FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

MEETING OF THE

ARTHRITIS ADVISORY COMMITTEE

8:00 a.m

Tuesday, September 30, 2003

Versailles Ballroom
Holiday Inn
8120 Wisconsin Avenue
Bethesda, Maryland

# ATTENDEES

COMMITTEE MEMBERS:

H. JAMES WILLIAMS, JR., M.D., Acting Chair
Department of Internal Medicine
Division of Rheumatology
University of Utah School of Medicine
50 North Medical Drive
Salt Lake City, Utah 84132

KIMBERLY LITTLETON TOPPER, M.S.
Executive Secretary
Advisors and Consultants Staff (HFD-21)
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Building 5630, Room 1093
Rockville, Maryland 20857

JENNIFER J. ANDERSON, PH.D.
Department of Epidemiology and Biostatistics
Boston University School of Medicine, A-203
715 Albany Street
Boston, Massachusetts 02118

LEIGH F. CALLAHAN, PH.D.
Department of Medicine
Division of Rheumatology
Thurston Arthritis Research Center
3330 Thurston Building, CB#7280
University of North Carolina
Chapel Hill, North Carolina 27599-7280

JOHN J. CUSH, M.D.
Rheumatology and Clinical Immunology
Presbyterian Hospital
8200 Walnut Hill Lane
Dallas, Texas 75231-4496

SUSAN M. MANZI, M.D., M.P.H.
University of Pittsburgh
School of Medicine
S722 Biomedical Science Tower
3500 Terrace Street
Pittsburgh, Pennsylvania 15261

ATTENDEES (Continued)

COMMITTEE MEMBERS: (Continued)

WENDY McBRAIR, R.N., M.S., C.H.E.S.
Consumer Representative
Director
Southern New Jersey Regional Arthritis Center
Virtua Health
1 Carnie Boulevard
Voorhees, New Jersey 08043

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

GRACIELA S. ALARCON, M.D.
University of Alabama
Division of Clinical Immunology & Rheumatology
830 FOT
510 20th Street South
Birmingham, Alabama 35294

JILL P. BUYON, M.D.
Department of Rheumatology, Room 1608
Hospital for Joint Diseases
301 East 17th Street
New York, New York 10003

JOHN C. DAVIS, M.D.
University of California, San Francisco
Department of Medicine-Rheumatology Division
533 Parnassus Ave., Room U-383, Box 0633
San Francisco, California 94143-0633

BETTY DIAMOND, M.D.
Department of Microbiology and Immunology
Albert Einstein College of Medicine
301 East 17th Street
Bronx, New York 10461

MARY ANNE DOOLEY, M.D.
University of North Carolina at Chapel Hill
Department of Medicine
Division of Rheumatology and Immunology
3330 Thurston Boulevard
Chapel Hill, North Carolina 27599-7280

ATTENDEES (Continued)

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

MICHAEL FINLEY, D.O.
Western University
College of Osteopathic Medicine
309 East Second Street
College Plaza
Pamona, California 91766-1854

ALLAN GIBOFSKY, M.D.
Cornell University Medical College
425 East 79th Street
New York, New York 10021

BEVRA H. HAHN, M.D.
UCLA School of Medicine
Department of Medicine
Division of Rheumatology
1000 Veteran Avenue
Rehab Center Room 32-59
Los Angeles, California 90095-1670

JOHN HARDIN, M.D.
Albert Einstein College of Medicine
Department of Medicine
1300 Morris Park Avenue
Forshheimer 713
Bronx, New York 10461

GARY HOFFMAN, M.D., M.S.
Chairman, Department of Rheumatic and
Immunologic Diseases
The Cleveland Clinic Foundation A/50
9500 Euclid Avenue
Cleveland, Ohio 44195

GABOR G. ILLEI, M.D.
National Institutes of Health
National Institute of AMS
Office of the Clinical Director
9000 Rockville Pike, Building 10, Room 9S205
Bethesda, Maryland 20892

ATTENDEES (Continued)

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

NORMAN T. ILOWITE, M.D.
Director, Division of Rheumatology
Division of Pediatric Rheumatology
Schneider Children's Hospital, CH197
269-01 76th Avenue
New Hyde Park, New York 11040

MATTHEW LIANG, M.D. (by teleconference)
Department of Medicine
Division of Rheumatology/Immunology
Harvard Medical School
Brigham and Women's Hospital
75 Francis Street
Boston, Massachusetts 02115

JOAN T. MERRILL, M.D.
Clinical Pharmacology Research Program MS 22
Oklahoma Medical Research Foundation
825 Northeast 13th Street
Oklahoma City, Oklahoma 73104

DAVID PISETSKY, M.D., PH.D.
Division of Rheumatology and Immunology
Duke University Medical Center, 151G
Durham VA Hospital
508 Fulton Street
Durham, North Carolina 27705

DANIEL J. WALLACE, M.D., F.A.C.P., F.A.C.R.
Wallace Rheumatic Study Center
8737 Beverly Boulevard, Suite 301
Los Angeles, California 90048

MICHAEL H. WEISMAN, M.D.
Division of Rheumatology
Department of Medicine
Cedars-Sinai Medical Center
8700 Beverly Boulevard, Suite B-131
Los Angeles, California 90048

ATTENDEES (Continued)

FOOD AND DRUG ADMINISTRATION STAFF:

JOEL SCHIFFENBAUER, M.D.
Office of Drug Evaluation V
Division of Arthritis, Analgesic, and
Ophthalmic Drug Products
5600 Fishers Lane, HFD-170
Rockville, Maryland 20852

JEFFREY SIEGEL, M.D.
Office of Drug Evaluation VI
5600 Fishers Lane, HFM-582
Rockville, Maryland 20852

LEE SIMON, M.D.
Office of Drug Evaluation V
Division of Arthritis, Analgesic, and
Ophthalmic Drug Products
5600 Fishers Lane, HFD-170
Rockville, Maryland 20852

ALSO PRESENT:

BILL FREIMUTH, M.D., PH.D.
KATHLEEN ARNTSEN

C O N T E N T S

Systemic Lupus Erythematous Concept Paper

* * *

P R O C E E D I N G S

(8:00 a.m.)

DR. WILLIAMS: We welcome you all to this session of the Arthritis Advisory Committee meeting. I'm Jim Williams and I've been asked to act as chair today.

We'd like to begin by introducing the members of the committee, and we'll start with Richard and move around this way.

DR. LOONEY: I'm John Looney, University of Rochester, rheumatologist.

DR. HARDIN: John Hardin, Albert Einstein College of Medicine, Division of Rheumatology.

DR. DOOLEY: Mary Anne Dooley, University of North Carolina, Chapel Hill, dermatologist.

DR. ALARCON: Graciela Alarcon, University of Alabama at Birmingham, rheumatologist.

DR. PISETSKY: David Pisetsky, rheumatologist, Duke University.

DR. GIBOFSKY: Allan Gibofsky, rheumatologist, Hospital for Special Surgery, Cornell.

DR. HOFFMAN: Gary Hoffman, rheumatology, Cleveland Clinic.

DR. ANDERSON: Jennifer Anderson, statistician, Boston University.

DR. WILLIAMS: Jim Williams, rheumatologist,

University of Utah.

DR. CALLAHAN: Leigh Callahan, outcomes researcher, epidemiologist, University of North Carolina, Chapel Hill.

MS. McBRIAR: Wendy McBriar, Director of Arthritis Services, Virtua Health, consumer rep.

DR. MANZI: Susan Manzi, rheumatologist, University of Pittsburgh.

DR. ILOWITE: Norman Ilowite, pediatric rheumatologist, Schneider Children's Hospital and Albert Einstein College of Medicine.

DR. DAVIS: John Davis, rheumatologist, University of California, San Francisco.

DR. DIAMOND: Betty Diamond, Albert Einstein College of Medicine.

DR. BUYON: Jill Buyon, New York University School of Medicine, Hospital for Joint Diseases, rheumatologist.

DR. WALLACE: Dan Wallace, rheumatologist, Cedars-Sinai, UCLA.

DR. SIEGEL: Jeff Siegel, Division of Clinical trials, FDA.

DR. SCHIFFENBAUER: Joel Schiffenbauer, FDA, Division of Analgesic, Anti-inflammatory, and Ophthalmic Drug Products.

DR. SIMON: Lee Simon, rheumatologist and Director of the same division, FDA.

DR. WILLIAMS: We'll ask Kimberly Littleton Topper to read our conflict of interest statement.

MS. TOPPER: The following announcement addresses the issue of conflict of interest with respect to this meeting and is made a part of the record to preclude even the appearance of such at this meeting.

The committee will discuss the proposed systemic lupus erythematous (SLE) concept paper, a preliminary discussion for creating a guidance for development of drugs, biologics, and devices for the treatment of SLE. The committee will also discuss the section concerning clinical trial design.

The topic of today's meeting is an issue of particular matter of broad applicability. Unlike issues before a committee in which a particular product is discussed, issues of particular matters of broader applicability involve many industrial sponsors and academic institutions.

All special government employees have been screened for their financial interests as they may apply to the general topics at hand. Because they have reported interests in pharmaceutical companies, the Food and Drug Administration has granted general matters waivers of broad

applicability to the following SGEs which permits them to participate in today's discussions: Drs. Jill Buyon, Betty Diamond, Mary Anne Dooley, R. John Looney, Susan Manzi, Joan Merrill, Daniel Wallace, and Michael Weisman.

A copy of the waiver statements may be obtained by submitting a written request to the Freedom of Information Office, room 12A-30 of the Parklawn Building.

Because general topics could involve so many firms and institutions, it is not prudent to recite all potential conflicts of interest, but because of the general nature of today's discussion, these potential conflicts are mitigated.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants' involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firms whose products they may wish to comment upon.

Thank you.

We also have a person connected by telecon. Dr. Liang?

DR. LIANG: Yes.

MS. TOPPER: Would you introduce yourself, please?

DR. LIANG: I'm Matthew Liang, a rheumatologist from Harvard Medical School.

DR. WILLIAMS: Thank you.

We'll now turn the time to Lee Simon who will give us our charge and an overview.

DR. SIMON: Thank you and good morning and welcome to our second day. We certainly had an entertaining day yesterday, although quite demanding in both time and attention. I hope you all had a good night rest and a good dinner so that you could prepare and be fortified for the discussion this morning.

We discussed and reviewed some of the issues regarding pivotal trial design, looking at some of the questions that we entitled "state of the art" yesterday, and then we also discussed and reviewed the issue of claims, as well as the issue of surrogate markers and how they might be applied as pivotal approvals for accelerated approval programs with phase IV commitments.

What became clear to some of us yesterday was that we all need to remember in discussing today when we revisit some of the issues, particularly related to trial design, that there are differences between the issue of regulatory approval and clinical practice. I cannot

underline how important it is for us to think in the context of regulatory approval and not how we practice medicine. Although it is nice when they are congruent, it is not required that they be congruent. The bar for regulatory approval cannot be set in a way that it is impossible to achieve and it is not necessarily standard of care.

I remind you all that the ACR-20 in its applicability to rheumatoid arthritis is not a very high bar. It was created at a time when the best we had were IM gold, not well studied, and nonsteroidal anti-inflammatory drugs. The reality is we're not in a dissimilar position today. Although we might want to have the ACR-50 presently be the bar for approval in rheumatoid arthritis, that is only because we've had the ACR-20 which allowed us to see the discriminate ways that drugs behave between what we achieve with the ACR-20 and what we might want to achieve with the ACR-50.

Of course, we all want to cause remission and to cure our patients, but we are very nascent in this particular arena and we need to remember what that bar needs to be so that we can actually precipitate, engender, and interest interested people in wading into the field.

Under those circumstances, I implore you and ask you to think about that as we discuss the trial design

issues and what it would really take for approval. So I ask you to think about the issues of pivotal approval. What we're looking at here is not phase I and phase II trials, although that is important, and in fact, we will talk a little bit about those issues because those are issues that decide dose and proof of concept and how one wants to look at certain issues in phase III. But it's the phase III design which actually is sent to us not in exclusion of the totality of the evidence, but it is the phase III designs that we use to determine whether or not approval will be awarded.

So certain things happened yesterday that we became confused about, and I'd like to highlight those and ask us to think about them as we go through the trial design discussions led by Joel Schiffenbauer and then the discussions afterwards.

The first that we are not clear about is the issue of signs and symptoms. We discussed the issues of lumping and splitting yesterday, but I'm still not sure and we're still not sure whether or not signs and symptoms are something that we want to pursue a la the signs and symptoms of lupus and you get approved for that. And it's not clear what the components of this indication would be. What would you have to prove to achieve that particular indication if in fact it should stand? And how would we

measure that?

In that context, there was a long discussion intermittently and repetitively about disease activity indices and their applicability. We became quite confused about that because some of us heard that a DAI could be a standalone and thus demonstrate overall disease activity and thus perhaps could be applicable for signs and symptoms.

But then we also heard that there's a hierarchy of the utility of these disease activity indices where BILAG seemed to be somewhat more flexible and better than SLAM and SLAM was somewhat better in certain circumstances than SLEDAI, but everybody seemed to have a different opinion about the SLAM and SLEDAI and where you would apply it and how it would be utilized.

Furthermore, we weren't sure that everybody concurred that perhaps there needed to be two disease activity indices used, not just one, although we heard that also repetitively through the day.

So I would ask us to think about that particular issue in trial design, and if that was the case, what would be the pivotal measure? What would be the primary measure? Would there be co-primaries or would there be one primary and one secondary and the secondary couldn't worsen? What would you have to win on to then win

approval?

Now, in the context of pivotal trial designs and pivotal measures for primary approval, we're unclear. We think we heard in a splitters' camp that whatever the sponsor would suggest, for example, the arthritis of systemic lupus, that that would distinguish it from systemic lupus. We heard that there was not a lot of enthusiasm for a drug to treat lupus as opposed to components of lupus, which may be a temporal issue. Perhaps we're not there yet that we're comfortable with understanding all of the biology of the disease, thus all of its manifestations, and we're not entirely sure that there is yet a drug that could address at the same time thrombotic issues, CNS lupus, nephritis, and the signs and symptoms such as arthritis and rash and fever all at the same time and thus getting the acronym, the treatment of systemic lupus.

So we'd like to reiterate and concur with you that in fact you do want to go the route of per whatever the sponsor wants and allow them to demonstrate what their measurements will be, determine what their methods of outcome would be, and if they win, they get that approval.

Then finally, in the discussion of surrogates and accelerated approval, we were not clear about what the outcome of that discussion was. Some of us heard that

there was enthusiasm for a composite outcome, perhaps for example, antibodies to double-stranded DNA in the context of proteinuria and an active urinary sediment and perhaps a change in urinary creatinine clearance that would not worsen, perhaps even improve, but certainly not worsen. And that, in association with a quality of life indicator and perhaps a disease activity index, could lead to an accelerated approval and then a phase IV commitment for clinical linkage.

