in patients who have chronic renal insufficiency? Does the Agency consider restoration of independence from dialysis to be a valuable endpoint the proportion of patients who enters the study dependent on dialysis? Does the Agency consider documentation of sustained improvement in blood pressure control, reduction of edema, normalization of lipoprotein abnormalities, improved serum protein/albumin status which would be expected to occur with better control of massive proteinuria to be acceptable surrogate evidence of reduction of proteinuria and renal function improvement? Such outcomes should be considered whether these are primary or secondary specific outcomes that would support an approval. Although it is preferable for LN patients to resolve all manifestations of disease completely as early as possible, there is a large body of literature and experience suggesting that complete resolution of proteinuria may take a very long time (up to x years?). Furthermore there are studies (EuroLupus) that suggest that more modest reductions of proteinuria	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		are associated with significant improvements in long term outcomes. Recent large, well-designed and controlled multicenter studies have failed to achieve primary endpoints based on stringent reductions in proteinuria such as what the Agency is suggesting in this draft guidance. While it is agreed that it is desirable to follow LN patients for longer than 12 months after beginning treatment in a clinical study to document continued improvement in proteinuria, it is unreasonable to expect a reduction in proteinuria to 0.5 g/24 hours for all subjects within a 1 year study (especially those who enter the study with high grade proteinuria). For example, a patient who enters a study with a urine protein of 15 grams/24 hours is deriving substantial benefit if his/her urine protein is reduced to 1 or 1.5 grams/24 hours. Does the Agency really want to consider such patients to be non-responders? Would the Agency consider more modest improvements in proteinuria as grounds for approval in 1 year studies, supplemented by longer term follow up information?	
Sections 4.3.1 (General considerations) and 5.2.1 (Decrease in cumulative	3	Can the Agency please comment on how they view corticosteroids used for conditions other than the targeted indication in the concerned study? Please consider whether the guideline should include text regarding handling of corticosteroids for conditions other than the manifestation of lupus targeted in the	An additional sentence to include the requirement to document and consider the use of systemic steroids for diseases other than SLE could be considered

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
steroid dose) (no specific lines)		study. If the Agency has preferences for how stress doses of corticosteroids (for example prior to surgery), inhalational, nasal, optic, otic or other topical corticosteroids are documented and handled in a lupus study, please consider adding this text.	

3. Specific comments on text

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
4-6 13-15	4	Comment: Recommend the guidance be specific to just systemic lupus erythematosus (SLE), as the others (cutaneous lupus, lupus nephritis), represent specific organ manifestations of SLE Proposed change: Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus	Partially agreed. LN still covered in the guideline. But due to limited regulatory experience, the guideline does not any longer address the cutaneous lupus.
Line 65	1	Comment: Don't use term "revised". It means little to reader especially as criteria are continually revised. If you wish to specify a specific set of criteria then state the version number and date, as done for WHO classification on line 117. Otherwise, keep the text general and therefore it will remain up to date. Proposed change: "…..American College of Rheumatology revised……"	Agreed - word removed
66	4	Comment: SLICC needs to be defined Proposed change:	Agreed – put in List of abbreviations and amend text such that full term described in text

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Systemic Lupus Erythematosus International Collaborating Clinics	
Lines 77-78	5	Comment: Manifestations of lupus are not restricted to connective tissues. Proposed change: Revise line 77 "Systemic lupus erythematosus (SLE) is a chronic autoimmune inflammatory ~~**connective tissue**~~ disease ..."	Agreed
80-81	4	Comment: Suggest that prevalence and incidence rates be provided using same units of measurement	Agreed – text added incidence rate 3.8/ 100,000/year (95% confidence interval 2.5–5.1/100,000/year prevalence rate 27.7/100,000 (95% confidence interval 24.2–31.2/100,000) in the population and 206.0/100,000 in Afro-Caribbean females

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			The reported prevalence of systemic lupus erythematosus (SLE) in the population is 20 to 150 cases per 100,000 [1-3]. In women, prevalence rates vary from 164 (white) to 406 (African American) per 100,000 [2]. Due to improved detection of mild disease, the incidence nearly tripled in the last 40 years of the 20th century [4]. Estimated incidence rates are 1 to 25 per 100,000 in North America, South America, Europe and Asia [3,5-7]. Taken from http://www.uptodate.com/contents/epidemiology-and-pathogenesis-of-systemic-lupus-erythematosus SLE affects women more frequently than men and is more common among Afro-Caribbean and Asian

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			compared to Caucasian subjects. Incidence rates in Europe vary, but generally fall between 2 to 4.7 per 100,000 per year and prevalence rates range from 20 to 150 cases per 100,000 in the overall population with higher rates in women
93-94	4	Comment: ACR classification criteria are used to define patients who are eligible for entry into clinical trials rather than used to make “the diagnosis” of SLE.	Agreed but no consequential changes required.
Lines 129 - 132	1	Comment: This text mentions use of glucocorticoids for treatment of acute SLE, but do not list them as in line 131 for use in disease modification in the induction and maintenance phases of SLE. However, in lines 281 – 286, the use of glucocorticoids for induction and maintenance treatment in SLE is discussed. Proposed change: Line 130 - “For disease modification in the induction and maintenance phase, various immunosuppressive or immunomodulating drugs, ***including glucocorticoids***, alone or in combination are used”.	GL re-worded in line with LRI comments which also addresses this point

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
129-132	4	Comment: The use of glucocorticoids for treatment of acute SLE is mentioned, but the guideline does not list them as in line 131 for use in disease modification in the induction and maintenance phase of SLE. However, in lines 281-286, the use of glucocorticoids for induction and maintenance treatment in SLE is discussed. Proposed change: [line 130] “For disease modification in the induction and maintenance phase, various immunosuppressive drugs, including glucocorticoids, alone or in combination are used”.	Agreed –text amended with removal of the term “disease modifying” – see next comment below.
Line 130	1	Comment: Clarify intended definition of the term “disease modification” in line 130.	GL re-worded in line with LRI comments which also addresses this point – the term “disease modification” has been removed
130	4	Comment: It would be helpful to define the term “disease modification”	Point accepted. Reduction in disease activity substituted
Line 133	1	Comment: In our view, it does not seem appropriate for compounds to be specifically referenced within a Guidance document, but rather a preferred strategy, approach or application that has been utilized, unless a specific compound/class is to be utilized as the standard	Agreed – reference to Belimumab removed

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		comparator.	
139	4	Comment: The use of the term “dramatically” implies greater benefit and perhaps, less remaining unmet need than actually exists. Proposed change: “Although recent improvements in treatment regimens and medical care have reduced mortality and morbidity…”	Point accepted – term removed
Lines 149 – 150	1	Comment: The text states that the current guideline does not address central nervous system and secondary antiphospholipid syndrome patients but that these patients are not excluded from trials. This does not reflect clinical practice or historical advice given to sponsors regarding CNS patients in particular. CNS patients would benefit greatly if sponsors were given better guidance in this context. Although the rationale may be implicit, e.g., unclear clinical endpoints for CNS, high complexity, it might be recommended that a rationale be supplied for this decision. In addition, it is recommended to incorporate into the guideline advice regarding development of therapeutics to include SLE patients with CNS involvement. Proposed change:	Accepted – further amendments to this section have been made to incorporate LRI and EFPIA comments which overlap. Amended text below: Whilst other subsets of SLE such as central nervous system (CNS) lupus and secondary antiphospholipid syndrome are not specifically covered by this guideline in view of either the difficulty in making a diagnosis and/or the absence of validated efficacy assessment tools, it is encouraged to include patients with these conditions in the trials. [GM1]. Results from

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		"***Whilst other subsets of SLE***, ~~C~~central nervous system lupus and secondary antiphospholipid syndrome, are not specifically covered by this guideline; ~~however these~~ patients ***with these conditions*** are not ***necessarily*** excluded from the trials."	systemic lupus could in principle be generalised to these conditions to the extent the inflammatory activity is concerned. Due to limited regulatory experience, cutaneous lupus is not under the scope of this guideline.
149-150	4	Comment: Although the rationale for exclusion of CNS lupus and secondary antiphospholipid syndrome may be understood (e.g., unclear clinical endpoints for CNS), it is recommended that the rationale for this exclusion be provided.	Accepted – additional text added – see above:
Lines 151 - 174	1	Comment: The guidance mentions dosing in adolescents can be extrapolated from adults where possible (line 625). As a Concept paper on extrapolation of efficacy and safety in medicine development (EMA/129698/2012) is available, it is suggested to reference it here too as a relevant guideline. Proposed change: Add: Concept paper on extrapolation of efficacy and safety in medicine development (EMA/129698/2012)	Partly Accepted A CP is not a GL and more a general discussion - reference to CP added into text: Juvenile onset SLE shares many pathophysiological features with adult SLE allowing extrapolation of efficacy from adult studies to paediatric population (see concept paper EMA/129698/2012). Such

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			studies …
153-154	4	The guidance refers applicants to other relevant European and ICH guidelines. As lines 626-626 mention that dosing in adolescents can be extrapolated from adults where possible it is suggested that an additional reference be added to the list of relevant guidelines [Concept paper on extrapolation of efficacy and safety in medicine development (EMA/129698/2012)]	Partially accepted –see above
178-179	4	Proposed change: "Participating patients should have a definite diagnosis of SLE based on the revised American College of Rheumatology classification criteria **or the** SLICC SLE classification criteria."	Accepted – text amended
180-182	4	As mentioned in line 135, heterogeneity can confound clinical trial results. Therefore, encouraging the enrolment of a broad spectrum of patients compatible with the objectives of the planned study (lines 180-182) is questionable. The science often dictates enrolling a more uniform or homogenous patient population. Proposed change: Considering that SLE can have a wide range of manifestations and affected patient populations can be diverse, it is encouraged that as broad a spectrum of patients compatible with the objectives of the planned clinical trial should be enrolled **unless a specific subset or subsets of SLE patients is planned to be**	Accepted –indication will reflect studied population and text amended

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		studied. In all cases study population characteristics	
Lines 181 – 182	1	Comment: Please clarify "it is encouraged that a broad spectrum of patients compatible with the objectives of the planned clinical trial be enrolled" in light of the reference to specific organ manifestations or specific organ system as written in lines 185 and 187. To this effect, a change is thus proposed: Proposed change: ".....planned clinical trial should be enrolled ***unless a specific subset or subsets of SLE patients are planned to be studied***. ~~Nevertheless~~ ***In all cases***, study population characteristics including ....."	Accepted - as for LRI comments – amended text: Considering that SLE can have a wide range of manifestations and affected patient populations can be diverse, it is encouraged that as broad a spectrum of patients compatible with the objectives of the planned clinical trial should be enrolled **unless a specific subset or subsets of SLE patients is planned to be studied. In all cases** study population characteristics including de
Lines 183 - 184	1	Comment: Text specifies "previous and concomitant therapies (including those not directly aimed at SLE, but which could for example alter the extent of organ damage), should be predefined in detail..." Although information on all concomitant medications will be collected, it may be challenging to predefine which may alter the extent of organ damage Proposed change: Please specify if there are specific categories of medications that should	Accepted – similar to LRI comments – see amended text and also removed the word "predefined" **In all cases** study population characteristics including demographics, duration of the disease, previous and concomitant

