INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic disease characterized by protean manifestations often demonstrating a waxing-and-waning course. While in the past a diagnosis of SLE often implied a decreased life span due to internal organ system involvement or to toxic effects of therapy, recent improvements in care have dramatically enhanced the survival of SLE patients with the most severe and life-threatening manifestations. Unfortunately, current treatments for SLE remain inadequate as many patients have incompletely controlled disease, progression to end stage organ involvement continues, and current therapies carry potential risks of debilitating side effects. Therefore, it is important to create an environment that will foster the development of novel therapeutic agents that potentially will be more effective and ideally less toxic.

Although many patients with SLE exhibit symptoms that involve the skin and joints, other symptoms of SLE vary widely between patients. There is no single demonstrated biological mechanism to explain the varied manifestations of disease. It is challenging to develop indices of disease response to therapeutic intervention. Disease activity scores allowing a comparison of disease severity in SLE patients whose disease affects different organ systems have been developed but may not always perform optimally as response measures in clinical trials. Using a variety of techniques, several different groups have developed validated indices that have now been shown to reliably measure disease activity in SLE patients in varied settings. Some of these indices have been shown to mirror the assessment of experienced clinicians and to sensitively measure changes in disease activity. One has also been demonstrated to predict the need for alterations or intensification of therapy; thus, these indices may be able to play an important role in future clinical trials of novel agents.

Although there are indices that measure disease activity in SLE, it is uncertain whether they will clearly delineate measurable important clinical responses to therapy in all situations. It is not certain that all agents with therapeutic benefit in SLE would lead to improvement in these measures. Some treatments may target a biologic mechanism that underpins some lupus manifestations, or only those related to a single organ system. This guidance addresses claims of improvement in overall activity of the disease SLE, as well as claims of improvement in organ-specific manifestations of SLE such as lupus nephritis. It is important that any therapy that claims to improve disease in one organ system not worsen disease elsewhere. The primary outcome measure selected for a given trial in SLE, should assess other aspects of the disease process, as it may be informative about the overall risk-benefit assessment (see RISK-BENEFIT ASSESSMENT). In this situation, the appropriate use of disease activity measures may be very useful.
I. MEASUREMENT OF LUPUS DISEASE ACTIVITY AND CLINICAL OUTCOMES

A. Disease activity indices

The clinical measurement of disease activity in SLE involves an assessment of either the presence or absence of the characteristic signs and symptoms of disease and the results of laboratory parameters. Recent discussions within the scientifically invested academic and clinical community of investigators interested in this disease have identified a series of important measures to be applied as outcomes and, taken together, are probably the critical assessments to be measured within a clinical trial. These include:

- A measure of disease activity,
- A measure of disease-induced damage,
- A measure of therapy induced damage,
- A measure of response as determined by the patient “a patient global response”,
- A measure of health related quality of life.

These should be measured either as co-primary outcomes for response or incorporated into a response index. Decisions regarding therapy are based on patterns of stable, increasing or decreasing disease activity. To measure disease activity in studies with groups of patients with varying manifestations of lupus, indices of disease activity that attempt to correlate these results with assessments of panels of expert clinicians have been developed. These indices identify disease manifestations using predefined criteria based on the presence or absence of different aspects of the disease, or in one measure, on the clinician’s assessment of the need to change therapy. Studies have attempted to validate these measured indices with regard to: the concordance of scores with expert opinion; inter-observer variability; correlation between individual patients’ scores on different indices; correlation between scores and changes in disease activity; correlation between increases in scores and clinical decisions to increase therapy.

Some of the available instruments have been validated in cohort studies as reflecting change in disease activity but not in prospective randomized clinical trials: the SLE Disease Activity Index (SLEDAI); the British Isles Lupus Assessment Group (BILAG); the SLE Activity Measure (SLAM); the European Consensus Lupus Activity Measure (ECLAM); the Lupus Activity Index (LAI) and the National Institutes of Health SLE Index Score (SIS) (ref - Strand V, Gladman D, Isenberg D, Petri M, Smolen J, Tugwell P, Outcome measures to be used in clinical trials in systemic lupus erythematosus, *J Rheumatol* 1999 Feb;26(2):490-7).

The BILAG has also performed appropriately by allowing the clinician to determine whether or not to increase or decrease therapy based on response. There is disagreement among lupus investigators about the appropriate weights that should be accorded to individual components of these scales and how to apply them if at all as responder indices.
When applying these disease activity indices to clinical trials, care must be taken to ensure that these measures accurately assess disease activity. If improvement is noted in one disease manifestation, it would be important to determine that other disease manifestations do not significantly worsen. Careful training of investigators is essential to ensure uniform scoring. Definitions of disease manifestations and levels of disease severity should be clearly specified. If there is a lack of reproducibility of these measures from clinician to clinician, it may seriously impair the interpretability of the trial results.