We also heard that people were uncomfortable with the more traditional measures that people have used such as serum creatinine and that the length of time it would take to lead to change that was consistent and then showing differences in end-stage renal disease development. I remind the committee that the agency in the past has considered doubling of serum creatinine as a link to increased risk for end-stage renal disease. One of the reasons why that shows up in the document is because that's been a tried and true methodology of studying that particular patient.

We don't believe that that's actually a good temporal approach. It takes a long time, as had been mentioned in the open public forum, and we were looking for some other measures that would allow us to gain an understanding in a shorter period of time to allow the

sponsors to approach trials that would not last 2 to 3 years. We were hoping we could do something in 6 months to a year and then link that to a subsequent postmarketing study that might go on longer.

I don't know how you all think about that today because some of us heard that you were not enthusiastic about that either, that even in the composite approach, that you were a little uncomfortable with the implications of that.

We were charged yesterday by some of the other speakers to think about taking risks. In the context of safety, of course, we don't want to take too many risks, but at the same time, we need to be at a place in our development programs to allow the sponsors some latitude so that we can understand and learn about the disease, we can stimulate risk-taking in our colleagues in industry and otherwise, and perhaps learn something about this disease.

So I ask you all to take off a little bit of your clinicians' hats, put on a little bit of your trial design hats as we go into the next part of this discussion and think about trial design development, the implications of pivotal trial designs, the implications of primary outcomes, how to identify them and what we will do with them in the context of drug approval.

Thank you, Mr. Chairman.

DR. WILLIAMS: Thank you, Dr. Simon.

We'll now hear from Dr. Joel Schiffenbauer, and he'll be our first presenter.

DR. SCHIFFENBAUER: Good morning. The topic for this morning's discussion is trial design issues in lupus, and my name is Joel Schiffenbauer.

SLE is a disorder that may wax and wane with and without therapy, making determination of the efficacy and safety of new therapies difficult. The use of potentially toxic medication requires rigorous study design to demonstrate clear evidence of efficacy and safety. The challenge this morning is to present approaches about study design to hopefully address some of these concerns.

This is a list of the topics that I'm going to try and get through. I won't read through these, but let me just go right into the first topic, choice of endpoints.

The primary consideration in any efficacy trial design is what is the trial design to show and therefore the design will depend on the claims sought. So, for example, some of the endpoints that were discussed yesterday include an organ-specific endpoint, signs and symptoms, a flare endpoint, and then other endpoints such as steroid-sparing or surrogate endpoints.

I've listed here some of the advantages and disadvantages to these approaches. Some of this was

discussed yesterday, so I won't spend too much time going over it, but I'd just like to make a few points in this regard.

The first endpoint would be some measure of disease activity using a disease activity index. The advantages to this approach is that it allows a recruitment of adequate numbers of patients. However, a disadvantage that I don't think was mentioned yesterday is that there is potential for imbalance in disease manifestations in treatment and control groups based on analysis by indices, and that would be of concern in data analysis.

The second endpoint is a flare design. Again, that would allow recruitment of sufficient numbers of patients and may also reduce time of under-treatment or partial treatment. Again, it's problematic for analysis if flares differ in the treatment and control groups.

The third endpoint and perhaps the most straightforward is the organ-specific endpoint analyzing a single organ in a single trial. This allows for a homogeneous population as well as well-defined outcomes, but of course, may make recruitment of adequate numbers of individuals more difficult.

And lastly, I'd like to propose the organ-specific outcome but stratified by organ. So in this trial design, a single trial could recruit individuals with

renal, skin, joint disease, and have each organ stratified. This will tend to improve the power while maintaining the homogeneity of the two treatment groups. However, it may increase complexity of analyses.

Having decided on the approach, the next step would be to decide whether you want to look at individuals with active or inactive disease, and then under each of those headings, whether the individual is treated and active disease such as a partial or a non-responder or untreated and active disease such as an individual naive to any therapy. Likewise, for inactive disease, whether that's inactive due to treatment on some dose of steroids or inactive and untreated.

This will then determine the endpoints that will be considered for the trial. So for an individual with active disease, one could study a disease activity measure, either an index or organ-specific endpoint. One can look at a responder index, and in this regard an example I give is some combination of disease activity measure, health-related quality of life, damage, and steroid dose, and any other measures so desired. Or alternatively, a steroid dose or concomitant medication dose could be the endpoint.

For inactive disease, most likely the endpoint would be flare, either time to, number of, or rate of, or

again, it could be a steroid dose or concomitant medication dose.

Whatever endpoints are chosen, there are two questions that need to be addressed. What changes are considered clinically meaningful and what constitutes a successful outcome? And we'd ask the committee to address some of those concerns in the questions this morning.

I've tried to summarize everything I just said in this relatively simple two-by-two table. So across the top, I have the disease activity active or inactive, and across the side, the two basic outcome endpoint measures, organ-specific or signs and symptoms. So for a study designed to look at an organ-specific outcome in active lupus patients, the endpoints could be a disease activity measure specific for that organ, a responder index or a steroid dose, or if the study is designed to look at an organ-specific outcome in inactive lupus patients, a flare design or maintenance design, which would be similar to the flare design, or a steroid dose or steroid-sparing would be appropriate outcomes.

For signs and symptoms in active lupus patients, the outcomes could be a disease activity index of your choice or steroid dose, and for signs and symptoms in inactive lupus patients, a flare, maintenance, or steroid dose would be the appropriate outcome measures.

I'd like to spend a few slides just mentioning some issues about flare design, and some of these questions were addressed yesterday. But the question is, what reduction in flare rate would be considered clinically meaningful in the context of adverse events? Are all flares equal, renal versus joints as an example? We touched on this yesterday. And lastly, should a new therapy be asked to address the treatment of active disease, in addition to preventing flares? Again, we touched on this issue yesterday.

There are some advantages and disadvantages to the flare design, which I'd just like to briefly mention here. A flare design could be considered, in a sense, a responder analysis in that it takes into account the individual response. It also reduces time of partial treatment or under-treatment of the individual. However, there are some disadvantages to the flare design. One is the heterogeneous outcomes that may occur in the treatment and control groups. It also does not demonstrate treatment of active disease and in some cases may be impractical in that there are relatively few flares, and so trials may take a much longer duration.

I've given two examples in the next two slides of some flare definitions and there clearly are many others. We talked a little about the SELENA flare

definition yesterday, but these are just two examples that I'd like to give. The first is for a flare definition, an organ-specific, in this case renal, attributed to lupus by a treating physician which may require one or more criteria, and the two criteria I've listed here are a reproducible increase in serum creatinine greater than 20 percent, accompanied by proteinuria, hematuria, and/or red cell casts and/or white cell casts; or reproducible increase in 24-hour urine protein. The question is by how much.

The second definition would be considered a general flare definition, and this is defined as at least one of the following: an increase in prednisone greater than 5 milligrams a day for at least 14 days since the previous visit; an SLE manifestation requiring hospitalization; or an addition of new medication or an increase in the dose of an existing medication to specifically treat a manifestation of increased lupus activity.

Let me now move on briefly to data to collect in trials of lupus. Again, we touched on this yesterday. This is a listing of the domains that have been suggested to look at in any trial of lupus proposed by the OMERACT group. This is one of the publications, Lupus 2000, volume 9, page 322.

The first domain is a measure of disease activity which can either be the disease activity index or an organ-specific definition here.

The second domain is a measure of damage. The ACR-SLICC Damage Index measures overall damage, although damage can certainly be defined on an organ-specific basis. In either instance, one needs to determine the toxicity from the drug versus damage due to the disease itself.

The third domain is a measure of health status or health-related quality of life, and we discussed the use of the SF-36 yesterday.

Then lastly, the economic costs and adverse events.

I've listed here some of the sample data that may be obtained for a trial in lupus nephritis. First would be renal pathology, and the question, does everyone need a biopsy? We've touched on that also. Urine protein, urine sediment, some measure of renal function, whether it's serum creatinine or an appropriate measure of glomerular filtration rate. And the question is, what threshold of GFR would be important to study? Then lastly, other adverse events.

But the question remains, what data is needed, let's say, for a trial in central nervous system lupus. Would we require trials to include MRIs with or without

gadolinium, lumbar punctures with cerebral spinal fluid analyses, EEGs, or what? And then the question is, what data is needed for other manifestations? For example, in a trial looking at the skin manifestations, certainly skin biopsies would be easy to do and should be required. But what, for example, should we look at in pulmonary disease or in other manifestations?

Let me move now on to some other trial design issues, controls and standard of care issues. I've listed here, for those interested, a web site that you can go to to look up information about trial design. This is the fda.gov/cder/guidance web site, which many of you may be familiar with. I've listed here some of the sources of information that you can find.

The first is the ICH E9. ICH is the International Conference on Harmonization. It's a group of U.S. and international regulators that get together to propose harmonized standards for trial design and trial conduct. The first document is the ICH E9, statistical principles for clinical trials.

The second is ICH E10, choice of control groups and related issues in clinical trials.

I'd also refer you to the Rheumatoid Arthritis Guidance which discusses many of the same issues that we are going to be discussing this morning, and then hopefully

in the future, there will be some guidance related to lupus.

Lastly, I would refer you to the CONSORT recommendations published in Lancet 2001, volume 357. CONSORT is Consolidated Standards of Reporting Trials. These are recommendations really for reporting trials in journals, but they discuss many of the important issues in trial design.

So controls. Ideally a study would have placebo and that could either be a standard of care plus placebo versus a true placebo plus an active control plus a dose response. What this allows for is a measure of the absolute effect size, that is, comparing the new drug versus placebo. It shows existence of an effect. It shows a dose response and allows comparisons of new therapy versus the standard, a comparator.

In looking at lupus trials, there are basically two approaches, either the superiority trial or an equivalence or noninferiority trial. I've provided here two examples of a superiority trial.

So, for example, the first one is a standard of care which could either be, as an example, steroids plus cyclophosphamide plus a new drug versus the same standard of care plus placebo. In this case, one would need to show that the new drug is superior to placebo.

The second example is the standard of care, which in this case I've given as an example steroids, plus the new drug versus standard of care plus cyclophosphamide. In this case the new drug would have to be shown to be superior to cyclophosphamide.

Alternatively, one can consider the equivalence or noninferiority trial and the example here is standard of care plus new drug versus standard of care plus comparator. Now, in this case, the new drug should be shown to be equivalent to or noninferior to the comparator by a predefined margin or delta and the comparator must have been shown to be effective compared to placebo in previous trials. And I'll come back to equivalence trials in a few slides.

The other consideration is can there be a period of placebo therapy or steroids plus placebo. This would certainly depend on the organ studied and on the severity of the disease, but it's important to use this at the beginning of an active controlled trial to establish assay sensitivity, that is, to show that the new drug is superior to the placebo. The question in this regard is, are there instances where steroids only are an acceptable treatment in lupus nephritis? And we'll come back to that.

I'd like to mention briefly just two other trial designs. The first one is called the randomized

withdrawal design. In this trial, subjects receive test treatment for a specified time and are then randomly assigned to continue treatment with the test treatment or placebo. I'll refer you again -- you've heard about this -- to the New England Journal article 1991. This is the Canadian hydroxychloroquine trial which is a variant of this randomized withdrawal design.

The second design is a replacement study. So in this design, a new drug or placebo is added by random assignment to conventional treatment, which is given at an effective dose, and then the conventional treatment is withdrawn, usually by tapering. The outcome measure is looking at the ability to maintain the patient's baseline status or, in other words, preventing a flare. This approach would be useful for any agent that's considered to be a steroid-sparing agent.

Is there a standard of care? This, of course, depends on the organ studied. I've already asked the question for lupus nephritis. Are there instances where steroids only are acceptable? What is the standard of care for central nervous system disease? How about for other organs? The caveat is that if we insist on using cyclophosphamide in all instances, for example, of lupus nephritis, it may be difficult to demonstrate an effect of a new therapy especially if the mechanisms of action are

similar. So we'd ask you to consider that in the questions later this morning.

Just a comment about the concept of add-on trials, and I've provided a reference in Arthritis and Rheumatism 2003. This is an editorial by Martin Bois. It was in reference to add-on trials in rheumatoid arthritis, but many of the issues are the same.

The first is that add-on trials will be performed in individuals who are nonresponders or partial responders to therapy and we're adding on a new therapy. The first issue is how do we define a partial responder in systemic lupus erythematous? The second is with any new therapy, we'd like to understand the toxicity of that therapy, but in add-on trials, we're concerned now about toxicity of not only the new therapy but about combination therapy. So the recommendation would be for investigators to consider the use of a factorial design which basically looks at the various combinations of therapy.

I already mentioned something about equivalence or noninferiority trials. Again, this trial design involves comparing a new drug to a standard comparator, and again, the comparator must show historical evidence of sensitivity to drug effect based on prior placebo-controlled trials. You then predefine a margin of difference between the new drug and the comparator, and

this margin cannot be greater than the smallest effect size that the active drug or the standard comparator would be reliably expected to have, compared with placebo in the historical trial.

Let me briefly move on to issues about blinding. Blinding is intended to minimize potential biases resulting from differences in management of patients or interpretation of results. The question is then, can trials with IV cyclophosphamide or potentially any new therapy be adequately blinded, especially if there are changes in laboratory results, symptoms such as nausea, or signs such as hair loss?

I would refer you to an old article, 1971 Annals of Internal Medicine, volume 75, by Steinberg for its trial design. In that trial he assigned therapists and observers. So, for example, the therapist made changes to the dose of medication without knowing whether they were changing placebo or cyclophosphamide based on the white count; whereas, the observer did not know anything about the laboratory data and was responsible for determining the clinical status of the patient. Pharmacists prepared medications, so it was unknown what the individual was getting, and he actually gave all the patients that came into the trial wigs so the issue of hair loss did not come up.

Why blind? Subjects on active drug might report more favorable outcomes because they expect a benefit or might be more likely to stay in a study. Knowledge of treatment could affect the vigor of attempts to obtain on-study follow-up. Knowledge of treatment could affect decisions about whether a subject should remain on treatment or receive concomitant medication, which is a big concern in lupus trials. And knowledge of treatment could affect decisions as to whether a given subject's results should be included in the analysis. We've asked you, the committee, to comment on the issue of blinding in trials.

The next issue is data analysis. In data analysis, it's important to prespecify how missing data will be handled, especially in relatively small trials. The standard approaches have been the last observation carried forward or the worst observation carried forward, but certainly other conservative methods of imputation could be appropriate such as imputing placebo or treatment and treatment values for placebo.

Alternatively, one could consider the use of a responder index which would obviate the need for imputation of missing data, and this could include a response at any time, response at the last visit, or response at each visit. The use of a responder index may also be useful to maintain power but reduce sample size.

One could stratify by any number of factors. We already talked about stratification by disease manifestation, but one could also stratify by dose of steroid or others, with the caveat that too many stratification factors leads to too small numbers of individuals in different treatment groups and may make demonstration of efficacy more difficult.