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		be considered here as potentially altering the extent of organ damage but not directly aimed at SLE.	therapies (including those not directly aimed at SLE, but which could for example alter the extent of organ damage, e.g drugs associated with photosensitivity, thrombocytopenia etc.), should be carefully recorded at the beginning of the study.
183-185	4	Comment: The current text specifies that previous and concomitant therapies (including those not directly aimed at SLE, but which could for example alter the extent of organ damage) should be predefined in detail. Although information on all concomitant medication should be collected, it may present a challenge to predefine which medications may alter the extent of organ damage. It would be helpful to specify if there are specific categories of medications that should be considered as potentially altering the extent of organ damage, but not directly aimed at SLE.	Partially accepted – some meds – such as those that cause photosensitivity, other meds that protect renal function – could impact on outcome of trials if not balanced. It is a matter for the applicant to address these. Text slightly amended as below: **In all cases** study population characteristics including demographics, duration of the disease, previous and concomitant therapies (including those not directly aimed at SLE, but which could for example alter the extent of organ damage, e.g drugs associated with photosensitivity,

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			thrombocytopenia etc.), should be predefined in detail and carefully recorded at the beginning of the study.
Lines 187-191	5	Comment: Proposed changes below to make the text more specific. Proposed change: Revise line 184, "In the case that a specific patient cohort with certain organ manifestation is planned to be studied, the measures of how the organ involvement has been diagnosed and severity of manifestations should be well described **and documented**. Patients whose disease is limited to specific organ system only (e.g. cutaneous lupus), should undergo additional tests including serological analysis for autoantibodies and antiphospholipid antibodies **to test for more systemic manifestations towards the end of assessing additional therapeutic activities and confirming diagnosis**."	Agreed Point taken. Text deeply revised due to changes in the target populations covered in the guideline. .
Lines 190 - 191	1	Comment: It is certainly possible that patients with only 1 organ system have definite lupus (eg nephritis). It should be clarified if the additional test is to support the general lupus diagnosis. Proposed change: "Patients whose disease is limited to specific organ system only (e.g.	Text revised taken comment on board.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		cutaneous lupus, ***nephritis lupus***), should undergo……"	
190-191	4	Comment: It is recommended that there be clarity on whether the requested additional tests are to support the general lupus diagnosis.	Accepted – text clarified
Lines 191 - 193	1	Comment: The guideline suggests that serological analysis should be performed for patients who are being evaluated for organ-specific manifestations of disease as this information is useful "in order to compare those who seroconvert". This implies that seroconversion may be used as a surrogate market of disease progression or activity. Is this the intent of the authors? Proposed change: Addition of text to clarify the utility of serological data collected longitudinally.	Partially accepted. Text revised.
Lines 201-209	5	Comment: Recommended timing of renal biopsy is within the 6 months prior to randomization. This inclusion would appropriately exclude subjects with chronic disease, and "fixed" proteinuria who are unlikely to benefit from therapeutic intervention. However, in clinical practice it is not unusual to decide to escalate therapy in a subject with known renal disease based on worsening proteinuria. The presence of new, active urinary sediment should be considered for subject entry. A six-	Partially agreed – longer window might be justified for slowly progressing courses of the LN, although for rapidly progressing this window would be far too long,

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		month window would lead to significant restrictions in the ability to recruit subjects, even those subjects with active, worsening nephritis. It might be helpful to propose a central adjudication committee composed of independent nephrologists and rheumatologists who may be able to review and adjudicate study entry in such situations. We would submit that a period of 1 or 2 years would be more acceptable, in conjunction with central adjudication. Also, in order to better determine the histologic "type" of lupus nephritis and the degree of renal and tubular-interstitial involvement, and possibly, response to therapy, consideration should be given to establishing a core renal pathology laboratory for classification and follow-up. Proposed change: Revise line 204, "The biopsy should .... ideally be performed as close to the start of the investigational therapy as possible and within 6 months of randomization. **For certain patients with pre-existing renal disease, it may be acceptable to perform a biopsy within 12 months of randomization if renal disease worsens, e.g., proteinuria increases or new urinary sediment is detected. In such instances, a central adjudication committee composed of independent nephrologists could be considered to determine study eligibility.** Combination of different classes..."	Text on the time window for biopsies has been revised in a more flexible way
204-206	4	Comment: The timing of a renal biopsy 6 months prior to	see comment above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		randomization would appropriately exclude subjects with chronic disease and fixed proteinuria who would be unlikely to benefit from therapeutic intervention. However, in clinical practice it is not unusual to base a decision on escalating therapy in a patient with known renal disease based on worsening proteinuria, especially in the presence of new, active urinary sediment. We suggest that this needs to be considered for study entry. A 6-month window may limit the ability to recruit subjects, even in those subjects with active, worsening nephritis. The requirement for biopsy within 6 months of randomization may be acceptable for a trial of induction therapy in newly diagnosed, active lupus nephritis; however, this requirement will be difficult for a study with patients who are more stable or are entering a maintenance trial. This is especially relevant when seeking flare prevention; decrease of stable, refractory disease; or, decreased corticosteroid or immunosuppressive dose. One option rather than mandate a renal biopsy within 6 months might be to establish a central adjudication committee of independent nephrologists, who could review and adjudicate study entry in these situations. We would suggest that a period of 1-year or more for a renal biopsy before study entry might be acceptable in conjunction with a central adjudication process. In summary, the timing of the renal biopsy should be flexible and related to the nature of the drug being investigated and the design and anticipated outcome of the trial.	
Line 206	1	Comment:	see comment above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Reference is made to obtaining biopsies within 6 months of randomization. This may exclude patients who are still relevant for participation in a clinical trial but have had a biopsy within 12 months. Please indicate the rationale for requiring a biopsy within 6 months versus 12 months. A proposal to use 12 months – it is an invasive procedure and if someone has significant disease and a biopsy 8 months ago it seems unethical to make them undergo another. Proposed change: "… as close to the start of the investigational therapy as possible and within ~~6~~ ***12*** months of randomization."	
214	4	Proposed change: In order to demonstrate a reduction **in** disease activity….	Accepted
Lines 214 - 216	1	Comment: It would be helpful if the EMA provided an example as to what is considered a clinically important and sufficient level of disease prior to enrolment in order to demonstrate a significant change e.g., baseline SLEDAI >6. Consider steroid resistant disease as an alternative measure of disease activity e.g., a patient who for a minimum of 6 months cannot reduce glucocorticoids below 15 mg/day of Prednisone (or its equivalent).	*Accepted* *Amended text* In order to demonstrate a reduction in disease activity (induction of clinical response) patients need to have a clinically important and sufficient level of disease activity

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		These patients may have controlled disease with 15 mg/day but would flare with tapering. Given both the short-term and long-term sequelae of glucocorticoids, a goal of ≤7.5 mg/day or lower is clinically meaningful. Proposed change: "In order to demonstrate a reduction *in* disease activity (induction of clinical response) patients need to have a clinically important and sufficient level of disease activity prior to treatment in order to demonstrate a significant change ***(e.g. baseline SLEDAI >6)***."	prior to treatment in order to demonstrate a significant change (e.g. baseline SLEDAI >6).
214-217	4	*Comment:* *It would be helpful if a definition or examples were provided as to what* is considered 'clinically important' and a 'sufficient level' of disease activity (e.g. baseline SLEDAI >6). We would suggest that steroid resistant disease be considered as an alternative measure of disease activity (e.g., a patient who for a minimum of 6-months is unable to reduce corticosteroid use below 15 mg/day of prednisone or its equivalent). These patients may have controlled disease with 15 mg/day, but would flare with tapering of the corticosteroid dose. A goal of 5 mg/day or lower is clinically meaningful.	*Accepted – see above* Not accepted Those unable to reduce GC <15mg a day for 6 months would need to be accompanied by evidence of disease activity below 15mg – prior to starting therapy. Such detailed retrospective data on steroid tapering and disease flare is

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			unlikely to be sufficient for defining disease status for trials entry.
		Additionally, there needs to be clarity on an appropriate measure to demonstrate a significant change. Will this vary by the organ involved? Will sub-population analysis be required?	The appropriate measure is tailored in each trial depending on the recruited populations. Advice on sub-population analysis is getting into a detailed discussion and as the trials will vary in recruited population(s) and design this is not considered helpful in a guideline. This could have the opposite effect of stifling innovation by sponsors as pharma take GLs as rules that have to be followed.
		Lastly the term “functional disability” is used with reference to the status of the patients resulting from “the course of disease prior to baseline”. Would it be sufficient to employ the SF-36 (PCS and MCS) at baseline to assess functional disability? Would the Work Productivity and Activity Impairment (WPAI) patient reported outcomes questionnaire employed at baseline be satisfactory? Would information	Text revised to allow the use of alternative scales provided are validated and generally accepted.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		related to change in these measures before baseline be necessary?	
Lines 221 – 222	1	**Comment:** It should be highlighted that request for evidence of flares 6-12 months prior to enrolment is restrictive and operationally challenging. Further clarification is requested for lines 221 – 222 "In order to demonstrate prevention or reduction of flares (maintenance of clinical response) patients will need to have evidence of well documented flares for a period of 6-12 months prior to enrolment.", and can the agency consider discussing prevention of flare and reduction of flares separately. Modify/clarify language so there is flexibility in this requirement and provide sufficient rationale so sponsors can better interpret the EMA rationale on this issue.	Not accepted. There is no requirement for a number of flares for entry into the CT, but just the need to record. This will always depend on the intended target population and the aim of the clinical trial. Text revised to distinguish between time to and frequency of flares as possible endpoints.
221-222	4	Comment: The requirement that patients have documentation of flares for 6-12 months prior to enrolment is a very high bar given the difficulty in recruiting lupus patients and the likelihood that their histories may not be well documented. Collecting flare data retrospectively is challenging. Demonstration of prevention or reduction in flares can be adequately achieved over the course of a trial from baseline, without	Not accepted. See above.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		the pre-requisite for flare data prior to enrolment. Section 5.1.2 outlines the option to continue patients who have enrolled in a trial with active disease, i.e., in flare, presumably treated for at least 3 months, remitted, and then watched for disease flare. This is a viable option. It is recommended that this be mentioned earlier, or a reference to this section be included under 4.2.1 General considerations, for presentation of alternatives. Proposed change: Delete statement – In order to demonstrate prevention or reduction of flares (maintenance of clinical response) patients will need to have evidence of well documented flares for a period of 6-12 months prior to enrolment.	
Lines 221-222	5	Comment: The requirement that subjects have a well-documented flare 6-12 months prior to enrolment is confusing. It is well-recognized that flares are unpredictable in lupus, and the rationale for this requirement needs to be elucidated or eliminated. Proposed change: Delete lines 221-222, "~~**In order to demonstrate prevention or reduction of flares (maintenance of clinical response) patients will need to have evidence of well documented flares for a period of 6-12 months prior to enrolment.**~~"	See above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Line 223	1	Comment: First use of these abbreviations – BILAG, ELCAM, LAI, SLEDAI and SLAM. These need to be explained here or added as proposed in our major comment (comment n°7).	Agreed – List of abbreviations extended.
Lines 223 – 226	1	Comment: "The serological markers such as positivity for anti-dsDNA and complement levels should additionally be considered at study endpoints" As some SLE patients will be positive for ANA but not anti-dsDNA antibodies, we suggest - Proposed change: Suggesting adding "…positivity for anti-dsDNA *or* ***ANA*** and complement levels ….."	Accepted - partially in line with LRI comment – text modified: The serologic markers such as positivity for ANA. anti-dsDNA and complement levels should additionally be considered at study entry
Line 225	1	Comment: In general numeric rating scales (NRS) tend to be easier to understand and preferred over visual analog scale (VAS) measures by patients, resulting in less missing data. Literatures comparing different types of single-item scales conclude that generally, NRS's are preferable due to their good psychometric properties, minimization to linguistic demands and ability to be completed verbally.	Accepted – text amended as proposed