It is important to note that there are situations where changes in scores do not accurately reflect a change in disease activity. Misleading conclusions may result from applying some of the currently available indices that are transitional instruments, i.e. they score positively with the new onset of disease manifestations but not with persistent disease. The SLEDAI, for example, gives a positive score for new onset of CVA, seizure, cranial nerve disorder, rash, alopecia or mucosal ulcers while continued disease activity in these organ systems does not give a positive score (Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH and the Committee on Prognosis Studies in SLE. Derivation of the SLEDAI: A Disease Activity Index for Lupus Patients. Arthritis & Rheumatism, 35:630-640, 1992). This method of scoring would produce an undesired result in a clinical trial when the SLEDAI score has positive values in the initial visit for disease in the affected organ systems, and on subsequent visits the SLEDAI score would decrease, despite persistence of disease. While drawbacks do not preclude the use of an instrument with these characteristics, indices that produce a straightforward readout of disease activity are preferred. If a clinical trial indicates a better outcome from a therapeutic agent than control, the results of the study would be scrutinized carefully to ensure that the apparent benefits of the study drug are not an artifact of the scoring system. Although the SLEDAI is discussed in detail, similar concerns exist for a number of the other disease activity indices.

In applying disease activity indices to clinical trials of SLE, the interpretation of changes may be confounded if organ system dysfunction is present and attributable to a concurrent non-SLE illness, or organ dysfunction occurs and is attributable to therapy given for the disease. It is unknown how great an impact these concerns will have on the interpretation of clinical trial results. The results of ongoing and future trials may help assess whether further refinements of the currently available instruments are required.

The clinical course of SLE is characterized in many patients by periods of relatively stable disease followed by flares of disease activity. A problem with relying on scores to measure disease activity in trials is that measuring disease activity at fixed time points may miss flares of disease activity in between the times that measurements are performed. Definitions of flare have been proposed and applied longitudinally to patient populations. In one study, rates of flare were measured at an average of 0.6 flares per year [Petri M, Genovese M, Engle E, Hochberg M. Definition, incidence, and clinical description of flare in SLE. A prospective cohort study. Arth. Rheum. 1991; 34:937-44.]. A definition of flare should be shown to validly measure episodes of increased disease activity and correlate with a clinically determined need for increase in treatment. Definitions of major flare might include the requirement for initiation of high-dose glucocorticoid therapy or the institution of change in dose of immunosuppressive therapy, or the need for hospitalization or the occurrence of death. Important differences in the frequency of flares may exist based on gender, menopausal status, treatment and other patient characteristics.
There has been considerable interest in the development of a responder index to measure response to therapy on an individual basis. Some proposed definitions of a responder specify a minimum improvement in a measure of disease activity with no worsening in other aspects of lupus. A responder index would allow a clinical trial to determine directly what proportion of patients had a clinically meaningful improvement from therapy. Such a responder index should be assessed for intra-subject variability, content validity, and sensitivity to change to be fully validated. Full validation would also include a demonstration of the ability to discriminate treatment with a known active agent compared to an inactive control in a clinical trial. For example, a candidate responder index would accurately categorize a patient who experienced general improvement in many aspects of disease with mild worsening in one. Application of responder indices in prospective studies will help determine the utility of these measures in clinical trials.

B. Damage

Patients suffering from lupus experience irreversible damage to internal organ systems due directly to the abnormal biology associated with the active disease. Accumulation of damage occurs over a period of years while there may be associated therapy-induced damage as well. An index of damage was proposed and validated as the Systemic Lupus Erythematosus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index. Validation studies show that high scores on the SLICC/ACR Damage Index are predictive of increased mortality and damage in the renal and pulmonary components are associated with poor outcomes [Stoll T, Seifert B, Isenberg, DA. SLICC/ACR damage index is valid, and renal and pulmonary organ scores are predictors of severe outcome in patients with SLE. Br. J. Rheumatol. 1996; 35: 248-54.]. The prognostic information derived from SLICC/ACR Damage Index scores suggests they may be useful as stratification variables for clinical trials. The SLICC/ACR Damage Index measures only changes that have been present for at least 6 months, therefore only longer term clinical trials could demonstrate reduction in the rate of progression of damage using this measure. Some of the components of the SLICC/ACR Damage Index are measures of toxicity related to current treatment modalities. Use of the SLICC/ACR Damage Index as outcome measures in clinical trials could be complicated if a new therapy were associated with toxicities not measured by the Damage Index.