Alternatively, one could do a covariate analysis on predefined covariates. I've listed just some, but there may be others, anti-DNA at baseline, number of organs involved or disease activity at baseline, by center, or in the future possibly by cytokine levels, IL-6 levels, complement levels, et cetera.

The issue of concomitant medications is a very important one. Certainly we need to define the allowable medications at baseline, but also we need to define medications that will be allowed during the trial, such as starting of ACE inhibitors.

We also need to address in trial design the issue of rescue medication. Do patients stay in the trial once they've received some form of rescue? How much rescue is allowed? If a patient is allowed to increase their prednisone by 5 milligrams per week, do they stay in the trial?

This is an important concern because subtle

changes in steroid dose could influence outcomes. Therefore, we should consider a run-in period to standardize the steroid dose. Dose adjustments should be specified in the protocol, and I think Dr. Liang will address this in more detail. Then lastly, whatever change in steroid dose we look at, if we use this as an endpoint, they must be clinically meaningful.

Duration of studies. Duration of studies may depend on the claims sought. I will refrain from using the constitutional changes, but change the question to mean could a trial for some manifestation of lupus be 3 months in duration rather than the 6 months or 1 year trial that we've usually considered? Trial duration in individuals with inactive disease could be just the time to collect adequate numbers of flares, however long that may be.

We've talked about trial duration in active disease, whether the indication sought is for acute or induction therapy versus maintenance therapy. Even in a case of induction therapy which might be identified within weeks to months, we need to consider the demonstration of maintenance or durability of effect, and so at some point a chronic or maintenance trial needs to be performed. This could be months or possibly even years, and it could take the form of either an extension study or a phase IV study.

There are some practical considerations. It

may be difficult to perform a chronic, well-controlled trial in lupus secondary to flares, changing medications, dropouts, and changes in medical practice. On the other hand, in a disease that waxes and wanes, short-term trials may not provide adequate demonstration of efficacy, safety, and importantly, durability.

As I said, extension trials could be used to demonstrate durability and safety, but considerations of extension trials -- and this question came up yesterday. Are comparators needed? Should these extension trials be blinded or open-label? And we've asked the committee to address some of these concerns. Or could the long-term trial be a phase IV commitment? How long should it be? I think that length depends on what needs to be demonstrated.

Lastly safety concerns. Again, I've provided some recommendations from the ICH group. 300 to 600 patients should be studied for 6 months and 100 for 1 year, but this is defined for a chronic, non-life-threatening disorder. What is the standard for a disorder as varied as lupus in which some manifestations are chronic and others acute and life-threatening? I think that this depends, at least in part, on the toxicity profile of the drug under study.

So the question, does one size or does one approach fit all? I think clearly the answer is no. I

hope what I've done this morning is present multiple possibilities for "wins."

These are just a summary of the concerns that I've discussed in determining trial design. Should it be an organ-specific versus non-organ-specific? Active versus inactive disease? Activity measure, whether it's a disease activity index or organ-specific or flare? Superiority versus equivalence trials? Induction or maintenance therapy? Short- and long-term safety? And the data to collect.

Lastly, I'd like to thank all the people who I've discussed these issues with and for their useful input.

I will turn the meeting back to the chair.

DR. WILLIAMS: Thank you, Dr. Schiffenbauer.

We now have the opportunity to hear from Dr. Matt Liang by teleconference. Dr. Liang?

DR. LIANG: Thanks very much. I hope you can hear me because all I'm hearing is a buzz with your voice very muted.

DR. WILLIAMS: We can hear you fine, Matt.

DR. LIANG: Great. I think that this builds on yesterday's presentation, and you should have the full manuscript that we have submitted to ANR on the subject. This was one of the three initiatives that the ACR asked

our committee to deal with. Unlike the material from yesterday, this did not go through the usual approval process and endorsement by the board. Nevertheless, we thought it was a valuable exercise and at least should be fuel for debate.

We tried to make explicit something that is maddeningly difficult and that is the use of steroids in SLE management. Many people yesterday talked about the treatment being worse than the disease sometimes, and I think that that 900-pound gorilla that everybody was referring to was steroids because steroids arguably are the dominant cause of latent morbidity and mortality. If there was any strategy that could reduce the amount of steroids that we almost always use in serious, life-threatening manifestations of lupus, that would be a blow for freedom.

In any case, I think the first slide is just the title, and the next slide is the sponsorship, which included many of the same organizations that funded the original project, with the exception of the Office of the Clinical Director where we received support in kind to complete the project.

What we tried to do in Dusseldorf with the attendees was to develop an explicit process to actually come up with a specific tapering schedule based on some assumptions about a design that could be used. We used a

technique for achieving consensus called the nominal group technique to define mutually exclusive, collectively exhaustive disease manifestations of SLE or the phenotype. We asked the participants one by one and until everybody was exhausted and could name no more manifestations. Presentation, where they as clinicians would use the most steroids to control the signs and symptoms, and we labeled this severe SLE. Then in another separate exercise, same process, we asked them to define the manifestations of lupus where they would be moderately severe, where they would use moderate doses of steroids to control the signs and symptoms. And the remainder, although we didn't discuss it, were viewed as mild, but not the real emphasis of the exercise.

Then we presented a randomization, withdrawal design or tapering design, and we asked each clinician to write, if they felt comfortable writing it, a prednisone taper schedule. What we're doing is basically presenting the descriptive statistics as a recommendation.

The next slide is "SLE Phenotypes." I doubt you can read this, but it's in the handout and it's also in the paper. We tried to do this by organ system. You can see that some manifestations might be very severe or moderately severe, so they could occur in all three categories technically. But these were the items that

people named in the nominal group technique. In all cases we assumed that on the ground, face to face with a patient, the clinician had excluded non-SLE causes for these manifestations.

The next slide I think would be the hypothetical study of how you might evaluate whether a drug A had steroid-sparing ability. I think I should just walk through this a little bit. So you take patients. They would be randomized into treatment A plus steroids or B plus steroids. Mind you, the assumption here is that it is unethical to have, in patients with very serious manifestations of lupus, a patient that was not treated with steroids to control the acute inflammatory manifestations.

In any case, after a patient has been given a dose of steroids to control these manifestations and the agent A or B, they would be either improved, same, worsened -- no. I'm sorry. There's a mistake here. But basically they would be improved, same, or worsened, instead of the "improved" in the last box. These would be built on either target organ a priori criteria which we talked about, but didn't present in detail, that would be explicitly defined or the deltas of the disease activity units that we developed with the exercise from yesterday.

At this point people who are worse would be the

basis of an analysis at that point, but if they were improved, they would begin a protocolized steroid taper. And then if you follow the patients subsequently, as both groups are given the standardized steroid taper, they could enter into one of the three states at the bottom of the slide.

I hope that's clear.

Here are the results from the attendees where we asked them to give us the initial dose for severe lupus, moderately severe lupus, or mild, and how they might give it, either orally or by bolus, and we've listed what the final results were from the participants who felt like they were experienced enough to make a vote, so to speak, and we also present the range. It, again, underscores the fact that reasonable clinicians, given approximately the same kind of data in a similar context of a protocol, have a tremendous variation in terms of what they would prescribe in their patients.

Now, this actually may be the solution to one dilemma that is frequently presented, and that is that patients and physicians are oft loathe to enter a trial where they're completely hampered by a paint-by-numbers steroid dosing. The range could be a way that a protocol could at least be explicit but allow some individualization for the patient and perhaps the physician as well.

We also asked the group how long you would try to maintain steroid doses to suppress inflammation, and we called that the induction period. You see in the row for severe SLE and moderately severe SLE, the duration of induction therapy that the participants prescribed, and then again how many weeks they would keep someone on steroids until they were completely off.

Now, the next slide is "Steroid Taper for Severe SLE After Induction Period." So for the most severe manifestations in which the clinicians said that they would use the most steroids in their therapeutic armamentarium, this was the tapering that was done by these 27 participants, and you can see the descriptive statistics. Again, the range might be incorporated into a protocol to allow a little bit of flexibility. We did this assuming prednisone milligrams per day for a 70 kilo lady.

Then my last slide is basically the same kind of information for the moderately severe SLE patient, and you can see the same kind of information.

It's interesting. This obviously was not an easy exercise to force clinicians to develop this. On the other hand, there -- I think this is interesting and informative. There were two committee members who felt that they couldn't really put their name on the manuscript, and both said that they did not want their names on because

they didn't agree with the tapering schedule, which is kind of interesting because I think this is what happens when you have reasonable clinicians assembled. They disagree but they sometimes can't allow themselves to be put into an exercise prescribing a tapering dose.

In any case, we thought the committee might be interested in this because the studies that have been done on the subject show that the steroid dosing, when you present clinicians scenarios, is less driven by what we might think, and that is the patient characteristics, than by the physician characteristics, length of training, their age, et cetera. This is, I think, the first explicit exercise where we actually have at least a database recommendation.

Thank you.

DR. WILLIAMS: Thank you, Matt.

We've now come to the open public hearing, and I have to read a paragraph here.

Both the Food and Drug Administration and the public believe in a transparent process for information-gathering and decision-making. To ensure such transparency at the open public hearing session of the advisory committee meeting, the FDA believes that it is important to understand the context of an individual's presentation.

For this reason, the FDA encourages you, the

open public hearing speaker, at the beginning of your written or oral statement, to advise the committee of any financial relationship that you may have with any company or any group that is likely to be impacted by the topic of this meeting.

For example, the financial information may include a company's or a group's payment of your travel, lodging, or other expenses in connection with your attendance at the meeting.

Likewise, the FDA encourages you at the beginning of your statement to advise the committee if you do not have any such financial relationships.

If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

We have some speakers who have requested time here, and the first will be Dr. Bill Freimuth. Dr. Freimuth, you have 10 minutes.

DR. FREIMUTH: Thank you for the opportunity to speak to the Arthritis Advisory Committee. My name is Bill Freimuth. I am the Senior Director of Clinical Research for Rheumatology, Immunology, Infectious Diseases at Human Genome Sciences, and I would like to present to you some aspects dealing with the issues of clinical development of a potential novel, new therapy for SLE called LymphoStat-B,

and I'd like to present this as a case study for the endpoints and issues of trial design in SLE.

I'm going to briefly review the biology of BLyS and the pharmacologic rationale and nonclinical and clinical data of LymphoStat-B, review its phase II trial design, and then deal with questions that our company and our investigators have been struggling with in trying to develop a clinical development plan for LymphoStat-B and particularly phase II trial designs and pivotal trials in the future.

BLyS simply stands for B-lymphocyte stimulator. It was identified in a high-throughput proliferation assay based on our genomics database. It is a member of the TNF family. It has multiple alternate names. It is biologically active in its soluble form as a 51,000 molecular weight homotrimer that is cleaved primarily for monocytes. It binds one of three membrane receptors on B cells, and particularly it acts as a survival factor by inhibiting B cell apoptosis, as well as it stimulates differentiation of B cells to immunoglobulin-producing plasma cells.

The rationale for developing a BLyS antagonist for SLE is based on both animal model data and human data. The mouse data links BLyS with autoimmune disease such that transgenic models of over-expressing BLyS develop an

autoimmune SLE-like phenotype, particularly glomerular nephritis. Genetic models of autoimmune disease such as MRL and NCBWF1 mice have elevated levels of circulating BLyS. And use of soluble BLyS receptors administered in these animal models have ameliorated the disease progression and improved survival.

In humans, elevated BLyS levels are evident in the serum of SLE and RA patients, and these BLyS levels have correlated with serum IgG and autoantibody levels, particularly anti-double-stranded DNA in lupus and Rheumatoid factor in RA.

This slide shows an example of the elevation of BLyS. The BLyS concentration is showed on this axis. The normal range is 2 to 10 nanograms per ml. And two cohorts of SLE patients and RA patients basically show that 30 to 40 percent of the patients have an elevation in BLyS, and strikingly, when one collects synovial fluid from RA patients, the average BLyS level is twofold greater than what is found in the plasma.

LymphoStat-B that we are developing is a fully human IgG1 lambda monoclonal antibody that's specifically recognizes and binds soluble human BLyS and inactivates its biological activity. To study LymphoStat-B in animal models, LymphoStat-B does not bind to murine BLyS but does bind to human and monkey BLyS. Therefore, to study

LymphoStat-B in mice, we had to give human BLyS which does bind to murine BLyS receptors and increases the spleen weight, splenic B cells and serum IgA. And when one adds LymphoStat-B, it will selectively inhibit the BLyS-induced effects.

An example of this is shown on this slide where on the y axis you see the serum IgA in the mouse, and if you focus on the yellow, when one adds four daily doses of human BLyS, one doubles the murine serum IgA. If one gives concomitantly during that 4-day period the control IgG, there's no effect on the increased BLyS levels, and when one gives increasing levels of LymphoStat-B from .5 to 5 milligrams per kilogram, one sees a significant reduction of the human BLyS-induced IgA back to the basal levels.

We have also studied LymphoStat-B for its activity and safety in cynomolgus monkeys, and in this case LymphoStat-B was well tolerated at doses up to 50 milligrams per kilogram given every 2 weeks for 6 months, plus an 8-month follow-up period. There were no study agent-related infections during the treatment and recovery period, and activity of LymphoStat-B was demonstrated in decreases in B lymphocytes in lymphoid tissue in the periphery. This was substantiated by flow cytometry, organ weights, and histologic findings with effects of a partial depletion of B cells. The PK was linear in the monkeys

with a terminal half-life of 11 to 14 days. And we will be presenting more of these results at the upcoming ACR meeting.

One example of LymphoStat-B's ability to reduce CD20 is shown in this slide. This is the percent baseline CD20 cells where all monkeys have their CD20 normalized to baseline. There was a 6-month treatment and 8-month recovery period. If you focus on week 26, one will see that at this time there was a 58 to 65 percent reduction in B cells. The depletion remained for 2 to 3 months and then gradually increased, so by 6 months after the last dose of LymphoStat-B, the B cells returned to their baseline.

We have recently completed a phase I clinical trial in LymphoStat-B where we have studied four IV doses, 1, 4, 10, and 20 mgs per kg, with a placebo in a randomized, blinded study giving LymphoStat-B either as a single dose or as two doses 21 days apart. Overall, the results showed that the drug was well-tolerated. There were no drug-related serious adverse events. There was no increase in adverse events or laboratory abnormalities compared to the placebo. And there was no increase in the incidence of infection.

The pharmacokinetics were linear suggesting a 14-day half-life, and biological activity was observed by a significant decrease in CD20 cells. And again, we will be

presenting the complete results at ACR.

We have recently obtained fast track designation from the agency.

More importantly and relevant to the discussion today is the phase II trial design, and this is just the basics of a very complex trial design, which is a multi-center, randomized, double-blind, placebo-controlled trial, dose-ranging with three doses of 1, 4, 10 mgs per kilogram. Some of the basic entry criteria are patients with active SLE, a SELENA SLEDAI greater than or equal to 4, and on stable medications. In other words, this is adding LymphoStat-B onto standard of care. A maximum of 350 patients and LymphoStat-B will be administered IV at day 0, 14, 28, and every 28 days for 1 year.