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Proposed change: "…disease with visual analogue scales***, numeric rating scales*** and health related quality of life. The serologic markers such as …"	
225-226	4	Comment: Some SLE patients will be positive for ANA but not anti-dsDNA antibodies. Proposed change: The serologic markers such as positivity for anti-dsDNA **or ANA** and complement levels should additionally be considered at study entry.	– Partially accepted ANA is not very specific and increases with age – see amended text above
Line 233	1	Comment: This section includes both cutaneous lupus and lupus nephritis but does not make any reference to the target population for SLE. Considerations should be given to providing comment on this in the guidance document. Furthermore, in this section overall there is an absence of any reference to some of the other organ specific manifestations of SLE that could potentially be studied in clinical trials such as lupus arthritis, anti-phospholipid syndrome, hematologic manifestations to name a few. It would be helpful to include the CHMP's view on studying other organ specific manifestations as it applies to the investigation of medicinal products relevant to SLE or at least put into context why only two organ specific manifestations are	The guideline has been revised to clarify that it focuses on SLE mainly with recommendations only for LN, given the limited regulatory experience with other specific subtypes. This is not to discourage investigation in any of the others, and this has been clarified.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		mentioned in the guidance as being most relevant.	
Lines 234-238	5	Comment: As SLE therapies progress, having completely failed a background therapy will no longer be considered an adequate rationale for additional treatment. Analogous to rheumatoid arthritis (RA), the standard is predicted to evolve towards patients who have completely failed, been intolerant to, OR have achieved inadequate control of disease activity on background therapy. Proposed change: Revise line 234, "The accurate diagnosis of the CLE subtype(s) included in the trial, together with the extent of active disease and damage at baseline, should be recorded. For an investigational therapy for a second line indication that is for systemic use, subjects should have failed, **been inadequately responsive to,** or have been poorly tolerant to previous adequate trials of topical therapies and/or hydroxychloroquine, despite adequate UV-protection and smoking cessation advice."	Agreed: text added
Lines 235-236	1	Comment: It would be helpful to clarify what would be considered an appropriate comparator for a second line indication after failure of topical therapies and/or hydroxychloroquine?	*Cutaneous lupus not any longer covered in the guideline*

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
238-239	4	Comment: We would like to have clarity on what would be considered an appropriate comparator for a second line indication after failure of topical therapies and/or hydroxychloroquine (HCQ). Additionally, it would be helpful to clarify whether a drug can be approved for CLE if non-inferiority to HCQ can be demonstrated.	See above
Lines 238-239	5	Comment: It would be helpful to clarify whether a drug can be approved for CLE if a non-inferiority comparison to hydroxychloroquine can be demonstrated. Proposed change: Revise line 238, "For an investigational therapy **to be authorized** for first line **treatment, ~~therapy then~~** comparison with hydroxychloroquine is recommended. **The study design may be powered for non-inferiority to hydroxychloroquine**."	See above
Lines 246 – 248	1	Comment: The draft guidance recommends " Exclusion criteria for subjects with only cutaneous lupus and no systemic disease should include topical or any local therapy known to affect CLE within 4 weeks of baseline and use of concomitant DMARDs except in the case of add-on trials to hydroxychloroquine.".	See above.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Clarification is required on the use of background immunosuppressive (topical/systemic) and whether immunosuppressive are allowed on add- on trials used with hydroxychloroquine. Furthermore, can the agency consider removing the “4 week” period as this will depend on the patient population, treatment and on case by case basis? We propose this paragraph could be written more clearly as it would seem appropriate to allow continued use of topical therapies in subjects who are inadequate responders, as long as such therapies are stable prior to randomization; suggestion below: Proposed change: “Exclusion criteria for subjects with only cutaneous lupus and no systemic disease should include ***newly added or changes to*** topical or any local therapy known to affect CLE ~~within 4 weeks of baseline~~ and use of concomitant DMARDs except in the case of add-on trials to hydroxychloroquine ***which should be stable prior to baseline***”	
246-248	4	Comment: Clarification would be helpful on the following: Background use of immunosuppressives (topical/systemic); should they be stopped within 4 weeks prior to the start of the study? Concomitant medication allowed in the case of add-on trials, in the	See above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		case of HCQ; Immunosuppressive allowed on add-on trials used with HCQ; Confirmation that in patients with CLE, all topicals need to be stopped 4 weeks before baseline. We propose that this paragraph be rewritten, as it would seem appropriate to allow continued use of topical therapies in patients who are inadequate responders, as long as such therapies were stable in the 4 weeks prior to randomization. Proposed change: Exclusion criteria for subjects with only cutaneous lupus and no systemic disease should include **newly added or changes to** topical or any local therapy known to affect CLE within 4 weeks of baseline and use of concomitant DMARDs except in the case of add-on trials to hydroxychloroquine, **which should be stable in the 4 weeks prior to baseline.**	
252-257	4	Comment: Stratification of patients for randomization; consideration should be made to stratify patients by related groups of the defined lupus nephritis classes I-VI, not all 6 classes separately. Otherwise, the groups may be too small for meaningful analysis or the resulting trials too large to be feasible. Similarly, stratification for other variables, such as disease activity, race, etc. may result in too many stratification categories for meaningful statistical analysis. As an overall principle, the guidelines should suggest consideration of stratification in these	Accepted. Text revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		important areas, but leave some flexibility to decide the level of stratification in a given trial.	
Line 253	1	Comment: Should active urinary sediment or urinary RBC and WBCs also be mentioned?	Accepted. Text revised.
Lines 261 - 262	1	Comment: The text “GFR could be clarified”. If the intent is to exclude patients with end stage renal disease, the text could be expanded to indicate the following: “ the GFR value that defines end stage renal disease for exclusion should be given.” In addition small clarifications are suggested in the text. Proposed change: “In the case that patients with end stage renal disease are excluded from the trial, this should be ~~recorded~~ ***specified*** in the protocol and ***the lowest limit of*** GFR ***value that defines end stage renal disease for exclusion*** should be given.	Accepted. Text revised. Exclusion criteria are not any longer included. This is left to sponsors to be decided on a case by case basis.
Lines 269-274	5	Comment: The regulatory implications of patients that are having inadequate control and need to be switched to a more active therapy should be	Seems OK with deletion of clause “as described below” Suggest sticking with rescue as non-responder and deleting lines

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		clarified. Proposed change: Revise line 272, “Certain common practice modifications of background therapy could be allowed; these modifications should be well defined and carefully documented in the protocol (this includes also non-SLE medication, e.g. ACE inhibitors). **If medication changes are substantial based on uncontrolled disease, then patients can be rescued as described below and considered for the purpose of the primary efficacy endpoint as treatment failures.** “	277-278 Revised text on background medication moved to Section 6: Changes in background medications that are used to treat patients with SLE can obscure detection of a treatment effect with the study drug. Therefore, background therapy should be standardized and stable as far as possible without compromising optimization. Patients’ needs during the trial should be addressed appropriately. Certain common practice modifications of background therapy could be allowed; these modifications should be well defined and carefully documented in the protocol (this includes also non-SLE medication, e.g. ACE inhibitors). If medication changes are substantial based on uncontrolled disease, then patients can be rescued and considered for the purpose of the primary efficacy

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			endpoint as treatment failures. The trial should include predefined escape conditions to allow switching to “rescue medication” when the patient fails to improve or the condition worsens. The choice and terms of rescue medication should be predefined in the protocol.
4.3. Concomitant medication 4.3.1. General considerations (lines 274-277)	3	Can the Agency elaborate on the following issues from a statistical perspective: • Is it enough that the data on rescue medication will be collected and presented? • Should patients who receive rescue medication be considered as protocol violators (although the protocol will allow for rescue medication) • It is not clear whether the analysis of responders vs. non-responders should be an exploratory analysis, and whether treatment arm should be taken into account in such an analysis?	Yes, in principle. Discussion on the potential impact on study results would be welcomed. – No – they are treated as failures Exploratory is acceptable taking into account treatment arms
277-280	4	Comment: As the guideline is currently written, it appears that patients who	Partially Accepted and in line with Pfizer comments – amended text in

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		switch to rescue medication are to be considered as protocol violations. It could be clarified within the study protocol that, in some cases, patients are to be switched to rescue medications. Information on the number/proportion of patients in the different study arms who require rescue medications would be of use in interpreting study results. This would not be the case if these switches to rescue medications were considered protocol violations. It is recommended that there be additional discussion included on how to handle the analysis of patients who receive rescue medications. Additionally, in a trial of long duration (e.g., 1 year), it is recommended that the analysis should not only focus on final background treatments, but also significant background treatment changes made during the study, which may impact efficacy outcomes. Proposed change: Comparative analysis of final background treatments in the responder and non-responder groups including "drop-out patient groups because of **switch to rescue medication"** could add additional value to interpret the results and help in future study designs.	Section 6.3.2 The trial should include predefined escape conditions to allow switching to "rescue medication" when the patient fails to improve or the condition worsens. The choice and terms of rescue medication should be predefined in the protocol.
Lines 278 – 280	1	Comment: The guidance currently indicates that "*It should also be made clear, how the use of rescue medication is going to be analyzed. Comparative analysis of final background treatments in the responder and non-responder groups including "drop-out patient groups due to protocol violation" could add additional value to interpret the results and help in*	

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		future study design." As it is written it seems that patients who switch to rescue medication are to be considered as protocol violations. However, it would be clarified in the protocol that in some cases patients are to be switched to rescue medications. Information on the proportion of patients in the different arms of a study who require such treatment would indeed be of use to interpret results from the study, however, not if the cases are considered protocol violations. Drop out of patients due to protocol violations and patients being switched to rescue medications are 2 different things they should both be handled as nonresponders. Proposed change: "Comparative analysis of final background treatments in the responder and non- responder groups including "drop-out patient groups due to ***switch to rescue medication or*** protocol violation" could add additional value to interpret the results and help in future study design." Also suggest also including more discussion on how to handle in the analysis patients who receive rescue medications. In addition, in a trial of long duration study (e.g.1 year), the analysis should not only focus on "final background treatments", but also significant background treatment changes made during the study which may impact efficacy outcomes.	Accepted – change made in response to LRI comment. Agree – suggested addition to text: See above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Line 281	1	Comment: Use of glucocorticoids in clinical trials: The guideline mentions that glucocorticoids are the accepted treatment (although not the only accepted treatment) for moderate to severe SLE but does not offer a quantitative definition of moderate to severe SLE anywhere else in the guideline. In addition, the use of glucocorticoids in clinical trials is the most confounding aspect of trial design. The need to pre-specify a plan for dosing and analysis is indicated but the implications of these on the final outcome is not clarified. For example, alternations in dose and timing of such versus exit from the trial. Proposed change: Define moderate to severe disease for purpose of clinical trial enrolment and clarify that glucocorticoids are one of the accepted treatments, not the only one ie "Glucocorticoids are *one* of the accepted treatments…"	*Text revised.*
281	4	Proposed change: Glucocorticoids are **one** of the accepted treatments for moderate to severe SLE.	Text revised
301-303	4	Comment: ACE-inhibitors and/or ACE receptor antagonists are considered	Accepted. Text revised