C. Organ specific indices

Organ specific measures of disease activity provide another approach to the study of therapeutic efficacy in lupus. This approach recruits a more homogeneous population of patients compared to the disease activity index approach, although it is recognized that patients will often have more than one organ system involved. Powering such a study may be problematic. Many patients with lupus are not given therapy for increase in disease activity, but for disease affecting specific organs, such as involvement of the CNS, the kidney, the lung, the skin, the joints or other organs. Responder indices could be applied in clinical trials by determining which organ system or systems have been most problematic for each enrolled subject, then measuring if subjects demonstrate improvement in those organ systems using prespecified criteria, such as components of validated disease activity indices. A responder measure of this type has the advantage of addressing the particular disease manifestations of most concern for individual patients. Interpretation of a clinical trial using this type of responder index could be problematic.
Lupus nephritis has been the most commonly studied organ specific manifestation of lupus. The presence of glomerulonephritis identifies a subset of lupus patients who may progress to end-stage renal disease (ESRD) and may have an increased mortality. Patients with severe lupus nephritis are often treated with high doses of immunosuppressive agents, including cyclophosphamide and high doses of corticosteroids. The outcome of lupus nephritis has improved markedly in recent years with 5-year survival rates of 90% or greater and 10-year survival rates of more than 80% reported (Urowitz MB, Gladman DD. Evolving spectrum of mortality and morbidity in SLE. Lupus. 1999; 8: 253-255). Measurement of renal disease in SLE clinical trials requires knowledge of the histologic description delineating the extent of inflammation or scarring, because the outcome and clinical features vary markedly among the various WHO categories of lupus nephritis.

A variety of outcome measures have been used in clinical trials of lupus nephritis. Mortality is an important outcome measure, but low mortality rates and necessary long observation times make mortality a relatively insensitive measure in clinical trials. Measures of renal function are used as outcome measures, including progression to ESRD, sustained doubling of serum creatinine, creatinine clearance, iothalamate clearance etc. The protein/creatinine ratio in urine may serve as an indicator of the need for further assessment of renal function with a study of creatinine clearance. The use of the doubling of serum creatinine is the most validated of these measures as it has been shown to reliably predict long-term renal outcomes; however, it is confounded by its insensitivity. Although less validated and requiring careful collection, the use of the protein/creatinine ratio followed by estimation of changes in GFR utilizing creatinine or iothalamate clearance, would be preferable in that, in the acute circumstance, these measures appear to be more reliable. Confounding variables (Boumpas DT, Balow JE. Outcome criteria for lupus nephritis trials: a critical overview. Lupus. 1998; 7: 622-629) may complicate interpretation of renal function measures, including serum creatinine, creatinine clearance. The sponsor should consider a measure that is clinically validated and provide data to support that choice. Changes in urinalysis may provide important information for the assessment of renal inflammation in lupus nephritis. The presence of cellular casts and hematuria, when measured accurately, are considered sensitive indicators of the level of activity of lupus nephritis. Central laboratories may be unreliable in assessing the presence of casts as they may break up during transport. Major flares of lupus nephritis, as assessed by urinary sediment, proteinuria and renal function, are an outcome measure. Patients who experience nephritic flares characterized by a nephritic sediment and an increase in serum creatinine or decrease in GFR may be at increased risk of developing a persistent doubling of serum creatinine. Renal remission in response to therapy has been defined as a return of an elevated creatinine and proteinuria to normal levels and normalization of nephritic sediment. Patients who fail to normalize an elevated serum creatinine in response to therapy may have an increased risk of progression to renal failure (Levey AS, Lan SP, Corwin HL, Kasinath BS, et al. Progression and remission of renal disease in the Lupus Nephritis Collaborative study: Results of treatment with prednisone and short term oral cyclophosphamide. Ann. Int. Med. 1992; 116: 114-123.). Assessment of proteinuria is particularly important in patients with membranous glomerulonephritis; however this is a less common form of lupus nephritis. Increases in proteinuria in patients with other forms of glomerulonephritis may not translate into unfavorable long-term outcomes.
D. Quality of Life/Fatigue

Health-related quality of life measures should be included in all trials of SLE. Instruments that assess health status and health-related quality of life can measure aspects of SLE and its impact on patients that are not fully assessed by other outcome measures. Trials demonstrating improvement in a specific organ or in disease activity should demonstrate no or minimal worsening in measures of quality of life. Patients with active SLE may have increased disability as assessed by the Health Assessment Questionnaire (HAQ) or modified Health Assessment Questionnaire (MHAQ). Health-related quality of life has been assessed in lupus patients using a number of generic instruments including the HAQ, MHAQ, Arthritis Impact Measurement Scale (AIMS), the Medical Outcomes Survey Short Form-20 (SF-20) and Short Form-36 (SF-36). Differences compared to controls have been observed in several domains and subdomains. Some instruments do not adequately assess fatigue, an important symptom for many lupus patients. Specific instruments have been studied for assessment of fatigue, (e.g. the Krupp Fatigue Severity Scale (KFSS)). Health-related quality of life instruments used in clinical trials of SLE should undergo validation regarding content validity (inclusion of all relevant domains), construct validity, sensitivity to change and other criteria. The use of these outcomes is a critical component to understand both the efficacy of an agent as well as its potential adverse events. If the measure does not improve with a specific therapy, it should not worsen.