In this trial design, we have two co-primary endpoints. The first one is the SELENA SLEDAI activity at week 24. The second one is the time to first flare defined by the SELENA SLEDAI flare index over 52 weeks. The sample size was based on 80 percent power and a .05 alpha to detect in one of more of the active LymphoStat-B groups compared to placebo either a 25 percent absolute or a 100 percent relative improvement in the percent change from baseline score in SELENA SLEDAI at week 24. That is assuming a placebo 25 percent response and being able to detect a 50 percent improvement in one of the LymphoStat-B

arms.

The second co-primary endpoint was powered to see a reduction in the percent of subjects having their first flare by week 52 and reducing it from 65 to 43 percent.

We are also looking at a variety of major secondary endpoints that have been discussed at this meeting, including week 52 SELENA SLEDAI and BILAG scores, time to first flare defined by BILAG, reduction in steroid dose, area under the curve of SELENA SLEDAI and BILAG over 52 weeks.

In addition, we're studying a variety of biological markers, including autoantibodies, complement, and subsets of B cells and plasma cells in immunoglobulin subclasses.

Most importantly, the background I just gave you is to deal with the issues and questions that we as a company, trying to develop a new, novel therapy in SLE, have been dealing with in discussions with our investigators. These questions are: would an effect in either SELENA SLEDAI at 24 weeks or time to first flare over 52 weeks be an adequate basis to move forward to a confirmatory trial?

Which endpoint is thought to be more clinically meaningful?

Is the magnitude of effect being tested clinically relevant, and would a lesser effect also be clinically meaningful?

Are there other endpoints that would be preferred or considered more clinically meaningful than the ones described? For example, would significant benefit in one or more of the SLE organ system manifestations such as defined in BILAG be a relevant primary endpoint?

Would a sign steroid-sparing effect, with or without a positive trend in disease activity and/or flare, be a sufficient primary endpoint?

Which endpoint would be the most compelling as a primary endpoint in a pivotal trial is one of the key questions.

Lastly, several other clinical endpoints and markers of biological activity are being explored. Which of these are believed to be the most meaningful, and is there currently sufficient evidence to consider any of these biological markers reasonably likely to predict clinical benefit?

We think it is vitally important that the committee and the agency address these questions and others that were brought up in the last presentation to help guide us in the development of new therapies in SLE.

I thank you for your attention and look forward

to a lively discussion on trial design.

DR. WILLIAMS: Thank you, Dr. Freimuth.

Our next speaker will be Kathleen Arntsen. She's given 7 minutes.

MS. ARNTSEN: Good morning and thank you. My family paid for my expenses to come here and speak in honor of my birthday on Sunday. I am honored to be here and hope to enlighten you with my patient perspective written solely by me.

22 years ago I was diagnosed with SLE. The ongoing pain, overwhelming fatigue, and recurrent infections I have suffered since childhood finally had a name. I can tell you from firsthand experience that living with lupus is like swimming in shark-infested waters. The danger and uncertainty is always present and we are armed with nothing but our will to survive. We try to stay afloat while anticipating the next attack and remain ever-hopeful that a rescue ship will soon appear on the horizon. Existing treatments for lupus are totally inadequate, toxic, and cause detrimental side effects with long-term use. Many treatments being used are off-label if a physician is even willing to prescribe them. This profoundly disturbs me. Like most lupus patients, this disease cut me down in the prime of my life and has drastically impacted my future. It has stolen precious

time from me, as well as the opportunities to have a successful career, independence, financial security, or that of being a mother, just to name a few.

My complex medical picture includes multiple autoimmune disorders such as Sjogren's, PA, Graves, Raynaud's, APAS, psoriasis, and myasthenia gravis, as well as GERD, Barrett's, gastroparesis, colonic inertia, and MVP. I take 26 medications daily, costing $3,800 a month. I have endured decades of destruction and disfigurement from 22 years of constant glucocorticoid use and other treatments, and I used to weigh over 200 pounds. My entire digestive tract is impaired and it takes five different drugs to allow me to eat each day. I haven't eaten fruits or vegetables in six years now, and I suffer from constant colicky abdominal pain throughout the day and night. Colostomy seems to be imminent.

Like most lupus sufferers, I take each day at a time, trying not to think of the unpredictable course of this baffling ailment or the potency or long-term effects of the multitude of medications I absorb each day. My treatment is individualized, and during my most recent flare, my physician finally made the compassionate decision to try CellCept as a steroid-sparing agent. This drug has allowed me the ability to function for the past two-and-a-half years when I could barely think, walk, or raise my

arms above my head. No one should have to spend months in bone-gnawing, soul-wrenching pain, going from physician to physician begging for help. It is a desperate place to be.

For 18 years I have been a volunteer leader in a lupus foundation and have attended the ACR's and NIAMS' events as a patient advocate. I have learned to listen from years of hotline counseling and monthly support group facilitation. I am strongly committed to maximizing the quality of life for those affected by lupus by providing programs designed to empower patients to actively participate in their own health care to improve their disease outcome.

Like many patients, I have educated myself on my medical conditions, treatments, and tests. I am part of my treatment team and I play a major role in the decision making process, coordinating results between my physicians. I am copied on all tests and procedures and have 22 years of lab results entered into an Excel spreadsheet to assist my physicians and streamline my care.

I have been involved in research studies for lupus and gastroparesis. I was part of a phase III study for cisapride prior to its FDA approval and am presently enrolled in the ongoing safety study since it has been pulled from the market and I work very closely with my physician. I cannot eat without this drug and feel that it

is the only thing preventing esophageal cancer. I was a subject in a lupus Arava study and have participated in other studies. I deeply believe that a cure for this disease will be forthcoming from research, but we must urgently discover more preferable treatments and improve diagnostic techniques to give patients a better quality of life now.

I feel very strongly that patients should be more actively involved in the research trial process from its inception. Americans have evolved into informed consumers. The world of knowledge is at their fingertips through present technology. Although our agency services rural upstate New York and the majority of people residing there have little higher education, I can assure you that they are very astute shoppers. The time has come to revolutionize the way we view patients. They must be better informed and educated regarding research trials. Placing an informed consent document in their face and asking for a signature is not sufficient. There is a significant step missing in the trial process that should include an informative education session involving the patient and advocate of their choosing and a trial educator, for lack of a better title. Patients are overwhelmed enough when first presented with trial participation and not given sufficient time or material to

make knowledgeable choices. Even airlines give consumers 24 hours to make a decision before a commitment. Any patient who cannot make an informed decision based on information supplied should be eliminated as a trial candidate. If we raise the bar to new heights, as well as the patient expectations, they will meet the challenge. Empowering patients and giving them back some of the control they have lost with disease can only result in a more favorable outcome for all involved. Allowing a patient to be a partner in the process allows them to take ownership of the study.

In conclusion, I would like to share a compelling call with you that I just recently received. A 25-year-old woman was diagnosed with SLE in May, presenting with joint pain, fatigue, and pericardial effusion. She was placed on 40 milligrams of prednisone and Imuran and continued her studies in the local residency program. She then developed shortness of breath and was diagnosed with anti-cardiolipin, started on Coumadin, and a filter was placed in her vena cava.

In July she saw her rheumatologist, complaining of fever and fatigue, and was sent to her primary care physician who did a brief exam and sent her back to work. Shortly thereafter, she was admitted to the hospital with sepsis, bacteremia, and gangrene of the bowel. Emergency

surgery was performed to remove part of her bowel and cultures revealed a Gram-negative infection. Antibiotic therapy was started and she was diagnosed with pulmonary hypertension.

Her family, which included a physician, decided to move her to a major teaching hospital where she continued to fail. She was intubated, a Hickman port was inserted, and Flovan therapy was initiated for her PAH. She went into shock and her organs began shutting down. Kidney dialysis was started and gangrene presented in her extremities. Her arms and legs were then amputated from above the elbows and knees down. Just as her family decided to take her off the respirator, she rallied and her organs began to function again little by little.

She still believes that she can be a physician and her family does not have the heart to tell her otherwise at this point. This young woman came to America several years ago with the aspiration of being a physician and now, because of lupus, she has not only lost that dream but also her independence and any promise of a productive existence.

Please do not think that this situation is rare. Every minute of every day another person is struck down in the prime of their lives by this devastating disease, placed on immune-compromising, toxic drugs and

treated by physicians who are grasping to find some sort of balance in their care.

We must not be complacent in thinking that we have progressed in treating this disease. I passionately implore you to move forward on this document before one more patient loses another piece of themselves to this horrible predator. Please improve the quality of life for those suffering from lupus by expediting the development of efficacious treatments and restore our hopes, dreams, and promise. Remember, lupus ends with us.

Thank you very much.

DR. WILLIAMS: Thank you, Ms. Arntsen.

MS. ARNTSEN: Can I ask if there are any questions?

DR. WILLIAMS: No, there isn't. We don't take questions.

Are there any other participants who would like to speak in this open hearing?

(No response.)

DR. WILLIAMS: Seeing none, we will move on then to the discussion. We've been given 11 questions to discuss in an hour. So we will need to move fairly expeditiously.

The first question is, in the context of a trial looking at multiple organs, stratified by organ, and

the outcome is statistically significant across all organs, but each organ only shows numerical trends, does this provide adequate data for improvement in each organ? If you agree, over what period of time should this be studied? That's a rather complex question.

The committee looks like they are still looking for the questions. There were some left at your position this morning, plus they were an extension from yesterday. The one this morning was left at your position with the page open to it. The other one were the questions you received yesterday that started off with "State of the Art," and it's on page 3 from yesterday. It's on page 2 from today.

Let me read it one more time now that you've all found it. In the context of a trial looking at multiple organs, stratified by organ, and the outcome is statistically significant across all organs, but each organ only shows numerical trends, does this provide adequate data for improvement in each organ? If you agree, over what period of time should this be studied? Joan and then Jack.

DR. MERRILL: No, it does not provide organ-specific information. It provides what it provides, but it does suggest that it's an effective treatment for lupus.

DR. WILLIAMS: Jack?

DR. CUSH: I think the design would be flawed because the person is going after multiple organs. It sounds like what they're really going for is signs and symptoms and they achieved it in some global fashion, but that they missed on multiple organ systems. So again, you can go for signs and symptoms and you can go for major organ involvement. There should only be a few, I think, that we can well study at this point, which is renal and heme and articular and cutaneous and maybe neuropsychiatric. But that needs to be studied up front and powered appropriately up front. But to go and say globally you're going to take care of all organs for lupus in a trial design makes no sense.

DR. WILLIAMS: John Davis?

DR. DAVIS: First, I wanted to congratulate Joel and his group for their presentation. I thought it was very clear, concise, very thoughtful, and thought-provoking and gives us a good platform to go from.

The second, I agree with Joan that this definitely does not give any organ-specific indications for us.

But again, that leads me back to where we are in our drug development and the molecules we have and the pathogenic mechanisms that we understand. It would very much specifically depend on the drug that we were testing.

And if I were to accept this, I would require at least a 6-month time period.

DR. WILLIAMS: Allan?

DR. GIBOFSKY: Well, I concur with Dr. Merrill and Dr. Davis. I'm not quite sure what the questioner was trying to get at. I think that the information that we would get from this would largely depend on what the primary endpoints are predefined and prespecified to be. As for the time period, I think that too would depend on what we were studying.

DR. WILLIAMS: Joan and then Dan.

DR. MERRILL: I want to make it clear that I do think that that would be a legitimate trial design. I disagree with Dr. Cush because -- I hate to do this to everyone -- if you can take multiple people from a BILAG A to a BILAG C, that's compelling information that you have a drug that does work for quite a few manifestations of lupus. I have no problem with treating different organs at the same time. That's what we do in practice.

DR. WALLACE: I think that anything that looks at an organ has to -- you just can't say numerically. You have to say what is the anatomy of the organ. What is the physiology of the organ? How much damage is there to the organ? How reversible is it? It's very, very complicated. And what are the influences of other medications that

aren't anti-inflammatory such as blood flow to an organ?

DR. WILLIAMS: David?

DR. PISETSKY: I think there's something implicit here in that we have outcome measures for individual organ systems, and beyond BILAG it's not clear to me that we do. So we've been talking about we treat arthritis of lupus, and yet I don't know there are any guidances as to what the criteria for a response would be in the arthritis of lupus comparable to ACR response in RA. And then I think you keep falling back to something like BILAG, which is someone's decision to treat, and I think it might be difficult for this kind of trial design unless you specify beforehand what you would consider a response for these different organs.

DR. WILLIAMS: Bevra?

DR. HAHN: I thought we discussed this thoroughly yesterday, and I thought that the majority of the panel concluded that this is acceptable. So I'm a little confused going around again. I guess we still are split in decision.

The DAIs have all been validated. They all work in this kind of situation. It gets you around the problem that for many organ involvements, the n isn't big enough to get enough patients to see a change in that organ unless it's fantastic. So if we get an ACR-70 type drug in

one of these organs, we'll be able to see it with a reasonable n, but until we have that, I think we have to settle for this number 1 based on the fact that it's not a real common disease, and organ manifestations are multiple, and all of the indices are pretty well designed to pick up change in organs. The response levels could be set beforehand to say what allows you to define BILAG B or C instead of BILAG A or SLEDAI scores going from 8 to 3 or something. All that can be set beforehand. It's not all that difficult actually.

DR. WILLIAMS: Based on Dr. Simon's introduction today, while we thought we might have been clear in our own minds, I'm not sure we've conveyed that to agency yet.

Dr. Simon?

DR. SIMON: Since we've returned back to the disease activity indices yet one more time and with Matt on the phone, I was wondering if we could take a moment and you could answer a question for us. We heard yesterday that the disease activity index measurement process is impacted by the physician who is performing it, and I thought I heard that that was the ideal circumstance, that there would be some input of the physician into the scoring based on using judgments. That's of some significant concern to us in trials because I don't understand how

objective these measures are then, if there are judgment calls about how to score or the interpretation.

So if you all could help us understand that better, and it also reiterates the importance of blinding of the trials in that context. So if you could help us with that, that would be great.

DR. WILLIAMS: Ciela?

DR. ALARCON: Yes. The subjectivity actually is not such because what we are asking the physician is to say whether a patient that has the manifestation thinks that it's really due to lupus or not, and if it's not due to lupus, you're not going to score that manifestation as being part of a disease activity index. This is really the training that goes into applying those instruments. So if you train all your centers that are doing this trial, that shouldn't be a problem.

DR. WILLIAMS: Joan?

DR. MERRILL: Yes, I really want to say what Ciela is saying. Let me try to give an obvious one. You put a patient on a medication and the lymphocytes go down. Is that lymphopenia from lupus or from the medication? And sometimes you don't quite know the answer to that, but often you do because you stop the medication and the lymphocytes come back up. You're not going to score that. That's a drug effect. That is not lupus. But that's what

we're talking about judgment. You must attribute to lupus.