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		standard of care for lupus nephritis. It is recommended that ACE-inhibitors and/or ACE receptor antagonists be used in all patients with lupus nephritis unless they cannot be tolerated. Additionally, the duration of stable dose of ACE-I/ARA needed before randomization should be specified.	
Lines 311-313, and 336 – 338	1	Comment: It is suggested to include the composite indexes that are referenced in line 335 as this is the first time composite indexes are mentioned. Additionally, line 334-335 should be modified accordingly. Proposed change: Lines 311-313 "…..validated composite indexes ***(ie SLE Responder Index (SRI) and BILAG-based Composite Lupus Assessment (BICLA))*** in which ….." Lines 336 – 338 "…that combine multiple DAI are considered acceptable i.e ~~SLE Responder Index (~~SRI~~)~~ and ~~BILAG-based Composite Lupus Assessment (~~BICLA~~)~~.	Accepted; text amended (same as LRI comment)
311-314 (336)	4	Comment: It is recommended that the composite indexes referenced in line 336 be included in lines 311-314, the first time composite indexes	Accepted – text amended

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		are mentioned. Proposed change: Efficacy should be demonstrated preferably through validated composite indexes, **(i.e., SLE Responder Index [SRI] or BILAG-based Composite Lupus Assessment [BICLA])** in which the effect…	
312-314	4	Comment: In the executive summary (lines 69-71), the text indicates that decrease of cumulative steroid dose is an acceptable endpoint for assessing efficacy: "Acceptable endpoints should be used in order to assess efficacy. These endpoints include reduction of disease activity/induction of remission parameters; decrease of the cumulative steroid dose, prevention of flares/increased time intervals between flares (maintenance of remission) and prevention of long-term damage". However, in section 5.1(315-317) on primary outcomes in SLE, it states: "The aim of any study drug intended for maintenance of the response could demonstrate either the prevention of flares (decrease frequency and severity) and/or the reduction in the glucocorticoid use while maintaining the control of the disease activity and/or the prevention of long term damage." This statement indicates that corticosteroid sparing only supports a 'maintenance of response' label and not a steroid-sparing indication. Additionally, in section 5.2.1 (Decrease in cumulative steroid dose, line 398-400) it states that a secondary endpoint could be to evaluate "the percentage of patients whose average prednisone (equivalent) dose was reduced by a clinically relevant magnitude according to different stringent pre-	Not accepted. The primary endpoint for a drug needs to reflect a direct reduction in disease activity. It is less clear-cut to assess GC usage in a trial and for this reason such GC reduction, while acknowledged as an important clinical outcome, should be retained as a secondary endpoint unless the primary endpoint is a combination of a reduction in disease activity PLUS a reduction in GC usage. Text revised for clarity.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		specified criteria". It should be clarified whether this can be a primary endpoint.	
Lines 315 - 317 And Lines 398 - 406	1	Comment: In section 5.1 on primary outcomes in SLE, it states: "The aim of any study drug intended for maintenance of the response could demonstrate either the prevention of flares (decrease frequency and severity) and/or the reduction in the glucocorticoid use while maintaining the control of the disease activity and/or the prevention of long term damage." This statement thus indicates that corticosteroid sparing only supports maintenance of response label and not a corticosteroid sparing claim. Later (lines 398-406) the draft guideline states that a secondary endpoint could be to evaluate "the percentage of patients whose average prednisone (equivalent) dose was reduced by a clinically relevant magnitude according to different stringent pre-specified criteria...". It should be clarified, however, whether this can be a primary endpoint that supports a label claim. **Proposed change (to be added after lines 312-314):** ***The aim of a study drug can also be to demonstrate reduction in glucocorticoid use (while maintaining the disease activity level) as a primary outcome measure.***	See above – not accepted

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
322	4	Comment: In our opinion, the claim that SLEDAI and BILAG are extensively used in clinical practice is incorrect. Data derived from using these instruments must be translated for the practicing physician to understand the anticipated clinical impact.	Agreed – text modified
Lines 329 - 330	1	Comment: "Partial clinical response" could be replaced by the term "clinical response". A SLEDAI decrease of at least 4 points has been validated to be a clinically meaningful improvement. Referring to this as a "partial" response under-represents the positive impact of the treatment. Proposed change: ***A*** ~~partial~~ clinical response could exemplify clinically significant improvement that is not sufficient for major clinical response/complete response***. An example of a clinical response includes SLEDAI decrease of at least 4 points, or improvement in all BILAG A scores to BILAG B. In addition, a decrease in the SLEDAI Responder Index-50 (SRI-50) score of at least 4 points can also be used to detect responses.***	Accepted – text amended – same as LRI comment
329-332	4	Comment: It is recommended that the terminology, 'partial clinical response' be replaced by the term 'clinical response'. Responses such as SLEDAI	

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		decreases of at least 4 points can be clinically meaningful and therefore the term 'partial response' may under represent the clinical benefit. Proposed change: A **clinical response** could exemplify clinically significant improvement that is not sufficient for major clinical response/complete response. **An example of a clinical response includes SLEDAI decrease of at least 4 points, or improvement in all BILAG A or B scores to BILAG B or C scores, respectively. In addition, a decrease in the SLEDAI Responder Index-50 (SRI-50) score of at least 4 points can also be used to detect responses.**	Agreed – text modified as suggested
		Comment: Line 330-332 defines complete remission by complete absence of disease activity measured by disease activity in patients who do not require any on going lupus specific therapy. Clarification is recommended as to where a patient using only anti-malarial drug treatment for SLE could be considered in remission. Additionally, it is not reasonable to consider a 'complete clinical remission' without a time frame. How long would one need to follow a patient to confirm she/he is in remission, and could this be achieved without on going study therapy? We would recommend a 12-month time frame for 'complete clinical remission'.	Not accepted Complete remission should be clarified as being complete remission off treatment – very rare and disease continues to recur - J Rheumatol. 2005 Aug;32(8):1467-72 Complete clinical remission is defined by complete absence of disease activity measured by disease activity indices in patients who do not require any ongoing

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Proposed change: Complete clinical remission is defined by complete absence of disease activity measured by disease activity indices in patients who do not require any on going lupus specific therapy (**hydroxychloroquine allowed).**	lupus specific therapy for at least 12 months.
Lines 330 - 332	1	Comment: The draft guideline states "Complete remission is defined by complete absence of disease activity measured by disease activity in patients who do not require any ongoing lupus specific therapy". Please clarify if a patient using only anti-malarial drug treatment for SLE could be considered in remission. Also, please clarify the time period since the last flare needed for a patient to be considered in remission. Proposed change: "Complete clinical remission is defined by complete absence of disease activity measured by disease activity indices in patients who do not require any ongoing lupus specific therapy ***(hydroxicloroquine allowed)***".	Not accepted – see above
333-334	4	Comment:	Partly accepted

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		We recommend that the evidence behind the statement that a single disease activity index will miss significant disease activity and that a minimum of two indices are required be provided. The BILAG can be a challenge to administer in global trials without significant training on how to consistently score the domains. We would recommend that the BILAG not be an absolute requirement. Proposed change: It is recommended, **although not required,** to assess the effect on disease activity by more than one single score to ensure that the whole spectrum of the activity of the disease is captured and that results are consistent.	There is no evidence to state that we need more than on outcome measure in such a multifaceted disease. As trials in SLE are limited there is no evidence base for this requirement but it is accepted by experts in the field -(see Pfizer comments which are supportive of this)
Lines 333 - 336	1	Comment: The draft guideline moves away from accepting a single disease activity measure as a primary endpoint in SLE, considering it inadequate and that composite index should be utilized. Previously, with no medicinal product approved in SLE, SELENA/SLEDAI and BILAG were considered acceptable endpoints, however from the draft guideline only SRI and BICLA appear sufficient now. With still limited knowledge, i.e. successful trials conducted and just one medicinal product approved for SLE, this is an unwanted limitation until further successful trials are available. Acknowledging that the draft guideline is providing a listing of all of the possible disease activity measures, there should nonetheless be consistency in indicating what is the appropriate measure to support the primary outcomes indicated	Partly accepted but text modified in line with LRI comments to remove this unwanted limitation but yet retain the need for demonstration that the overall condition of the patient is not compromised by assessing a single index as not all indices measure all aspects of the condition. (see above)

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		in the document or under what circumstances would a single instrument be appropriate (measure flare ?) It is proposed to revise the section 1) not to restrict the accepted primary endpoints to SRI and BICLA, but 2) to also accept other indices including single standardised disease activity indices, such as SLEDAI and BILAG as potential primary endpoints of efficacy assessment in SLE. Proposed change: It is recommended ***although not required*** to assess the effect on disease activity by more than one single score, to ensure that the whole spectrum of the activity of the disease is captured and that results are consistent.	
Lines 333-338	5	Comment: It is understood, and fully agreed, that this is an enormously complex disease, and that therapeutic efficacy cannot be limited to one outcome measure at the current time; we therefore support the use of multiple instruments. However, pivotal trials need to have a pre-defined efficacy objective based on a single, selected, primary endpoint. Proposed change: Revise line 333, “In the view of the complexity of SLE, measurement of disease activity	Accepted and text modified Note the term “regulatory” was removed (see above)

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		by a single index alone is considered insufficient to **fully** describe the therapeutic effect in individual patients. It is recommended to assess the effect on disease activity by more than one single score, to ensure that the whole spectrum of the activity of the disease is captured and that results are consistent. Validated composite indices that combine multiple DAI are considered acceptable ~~i.e.~~ **e.g.** SLE Responder Index (SRI) and BILAG-based Composite Lupus Assessment (BICLA). **However, for the purpose of defining the primary regulatory endpoint, and enabling adequate statistical design, it is understood that a single primary outcome measure is required, and this should be defined as the primary objective of the trial, as supported by secondary outcome variables.**"	
Lines 342 - 343	1	Comment: Percentage change from baseline may not be able to be analyzed depending on the index chosen. For example, SRI is a responder index, so cannot be analyzed as a percent change from baseline. For SELENA SLEDAI, the global disease activity portion of the SRI, both absolute and percent changes can be shown. This issue likely applies to other indices and therefore the approach to analysis should be left somewhat open. Proposed change : "The results should be presented by both the absolute and the percentage change of the selected index/composite between baseline	Accepted Text amended as proposed as same as LRI comment

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		and the end of the trial ***when possible, depending on nature of the index/composite***."	
342-344	4	Comment: The guideline indicates that results should be presented by both the absolute and the percentage change of the selected index/composite between baseline and the end of the trial. For binomial/categorical composite endpoints such as those specifically mentioned (SRI, BICLA), this type of analysis would not be appropriate. It is not possible to measure absolute and percentage of change of a composite index like the SRI or BICLA. The SRI is a responder index and cannot be analysed as a percent change from baseline. For SELENA SLEDAI, the global disease activity portion of the SRI, both absolute and percent changes can be shown. Implicit in the guidelines, although not clearly stated, results should be expressed as differences between treatment groups. Proposed change: The results should be presented by both the absolute and the percentage change of the selected index/composite between baseline and the end of the trial, **when possible, depending upon the nature of the index/composite.**	Accepted
Lines 347 - 346	1	Comment: Text indicates "The proper timing for the evaluation of the effect on disease activity will depend on the time it takes the study drug to	Accepted – text clarified (same as LRI comment)

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		achieve its optimal stable effect, on the severity of the disease and its intended place in therapeutics. For induction of response the minimum would be 3 months". Does this imply response must be maintained for 3 months, or that response should be achieved at the end of a study which is of a minimum 3 months in duration? Please clarify.	
347-350	4	Comment: The guideline states that the proper timing for the evaluation of the effect on disease activity will depend on the time it takes the study drug to achieve its optimal stable effect, on the severity of the disease and its intended place in therapeutics. For induction of response the minimum would be 3 months. We would suggest the need for additional clarity as to whether this implies response must be maintained for 3 months, or that response should be achieved at the end of a study, which is of a minimum 3-month duration.	Accepted – text clarified: See above
Lines 356 - 358	1	Comment: A stable dose of background treatment is another way of telling that a patient's disease is stable. This might be preferred over the example provided because physicians do not use SLEDAI in everyday practice. So confirming stable disease by stable SLEDAI scores over 2 months would require a screening period of at least 60 days. Proposed change : "Trials assessing flares should randomize clinically stable patients (e.g. stable SLEDAI score for at least ~~two consecutive visits with a minimum~~	Not accepted – see above proposal (page 5).