E. Serologies

Serologic markers play an important role in the assessment of disease activity in SLE, including assessment of anti-double-stranded DNA, complement levels, and others. Serologic markers are critical for understanding the pathogenesis of disease. Serologic markers have an imperfect correlation with disease activity and cannot substitute alone for a direct assessment of clinical benefit. Serologic marker data should be studied in clinical trials, and in conjunction with clinical measures may play a role in assessing clinical outcomes and identifying potential clinical benefit from new therapies. Long-term clinical outcome studies after registration may help to demonstrate clinical benefit associated with changes in serologic markers (see section V: Surrogate markers as endpoints).

III SLE CLAIMS

There are a number of claims that may be considered for SLE. Organ specific claims or clinical remission/complete clinical response are the most straightforward to define from a purely clinical standpoint. However, additional claim(s) may be considered. A consensus needs to be reached as to whether a claim for “reduction of signs and symptoms of SLE” or a “reduction in the constitutional aspects of SLE” is most appropriate and should be included in this document. A “reduction of signs and symptoms” is meant to designate a change in overall disease status or activity, may be measured conventionally by a disease activity index such as SLEDAI, SLAM, BILAG etc., and may include any and all manifestations of disease. A “reduction in constitutional aspects” as defined here means improvements in the following manifestations: arthritis, rash, fever, fatigue, and serositis, but would not include major organ system
involvement. Major organ system involvement would be covered by an organ specific claim (and
would be studied specifically in each organ). In either instance, a measure of health-related
quality of life should be performed. In addition, a question to determine the patient’s assessment
of their clinical response should be designed. This “patient global” assessment will allow for an
overall determination of the “state of the patient,” which may help infer whether other aspects of
the disease have improved or worsened as well identifying possible drug induced adverse events.
In addition, the agency is considering a claim that would require a meaningful change in a
health-related quality of life measure that has been validated in SLE (e.g. SF-36) in the context
of a positive improvement in a question that would reflect the state of the patient (patient global
analysis), and a concomitant measure of disease activity that could be one of the presently
available disease activity indices or some other measure, and once validated would be
appropriately applied. (The committee is asked to specifically discuss and provide feedback to
the Agency as to the most clinically meaningful claims and most important, the appropriate use
of measures to establish these claims.)

This document proposes the following claims for SLE: (1) Reduction in the constitutional
aspects of the disease; reduction in the signs and symptoms of lupus; (2) indicated for the
treatment of lupus involving a specifically identified organ, for example, lupus nephritis; or (3)
complete clinical response/remission. These proposed claims are discussed in the paragraphs
below.

A. Reduction in constitutional aspects of SLE/ Reduction in Signs & Symptoms of SLE

This claim is intended to reflect the demonstration of a benefit in reducing the signs of disease
activity in SLE as well as in reducing the associated symptoms. As part of this claim, changes in
skin disease, joint involvement, fever, weight loss, and serositis would be considered. SLE is a
disease of long duration, with a waxing and waning course; therefore this claim would ordinarily
be established by a trial of at least 1 year in duration. For products that may elicit the formation
of antibodies, the duration of the clinical trial should be adequate to assess whether antibodies
are formed and if they adversely effect efficacy and safety. Methods to assess the activity of
disease over the duration of the study are preferable to methods that measure disease activity
only at the beginning and end. As part of any trials in support of this claim, measures of damage
and health-related quality of life should be included. A patient global assessment should also be
determined. Acceptable outcome measures to demonstrate a reduction in signs and symptoms of
SLE include:

1) A validated disease activity index

A disease activity index could be a measure to demonstrate that treatment was associated with a
decrease in overall disease activity during the course of the study. Careful consideration should
be given to the optimal choice of comparator arm (placebo vs. active control, see Trial Design).
For example, the SLEDAI, the SLAM, the BILAG, the ECLAM or other validated index could
be utilized to measure disease activity.

Illustration: A success in a 1-year trial could be defined as a decrease in the area under the
curve for monthly measurements of SLEDAI scores. It could also demonstrate that changes are
clinically meaningful.
2) **A validated measure of flare**

A validated definition of flare could be used in a trial to demonstrate a decreased frequency of, or decreased severity of, flares. Currently, no measure of flare has been fully validated.