DR. WILLIAMS: Dan?

DR. WALLACE: The most obvious one is headache in somebody. Is the headache a lupus headache or is it a migraine? That's 8 points on the SLEDAI, which is a huge number, and that needs physician input.

DR. WILLIAMS: Jill?

DR. BUYON: Also, I would say that in the SELENA trial where we had 13 centers, it was very important along the way to do validation studies. So, in fact, what we did was give feedback so that we had patient cases, and patient cases that were real would be sent back to physicians and scored. So one of the reassurances that would be provided during trials is that there would be continued validation using real patients that each physician then could have input, and that would further validate that you were getting very good data coming in.

DR. WILLIAMS: This kind of leads us into question number 2 which is, are statistical changes in disease activity indices, such as a change in SLEDAI, considered robust evidence of efficacy? What change in disease activity indices is considered clinically meaningful?

Jeff?

DR. SIEGEL: Sorry. The answer to question

number 1 is really quite important to some of the issues we're struggling with, and we heard Jack Cush say this would not be acceptable and Bevra Hahn say clearly it would be acceptable. There are a lot of people on the panel who didn't comment. It would be helpful to us to know if there really is a consensus that this kind of design, even if it is a compromise, would be acceptable. Could we perhaps just get a little bit more?

DR. WILLIAMS: Yes.

Mike?

DR. WEISMAN: That's exactly what I was concerned about, going on to question number 2. I was a little confused by this. It seems to me that David's question about not knowing exactly what the specific outcome measures are for different organ systems in lupus is something that we've struggled with for a long time, and that's what the composite measures came from. That's why the composite measures were developed. So this is becoming a circular argument, and that's where the confusion, to me, is here.

Yesterday we heard conceptually, well, it would be fine if in fact we just leave it to the companies to come up with a design that was specified for an organ system, and as long as it was tight and as long as the statistical analysis was done properly and the primary

outcome measure is defined and there's concurrence and agreement on what that is. But nobody has ever done that. So we all agreed that that was a wonderful idea, but nobody has ever done it.

DR. MERRILL: Yes, they have.

DR. WEISMAN: Well, they've done it in renal disease.

DR. MERRILL: Yes.

DR. WEISMAN: But I'm separating that from renal disease. I'm separating that to everything else in lupus. It hasn't been done, and that's where the composite measure came from.

So I think we ought to just make a decision here or at least focus on the value of these composite measures or we're going to get rid of the composite measures and go back and redesign and reinvent this whole process. I think that's what I'm trying to get this group to focus on. And we need to do that. If we're going to stay with composite measures, we ought to pick the one that's most appropriate or we're going to drop it.

DR. MERRILL: I don't think we should pick one. I'm sorry.

DR. WILLIAMS: Jennifer?

DR. ANDERSON: Well, if we're still talking about question 1, I'll wait.

DR. WILLIAMS: Jill?

DR. BUYON: I think that we would be reinventing the wheel, and I would really suggest not. If we want to take a vote -- what I think is confusing here is you had two questions. One was would you accept a global change based on one of these instruments, and yes, we might do that. And the other was, within the specific organs, if they did not achieve a particular significant improvement, as you say, it's not that the labeling would be for that organ, but it might in fact be for what it was, which was a change in that instrument that a priori was considered to be a meaningful change, which will lead into question 2.

DR. WILLIAMS: Joan?

DR. MERRILL: Yes. I don't think we should eliminate any of the instruments at this time. I think that's premature. I think we're faced with a number of new biologic agents. Some of them may have widespread effects on lupus. Some of them may really be organ-specific. There may be a treatment for discoid. There may be a treatment for fibrosis in an organ. There may be a treatment for nephritis. So I think at this point we really need to leave people enough tools so that people can try and design a trial that will reflect the biologic effect that their trying to achieve.

DR. WILLIAMS: Betty?

DR. DIAMOND: Can I just suggest that maybe we should take a vote on this? Because I believe with Bevra that there's a great deal of consensus on this and that most of us would accept a global assessment as a global assessment of lupus activity, also acknowledging that other study designs to look at organ-specific disease are possible. But I don't think most of us share the concern that you can't do a global assessment using the instruments we have. So I think it would be just easiest to take a vote.

DR. WILLIAMS: Lee?

DR. SIMON: In thinking about the vote, please think about one global measure or is it several global measures? Yesterday I think Bevra had suggested perhaps we should be using two or three and not just one, and we do need that information as well. So please think about that.

DR. WILLIAMS: Mary Anne, then Jack.

DR. DOOLEY: Can we, as Jill suggests, make the vote whether or not we would accept the change in disease activity as a global change in lupus and divorce it from the issue about whether that would give approval for a specific organ?

DR. CUSH: That's sort of my point exactly. I don't think my point was any different than Joan's or Bevra's in that if you meet the disease activity

requirement, is that the same as signs and symptoms? I feel that it is, and it's treating the disease globally and you're controlling signs and symptoms just as you would with an ACR-20 for RA.

So I think that a disease activity measure meets a signs and symptoms definition. At what level? That has to be decided upon. How many? I think we could talk about that, but I agree more than one, and you have five or six to choose from. Meeting two out of those as a minimum requirement at a certain level seems prudent in going for a global indication for signs and symptoms.

DR. WILLIAMS: Jennifer?

DR. ANDERSON: We seemed to have moved into question 2, so it's not just about the stratified study but about the outcome measures. So I'd like to say something about the outcome measures.

The question of which one to use and what to consider as -- the amount of change that would be acceptable is what I was going to address. Is that premature to do that?

DR. WILLIAMS: Let's first get this first question because we're going to come to some sort of a vote.

Betty?

DR. DIAMOND: I was just going to say I think

that these global assessments are just that, and to say whether there are one, two, three, four signs and symptoms is to remake them. I think it would be a claim of reduces disease activity, and it wouldn't be for stipulated signs and symptoms unless it was powered to address those particular signs and symptoms. But I think within that, we're all in agreement.

DR. WILLIAMS: Lee, do you want the agency to pose the questions you'd like us to vote on, or do you want me to pose them?

DR. SIMON: I think you should go ahead and pose them.

DR. WILLIAMS: Thank you very much.

(Laughter.)

DR. WILLIAMS: Based on that first question, I would say that based on the information we have here, we ask whether this would be an indication that there is improvement in signs and symptoms versus specific organ improvement, with the second part being, would you accept a single disease activity index or would you require multiple. And thirdly, if you required a single, which one would it be, or does it matter?

DR. WILLIAMS: Ciela, you had a comment?

DR. ALARCON: Yes. I think that whether you do one or two or three depends on whether you designed the

trial for that. You have to specify what's your primary outcome and then go ahead and measure that. I think that you cannot go and say, well, now I'm going to also measure the SLAM or the SLEDAI when initially I saw that I'm going to do just the BILAG.

DR. WILLIAMS: Are those questions fair for the agency?

DR. SIMON: Yes.

DR. WILLIAMS: I think we'll go around the table and ask us to address those, and we'll start with you, John.

DR. LOONEY: Could we vote on them one at a time just to keep clarity?

DR. WILLIAMS: Okay. Let's take the first one. Do we see this as evidence of efficacy for signs and symptoms or for specific organs?

DR. LOONEY: So let's rephrase that question. Do we think that we can use the disease activity index for global signs and symptoms? And I would say yes.

DR. ILLEI: Yes.

DR. HARDIN: Yes.

DR. HAHN: Yes.

DR. DOOLEY: Yes.

DR. ALARCON: Yes.

DR. PISETSKY: Yes.

DR. MERRILL: Yes.
DR. GIBOFSKY: Yes.
DR. HOFFMAN: Yes.
DR. CUSH: Yes.
DR. ANDERSON: Yes.
DR. WILLIAMS: Yes.
DR. CALLAHAN: Yes.
MS. McBRIAR: Yes.
DR. MANZI: Yes.
DR. ILOWITE: Yes.
DR. FINLEY: Yes.
DR. DAVIS: Yes.
DR. DIAMOND: Yes.
DR. BUYON: Yes.
DR. WALLACE: Yes.
DR. WEISMAN: Yes.
DR. WILLIAMS: Do we see this improvement as in question 1 as signs of specific organ involvement? John?
DR. LOONEY: No.
DR. ILLEI: No.
DR. HARDIN: No.
DR. HAHN: No.
DR. DOOLEY: No.
DR. PISETSKY: No.
DR. MERRILL: No.

DR. HOFFMAN: No.
DR. CUSH: No.
DR. ANDERSON: No.
DR. WILLIAMS: We skipped Ciela.
DR. ALARCON: No.
DR. WILLIAMS: No.
DR. CALLAHAN: No.
MS. McBRIAR: No.
DR. MANZI: No.
DR. ILOWITE: No.
DR. FINLEY: No.
DR. DAVIS: No.
DR. DIAMOND: No.
DR. BUYON: No.
DR. WALLACE: No.
DR. WEISMAN: No.
DR. WILLIAMS: Matt, I keep skipping you. Matt?
DR. LIANG: The first was yes and the second was no.
DR. WILLIAMS: Thank you.
Do you require further questions? Would you like to know if they require one or more?
The next question is for improvement in these signs and symptoms, would we require one or more disease

activity measures? I understand some of the concerns Ciela has, but that will be the question. John?

DR. LOONEY: I guess I would say that, assuming that the people can prespecify which one they would take as their primary outcome, I would say one.

DR. ILLEI: One.

DR. HARDIN: One.

DR. HAHN: More than one.

DR. DOOLEY: I would specify two, with one being BILAG.

DR. ALARCON: Two.

DR. PISETSKY: Could I ask clarification? If you're doing more than one, is it either/or or both? If you do two --

DR. WILLIAMS: The question is do you require one or do you require more than one.

DR. PISETSKY: To be positive on more than --

DR. WILLIAMS: To be considered as positive for signs and symptoms for --

DR. PISETSKY: So if you do one, you are only doing one, not that you're positive in one.

DR. WILLIAMS: No. You do one and you show positivity. Therefore, you have benefit in signs and symptoms of lupus, or you require two.

DR. ALARCON: Jim, you have to prespecify that.

DR. SIMON: Let me just clarify that from a trial design point of view, from our point of view. We have done this before. You are all aware that in osteoarthritis we required three co-primary outcomes that have to win. The trial has to be powered to do that. We don't have a responder index like we do in the ACR rheumatoid arthritis trial designs. So it is possible that you can power a trial that would have two co-primary outcomes. Each you have to win on. A score like this would lend itself very nicely to that in particular.

So with those caveats -- and I would ask the chair to ask the question -- with the proviso that the trial was designed appropriately to consider the possibility of more than one co-primary outcome where you would have to win on both or more for a success, then that would be the question that would be applicable, fully recognizing that the power issue of a trial that requires several co-primaries becomes much more complicated and if you go above three co-primaries, you might as well shoot yourself because you basically can't interpret the results.

DR. MERRILL: Clarification. Are we requiring more than one activity index or allowing it?

DR. SIMON: Okay. That's the other question, and that's an excellent one. We're asking the question from the point of view, since they appear to measure

different things and they somewhat ask different questions, so that's a different input into the response, we would ask the question in the context of requiring them.

However, let's be clear about the entirety of this. You could also require them to be secondary outcomes, but you would not make a pivotal decision on the secondary outcomes. They would inform you. They could be in the label describing experiences for the patient and the treating caregiver, but they would not be what you would make your decision on for win or not win for approval.

So the question really should be, given all the caveats and all the other things about the trial, would you want one or two or more co-primaries for pivotal approval, not really whether or not you want the information, because you want the information. So we would assume they would be otherwise secondary outcomes to be measured.

DR. MERRILL: May I make a clarification here as a part of that? There have been published studies, a number of published studies, that show that these diseases do get the same results.

VOICES: Indices.

DR. MERRILL: Yes, the indices do get the same results. They are, therefore, to some extent redundant.

DR. WILLIAMS: Dave, do you have a question? Your microphone is on.

DR. PISETSKY: No.

DR. WILLIAMS: I'm not sure what the question is myself right now.

Mary Anne?

DR. DOOLEY: I was just going to clarify one reason why some of us may want two rather than just one is that although they get at the same thing and that if you look at a group of patients, that these things do correlate well. If you have a particular organ focus or your group of patients has a particular disease manifestations you may heavily weight on one of the instruments. So, for example, in nephritis, as Jill had mentioned yesterday, you get points for having proteinuria, for having red cell casts, for having white cell casts so that you get a preponderance of points on one organ system. So for that reason, if you're going to look at global lupus activity, you may wish to look at more than one instrument. That would be my rationale for looking at two.

DR. LOONEY: I guess if we're going to focus on a specific organ, though, I would like an organ-specific instrument and not a global one. I think for people who want to look at a more global picture of lupus, what particular kinds of patients they're recruiting may determine which of the scales is the best one for them to use. For that reason, I would like them to be able to have

the flexibility to do that. Especially since one of the goals here is to really encourage the development of these products, I don't really want to make it more difficult for people to get approval because we were expecting them to power it for two different indices which overlap in what they're measuring.

DR. WILLIAMS: Mike?

DR. WEISMAN: Each of these instruments has a certain sensitivity to change based upon some selectivity for the populations that are being studied. They're different in that sense. We've heard all that yesterday and we know this. It's going to be very difficult to require improvement in two of these instruments because the companies, or whoever, is going to select the instrument based upon a particular group of lupus patients that that particular drug is going to be most effective in. So I think that's all we can go. That's all we know at this point. I can't see how we're going to require two instruments. Who's to decide which two, for example. So I have a lot of difficulty with that. That's the problem that I have in your question, Lee. So I would vote for one.

DR. WILLIAMS: Susan.

DR. MANZI: I'm pretty much agreeing with a few people, but in response to Mary Anne's comment, I really

think this is more of dialogue and education of the sponsors when they're designing their trials as to which instrument makes more sense. It's the design of the trial. It's what they're trying to show. There are a lot of factors. I think requiring two is not the answer to that. I think it's understanding the design of the trial, the nuances of the instruments, because they all work and they can all show change. It's just a matter of which is appropriate for that study.

DR. WILLIAMS: Gabor?

DR. ILLEI: Yes. I just want to say that at least we have data for how each individual instrument works, and although it makes intuitive sense that two may be better, we don't have any data for that. So that's why I voted to accept one instrument.

DR. WILLIAMS: Betty?

DR. DIAMOND: I think the issue is not two instruments. It's setting the standard. It's question 2. It's what's a significant difference within any one instrument, and I think if you achieve that, there's no question that you've achieved efficacy.

DR. WILLIAMS: Bevra?

DR. HAHN: I was just thinking of a study design which I thought we were talking about which the primary outcome is reduction in disease activity, and I was

thinking that if you could show it by more than one instrument, that people will believe you, and that if you have only one instrument, then there will be all of the concern that it depends entirely on the patient population and it may not apply to everybody else. And there's a little more believability if there are changes in two of the instruments and a little more general applicability. That's what I had in mind.