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		~~interval between visits of~~ 2 months ~~or~~ ***patients who have been on a stable dose of standard of care for at least 2 months*** ~~)~~."	
Lines 362 - 365	1	Comment: The text states that using the BILAG index in a prevention of flare study design, a flare can be considered to be "1 new category A or 2 new category B items." Please consider that the following endpoint also be acceptable: ≥1 new category A and/or ≥1 new category B item. BILAG B scores are clinically significant, as the BILAG is based on physician's intent to treat. BILAG B is considered to be: disease which is less active than in 'A'; mild reversible problems requiring only symptomatic therapy such as antimalarials, NSAIDs or prednisone <20 mg day. It is important to also include the example criteria that may represent an increase in disease activity considered a severe flare. The text should indicate that whilst all flares may be measured more emphasis, in terms of clinical importance, should be placed on severe flares. Proposed change: "The flare is reflected in an increase in the disease activity score, for example an increase in SLEDAI-2K score ≥4 points, an increase in SELENA-SLEDAI score of ≥3 points, ~~or~~ **≥**1 new category A or **≥** 2 new	Accepted. Text revised

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		category B items on the BILAG score***, or ≥ 1 new category A or ≥ 1 new category B item on the BILAG score. Emphasis could be placed on severe flares, such as increase in SELENA SLEDAI score of at least 3 points to >12 or hospitalisation for SLE activity."***	
362-365	4	Proposed change: The flare is reflected in an increase in **an accepted disease activity score demonstrated to reflect a clinically meaningful change, for example an increase in SLEDAI-2K score ⩾ 4 points, an increase in SELENA-SLEDAI score of ⩾ 3 points or 1 new category A or 1 new category B items on the BILAG score.**	Agreed – see above
Lines 366 - 373	1	Comment: In section 5.2.1 the draft guideline positions a reduction in frequency of flares as the preferred outcome for measuring maintenance of response. However, the frequency of flares observed in the clinical treatment of SLE is such that clinical trials in this population will by necessity be very large and longitidunal. The 'gold standard' for treatment of SLE by physicians is to reduce background medications (addressed as a secondary outcome in section 5.2.1) Proposed change: ***Reduction in concomitant SLE medication without flaring can also be an alternative measure for assessing maintenance of***	Not accepted as primary. The drug needs to have a direct effect on disease activity.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		clinical response.	
371-373	4	Comment: The guideline indicates that a maintenance of response can be met by expressing the differences in proportions of patients in different study arms who remain flare free over at least 12 months. We would suggest a 6- month time frame, as maintenance of response is not as stringent as remission.	Guideline revised substantially to make it clear appropriate treatment duration
Lines 375 – 378	1	Comment: The draft guideline suggests, when evaluating prevention of long term damage as an efficacy endpoint, including patients without pre-existing damage as it is hard to evaluate differences in damage accrual if the population enrolled has highly variable baseline damage. In practice subjects with disease of short duration and without pre-existing damage are not the population typically enrolled, particularly in studies of lupus nephritis. In addition, this requirement may not be feasible as patients without damage have very low transition rates to damage and conducting a damage study in such a patient population would require an unfeasibly large number of subjects for an unfeasibly long duration of study. It is recommended that rather than requiring inclusion of only patients without preexisting damage, the draft guideline recommend that stratification for baseline damage is used to control for variable baseline damage.	Accepted - text amended (in line with LRI comments) Accumulated multi-system chronic organ damage as measured by the SLICC/ACR damage index is suitable to use in studies enrolling patients particularly those with short duration of disease and without pre-existing damage as it is more difficult to evaluate differences in damage accrual if the population enrolled has highly variable baseline damage. As the evaluation of damage accrual will be clearer in those with low baseline damage, it is recommended to stratify by baseline damage.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
375-378	4	Comment: The guideline suggests that when evaluating prevention of long-term damage as an efficacy endpoint, only patients without pre-existing damage should be included as it is hard to evaluate difference in damage accrual if the population enrolled has highly variable baseline damage. This requirement may not be feasible. Patients without damage have a very low transition rate to damage and conducting a damage study in such a patient population would require an extremely large number of patients for a very long study duration. We recommend that stratification for baseline damage be used to control for variable baseline damage.	Accepted – text modified: See above
Lines 379 - 384	1	Comment: It is not clear from the preceding sentences why a minimum 18 month trial duration is needed for damage to occur or progress and to stay present for another 6 months. Potentially one could demonstrate that in a study of 12 months duration given that lupus is diagnosed at study entry. In addition, SLICC/ACR measures damage in a 6-month period. Two 6-month periods of measurement is sufficient, three seems excessive. It is thus suggested to change the requirement of damage measurement from 18 months to 12 months. Proposed change: “Therefore to measure the damage that has accrued during the clinical trial, the trial has to be long enough (for at least ~~18~~ ***12*** months for	Accepted – text changed. Damage items are usually recorded if the clinical item has been present over 6 months or associated with immediate pathological change indicative of damage. Therefore to measure the damage that has accrued during the clinical trial, the trial has to be long enough (for at least 12 months) for damage to occur and remain present for 6 months.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		damage to occur and remain present for 6 months."	
Line 388	1	Comment: The section on prevention of long term damage references the organ specific outcomes section which discusses renal outcomes primarily in the context of a LN study. It could also be relevant to consider outcomes to demonstrate lack of progression of renal disease in a systemic SLE study. Proposed change: Clarify renal outcomes that would provide meaningful demonstration of the lack of progression of renal disease in a systemic SLE population who do not have severe renal disease at baseline	Text revised in the corresponding section.
Lines 394 - 406	1	Comment: Examples are given for demonstration of steroid reduction by a clinically relevant magnitude according to predefined criteria. For all examples, a decrease to less than or equal 7.5mg/day is mentioned. However, subjects with high starting doses could achieve a clinically relevant reduction with reduction of associated toxicities and still be at doses above 7.5mg/day. It should be clarified if alternative measures of steroid reduction are considered clinically relevant in patients with disease controlled, eg, a comparison between treated and placebo	Point taken and text revised to make it clear that this is just an example.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		patients in the reduction in steroid dose from baseline dose could also demonstrate a clinically relevant reduction. Decrease in “cumulative steroid dose” is defined as a potential efficacy endpoint. However, the definition of the efficacy endpoint defined on lines 400-402 describes endpoints based on proportion of subjects achieving a clinically meaningful reduction in steroids, and not a cumulative reduction in steroids. Please consider modifying the term “cumulative steroid dose” to a term that better reflects the description of how the endpoint should be measured.	Accepted -removed the term cumulative from line 396
Lines 395 - 403	1	Comment: An additional section should be added to providing guidance on discontinuing immunosuppressants and other treatments.	Not accepted While the side effects of steroids are well established and the clinical benefit from a reduction in dose for those on long-standing steroids above 7.5mg is accepted, less is known about other immunosuppressive. In addition as few are licensed specifically for SLE it seems unhelpful to mention these here –comments please.
395-406	4	Comment:	Agreed

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		The guideline provides examples for demonstration of steroid reduction by a clinically relevant magnitude according to predefined criteria. For all examples, a decrease to ≤ 7.5 mg/day is mentioned. However, patients with high starting doses could achieve a clinically relevant reduction, with reduction of associated toxicities, and still be at doses > 7.5 mg/day. Clarification is needed as to whether alternative measures of steroid reduction are considered clinically relevant in patients whose disease is controlled. (e.g. comparison between treated and placebo patients in the reduction in steroid dose from baseline dose could demonstrate a clinically relevant reduction). We suggest that a minimum % reduction be cited which would avoid minor reductions in steroids from being deemed significant. Proposed change: The efficacy evaluation for steroid tapering should be based on the percentage of patients whose average prednisone (equivalent) dose was reduced by a clinically relevant magnitude according to different stringent pre-specified criteria, i.e. subjects whose prednisone equivalent dose was > 7.5 mg/day at baseline and reduced to ≤ 7.5 mg/day (**and at least a 20% reduction**) without any flares for….	
408-428	4	Comment: This section on quality of life instruments discusses how these instruments are problematic, i.e., they do not correlate with disease activity measures or, they are non-validated. Their utility in assessing a drug's efficacy or implications for drug approval is not clear. The recommendation that these measures be performed appears to be a request to gather additional data, perhaps for exploratory	Partially accepted. Text simplified and clarified – Accepted: Other validated instruments are

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		analyses. Further clarification would be helpful. The limitations of fatigue assessments are discussed however the FSS is the only assessment mentioned. Does this imply endorsement for only this assessment or are other instruments such as the FACIT-F also acceptable?	acceptable – text modified.
Lines 409 - 416	1	Comment: The statement "*no single tool exists that measures all the aspects that influence health related quality of life*" is not accurate. HRQoL measures such as the Lupus QoL and SLE QoL have been developed to assess the impact of disease, treatment, and co-morbid conditions on multiple dimensions of a patient's life; it is important to consider their added value when assessing patients in Lupus (Agarwal et al., 2009; McElhone et al., 2007). Proposed change: "Health related quality of life (HRQoL) is known to be ~~impaired~~ ***compromised*** in lupus patients and appears to be an independent outcome measure ~~. As at the time of writing this Guideline, no single tool exists that measures all the aspects that influence health related quality of life (fibromyalgia, fatigue, cognitive dysfunction, depression, other co-morbidities and concomitant medication) in lupus. Therefore, although HRQoL is important to consider from a patient's perspective, the measure~~ ***that*** does not	Accepted – text amended in line with proposal with additional clause on which measure to use

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		necessarily correlate strongly with disease activity or organ damage. ***Given the fact that HRQol is an independent outcome in Lupus patients, it is therefore important to consider assessing the impact of the condition on the patient's HRQoL to have a complementary source of information beyond what is captured by measures of disease activity and organ damage.*** As HRQoL is of central relevance from the patient's perspective, particularly in cutaneous lupus, supportive data from HRQoL is strongly recommended. ***In essence, a comprehensive assessment of patients with Lupus should include disease activity, organ damage, and HRQoL."***	
Lines 417 - 423	1	Comment: We think that the statement "*As the SF-36 in SLE patients with established disease changes little over a longer period (8 years), the SF-36 is more sensitive to change over short time periods and in cases of earlier disease where there is less damage*" is not accurate. The sensitivity of SF-36 to improvement or decline in disease activity over short time periods is poor, with small to absent effect size (AGGARWAL et al., 2009). If the CHMP believes that a generic measure of HRQoL needs to be used alongside lupus specific measures of HRQol, we suggest that the agency consider the EQ-5D as an alternative to SF-36 (Wolfe et al., 2010; AgarwalL et al., 2009). We believe that having 2 lengthy	Text revised and simplified.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		measures of HRQoL in a clinical trial setting can be burdensome to the Lupus patient. We thus, recommend that the CHMP considers LupusQol as a standalone measure of HRQoL or alongside EQ-5D in a clinical trial setting (MCElhone et al., 2007). Proposed change: "Medical Outcome Study Short Form 36 (SF-36) has widely been used to assess physical, psychological and social impact of chronic diseases like lupus. ~~As the SF-36 in SLE patients with established disease changes little over a longer period (8 years), the SF-36 is more sensitive to change over short time periods and in cases of earlier disease where there is less damage.~~ ***In addition, the EuroQol-5D (EQ-5D), a recommended measure of HRQoL by the European HTA organizations, has widely been used to assess HRQoL across rheumatic diseases and has demonstrated satisfactory psychometric properties in Lupus. EQ-5D has been shown to correlate strongly with SF-36 in lupus; however the sensitivity of both measures to improvement or decline in disease activity is poor, with small to absent effect size***. Lupus specific instruments include the Lupus Quality of Life (Lupus QoL), SLE symptom checklist and SLE Quality of Life (SLE QoL). As ***most of the Lupus-specific*** ~~these~~ instruments have not been validated in clinical trial settings and their correlation with SF-36 is variable, it is prudent to use these instruments together with SF-36 ***or EQ-5D. Evidence exists to support the validity of the LupusQoL, a disease specific measure of HRQoL that has been developed***	