**Illustration:** A success could be defined as a decrease in the time-to-flare or as a decrease in the number of flares over the course of a 1-year trial.

**B. Effectiveness in the treatment of a specific organ** (for example lupus nephritis)

Trials intended to study clinical benefit for specific organ systems could enroll subjects with disease affecting a single organ system (e.g. lupus nephritis). Patients enrolled with multiple organ systems identified as the major clinical concern can be stratified for the different organ systems for randomization and analysis. The definition of a response should be specified for each organ system under study.

**Illustration:** Trials of patients with disease activity affecting specific organ systems can define success as an increase in the proportion of responders receiving study drug than in controls.

Trials designed to assess efficacy of a product for the treatment of lupus nephritis would be expected to demonstrate an improved outcome for patients with biopsy-proved severe glomerulonephritis (WHO grades III or IV), or membranous glomerulonephritis. Short-term benefits may not reliably predict long-term outcomes, therefore trials of lupus nephritis are expected to be at least 1-year in duration. The following outcome measures could establish efficacy in lupus nephritis:

1. **Incidence of mortality and progression to end-stage renal disease (ESRD)**

2. **Sustained doubling in serum creatinine or other measure that has been validated including approximations of GFR such as iothalamate clearance or creatinine clearance studies**

Doubling of serum creatinine has been shown to be associated with progression to ESRD, so a decrease in the proportion of subjects meeting this endpoint in the treatment group compared to controls could be interpreted to define a patient benefit. Lesser degrees of change or changes in other measures could be considered. Similarly a significant change in GFR, which has clinical importance, can be considered. Sponsors should provide data to demonstrate that these changes are associated with benefit or a significant reduction in the rate of progression to ESRD.

**Illustration:** A success in a trial using this outcome measure would be defined as a decrease in the proportion of subjects whose serum creatinine attains a level double that of the baseline value and remains doubled for at least 6 months. Alternatively, a success in a trial that demonstrates a sustained change in GFR, such as preventing a fall in GFR of 50%, or demonstrating a rise of GFR by 50%, can be considered.
3. A validated surrogate marker for lupus nephritis

21 CFR 314, subpart H (Accelerated Approval of New Drugs for Serious or Life Threatening Illnesses) and 21 CFR 601 subpart E (Accelerated Approval of Biological Products for Serious or Life Threatening Illnesses), provides for FDA approval of drugs intended to treat serious and life-threatening diseases. Approval is based on the effect on a surrogate marker, provided specific criteria are met, and there is a commitment to define the actual clinical benefit of the agent in studies completed after approval. Demonstration of marked and sustained improvement in renal function and renal inflammation in a seriously affected population of patients with lupus glomerulonephritis would qualify for consideration under these regulations, provided that the measure of improvement was previously demonstrated as associated with improved patient outcomes. Sponsors are urged to consult with the relevant FDA staff before embarking on a clinical program based on these regulations.

Use of the accelerated approval pathway for a product for lupus nephritis, for example, would necessitate the timely completion of studies of long-term clinical outcomes post marketing.

4. Induction of renal remission

Active lupus nephritis is associated with evidence of renal inflammation, including cellular casts, proteinuria, and decreases in renal function. Serious lupus nephritis is frequently treated with cyclophosphamide and high doses of corticosteroids, agents that are associated with an increased risk of significant toxicity. A treatment that induces a sustained remission in lupus nephritis would confer a clinical benefit. Clinical studies of lupus nephritis use varied definitions of renal remission, but generally specify decreases in hematuria and cellular casts, decreases in proteinuria, and stabilization or improvement in renal function. A clinical trial intended to demonstrate induction of renal remission would specify a definition of renal remission that includes all relevant parameters. Evidence supporting an association with improved clinical outcome (e.g. decreased likelihood of developing end-stage renal disease or need for dialysis) should be provided to support the selected definition of renal remission.

5. A reduction in the number of renal flares

A validated definition of flare could be used in a trial to demonstrate a decreased frequency of, or decreased severity of, flares. Currently, no measure of flare has been fully validated.

Illustration: A success could be defined as a decrease in the time-to-flare, or as a decrease in the number of flares over the course on a 1-year trial.

Trials to demonstrate effectiveness in the treatment of specific organs should include measures of damage and health-related quality of life. Ideally these measures should improve in an important fashion.

C. Complete clinical response/remission

Complete clinical response/remission claim applies to products that demonstrate the ability to induce a clinical response, characterized by the complete absence of disease activity for at least 6
consecutive months. This response is termed complete clinical response if the subjects continue to receive lupus-directed therapies. Remission occurs if subjects were receiving no ongoing therapy for their SLE. A trial in support of the claim of Complete Clinical Response would be at least 12 months duration and demonstrate an increase in the proportion of subjects in whom a disease activity measure achieves zero.