DR. WILLIAMS: Dave?

DR. PISETSKY: If it's one instrument, does the trial designer have the option to select them from any of the group out there, or will there be a certain one that's chosen, so different people could use different instruments amongst that? I would have concern about that just in terms of trying to understand amongst agents if everybody is using a different outcome measure. You do need some standardization.

DR. WILLIAMS: Dan?

DR. WALLACE: I agreed with Mary Anne. I think you need really two instruments. You can argue, for example, that the SLAM doesn't different from fibromyalgia symptoms, that the SLEDAI is too heavily weighted in CNS, and I think that if you have two, you really cover all the bases and answer all the questions.

DR. WILLIAMS: Joan?

DR. MERRILL: I think that if you use the BILAG, you've covered all your bases.

(Laughter.)

DR. WALLACE: I agree with you.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: I was going to tell Dan that I actually have changed my opinion. I'm sorry.

(Laughter.)

DR. DOOLEY: But I am persuaded by the argument that the sponsors will appropriately choose the instrument to reflect the population that they're doing, and I don't think that any of us would read a study and say, well, I don't believe this because they used the SLAM rather than the SLEDAI. I think the data are going to be presented on the patients in summary form, as well as the outcome on activity measures. I would accept an outcome on the SLAM, the SLICC, the BILAG, the SLEDAI without any prejudice.

DR. WILLIAMS: Gabor, then Jack, then Lee. Then we're going to vote.

DR. ILLEI: What I wanted to say was said already.

DR. WILLIAMS: Jack?

DR. CUSH: Call the question.

DR. WILLIAMS: Lee?

DR. SIMON: Before you call the question, Joel

in his presentation raised a question about a single instrument use having the risk that there could be imbalance in manifestations between one group versus the other group. Depending on the instrument, if one group had a predominance of hemolytic anemia patients through a randomization, which can happen, and the other group has a predominance of nephritis and not the same manifestations, through randomization -- we're talking about a randomized trial -- would it not be more likely then that more than one instrument would allow better understanding of the responses in that any one therapeutic may not be able to treat both of those manifestations equally?

Our concern is that, as it relates to the choice of one instrument for a pivotal outcome, fully recognizing that one would assume that there would have been data accumulated before in phase I through phase II to suggest that, but at the same time anybody who's done a lot of trials knows that in designing a trial, you can go awry in that one particular trial.

So could you comment on the potential imbalance of recruitment in patients that would then lead to one of these disease activity indices not performing technically appropriately based on the intervention and the distribution of patients to one arm versus another?

DR. WILLIAMS: You're calling for more

discussion and I've had others who have called for the question. It's your meeting.

DR. SIMON: It's your meeting, number one, and number two, I'm not sure there's an answer but I wanted to be sure that when people voted, they were thinking about this particular problem.

DR. LIANG: Mr. Chairman?

DR. WILLIAMS: Matt?

DR. LIANG: Can I just throw something out on the table? I think that my judgment is that in the ideal world we would have finished off the ACR initiative, and one of the central pieces was to develop a repertoire of target organ response criteria that would be done a priori using available metrics and clinical sensibility really, because I don't think we'd ever get enough numbers to either generate or validate these response criteria. And these would be used as the primary endpoint for sample size calculations if someone was looking at a homogeneous group, but in all instances, the measures that are used to capture activity in these organ systems could be treated as covariates measured in all trials, depending on whether the manifestation was present or not, and used in the analysis.

I think that plus the disease activity measure would be my preference. But the sample size would be driven by what the designers were trying to answer, and I

would think in large part, depending on whether it's phase I or II, it could be preferentially a major target organ and secondarily the disease activity measures.

I don't think we need treatments for mild lupus. We need treatments for severe lupus, and that was another one of the assumptions that we were predicating our work on.

DR. WILLIAMS: Mary Anne and then Joan.

DR. DOOLEY: I think any one of the instruments would allow you in a very transparent way to see if there was an imbalance in patients in a particular manifestation, and that certainly if you were going to include patients with nephritis or a major manifestation that would have a significant impact on outcome, that you would stratify your groups. So I would say that any instrument that you chose would allow you to determine if there was an imbalance in a particular manifestation and that you could, in fact, account for that statistically.

DR. WILLIAMS: Joan. Then we have 10 more questions, so we're going to finish this one up.

DR. MERRILL: I don't think any of the instruments are particularly flawed in the way that you fear, Lee. Having said that, I think we have to just trust the designers of the study. Who are you going to enroll? What are you treating with? And what do you expect? I

think the studies will be designed keeping in mind -- and some studies are designed with stratifications and randomization. If that's necessary, that should be built in from the beginning.

DR. WILLIAMS: I'll remind us for the first two votes, we voted that this study was for signs and symptoms and not for organ-specific. The third question now is do we require one primary or more than one primary variable. John?

DR. LOONEY: One.

DR. ILLEI: One.

DR. HARDIN: One.

DR. HAHN: More than one.

DR. DOOLEY: One.

DR. ALARCON: Two.

DR. PISETSKY: More than one.

DR. MERRILL: One.

DR. GIBOFSKY: More than one.

DR. HOFFMAN: One.

DR. CUSH: One.

DR. ANDERSON: One, but several indices as secondary.

DR. WILLIAMS: One.

DR. CALLAHAN: One.

MS. McBRIAR: One.

DR. MANZI: One.
DR. ILOWITE: One.
DR. FINLEY: More than one.
DR. DAVIS: One from a recommended list from the FDA.
DR. DIAMOND: One.
DR. BUYON: One.
DR. WALLACE: One if it's the BILAG; two if not.
(Laughter.)
DR. WEISMAN: One.
DR. WILLIAMS: Matt?
DR. LIANG: One.
DR. WILLIAMS: Jeff, Lee, is that okay?
DR. SIMON: Thank you.
DR. WILLIAMS: Moving on to question 2, are statistical changes in the disease activity indices, such as a change in SLEDAI, considered robust evidence of efficacy? What change in a disease activity index is considered clinically meaningful? And Jennifer has been waiting a long time for this one.
DR. ANDERSON: The trial design that was presented in the open part of the session suggested an outcome measure which would be 25 percent improvement in SELENA SLEDAI. Then yesterday in the presentation that

Matt Liang made, among the experts 70 percent or more agreed that a change in SELENA SLEDAI, an improvement of 7 was clinically meaningful. And yet, the entry criteria for the proposed trial suggested that the SELENA SLEDAI be at least 4 at the beginning.

So I don't know what the usual distribution including the observed range and then also the possible range of these instruments is, but it would seem that it's likely that both SELENA SLEDAI and BILAG have a similar range because the experts came up with exactly the same changes for improvement and worsening -- well, improvement of 7 and a worsening of at least 8 for each of those. So I don't know whether that's true or not, but that's sort of like the implicit scale that they're putting on them.

So all of this is preamble to saying that it's possible that a 25 percent improvement is a good improvement, but I think there has to be a minimum change added to that. I don't know whether it has to be 7 because then that would mean that you've got -- if you're starting off -- if the typical value at the beginning is, say, 15, you'd have to improve by almost 50 percent to improve 7.

I don't have any idea what these distributions are. So maybe if somebody does have some idea, that would be helpful in deciding what kind of percent change and how much change would be considered meaningful.

DR. WILLIAMS: Jill?

DR. BUYON: Well, I think first the problem is designing the type of trial you're doing because if you're going to enter a patient where you require that patient to have a SLEDAI of 4 or greater, there's no way you can make a change of 7. So, obviously, it really depends on what is the question being asked, and I think the difficulty in addressing question 2 is the type of trial design. Is it time to flare, and how do you use the instrument? Is it starting off with a certain number in the instrument? But I would submit that it would be unlikely -- we'd be looking at a trial where we're asking a patient to come in with 4 or greater and then expecting to see a change in that as the final outcome. So this is a very difficult context in which to answer this question because we don't know what the trial design is, and I think that's one of the biggest problems.

But in the SELENA SLEDAI, changes of 3 were not consistent with flares. So when we defined flares as mild, moderate, or severe and even looking at mild-moderate flares, it didn't perform well with a change of only 3. We missed flares or didn't see them.

DR. WILLIAMS: Joan?

DR. MERRILL: Yes. I have to agree that a flare index is a very difficult thing. Will the SELENA

SLEDAI flare index be validated soon and published? A question to Jill.

DR. BUYON: I'm not sure how to answer that.

DR. MERRILL: Because otherwise we have no validated or published flare index, which is a problem per se, unless you can use an instrument and define flare as numbers in that instrument.

I don't quite understand this 7. You mean people are expected to improve by 7 points?

DR. ANDERSON: This was part of Matt Liang's presentation yesterday on the ACR SLE response criteria initiative. The slide on clinically meaningful differences for specific instruments.

DR. MERRILL: In the SLEDAI.

DR. ANDERSON: SELENA SLEDAI was 7, as was BILAG, and SLEDAI was 6.

DR. MERRILL: All right. I think that that would be untenable if you were treating moderate lupus.

DR. WILLIAMS: You're being quoted, Matt. Do you have anything you want to say?

DR. LIANG: The answer would be too long. That's the data. I think that what is being talked about is really to express the change, whether it should be a percent change or an absolute change. I think that's a decision of an investigator, but I think a change in

someone who's got little activity has a different kind of significance than someone who's got a lot of disease activity. I think that that's more an issue of reporting than anything else. The data is there and it can be expressed in different ways to get into that.

I think the other thing that our data suggests is that you're not going to do a trial in people with little activity. I think we're all talking about patients with either very severe or moderately severe disease with a lot of activity. Therefore, these changes reflect where we would want new agents.

DR. WILLIAMS: Thank you.

Jack?

DR. CUSH: I want to ask Matt and Joan and Jill and anybody else who wants to comment on their experience with using these tools, but is a 25 percent improvement in SLEDAI or SLAM or BILAG enough, or do you need 50?

DR. MERRILL: I think it depends on the drug, and I think we sometimes are treating mild to moderate lupus. I would like to be able to capture the differences for a person who improves in arthritis, which is 4 points on the SLEDAI, or who improves on arthritis and rash, which is 6 points on the SLEDAI. And if that person's pretty severe arthritis got better, I would like to see that 4-point change, and I'd like to know in a published paper

that there was a difference there. So I think trying to enforce numbers when there are so many different drugs and so many different ways that they might work is not going to work. I think that a trial design has to come before the committee and it has to be figured out on a case-by-case basis.

DR. WILLIAMS: Jill?

DR. BUYON: I fully agree with that. I want to clarify, I was actually the person who did the SELENA SLEDAIs on 350 paper patients. Part of the problem was that you couldn't really identify change in patients who came with low levels of activity. So if they started with SLEDAIs that were less than 5, you could not really ascertain meaningful changes because in many cases that might have been a C3 that normalized or DNA and everything clinically stayed the same. On the other hand, when patients came in with high SLEDAI scores, then the meaningful change was 7.

So I want to clarify, and I hope Matt will concur. But that basically needed to be told to you so that you could understand the context of that change. It's harder to ascertain change with these instruments when patients come in with lower scores. So, again, we're voting by instrument. I take the good faith that the company who is sponsoring the trial will, a priori, know

that if they're looking at a patient who's mild, SLEDAI would not work in that particular situation.

DR. WILLIAMS: Lee?

DR. SIMON: So, Joan and Jill, help me understand this. Are you suggesting then that you would actually parse out a change that would be perhaps small in a SLEDAI score that would interpret an important event in the improvement of arthritis for an approval as opposed to a publication?

DR. MERRILL: If I had a medication that improved lupus arthritis significantly, I would like to capture that, and I think maybe Jill has made the point that the SLEDAI might not be a good instrument to use for that. The SLEDAI might be a much better instrument applied to more severe lupus.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: Forgive me if someone has already made this point, but I think the degree of improvement would also depend on the toxicity of the drug. If I was using Cytoxan, sure, I'd want at least a 7-point improvement. But if I'm using something with far less toxicity, I would accept a lower amount of improvement. So to some extent, it does depend. That's also related, obviously, to the severity of disease and, therefore, the entry scores that patients would be coming in with. But

the toxicity of the drug that you're proposing and the severity of illness of the patients would make a difference in terms of what a meaningful change would be.

DR. WILLIAMS: Ciela?

DR. ALARCON: The design for the patient with very low lupus activity will be really time to flare. It will not be really improvement or a decrease in the number in the instrument.

DR. WILLIAMS: Joel?

DR. SCHIFFENBAUER: I just wanted to get clarification. The question was referring to a disease activity index or a measure of global activity, but the issue of measure of flare came up. My understanding would be that any statistically significant difference in flares, rates of flares, number of flares, would be considered clinically meaningful. Can I get some agreement on that aspect of it and then go back to the disease activity index issue?

DR. WILLIAMS: Joan?

DR. MERRILL: Yes, I think the numbers of flares is definitely clinically meaningful. I have some possibly piddling concerns about the use of flare indices. For example, it's summertime and people go out in the sun and they get a skin flare. That's a minor flare, but it still counts. So it depends on the kind of flare you're

counting and you really have to differentiate between these mild ones and the really significant flares.

DR. WILLIAMS: My understanding of question number 2 from the discussion is that we can't give you a specific answer. It depends on the severity of the disease, the toxicity of the medication.

Question number 3. Please discuss the data that should be collected for a study of lupus nephritis. Please discuss the sensitivity to change and clinical intepretability of change in GFR versus doubling of serum creatinine versus 50 percent increase in serum creatinine. What is clinically meaningful change in hematuria and proteinuria? Can resolution of hematuria/proteinuria be considered evidence of an important clinical benefit in the treatment of renal disease? Is the measure of RBC casts more useful for this?

DR. WALLACE: I think we should hear from Matt because his committee has come out with summary recommendations on that.

DR. WILLIAMS: Matt, do you want to start off?

DR. LIANG: (Inaudible) physiology or data to really make an informed choice, and when you review the literature, people have defined it so many different ways that it's impossible to do any qualitative or quantitative synthesis in a meaningful way.

So having none, the committee took the low road and said that it's better to be consistent than to be right, and we have put together recommendations in writing based heavily on how the nephrology community has moved towards measuring renal function, but basically using clinical judgment to a priori define what we think are improvements, stable, and worsening renal disease for the glomerular nephritides in lupus.

That manuscript is being finalized, but delayed because I've been out, and it's going to work its way through the ACR committee structure. It's the first of the seven target organs that we have dealt with in various forms.

DR. WILLIAMS: Bevra.

DR. HAHN: Could you give us an idea of what the conclusions are, Matt? Is it a composite?

DR. LIANG: At the end of the day, the groups that have met have felt that you needed to have a measure of renal function, and basically the nephrologists in a very extensive documentation have said that clearances based on the serum creatinine and other easily obtainable information is good enough. That would be one parameter, and we basically said -- I've forgotten exactly what the percentage was at the end of the day would be an improvement. Another would be stable and another would be

worsening.