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		and psychometrically validated in lupus patients in the UK. Discriminant validity of the LupusQoL has been shown for different levels of disease activity as measured by BILAG or SELENA-SLEDAI. LupusQol can be a standalone measure of HRQol or alongside EQ-5D to help decrease the patients' burden of having to complete two lengthy measures of HRQoL."	
Lines 424 - 428	1	Comment: Despite the text indicating the limitations of the various fatigue assessments, ultimately the FSS is only mentioned. Does this imply endorsement for just this assessment, or are other instruments such as FACIT-F or BFI (Brief Fatigue Inventory) also acceptable? Please clarify if FSS is the preferred assessment tool for fatigue or just an example and if other instruments are also accepted. Proposed change: "Fatigue is a major concern for adults with SLE ***affecting more than 90 percent of the lupus patient population. Many instruments have been developed to measure fatigue severity and its impact in SLE. Some disease specific instruments are considered relevant outcomes for the characterization of response to treatment such as*** the ~~and the scores of fatigues domain tend to be poor regardless of levels of disease activity and damage. Despite of its relative importance, consensus of which scale possesses the most suitable properties is lacking. F~~ ***f***atigue severity scale (FSS) ***which*** is	Accepted – none is preferred. Text amended in line with proposal

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		most commonly used and correlates moderately with the 8 scales of SF-36, ***the FACIT fatigue or the Brief Fatigue Inventory (BFI). Other alternatives might be used provided validated and generally accepted "***	
Lines 432 - 435	1	Comment: Use of biomarkers may also establish early evidence of target engagement and thus enhance confidence in the drug having appropriate biological activity Proposed change: "It is therefore advised that identification and ~~subsequent~~ inclusion of biomarkers is incorporated as an integral part of the drug development programme."	Agreed
Lines 454 - 466	1	Comment: Clarification of activity and damage scores and their significance: Damage happens as activity improves. Separation of activity and damage allows for seeing the improvement, which relates ONLY to the activity score. Damage helps to make sure people think about what they are looking at. These should not be confused. CLASI activity score: only the CLASI activity score is used for entry criterion; it is always separated from the damage score. Drug effects are only related to change in activity.	Cutaneous lupus not any longer specifically covered in the guideline

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		How was the reduction of CLASI by 50% chosen as the appropriate clinical endpoint metric? A CLASI40, i.e., reduction of 20% is clinically meaningful (minimally significant change). In addition has the alternative of a 4-point or greater decrease from baseline been considered – especially if patients enter with scores that achieve at least moderate activity, i.e. >10? Please clarify.	
455-461	4	Comment: Only the CLASI disease activity score is used for entry criterion; it is always separated from the damage score. The effects of drugs are only related to change in activity. Damage happens as activity improves and separation of activity and damage makes it possible to see the improvement, which relates only to the activity score. Further clarification on how the reduction of CLASI by 50% was chosen as the appropriate clinical endpoint metric would be helpful.	Cutaneous lupus not any longer specifically covered in the guideline
468	2	Comment: Presume that 'DQLI' should be 'DLQI' Proposed change: 'patient's QoL and dermatology quality of life indices e.g. DLQI,	Cutaneous lupus not any longer specifically covered in the guideline

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		patient's global assessment and VAS'	
Line 477 - 478	1	Comment: 'Rebound on withdrawal': To our knowledge this is not standard. It is not clear why this is mandated, i.e. 'needs to be investigated in a randomized withdrawal phase.' This is different from SLE where documenting improvement has been difficult and maintenance of effect, prevention of flares, on treatment, is standard. However, a similar suggestion is not offered for the SLE or Renal Lupus patient populations in the guideline. Please provide a justification for requiring a rebound-on-withdrawal phase in studies in CLE patients.	Accepted – information on rebound on withdrawal no longer requested.
477-478	2	Comment: The proposals for demonstration of duration of efficacy (for disease activity) and investigation of rebound on withdrawal would fit better under section 5.1.1 rather than only under 'cutaneous outcomes' (section 5.3.1). It would not only apply to cutaneous outcomes. (Or even under section 6 'strategy and design of clinical studies' – specifically section 6.2)	Update – remove requirement for rebound on withdrawal.
477-478	4	Comment: Additional information on the requirement to investigate rebound on withdrawal in a randomized withdrawal phase is needed.	See above

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Lines 479 - 480	1	Comment: The guidance should not prescribe a specific mechanism for the long-term follow-up of patients because such data can be successfully collected by patient registries as well as an open label extension studies. Proposed change: "For a therapy that has efficacy in reducing disease activity, long-term follow-up of patients ~~in an open label extension~~ will be required to demonstrate efficacy for reduction of damage."	Accepted
479-480	2	Comment: Would the proposals for demonstration of efficacy for reduction of damage in a long-term follow up in an open label extension fit better under section 5.1.3 'prevention of long term damage' rather than only under 'cutaneous outcomes' (section 5.3.1). It would not only apply to cutaneous outcomes. (Or even under section 6 'strategy and design of clinical studies' – specifically section 6.2)	Sentence moved to section 5.1.3
Line 483 - 484	1	Comments: *"Primary renal specific endpoints in a trial, conducted specifically among lupus nephritis patients, should include SLE endpoints as co-primary endpoints"* It should not be required for a LN study to include SLE endpoints as co-	(see below)

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		primary. The possible inclusion in a trial on lupus nephritis patients of SLE endpoints as co-primary endpoints, beside primary renal specific endpoints, should be further justified. It should also be more clearly indicated which SLE endpoints would be relevant as co-primary endpoints (e.g. relevant biomarker endpoints). Many LN patients will not have significant extra-renal disease activity at entry and the trials will be underpowered to evaluate the effect of intervention on extra-renal disease activity. Clinical indices of systemic SLE are already listed on line 503 as Secondary specific outcome which is not consistent with a requirement for a co-primary SLE endpoint. Given the size of the population (orphan disease) it may be difficult to power for an SLE (e.g., SRI responder index, BICLA) co-primary endpoint. The expected delta in the renal endpoint (complete renal response) will be relatively small (e.g. 10-20%), thus a samples size of >150 patients/group will be needed to adequately power the test involving this endpoint. Hence, the overall sample size will already be greater than 300 patients with LN and will be difficult to recruit. In case the general SLE endpoint is a co-primary endpoint and the delta on the general SLE endpoint is smaller than the delta for the renal response (e.g delta for Benlysta <15%) an even larger sample size would be needed to power the trial adequately for both endpoints. Therefore, we recommend that the general SLE endpoint should be	

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		considered as an important secondary endpoint for Lupus Nephritis tbetter understand how other manifestations of the disease behave during the trial. Proposed change: "Primary renal specific endpoint in a trial, conducted specifically among lupus nephritis patients, should ***be renal specific and*** include SLE endpoints as ~~co-primary~~ ***secondary*** endpoints.	
483-484	4	Comment: It is suggested that there not be a requirement for a lupus nephritis study to include SLE endpoints as co-primary. The primary endpoints of a lupus nephritis study should focus on renal outcomes (e.g., renal response). Given the size of the lupus nephritis population, it may be difficult to power for a SLE co-primary endpoint (e.g., SRI responder index, BICLA). However, we do believe it is important to include other SLE endpoints as secondary endpoints to increase understanding of how other manifestations of the disease behave during the trial, and specifically, whether there is deterioration in extra-renal manifestations of SLE. Proposed change: Primary renal specific endpoints in a trial, conducted specifically among lupus nephritis patients, should include SLE endpoints as **secondary**	See below

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		endpoints **or to determine whether there is deterioration.**	
Line 483-500	5	Comment: The text intimates that even when a protocol is specifically aimed at lupus nephritis, systemic SLE endpoints should be used as co-primary endpoints. This functionally would mean that a trial that showed a profound effect on renal endpoints (one of the most common life-threatening aspects of the disease) that failed to improve the more diverse SLE endpoints could be considered a failed trial from a regulatory perspective and, thus, may not support approval of a drug that improved renal outcomes. Such an outcome would seem unintended from a public policy perspective. The final guideline should reflect that for trials aimed at improving renal outcomes, broader disease activity should be measured as an important secondary endpoint, and need not be considered a co-primary endpoint. Proposed change: Revise line 485, "....results obtained from certain classes cannot generally be extrapolated to other classes. **For trials aimed at improving renal outcomes, broader disease activity should be measured as an important secondary endpoint, and need not be considered a co-primary endpoint**."	Point taken on board and text revised accordingly

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Lines 487 - 489	1	Comment: Consideration should be given to allowing partial response as the primary endpoint in a lupus nephritis trial (with complete response as a secondary endpoint) or analysis of renal response as an ordinal endpoint (no response=0 points, partial response= 1 point, or complete response = 2 points). A partial response is indeed a clinically important and meaningful outcome which can be considered, and similarly to complete response, a relevant primary endpoint; for example, a partial renal response has a significant effect on renal and patient survival over 10 years compared to non-renal response in subjects with diffuse proliferative LN. [Chen YE, Korbet SM, Katz RS et al for the Collaborative Study Group. Value of a Complete or Partial Remission in Severe Lupus Nephritis. Clin J Am Soc Nephrol 2008;3: 46–53]. Moreover, could you please clarify: • response criteria for partial response, complete renal response or remission. Also, it would helpful to provide guidance on the use of spot urine, 24 hour urine collection, and the formulas to calculate GFR since the text specifically mentioned measured GFR • if all components (GFR, proteinuria, urinary cells and sediment) need to be assessed for a major/complete renal response since only GFR and proteinuria is mentioned • what does “clinically significant” improvement of renal function	Accepted. See below

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		mean? Finally, please consider either removal of findings in active urine sediment as a requirement for demonstrating complete renal response or provides a very specific definition of abnormalities. It is recommended considering only cellular casts as an abnormality. Proposed change: The guideline should allow flexibility for sponsors to define primary and secondary endpoints in lupus nephritis trials, and not require complete renal response as the acceptable primary endpoint in a lupus nephritis induction trial. The CHMP is encouraged to retain language describing the definition of a major/complete renal response.	
487-499	4	Comment: The requirement for normalization/return to baseline of measured GFR or proteinuria (< 0.5 g/24 hours) may result in excluding patients with pre-existing renal damage from clinical trials. A complete response is seen in only a minority of lupus nephritis patients. Consideration should be given to allowing partial response as the primary endpoint in a lupus nephritis trial (with complete response as a secondary endpoint, or analysis of renal response as an ordinal endpoint (no response = 0 points, partial response = 1 point, or complete response = 2 points). It is not uncommon to have residual proteinuria > 500	Accepted. Text revised. See below