Claims using the organ specific approach may be either for the treatment of each organ studied (for example, lupus nephritis), or for the treatment of lupus, depending on the numbers and types of organs studied. This would also require that there would be no worsening in terms of a patient global assessment as well as health-related quality of life.

IV TRIAL DESIGN AND ANALYSIS

Careful consideration should be given to choosing endpoints that will accurately assess the clinical benefits of the product when designing a trial for SLE. This may involve a focus on one aspect of disease (e.g. lupus nephritis) over other important aspects. Adequate information should be collected about other aspects of disease to adequately assess the overall risk-benefit ratio. Clinical trials in SLE generally are expected to collect information about disease activity; irreversible damage due to SLE and its treatment; and valid health-related quality of life measures. Serologic studies may also provide important information about the mechanism of action of the product under investigation.

A. Phase 2 trials

Phase 2 trials are used to better define dose and exposure-related activity and toxicity of products under development. The safety of concurrent use of a new product with commonly used concomitant therapies should be established. Many of the outcome measures under consideration for trials of SLE have not been tested in large-scale randomized trials. Some outcome measures may prove less sensitive than expected. Unexpected confounding variables may complicate the interpretation of trials using these endpoints. These are reasons for careful consideration in phase 2 trials to ensure validation of clinical outcome measures used in confirmatory phase 3 trials.

B. Efficacy trials

1. Disease activity trials

The chronic nature of SLE and its waxing-and-waning course requires clinical trials intended to show a decrease in disease activity and be of sufficient length to assess the durability of benefits. A trial of 1-year duration is typical. One endpoint that measures the effect on disease activity is to compare the scores on a disease activity index at the outset and endpoint of the trial in subjects taking the new product, with those of subjects taking the control regimen. Another measure of disease activity is to use an area under the curve (AUC) dimension at regular intervals throughout the trial. This may more accurately measure disease activity during the study than at a single time point. Trials provide analyses of both landmark and time-weighted averages, to
better define measures of efficacy. A trial showing a treatment effect demonstrates a larger
decrease in the disease activity scores in the treatment arm compared to control. Several
confounding factors could complicate the interpretation of such a trial. First, many SLE patients
have frequent low scores on disease activity indices, but experience intermittent flares of disease.
If a new product decreases the frequency and severity of disease flares, but has only a small
effect on background disease activity, this may not be reflected in a clinical trial that measures
disease activity only at the end of the trial. Another confounding factor is the likelihood that
subjects who flare during the trial will be treated with additional medications (e.g. corticosteroids) potentially reducing their disease activity scores for reasons unrelated to the
study drug.

Another measure of decrease in disease activity is to assess the incidence of disease flares during
the course of a clinical trial. This type of trial might use a validated measure of SLE flare as its
primary outcome measure. As not all SLE patients experience flares in a given time frame, the
size and duration of the trial must be adequate to capture a sufficient number of flares in the
treatment and control groups to assess a decrease in the treatment arm. Collection of complete
information on concomitant medications is essential to ensure that a difference in the number of
SLE flares is attributable to the study drug. A trial that shows a treatment effect of study drug
demonstrates a decrease in the number of flares, or in the time to flare, in the treatment group
compared to control.

Considerations in determining the appropriate regimen for the control arm of a trial in SLE are
important. No subject should be denied recognized effective treatment for aspects of the disease
that may lead to irreversible harm. A potential design consistent with this principle randomizes
subjects to the addition of placebo or study drug to a generally acceptable standard of care
regimen. This seeks to demonstrate that disease activity is decreased in the treated subjects. A
study could randomize subjects to the receipt of a known active agent or the study drug, then
assess if there is a larger decrease in disease activity in subjects receiving the new product. It
may be appropriate to include early escape provisions for subjects who worsen on study to
ensure that no subject is denied potentially effective therapy.

2. Lupus nephritis trials

Diffuse proliferative (WHO class IV) and severe focal proliferative (WHO class III)
glomerulonephritis in patients with SLE who have measures of inflammatory activity and
damage is associated with increased long-term risk of progression to end-stage renal disease and
high mortality. Severe lupus nephritis is commonly treated with high-dose immunosuppressive
regimens including cyclophosphamide and high-dose corticosteroids. These regimens are based
on non-prospective cohort studies that suggest a decrease in the long-term risk of progression to
end-stage renal disease. There is a need for additional regimens as current treatments may be
highly toxic and not effective in all subjects.