They felt that a measure of urinary protein excretion would be another metric, and a convenient way to do that would be to get a (inaudible) urine protein/urine creatinine ratio, and we stated what we thought was an improvement, stable, and worsening renal disease.

Urinary sediment, even though everyone is in love with it, there's little data on reproducibility, but we felt that if a sponsor could commit the resources and guarantee quality and reproducibility, that urinary sediment would also be a parameter of active inflammatory disease. And we tried to state what we thought was explicit criteria.

And then the final one was -- I'm forgetting actually. We tried to make a statement on renal pathology which was that it's nice if you can get it, and we strongly urge it. We also urged that a repeat biopsy be done especially if one of the endpoints was remission at an appropriate interval after the treatment.

Those are the highlights, but the full document is working its way through.

DR. WILLIAMS: Did you have any specific comments on hematuria?

DR. LIANG: Well, hematuria was included in that urinary sediment. I think we all use it clinically,

but in a trial situation where you have multiple labs, multiple investigators, we thought that the quality assurance had to be guaranteed before one used it. Again, it's one axis of describing response.

DR. WILLIAMS: Jeff?

DR. SIEGEL: Matt, at the Dusseldorf meeting there was a lot of discussion about what change in proteinuria would be clinically meaningful.

DR. LIANG: Yes.

DR. SIEGEL: And there was some thought that you should really move from nephrotic range to below 1,000 or below 500 milligrams.

DR. LIANG: Yes.

DR. SIEGEL: Can you just discuss how that ended up in the final discussion?

DR. LIANG: Actually if it would please the committee, I'm away from the paper, but I can get it and come back with you when you're ready for it. I can give you more specifics. I can't do this from memory anymore.

DR. WILLIAMS: If you'd do that, we'd appreciate it, Matt.

DR. LIANG: I'll be back in 5 seconds.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: Matt, before you head out, we also distinguished between proliferative and membranous disease

so that the response would be different based on the lesion. That would imply that a biopsy prior to study entry would be required obviously.

DR. WILLIAMS: While we're waiting for Matt to come back, one of the questions is the sensitivity to change and clinical interpretability of change in GFR versus doubling of creatinine versus 50 percent increase in serum creatinine. Any comments on that?

DR. DOOLEY: As Matt has already described, this remains a contentious issue among the nephrology community as well, and I think that looking at the formula to calculate creatinine clearance was highly regarded, and that would be the Crockoft-Gault in adults, and correct me if I'm wrong, I think it's the Schwartz in children. So you would apply the appropriate instrument for the age of the patient, and that was accepted as a measure of creatinine clearance, recognizing the difficulty of doing iothalamate clearances or the concern about the patient's ability to complete 24-hour urine collection.

DR. WILLIAMS: Lee?

DR. SIMON: Could you just comment a little bit more about the difficulty in performing iothalamate clearances? Is this just a technical structural issue of bringing the patients in to do that and then, thus, not enthusiastic to be in a clinical trial, or is there some

other component to its difficulty?

DR. DOOLEY: I'm not an expert on this but my understanding of the difficulty is it's a radio-labeled study. Therefore, you have to be able to give the patient a radioisotope and you have to be able to collect the urine the patient passes and dispose of it appropriately. Many GCRCs don't offer that as a procedure. So the major concerns that I have seen have been in the use of the radioisotope and then the availability of the test.

DR. WILLIAMS: Jack, then Norm.

DR. CUSH: Glo-fil or iothalamate determinations are very reproducible and very reliable. They are easy to do. The biggest hassle is that the patient has to go somewhere else to have it done, number one, and then the availability in any center or any city is quite suspect. In Dallas, it has moved around to a few different places. It used to be at the medical school. Now it's over at Baylor. So it's a moving target. In a city as big as Dallas is, right now there's only one site that does glo-fils for our patients. So it is available but it can be hard to find even in big centers.

DR. ILOWITE: Noninvasive methods for determining glomerular filtration rate and degree of proteinuria have been validated in children, and it's even more extraordinarily difficult to get 24-hour urines in

adolescents, even in in-patients. Thirdly, our children's hospital IRBs I would expect to consider nuclear medicine scanning for creatinine clearance or glomerular filtration rates unethical if there was a noninvasive method that had been relied on and is validated.

DR. WILLIAMS: Gabor?

DR. ILLEI: In the literature, there are data that some of these estimates of GFR correlate with the true measure of GFR over 90 percent and they are actually more reliable than the creatinine clearance. There are different formulas from the diabetic renal disease studies, and the Crockoft formula is also about 90 percent in terms of correlation with measures of GFR.

DR. WILLIAMS: Joan?

DR. MERRILL: Yes. I want to point out that any nephritis trial at this point, especially with multiple agents being tested, is going to have to be a very multi-center study, and so it's probably impractical to rely on methods that may not be available in most cities.

DR. LIANG: Mr. Chairman?

DR. WILLIAMS: Yes.

DR. LIANG: Anytime you're ready.

DR. WILLIAMS: I'm ready now.

DR. LIANG: I could tell you about some of the definitions we had for complete renal remission, end-stage

renal disease, and nephrotic syndrome. I can tell you about what the recommendations were for calculated GFR, urinary sediment, and urinary protein. Also, we tried to list, in terms of adding to the CONSORT recommendations, what we thought were the essential covariates for the conduct and reporting of renal trials in SLE. So I'm prepared to give you any or all. I don't know if you want to spend all the time.

I think Jeff's comment was proteinuria?

DR. WILLIAMS: Yes.

DR. LIANG: Here we said a spot urinary protein ratio over urinary creatinine was the preferred measure, and it's documented in the kidney community with extensive documentation. We said that an improvement was at least a 50 percent reduction in the UP over urinary creatinine. A partial response was at least 50 percent reduction and the UP over UC equal to .222, and a complete response was a UP over UC equal to 0.2 to .2 and less than 0.2. Stable would be unchanged UP over UC, and worsening was 100 percent increase in the UP over UC and greater than 1 gram of protein per 24 hours.

Was that the question you had, Jeff?

DR. SIEGEL: Yes, thanks.

DR. LIANG: Okay.

DR. WILLIAMS: Are there other questions for

Matt regarding the data he has? Bevra.

DR. HAHN: Yes. Matt, what was the discussion about using creatinine clearance as opposed to creatinine or reciprocal of creatinine or something like that?

DR. LIANG: That was very interesting, Bevra. Based on the two committees' deliberations, I think the most experience in that ratio, 1 over creatinine, has been in diabetic nephropathy, and I think it was held out as a promise. Everyone is going to collect the creatinine. So I think that it's sort of moot. People could express it, and whether that is a better predictor of end-stage renal disease I think is a jump ball in renal nephritis, but there is some suggestion that it is. But I think everybody would be collecting the creatinine anyway, and that could be deduced from future data.

DR. WILLIAMS: David?

DR. PISETSKY: In terms of the renal improvement, if you have both renal impairment and proteinuria, do you have to meet criteria for improvement in both to be considered a responder?

DR. LIANG: Actually we did not deal with that. We were just trying to establish the essential key parameters that one should collect, but there was strong interest in someone doing that work, which is to create a one-number renal index. That obviously we couldn't do with

the kind of funding we had for these committee meetings. But that's certainly a worthwhile research goal I think.

DR. WILLIAMS: Bevra, did you have another question?

DR. HAHN: No.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: Well, I think it would be essential that you could not worsen your renal function and be counted as a success because as your creatinine clearance falls, your proteinuria will fall as well, so that you would have to have at least stable renal function to have a fall in proteinuria count as a success.

DR. WILLIAMS: Dan?

DR. WALLACE: One of the major concepts we discussed at this committee is that renal function per se rarely improves. Yet, preventing it from getting worse can be considered a success, and that has to really be factored into things.

DR. WILLIAMS: I'm not sure we've given you a lot specific help, but some generalized help. Are there any further questions the agency has?

Jill?

DR. BUYON: One clarification I would ask Matt. Did you have any time to have these changed because is a year good enough, is it 2 years? Because we've certainly

seen accomplishment of those goals and then 6 months later things relapse. So my question has to do with stability.

DR. LIANG: Yes, we did. Basically I think the committee recognized that short trials are -- you know, using these parameters are necessary and practical, but they thought that the minimum optimal length for assessing meaningful outcomes in trials of lupus GN would be at least 2 years and for membranous disease, even longer than 2 years. But I think this has to do more with -- well, this is the clinical sense of the kind of trajectories and the durations that you would need to do.

DR. WILLIAMS: Jack?

DR. CUSH: Matt, could you comment on whether the discussion at all migrated into -- instead of looking at improvement, which may be difficult and hard to agree upon, to rather look on failure as the outcome measure, so more hard and fast rules like end-stage renal disease or doubling of creatinine or worsening of proteinuria? Were those felt to be at all less preferable or equally useful?

DR. LIANG: There were other people who were at that committee meeting. I don't think I really nailed that with the committee. We were basically trying to develop the parameters and to define the parameters of improvement, stable, and worsening within those parameters, but not as deeply as you're asking.

DR. WILLIAMS: Since we've been of such specific help on nephritis, we'll now move to CNS lupus.

(Laughter.)

DR. WILLIAMS: Please discuss data to collect for trials in CNS disease.

Dan?

DR. WALLACE: I think that any CNS trial would have to include spinal fluid because you have cell count, protein, oligoclonal DANS, IgG synthesis rate, neuronal antibodies, even LE cell preps on Wright's stain of the spinal fluid. There's no other parameter for a CNS lupus other than imaging, functional imaging, that's as precise.

DR. WILLIAMS: Gabor?

DR. ILLEI: I think it should be clarified a little more exactly what we understand as CNS disease. Is it all neuropsychiatric manifestions of lupus or is it just lupus cerebritis inflammatory brain disease? Because I think that the data you collect for a neurocognitive study is different from one that you use for cerebritis.

DR. SIMON: Are you sure? I think we're starting from scratch. I don't think we have a real good understanding here, so we're trying to be as broad as possible without any assumptions that we understand that neuropsychiatric symptomatic manifestations -- so the psychiatric manifestations -- really don't have good

clarity about what any other kind of objective measures might have.

DR. WILLIAMS: Bevra?

DR. HAHN: Before we start this, are we going to adopt the international committee's classification of CNS lupus to base this discussion on, where we wouldn't have a word like lupus cerebritis, for example? There are something like -- I don't remember -- 17 or 21.

DR. WALLACE: 18 or 19 different types. At the SLICC meeting, when we actually broke it down, we figured out that 4 of the 18 were responsible for 95 percent of all the cases.

DR. HAHN: Do you remember what those 4 were, Dan?

DR. WALLACE: I think it was whatever we have as vasculitis, phospholipid-mediated, the vascular, which is the lupus migraine and cognitive impairment, and I can't remember.

DR. WILLIAMS: As verbal as this committee has been, there are few hands on this discussion.

(Laughter.)

DR. WILLIAMS: Bevra?

DR. HAHN: I brought it up because I honestly don't think we can discuss this until we decide. If we're going to use that classification, then we can decide only

certain categories are studiable, and how those could be studied. Without that, if we're just going to use just seizures and psychosis, then we're pretty limited.

DR. WILLIAMS: Jack?

DR. CUSH: As the diagnosis is so difficult in itself and the classification is hard to get everyone to agree upon, although I think that the international guidelines probably should rule at this point, pending further work from a guidance document like from Matt's group on end organ involvement with the brain, I don't think that trials in CNS disease can be done at this time.

DR. WILLIAMS: Betty?

DR. DIAMOND: I agree with Bevra that one should adopt that for CNS trials, and there are 19 different syndromes. I think that this is one of those situations where the data that you collect depends on your claims, and it's as Gabor said. If you're trying to treat vasculitis, you certainly need an LP. If you're trying to treat neurocognitive changes, it would be interesting research, but it's not clear that it's going to be an outcome measurement that you would need. So I think it's important to use the 19 syndromes and that studies have to clarify what their claims are and what they think they're treating.

DR. WILLIAMS: Joan?

DR. MERRILL: And I hope that whoever is sitting out here in this room who would love to study neuropsychiatric lupus or have it studied -- I don't mean to thwart your aspirations, but I agree with Jack. I think we can't have this discussion right now. I think there are too many etiologies involved that we don't really understand. There's a crying need for research on a clinical level to try to sort these patients in some way and measure outcomes. But there's no instrument -- and I'm including all the good instruments that we have for global lupus -- that really can capture before and after improvement/not improvement in neuropsychiatric lupus in any way that I think has been pulled together. So if Matt wants to fight for some more funding to do his kind of work, this is a crying need, and I don't think our discussion right now is going to be very productive.

DR. WILLIAMS: Wendy?

MS. McBRIAR: I would just like to encourage you if testing is done in this area that it be the least invasive possible.

DR. WILLIAMS: Bevra?

DR. HAHN: I suggest that for our next meeting that maybe this be tabled -- I don't know if we work that way on this committee -- and the international classifications be circulated to members of the committee.

There will be some we probably don't want to include, like anxiety is one and depression is one. We may not want to include those as they relate to lupus specifically. So maybe we need to have a look at them before we take this up.

DR. WILLIAMS: My sense of the committee is that we're not going to be much help on this question.

Question number 5. What is the standard of care for lupus nephritis? Are there circumstances where steroids alone would be the appropriate therapy for lupus nephritis?

Lee?

DR. SIMON: I just want to make one little caveat here. The way this question is designed is to tease out what we alluded to yesterday and just want to make clear to everybody the regulatory perspective of standard of care.

If Cytoxan is what you think is standard of care, along with some other drugs, because it has not been proven nor approved to actually do what we think it might by standard of care, it cannot be a comparator other than it serving as placebo. You can't do a noninferiority trial against cyclophosphamide at this stage of the game. Glucocorticoids, however, are approved and could be a comparator that you could beat or be not inferior than to

be able to be approved. So from a regulatory perspective, that's part of this question, and we wondered if you would think about it in that way.

DR. WILLIAMS: Dan?

DR. WALLACE: According to the NIH trials, steroids alone were equivalent to Cytoxan up to the first 5 years. After the first 5 years, they were associated with more morbidity and mortality. But one little thing that's not appreciated about the NIH study is that they mixed membranous with proliferative patients, which we would never do now. So the answer is we really don't know.

DR. WILLIAMS: Gabor?

DR. ILLEI: Well, I just have to voice my reservation in terms of the approach of not accepting cyclophosphamide as standard of care. I think cyclophosphamide is the standard of care for proliferative lupus nephritis. I think conceptually we do clinical trials, even if they are not optimal, to assess a chance of a drug, how they will work in practice, and even if a drug was accepted as standard of care and performs fairly well in practice, even if the studies that served as the impetus to use it in everyday care, I think it should be accepted as a comparator.

DR. WILLIAMS: Jack?

DR. CUSH: Lee, you're saying because

cyclophosphamide is not approved, it can't be the active comparator in a standard of care trial. Is that right?