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		mg/day, despite lack of evidence of disease activity in patients with Class III/IV lupus nephritis and some chronicity. Although the 10-year renal prognosis and survival is less in these patients than for those who achieve complete remission, it is still clinically meaningful. Given the relatively low rates of complete renal response achieved, partial renal response remains a clinically important outcome in patients with active lupus nephritis, reflecting improvement in renal disease activity, one goal of therapy. It would be helpful to specify the duration of the maintenance of major/complete renal response. Our recommendation for duration would be 6 months. Lastly, we suggest clarifying what would fulfil criteria in lupus nephritis for a 'partial clinical response' [clinical response] and maintenance of response. This might be stated as follows: Clinical response can be considered to be a > 50% decrease in proteinuria while GFR does not worsen > 10%. Maintenance of response can be considered to be ≤ 15% decrease in GFR on two successive measurements at least one week apart, and < 50% increase in proteinuria, over a one year time period.	
Lines 487-500	5	Comment: The text reflects an expectation that the improvement in renal function will be major or complete, as defined by normalization or return to baseline. If the renal damage due to lupus nephritis were completely	Accepted. Text revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		reversible, this would be a realistic therapeutic goal. However, in many cases the renal damage is not completely reversible, and therefore normalization or return to baseline is not a feasible endpoint in patients with long-term disease. The final guideline should be modified to reflect additional flexibility, and to accept clinically meaningful differences between groups at study endpoint as acceptable regulatory endpoints. Proposed change: Revise line 500, "Study endpoints must be appropriate to show efficacy for the indication sought. **Clinically meaningful differences between groups at study endpoint(s) would be considered acceptable regulatory endpoints**."	
Lines 489 - 492	1	Comment: The requirement for normalization/return to baseline of measured GFR or proteinuria (<0.5 g/24-h) may result in excluding patients with pre-existing renal damage from clinical trials. This might have the unintended consequence that the benefit of novel compounds in subjects with moderate to severe renal damage (clinical or histological) might not be tested prior to granting of a Marketing Authorisation. It would be helpful to clarify that return to normal levels is only an example of a treatment goal and not an absolute requirement. It is suggested that use of UPCR (spot urine) to measure protein	Accepted. Text revised

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		excretion as a surrogate for 24-hr urine collection. It would also be useful to have a GFR range that would be demonstrative of an improved GFR.	
Lines 494 - 496	1	Comment: The guideline discusses prevention of renal flare as an appropriate measure for assessing maintenance of response, but does not offer any definitions of a renal flare in this context. Please provide guidance on measuring / defining renal flare. It is proposed to specify the criteria for prevention of renal flares. In particular, it is proposed to state that a relative reduction in renal flares is considered sufficient, especially in light of chronic kidney damage. In addition, to be consistent with our previous comment (Line 487), knowing the clinical relevance of partial response in induction or maintenance of remission, this endpoint should be also considered a primary specific outcome and not solely a secondary specific outcome.	Accepted. Text revised.
Lines 498 - 499	1	Comment: The Draft Guidance specifies “prevention of long term damage, i.e. slowing progression of CKD” as one of 3 potential primary specific outcomes in a lupus nephritis trial. This endpoint is problematic as the rate of progression of long term damage occurs over many	Accepted. Text revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		years. Therefore, it is not feasible to conduct a well-controlled trial for the duration of time needed to establish sufficient progression in the control group. For example, the 10-year follow-up data from the Euro-Lupus Nephritis trial showed the rate of death, doubling of serum creatinine, or end-stage renal disease ranged from 5% to 11% in various groups studied (Houssiau 2010). The cross reference to "other EU guidance options" is also problematic as it is non-specific. Proposed change: Please define expectations for duration of observation for both control and treatment groups needed to achieve the outcome of prevention of long term damage in lupus nephritis	
Lines 501 - 512	1	Comment: Onset of action is an important clinical outcome for SLE/LN studies because a fast clinical effect can reduce irreversible damage to the kidney and other organ systems. Steroids are also considered as one of the most important cause of long term damage in SLE based on a recent SLICC report. It is recommended that "onset of clinical effect" and "cumulative steroids doses (steroids tapering)" being measured during a study and be included as a secondary endpoint.	Accepted – text to be modified

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Line 502	5	Comment: The requirement that renal outcomes be limited only to major clinical response can be problematic, both from the standpoint of adequately powering clinical studies, and that partial clinical response may be considered acceptable in renal subsets who are failing standard immunosuppressive treatment. Partial response (defined as a clinically meaningful difference in treatment groups) should be considered an appropriate primary endpoint, lest development be forced to focus on patients likely to have fully reversible disease. Proposed change: Revise line 502, "Partial response**, defined as a clinically meaningful difference in treatment groups,** in induction or maintenance of remission **(Note that a partial response may also be considered an appropriate primary endpoint)"**	Accepted. Text revised.
501-512	4	Comment: The secondary endpoint of 'histological results of renal biopsy (such as changes in Activity and Chronicity indices over at least a 6 month period) is not considered standard of care and could be overly aggressive for patients who are improving. It is not clear how this outcome can be used for a secondary endpoint. Can a renal biopsy result be used with renal clinical outcomes (proteinuria and eGFR) as a	–Not accepted. These are examples of what would be relevant clinical outcomes (secondary).

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		composite primary endpoint for definition of complete response? The long-term renal outcomes: development of ESFR (CKD 5D) with requirement of chronic renal replacement therapy and/or transplantation is questionable. It does not seem ethical to include a patient in a study and knowingly let her/him get to ESRD. Investigators should have the ability to switch the patient to a rescue medication before the patient reaches ESRD. This would be supported by an analysis of the proportions of patients in both groups having required this switch. Proposed change: Long term renal outcomes: development of ESRD (CKD 5D) with requirement of chronic renal replacement therapy and/or transplantation, **or proportion of patients having been switched to a rescue medication because of their deteriorating renal function.**	
Line 502	1	Comment: Remission is used for the first time here in this section. Please clarify if it is similar to what is referred to as complete renal response. Comment: Please clarify what would fulfil criteria in Lupus nephritis for	Accepted. Text revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		1. a partial clinical response 2. maintenance of response In addition, the term “partial clinical response” can be replaced by “clinical response”. A complete response is seen in only a minority (e.g., 40%) of lupus nephritis patients. Moreover, it's not uncommon to have residual proteinuria >500 mg/day, despite lack of evidence of disease activity in patients with aggressive lupus nephritis (e.g., Class III/IV) and some chronicity. Although their 10 year renal prognosis and survival is less than for those who achieve complete remission, it is still clinically meaningful. Consider adding a partial response in proteinuria (e.g., <1 g/day) as an acceptable component of the primary endpoint during induction and/or maintenance of major/complete renal response.	
Lines 508 - 509	1	Comment: “Secondary endpoints: Histological results of renal biopsy (such as changes in Activity and Chronicity indices over at least a 6-month period).” This is an invasive procedure which presents risk to the patient, is not standard of care and might be considered overly aggressive for patients who are actually improving. While this could be of interest, we would not favour advocating this routinely. Proposed change: *“Histological results of renal biopsy ~~(such as changes in Activity and Chronicity indices over at least a 6-month period)~~ –* ***data should be***	Partially accepted.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		collected where possible however, repeat biopsies are not a requirement. ”	
Line 510	1	Comment: The guidance states that “Long term renal outcomes: development of ESRD (CKD 5D) with requirement of chronic renal replacement therapy and/or transplantation”. Similarly to the comment related to lines 382-384, it does not seem ethical to include a patient in a study and knowingly let him/her get to ESRD. PIs should be able to switch the patient to a rescue medication before the patient reaches this critical condition, supported by an analysis of the proportions of patients in both groups having required this switch. Proposed change: “Long term renal outcomes: development of ESRD (CKD 5D) with requirement of chronic renal replacement therapy and/or transplantation, ***or proportion of patients having been switched to a rescue medication due to their condition deteriorating***.”	Not accepted. It is suggested as a possible secondary endpoints. Others can be included if considered relevant.
Lines 522 - 524	1	Comment: Meeting this recommendation as written would be onerous and not	Accepted text modified as proposed

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		standard practice. Proposed change: "***<u>Should specific indications targeted at specific patients be sought.</u>*** efforts should be made to find different doses and treatment intervals according to the respective targeted patient characteristics (i.e. severity, organ involvement)."	***<u>Should specific indications targeted at specific patients be sought</u>*** efforts should be made to find different doses and treatment intervals according to the respective patient characteristics (i.e. severity, organ involvement).
Lines 525 - 532	1	Comment: For Phase 2 trials would be interesting to have guidance on minimum background medication required or the EMA position if study designs like BOLD (induce remission then taper background medication to a minimum or nothing then compare rate and time of flares with investigational drug or placebo) are considered acceptable. Many biologicals are being investigated for the use in 2 or more autoimmune diseases in parallel or sequentially. The guideline does not provide for extrapolation of dose finding information across indications and seems to request data from a dose response study in SLE patients. We request that extrapolation of dose finding information across indications is accepted as an alternative if a suitable justification is provided. Finally, it is suggested not specifying "phase II" for dose finding. Since	Agree, text modified. Agree – text modified to delete "phase II"

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		the optimization of dose with respect to risk/benefit may be determined from the totality of a clinical program and not specifically based on a Phase II study. Proposed change: "Duration of the ~~phase II~~ dose finding study depends on the SLE patient profile (e.g. severity of organ manifestations), chosen endpoints and mode of action of the medication, but it should not be shorter than 3 months."	
Lines 529 - 532	1	Comments: *"For lupus nephritis patients separate appropriate dose finding needs to be undertaken for both the induction and maintenance phases."* Separate dose ranging in lupus nephritis should not be required in any situation. Depending on the drug mechanism of action and availability of predictive biomarkers or PD markers, extrapolation of dose finding information across indications may be acceptable in some cases as an alternative. Proposed change: "For lupus nephritis patients ~~separate~~ *appropriate dose finding needs to be undertaken* for both the induction and maintenance phase. ***Depending on the drug mechanism of action and availability of predictive biomarkers or PD markers,***	Accepted. Text fully revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		<u>extrapolation of dose finding information across indications (e.g. SLE) and from "induction" in lupus nephritis to" maintenance" in lupus nephritis may be accepted".</u>	
529-532	4	Comment: It is suggested that separate dose ranging in lupus nephritis should not be required if a safe and effective dose has already been determined for general SLE, which included patients with some degree of renal involvement, and the effect of renal function on the clearance of the product is well understood. Proposed change: **Appropriate** dose finding needs to be undertaken. For lupus nephritis patients, **separate dose finding studies may not be required if a safe and effective dose has already been determined in general SLE which included patients with some degree of renal involvement, and the effect of renal function on the clearance of the product is well understood.**	Accepted. Text revised.
Lines 529-532	5	Comment: This section of the draft guideline is not realistic for early-phase development for a complex, low-prevalence disease that has a large, unmet medical need. Specifically, lupus nephritis is a variety of lupus that has received orphan designation in some jurisdictions, and the conduct of separate dose-ranging trials (i.e. two simultaneous Phase 2 trials) is a requirement that would delay development substantially. Further, in other inflammatory diseases in which there are separate induction and maintenance trials (e.g. Crohn's Disease and ulcerative	Accepted. Text revised.