Studies that demonstrate improvement in mortality in lupus nephritis could document the
efficacy of a new product. Studies using mortality alone as the primary outcome may be
insensitive indicators of clinical benefit. Recent data shows that mortality is low in the majority
of patients. A study of mortality in lupus nephritis should be of adequate duration to document
benefits.
A study demonstrating a decrease in progression to end-stage renal disease would clearly document efficacy in lupus nephritis. Such a study should be years in duration as progression to end-stage renal disease occurs slowly. Another possible approach uses a doubling of serum creatinine as the primary outcome measure, based on studies that indicate a doubling of creatinine as highly correlated with subsequent progression to end-stage renal disease. As doubling of serum creatinine is a surrogate marker of clinical benefit, studies using this as an outcome measure should carefully collect information regarding progression to end-stage renal disease as well as in follow-up, to directly document the clinical benefits of the new agent. Validated changes in GFR can also be used as a measure of disease progression. Other possible measures that would prospectively predict disease progression should be validated.

After a diagnosis of lupus nephritis is established, disease activity is assessed by examination of the urinary sediment and by measures of renal function. Various measures of remission of lupus nephritis have been used to define patients with a substantial response to treatment, including measures of renal function, urinary sediment, and proteinuria. Attainment of remission is defined in terms of laboratory assessments, and patients with renal remission can be expected to experience a clinical benefit to the extent that they are: a) spared treatment with potentially toxic agents; and b) spared from ultimate progression to end-stage renal disease. Sponsors proposing to use attainment of renal remission to demonstrate efficacy of a product for lupus nephritis are encouraged to discuss their clinical development plans with the responsible review division at the agency. Proposals for clinical trials using renal remission as an endpoint would be expected to: a) provide a clear definition for renal remission and data supporting the choice of that definition; b) provide evidence that attaining a renal remission would be expected to translate into a clinical benefit to the patient; c) assess the durability of the renal remissions.

An increase in the frequency and severity of flares of lupus nephritis is correlated with worsening outcomes. Efficacy could be established by a reduction in the number of flares during a specific time period. Proposals for clinical trials using renal flare as an endpoint would be expected to: a) provide a clear and accepted definition for renal flare, and data supporting the choice of that definition; b) provide evidence that reducing renal flare incidence, by that definition of renal flare, would translate into a clinical benefit to the patient; c) assess the durability of the renal benefit.

Consideration should be given to the use of concurrent medications, including ACE inhibitors and anti-hypertensive agents, levels of blood pressure, and control of diabetes, for studies of lupus nephritis. If urinalysis is used as a measure of active inflammation, the investigator should demonstrate reproducibility and validation of the methods used and the results.

3. Other organ specific claims:

Responder measures for each organ system studied can be proposed, or organ specific measures from a validated disease activity index can be used. A responder measure of this type has the advantage of addressing the particular disease manifestations of most concern for individual patients and might provide a more homogeneous population for analysis. Trials can consist of a single organ or might involve more than one organ, with stratification by organ. Therefore it may be possible to study several individual organs within a single trial. The advantage of this approach is the ability to carry out large clinical trials while maintaining the homogeneity of the
population studied. Stratification by extent of organ damage at baseline may be advantageous. Consider restricting baseline glucocorticoid use to reduce the variability seen in studies that may introduce bias and make interpretation of results more difficult. Clinically important outcomes should be defined for each organ system and composite endpoints can be considered. In disease activity trials, multiple time points should be measured and may improve efficiency of the trial.

A successful trial may demonstrate a statistically significant number of clinical remissions in the treated group vs. the control group. Trends for improvement in each organ system can then be identified. The interpretation of a clinical trial using this approach could be problematic if improvement in the organ system measured were counterbalanced by worsening in other manifestations of lupus. If changes in treatment regimens are required, such as an increase in immunosuppressive agents, the results in the designated organ might be confounded.

For organ-specific trials, 3 to 6-month studies may be appropriate for those therapies considered remittive (induction therapy), with a longer term follow up for safety and durability of response. Maintenance therapy studies as short as 1 year can be considered.

Appropriate outcome measures in organ specific trials include: 1) maintenance, not worsening of, disease activity in the designated organ; 2) partial response; 3) complete response, still on medications; 4) complete remission; 5) flares (time to flare and/or numbers of flares); 6) total corticosteroid dose (defined as dose at the end of study and AUC). If corticosteroid dose is chosen as the endpoint, use of flexible dosing vs. forced tapering should be addressed. The potential need for rescue medication should be addressed in the analysis plan.

The organ specific measures, in an organ specific approach, call for additional data, including disease activity indices, damage indices, HRQOL, patient and physician global assessments, and toxicities.

Clinical trials of new therapies, both organ specific and not, may use add-on studies, or head to head comparisons with standard of care. Corticosteroids plus cyclophosphamide compared to corticosteroids plus new drug to demonstrate efficacy is an example. However, careful determination of baseline disease activity at cohort inception should be accomplished.

C. Safety trials

Studies to demonstrate the improved safety profile of a new drug compared to standard therapy can also be considered. These trials should be of adequate duration to establish efficacy as well as a clinically meaningful benefit in terms of safety. Steroid sparing agents should demonstrate not only that reduction in steroid use is statistically significant, but also that these reductions translate into an improved safety profile. Powering a trial to demonstrate improved safety may be problematic in lupus, although studying a collection of adverse events may help in this regard. Other trial designs can be considered but it is recommended that these be discussed with the appropriate reviewing division before initiation.
V. SURROGATE MARKERS AS ENDPOINTS

Surrogate or early markers of disease activity to assess efficacy in lupus trials can be considered. Surrogate endpoints should be proposed and validated for the treatment under study. Approval can be based on this validated surrogate endpoint. If the surrogate is not validated, but appears to be associated with a clinical benefit, subpart H or E of 21 CFR Part 314.500 can be invoked with a phase 4 study to demonstrate the clinical benefit required. Trends toward clinical improvement that are supported by improvement of the surrogate marker will be considered during the review process. As an example, early approval can be considered if both a measure of clinical disease activity as well as a surrogate marker improves. Additional efficacy as well as safety data can then be collected after approval to support the continued marketing of the product.

Surrogate markers may be laboratory studies involving biological markers or pathological changes identified in the organ under study. For example, a sustained doubling of serum creatinine has been proposed as a surrogate marker for the clinically important outcomes of end-stage renal disease (ESRD), and the need for dialysis or renal transplantation. Validated changes in creatinine clearance or iothalamate clearance can also be considered as surrogates for ESRD. Other markers might include, but are not limited to, T and/or B cell profiles, as assessed by flow cytometry, autoantibody subsets, and immune complexes, which are specifically defined, presence or absence of procoagulants, complement or its products. “Proof of concept” studies can be useful to support subsequent designs leading to consideration of approval. Sponsors can consider measuring the effects of a study drug against the effect of true placebo on T and/or B cell profiles in short term trials to determine a measure of potential efficacy, possible dose, and treatment duration for subsequent study in pivotal trials for approval.

The ability of the surrogate endpoint to predict clinical outcomes will be weighed against the risks associated with treatment. Sponsors are urged to consult with the relevant FDA staff before embarking on a clinical program based on surrogate endpoints.

VI. RISK-BENEFIT ASSESSMENT

Approval is predicated on controlled evidence demonstrating efficacy and an acceptable risk-benefit assessment. Assessment of risks and benefits includes an appraisal of the effects of the product on all aspects of the disease process, including disease activity, irreversible damage due to SLE and its treatment, and quality of life [ref]. The size of the safety database at approval should be consistent with the recommendations made by the International Conference on Harmonisation (ICH), but particular attention should be paid to the assessment of known toxicities, or to suspected pharmacologic effects that might imply delayed toxicities. The recommended size of the safety database may be lower for orphan indications, as it may be impossible or impractical to study large numbers of subjects. Although SLE is not an orphan indication, there may be subsets of patients with specific manifestations of SLE who represent an
orphan population indication. Sponsors may wish to discuss these issues with the appropriate
FDA staff early in the development of a new treatment. Sponsors with questions about the
expected size of the safety database should consult with the appropriate review division for
advice. Finally, if there is concern about rare but serious adverse events (e.g., from the
mechanism of action or experience with similar agents), a phase-4 commitment would be
appropriate to enable additional safety information to be gathered post marketing.

VII  LUPUS AND PHARMACOKINETICS

A. General
There have been few pharmacokinetic studies done in a prospective manner in the lupus
population. The bulk of the pharmacokinetic experience in these subjects has been anecdotal in
nature. Patient enrollment in pharmacokinetic studies should reflect the population for which
the drug is intended. As women represent the primary population afflicted with lupus,
enrollment in pharmacokinetic studies should incorporate a preponderance of women. Due to
the multi-symptom and body system nature of lupus, subjects enrolled in pharmacokinetic trials
for lupus should have the organ system involvement to obtain organ specific recommendations.

B. Special Studies
A characteristic feature of lupus is the associated change in the kidney, both structurally and
functionally. These kidney changes make it difficult to determine whether the standard renal
transplant model is adequate for the assessment of declining renal function in the lupus patient.
It is recommended that separate pharmacokinetic trials be conducted in lupus patients with
varying degrees of proteinuria to assess the impact on drug disposition and binding.

C. Drug Interactions
Drug interaction trials should be conducted with those agents commonly used in the treatment of
lupus. The potential for interactions with hormonal contraceptives should be assessed. These
assessments can include either in vitro or in vivo methodologies or a combination. The reader is
directed to the published FDA guidance's on in vivo and in vitro drug interaction studies.