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		colitis) it has been possible to conduct two-stage trials that can address appropriate dose-ranging in both settings. Finally, in clinical practice patients with excessive disease activity are induced into an acceptable level of disease and then maintained, which is reproduced in the setting of combined induction/maintenance designs. The final guideline should reflect some flexibility to conduct a single dose-ranging trial that adequately addresses both induction and maintenance activity, i.e., through a defined primary induction phase and a supportive maintenance assessment period. Proposed change: Revise line 532, "..... maintenance of the remission is advised. **However, it may be possible to conduct a single dose-ranging study that addresses both induction and maintenance activity, i.e., a study designed to cover a defined primary induction phase and an exploratory maintenance assessment.**"	
Line 534 - 536	1	Comment: As in comment on lines 246 – 248, guidance on withdrawal of background corticosteroids and immunosuppressive medications would be desirable.	Not agreed as this is the interaction section –It is not considered helpful to be prescriptive on how and when withdrawal of GCs should be performed. Also as noted above an indication for reduced immunosuppressives other than

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
			GCs is not considered acceptable in the guidance
Lines 561 - 569	1	Comment: A bullet list of trial designs and potential primary outcomes is started. Bullet A, A.1 and A.2 are given, what is B? It does not appear in the guidance.	Accepted . Text revised
Lines 561-565	5	Comment: The appropriate period for showing induction of remission should be based on the onset of action of the drug and the disease setting. Some drugs may require 3 to 6 months to induce a meaningful response, but others may show such responses in shorter time periods. Flexibility should be reflected in the final guideline. It is understood, as reflected in the subsequent text of the guideline, that lupus nephritis induction trials may require longer response periods. Proposed change: Revise line 565, "…. rebound should be addressed in the long term. **However, the appropriate period for showing induction of remission should be based on the onset of action of the drug and the disease setting. Some drugs may require 3 to 6 months to induce a meaningful response, but others may show such responses in shorter time**	Agreed – section deeply revised

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		periods. Thus, flexibility is required in study design to reflect this as well as the longer response periods that may be required for lupus nephritis induction."	
Lines 563-565	1	Comment: The guideline describes that induction of major response or remission should be assessed after a minimum of 3-6 months, and that lack of rebound should be assessed in the long-term. Proposed change: Provide more specific guidance on the structure of a primary endpoint in this setting – could a sponsor file for a MAA based on successful induction data without long-term efficacy and or / safety (section 7.3)?	Accepted **Text revised**
Lines 563 - 656	1	Comment: 'Absence of rebound, should be addressed in the long-term': While it is the ultimate goal to have patients in remission off medication, approval of therapies for remission induction and maintenance does not usually involve trials dedicated to taper and discontinuation. Rather patients are maintained in long-term safety extensions. The concepts of maintenance of effect and absence of rebound might be further clarified. If this refers to continued therapy, absence of rebound could be equated with prevention of flare. Could 'long-term' be more clearly specified?	Accepted. Not any longer a requirement

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Line 578	1	Comment: It is said in the guidance that "**For an agent used for both induction and maintenance** an additional 1 year to 2 years are needed after achieving the remission for observing the maintenance of the effect. For the maintenance only claim a 1 year period is reasonable." The reason why in the first case an additional 1 to 2 years after achieving remission would be needed whereas in the second, only 1 year would be sufficient, is not clear. The two should be aligned, bearing in mind product specificities. Proposed change: "***For the maintenance only claim and*** f**or an agent used for both induction and maintenance** an additional 1 year ~~to 2 years are~~ is needed after achieving the remission for observing the maintenance of the effect. ***Longer observation periods may be needed depending on product specificities*** ~~For the maintenance only claim a 1 year period is reasonable~~."	Text has been revised for clarity
578-580	4	Comment: The guidance states that for an agent used for both induction and maintenance, an additional 1 year to 2 years are needed after achieving the remission for observing the maintenance of the effect. For the maintenance only claim a 1 year period is reasonable. Clarification on why in the first case, an additional 1 to 2 years after achieving remission is needed whereas in the second, only 1 year would be sufficient, is needed. We recommend that the two should be aligned, keeping in mind product specificities.	Text clarified

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Proposed change: **For the maintenance only claim and** for an agent used for both induction and maintenance, an additional 1 year **is** needed after achieving the remission for observing the maintenance of the effect. **Longer observation periods may be needed depending on product specificities.**	
Lines 593 - 596	1	Comment: As part of the recommendations regarding specific assessment instruments for paediatric patients, we recommend to include a discussion about the need to adapt visual scales (e.g. VAS) for age-appropriate interpretation, especially as the Guidance recommends studies in children as young as 5 years of age (Line 625). We propose the following after Line 596. Proposed change: ***Patient Reported Outcome (PRO) measures may require validated adaptations for use in paediatric subjects, such as incorporating symbolic visual anchors and/or age-appropriate text in symptom rating scales (e.g., visual analogue scale, Faces Pain Rating Scale).***	Text modified as proposed
Lines 597 - 602	1	Comment: Reference should be added to the Draft Guidance to provide context on the domains for evaluation to assess overall response to therapy in	References has been deleted from the final guideline

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		juvenile-onset SLE Proposed reference: Gutierrez-Suarez R, et al. A Proposal for a Paediatric Version of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index Based on the Analysis of 1,015 Patients With Juvenile-Onset Systemic Lupus Erythematosus Arthritis & Rheumatism (2006). 54(9): 2989-2996. http://onlinelibrary.wiley.com/doi/10.1002/art.22048/pdf	
Line 610	1	Comment: Adolescent patients would be able to complete the assessment. It is thus suggested to modify the statement as suggested in the article of Ruperto et al 2011b describing the PRINTO criteria and assessment tools. Proposed change: 4. ***Patient's/***Parent's global assessment of the overall patient's wellbeing.	Text modified as proposed
Lines 614 - 615	1	Comment: It is suggested to discuss other endpoints for consideration as secondary end-points, as appropriate. Please consider adding the following after Line 615:	Text modified in line with the suggestion made

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		Proposed change: ***<u>Other composite indices that account for both improvement and worsening of disease manifestations in different organ systems, such as the SRI or BICLA, may be appropriate secondary efficacy endpoints, if adequately validated in pediatric patients.</u>***	
Lines 618 - 620	1	Comment: It would be useful to clarify the use of adult efficacy data to extrapolate to paediatric efficacy. The Guideline (Lines 587-590) acknowledges some differences between juvenile-onset SLE and adult-onset SLE, and therefore, clarification as to whether reference adult data can be viewed collectively (e.g., pooled data across studies) and/or by specific gender distribution would be welcome. Additionally, it would be useful to clarify whether adult studies can be looked at collectively regardless of baseline renal or CNS status for consideration of extrapolation?	There is no universally accepted way. Both collective and specific approaches are probably needed depending on the indication.
Lines 625 - 626	1	Comment: It would be useful to clarify the wording. Proposed change : “Safety cannot be extrapolated ***<u>from adult studies</u>***; however, it is not realistic to accumulate sufficient information on safety in pre authorisation studies in children.”	Accepted - text modified

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
Line 629	1	Comment: The guideline requests that long term post authorisation studies and patient registries are established to evaluate safety in paediatric populations. However, the paediatric population is already scarce and difficult to recruit. It would be preferable to have a study or a registry rather than both. Proposed change: "Long term post authorisation studies ~~and~~ ***or*** establishment of patient registries are necessary."	Agreed and revised
629-630	4	Comment: The guidance requests that long term post authorization studies and patient registries be established to evaluate safety in pediatric populations. Knowing that the pediatric population is scarce and difficult to recruit, it would be preferable to have either a study or a registry rather than both. Proposed change: Long term authorization studies **or** establishment of patient registries are necessary.	Accepted – text amended
Line 634	1	Comment: 'Available data' is a very general term and implies that 100% of study analyses must be repeated based on individual age subgroups for a sponsor to comply with the guidance. If this is not the intention, then	Text amended to clarify Study data should be reported

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		the wording of this text should be modified accordingly to clarify the scope in which analysis should be repeated for these patient groups as there may be cases that the number of subjects for an age category is too low for a meaning full analysis. Suggest adding "When appropriate" to the beginning of the sentence Proposed change: "***When appropriate***, ***a***~~A~~vailable data should be reported separately for patients aged 65-74, 75 and older."	separately for patients aged 65-74, 75 and older where available.
634	4	Comment: The term 'available data' is a very general term and could imply that 100% of study analyses must be repeated based on individual age subgroups for a sponsor to comply with the guidance. This may not be the intention of the guidance. We recommend the wording be modified to clarify the scope in which analyses should be repeated for the specific patient groups. There may be situations where the number of patients within a specific age category is too low for a meaningful analysis. Proposed change: **When appropriate,** available data should be reported separately for patients aged 65-74, 75 and older.	Partly accepted – wording modified: While onset of SLE is generally between the ages of 15-45 years, the improved survival of patients with SLE over the last 20 years and in addition cases of late onset SLE means that older patients should be included in clinical trials of adult SLE. Study data should be reported separately for patients aged 65-74, 75 and older where available
Lines 651-655	5	Comment: The final guideline should note in this section that drugs focused on	Accepted - added

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		addressing serious and life-threatening aspects of lupus, such as lupus nephritis, and that have strong treatment effects, will appropriately be subject to flexibility as regards the extent of the required safety database. We encourage the Agency to participate in additional dialogue in the design of specific studies, particularly regarding those aspects that relate to clinical safety evaluation. If relatively small numbers of subjects are enrolled, a pragmatic approach to long-term safety will be needed at authorisation to protect patient safety while providing access to a new treatment, particularly a treatment that has a strong clinical response. Proposed change: Revise line 652, "The safety database to be submitted for assessing a new product should comply with the corresponding guidelines, **but the weight of evidence should be considered on a case-by-case basis in the context of severity of disease and treatment effect. If relatively small numbers of subjects are enrolled in the development program, a pragmatic approach to long-term safety will be needed at authorisation to protect patient safety while providing access to the new treatment. For a product that has other indicated uses and is already marketed, routine pharmacovigilance will supplement the existing safety profile.** For substance groups..."	
Lines 659 - 660	1	Comment:	Accepted – text modified

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		The statement "Importantly long term data to assess the development of related malignancies should be provided." should indicate that this can be the subject of post-authorisation measures and the risk and action plan should be included in the RMP at time of submission. Proposed change: "Importantly ***a risk management plan should be submitted that includes measures for provision of*** long term data to assess the development of related malignancies ***post authorisation*** ~~should be provided~~"	Amended text Importantly, **a risk management plan should be submitted than includes measures for provision of** long term data post authorisation to assess the development of related malignancies.
659-660	4	Comment: The guidance states that long-term data to assess the development of related malignancies should be provided. We recommend that this requirement be the focus of post-authorization measures and the risk and action plan should be included in the RMP at the time of product submission. Proposed change: Importantly, **a risk management plan should be submitted than includes measures for provision of** long-term data to assess the development of related malignancies **post-authorization.**	Accepted- text modified Importantly, **a risk management plan should be submitted than includes measures for provision of** long term data **post authorisation** to assess the development of related malignancies.
664-672	2	Comment: Why are only some abbreviations included in section 8? Also not all	Accepted – admin check through GL to be done

Line number(s) of the relevant text	Stakeholder number	Comment and rationale; proposed changes	Outcome
		abbreviated terms have been written in full in the document when first referred to (e.g. ANA has not been explained on line 100).	