DR. SIMON: No. What I'm saying is that it can always be used as an active comparator at any time you want, but to be able to achieve your proof of evidence that your study drug works, you'd have to show that you are better than cyclophosphamide because cyclophosphamide, from a regulatory point of view, regardless of its use as standard of care, has not been approved for the treatment of lupus nephritis if that's what you're studying.

DR. CUSH: But do these rules apply to orphan situations such as this? The reason there's no data and it hasn't been studied is because, A, the drug is very old and its use is not really that great. I think everyone would say that this is clearly the standard of care. At least, that's what I'm going to say.

DR. SIMON: Based on what?

DR. CUSH: Based on its use.

DR. WILLIAMS: You're asking what the standard of care is. The standard of care is cyclophosphamide. It doesn't necessarily mean it's evidence.

Joan?

DR. MERRILL: Lee, is there any appropriate mechanism that this committee could communicate to the FDA the opinion, if we have it, which we would have to vote on,

that this rule is unresponsive to what we need to accomplish in lupus? Just to communicate our opinion.

DR. SIMON: This isn't a rule. This is much more than that. It is one of the fundamental issues of the establishment of efficacy within the construct of the agency. That's one.

Two, you can obviously make a consensus opinion here, whatever that might be, and we will be happy to convey that opinion to the powers that be.

DR. MERRILL: I don't think anyone is comfortable with this. I think all of as physicians would be thrilled to get a drug that's equal to cyclophosphamide and safer and doesn't cause sterility.

DR. SIMON: Can I ask another question, though?

DR. MERRILL: Yes.

DR. SIMON: We're all very opinionated about this. This is one of the more emotional issues within the field. Where does the emotion come from? What data? I'm not asking your personal experience. I have the same personal experience that you have of being a rheumatologist for 25 years and taking care of patients with lupus nephritis. But there is an enormous amount of emotion that is based on no or very little data. And please do not quote the NIH trials, which are nonexistent, because they're retrospective analyses.

DR. MERRILL: No, no. Hold on a second. The emotion is not based on data, but we haven't got any alternative. So you can't ask me to produce data. I would be happy to have something to offer my patients as good as cyclophosphamide, because that's all I have. Of course, I wish I could get something better, but if I learned that there were a drug that was equal to cyclophosphamide in a trial that would not cause a 22-year-old to become sterile, I'd want to use it.

DR. WILLIAMS: Gabor?

DR. ILLEI: Just a comment on the NIH studies although I was not personally involved in any of those. The first that was published by Austin back in the early '80s was a summary of five different studies, and those are all perspective. But the others published by Boumpas and Gourley subsequently were all perspective, randomized, controlled studies. They were not retrospective analysis of data. They were not placebo-controlled but they were prospective and randomized.

DR. WALLACE: The NIH-funded Ed Lewis multi-center trial study with Cytoxan apheresis was also prospective on over 100 people.

DR. WILLIAMS: Betty?

DR. DIAMOND: I just want a clarification. You're saying what Joan thinks you're saying. Right? That

noninferiority to Cytoxan with less side effects is not an approvable indication. It's not an approvable claim. Is that correct?

DR. SIMON: The question is in that Cytoxan is not approved for this indication, a trial against it as the comparator, you would have to be superior for approval. There is no mechanism to provide a noninferiority claim to a drug that is not approved in the indication even if it is more safe.

DR. WILLIAMS: Jeff?

DR. SIEGEL: I wanted to respond to Joan's question about what she and other people could submit to the agency that could be helpful, and this is by way of fleshing out some of the concerns that Lee has expressed.

Investigators that I've talked to who want to be able to have a drug approved based on being as good as cyclophosphamide or almost as good as cyclophosphamide but less toxic say that when a drug works as well as cyclophosphamide, they know. Well, what would be helpful is for us to know how you know, exactly how you'd measure it.

So the reason that we ask for either superiority or noninferiority is that the agency does not want to approve drugs that don't work.

DR. MERRILL: You don't know with a head-to-

head trial. I am not suggesting that we're going to somehow emotionally get a drug approved. I'm asking for some --

DR. SIEGEL: Joan, let me -- can I just finish?

So in a noninferiority trial, the way we make sure that we're not approving a drug that doesn't work is to look at the active comparator -- in this case it would be cyclophosphamide -- and ask what its effect size is, how effective is it.

So what we would ask you to do, if you wanted to submit your opinions to the agency to help us in our decision making, is to decide what is the effect of cyclophosphamide. And it would be, for example, in such and such a group of patients, the effect of cyclophosphamide is to cause resolution of nephritis in 50 percent, 25 percent, 75 percent of patients as defined by thus and such within such and such a time frame. I haven't heard that yet, but knowing what people believe the effect of cyclophosphamide is would be helpful.

The NIH studies established, or at least indicated, that over a 5- to 10-year time frame, the progression to end-stage renal disease was lower than with an active comparator, corticosteroids. I don't think you all are saying that you think a drug is as good as cyclophosphamide because in 5 years you have less

progression to end-stage renal disease. There's some other effect of cyclophosphamide you're basing your presumption on. It's presumably resolution of nephritis, urinary sediment, normalization of creatinine, something. Defining what that is and what the effect is you think you're seeing would be very important and helpful to let us be more specific about what we're talking about.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: I think one of the difficulties that we all face is it's going back to the NIH trials and trying to interpret them because, if we remember, the original NIH trial, that then took 15 years to show a difference, didn't use the regimen of Cytoxan that we currently use. So patients were only given a dose of Cytoxan every 3 months from the very beginning. So along the way, this so-called NIH regimen has changed several times.

Additionally, in at least the first two trials, patients with severe renal disease were excluded. So in the original trial, you couldn't come in with a serum creatinine above 2. In the subsequent trials, you couldn't enter presenting with acute renal failure, which is a not uncommon presentation, at least at our institution. And if you required dialysis, you could not come in.

So the reality is we look at these studies that

were done at least initially in caucasian patients, the lowest risk group, and try to interpret them in light of the patients that we actually see.

If we look at our data, that is, southeastern United States, two-thirds African American, Cytoxan works, if you look at the group overall, about 70 percent of the time, similar to many of the older RA medications. So a highly toxic drug that does produce a benefit, but certainly not for 100 percent of the patients.

And then if you look at subgroups, particularly African Americans, we see a much lower rate of efficacy.

So I think one of the difficulties in your question is that the drug has not been appropriately studied in a clinical trial situation for us to be able to state what we believe the response would be.

DR. WILLIAMS: Mike Weisman?

DR. WEISMAN: The question that Lee is posing to us is pretty straightforward. The message starts out that the agency will not permit an advertisement of a drug that is equivalent to another drug that has not been approved for the disease. That kind of makes sense to me. I don't see why we're hung up on that. That's the rule. Right? That's the rule and we can't get around that.

So he's asking the separate question here which is, what are the circumstances where steroids alone would

be appropriate therapy for lupus nephritis to allow or permit the possible claim of an effective new agent for the disease? Let's answer that question instead of just kind of going back over this same issue.

Is it possible? Is there a form of lupus nephritis where steroids alone over 3 to 6 months would be an appropriate comparator to a BLyS agent or a CellCept or something else that might be investigated as a superior drug or even equivalent to steroids and safer in lupus nephritis? Is there a period of time, 3 to 6 months?

DR. WILLIAMS: I have myself down, and I don't think that I would be successful in recruiting patients to a trial that allowed steroids only for treatment of lupus nephritis in our area.

Norm?

DR. ILOWITE: I wanted to tweak Betty's hypothetical question. Would a company be able to come in with a claim based on noninferiority if it was comparing steroids plus Cytoxan plus placebo to steroids plus Cytoxan plus active drug, where steroids is the approved agent?

DR. SIMON: What's the primary outcome that you're measuring?

DR. ILOWITE: Well, before I dig a hole, can you think of an outcome that would be approvable under that design?

DR. SIMON: Well, it really turns on the issue that the consistent therapeutic is the glucocorticoid, and the study drugs are really cyclophosphamide versus the new therapeutic that you're talking about. Under those circumstances, if your new therapeutic was better than the combination of glucocorticoid and Cytoxan, then there's no problem. If the new medication is designed to be not inferior to the glucocorticoid and Cytoxan -- and I think Michael's previous statements are the issue at hand -- the label and otherwise would look like that this new drug was not different than glucocorticoids. It could not really reflect the benefit or lack thereof of cyclophosphamide in that context, and cyclophosphamide does not have a proven role clearly in the treatment of lupus nephritis.

DR. WILLIAMS: Joan?

DR. MERRILL: I would be willing to say that it would probably be considered ethical at my institution to do a trial in which people were started on glucocorticoids plus placebo or glucocorticoids plus agent with a 2-month check, and actually a continuous check. If they get worse at any time, they're going to have to switch. If they stay stable at the 2-month check, if they're not improving, that might be time for something or maybe you could go from there to 3 months where you know you're a failure. You don't see improvement. Then you're a failure and you've

got to do something. You could have a little something and a big something, something like that. You could design a trial like that, and I actually think it would be ethical.

We did the CellCept trial, as you're aware, and if CellCept had not worked at all, we would have had patients stuck on steroids for up to 3 months if they weren't getting worse. At any point we could have jumped out and saved them.

DR. WILLIAMS: Joel?

DR. SCHIFFENBAUER: Yes. I just wanted to follow up with Dr. Dooley's comment there. Clearly in the NIH trials, there were subsets of individuals that had relatively stable disease, even though they had diffuse proliferative, and the question is would that be a population that could be studied with steroids alone with an early escape, as Dr. Merrill has pointed out, for worsening disease. They then could be treated with a more aggressive therapy. The benefit to doing that would be to simplify the analysis and also eliminate the cyclophosphamide which, as I said, may actually make it difficult to demonstrate effect of any new therapy that we want to look at.

DR. WILLIAMS: Susan?

DR. MANZI: Well, I think everyone is gradually coming to the table with what I was going to pose as a

question. But first, I wanted to comment.

I don't think you have to go head to head with Cytoxan to show efficacy if you define what response is up front, what is an agreeable improvement, which is what Matt's group is doing. And if the drug does that, that's fine.

Then I was going to pose the exact question that Michael did. Can we conceptualize a trial with an escape clause so that we felt comfortable with that to just treat short-term steroid alone in proliferative disease? My contention is you could. I wouldn't see IRB issues and patient issues as barriers to that. I'm not talking about aggressive creatinines coming in at 2. Those are the kinds of exclusions that I think sponsors are aware of. It is just safety nets built in and just look at the efficacy of the drug based on a priori response. And that seems feasible to me.

DR. WILLIAMS: Jack?

DR. CUSH: I think, despite Michael's comments, which I agree with, it seems pretty simple. But there is an unfortunate discordance between what's obvious as far as the FDA regulations say, we can't have a noninferiority claim because of the shortcomings of what's been done thus far, but nonetheless, the fact is what has been done without much data is that the standard of care really is IV

Cytoxan for people with class 3 and 4 disease.

But knowing that we can't do that, we could go ahead and we could do a glucocorticoid head-to-head trial and prove at least equivalence, if not superiority, and have certain safety outs for toxicity reasons. But, unfortunately, that's very, I think, inhumane to many of our patients because 6 months of high doses of steroids they will hate and they will hate us for it and they will hate themselves. It's really unfortunate we can't do that.

To get to Jeff's question, I'll answer his by saying, how do I measure the outcomes here? I would want improvement or resolution in proteinuria/hematuria, a rise in creatinine, and some sort of serologic measures at least in 2 out of 4 for at least 6 months, and that would be my improvement in a trial.

DR. WILLIAMS: Gabor?

DR. ILLEI: I think that it's feasible to do a lupus nephritis study with steroids being the comparator or control, especially if you use pulse, mostly pulse Solu-Medrol. I think the last NIH studies has shown that at 6 months the response rate is fairly similar to Cytoxan. I think that there is a way to choose patients who have active proliferative disease but do not have bad prognostic factors and setting up strict withdrawal criteria. I think it can be done safely.

DR. WILLIAMS: We have other questions. We still have six more people on this one. Bevra?

DR. HAHN: No.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: I think again, to go back to the original NIH trial to try to get a sense of is it safe to treat patients with proliferative nephritis with steroids alone, remember that those patients had an average duration of nephritis of 11 months before they came in, and they were 100 percent caucasian. So I would say if you're going to look for lupus nephritis that is reasonably stable, then look at membranous nephritis.

But then you present the sponsors with a difficult task. The outcome of lupus membranous in general is going to be good, and you're going to change on one primary parameter which is going to be proteinuria. So you set a much more difficult task to show efficacy.

I think that John Esdaile has shown that the longer that you delay the initiation of cytotoxic therapy, the worse the long-term outcome in renal failure is.

So I guess I would take the opposite point. We're not trying to develop a drug for mild lupus nephritis. I think what we're trying to do is to develop a drug ideally better than Cytoxan with less toxicity. We're not trying to develop a drug for milder forms of the

disease. At least I'm not interested in that.

I think that we would be, knowing that race is one of the major predictors of poor outcome of lupus nephritis, in the position, if we're going to exclude high-risk patients, of depriving African Americans who, after all, have three times the incidence of lupus, of participation in such trials. And I would not ethically randomize an African American patient with proliferative nephritis to a steroid-only arm.

DR. WILLIAMS: Jill?

DR. BUYON: I was just going to say that I think it does completely depend on what's going to be the entry criteria, but at a meeting that several of us were at not really more than 6 months ago, I was the one heretic that proposed we have a head-to-head against prednisone. Just looking at practices, which you have to evaluate, nobody agreed with me that we could do that.

One thing we have absent here are any nephrologists. I don't think there are any nephrologists among us. I would submit that it would be difficult to do this in isolation without the opinion of a nephrologist because it was such individuals that felt that my proposal was unethical, and I think we do have to address that.

DR. WILLIAMS: Gary and then David, and then we're done with this question.

DR. HOFFMAN: I would be one to speak for a randomizing to a steroid-only arm, given the following constraints. I think you have to know going into the study what the damage and chronicity factors are. I think you need to know what the degree of global sclerosis is. I think people who have high damage indices are not going to be able to be enrolled in a study of this type, in part because their margin of safety, their opportunity for reversibility is modest, if at all existent. I think people would have to be enrolled based upon activity scores and opportunity for reversibility. I'm not aware of any study that has done that specifically and looked at steroids alone versus steroids plus a cytotoxic agent or any other immunomodulatory agent.

I do think that you can take patients such as that and randomize them to standard of care, which I think there's a consensus, although not approval through the FDA recognizing that as standard of care. I think you can have a Cytoxan arm under that scenario compared to a test agent. And I think your endpoints would be then outcomes that would measure improvement, and we've mentioned a number of those.

Reversibility, because we do know -- I don't think Dan meant this when he said it, that renal lesions are irreversible. I think certainly those that have high

activity indices and people presenting with RPGN with creatinines of 3 or even people on dialysis have reversibility. We have a number of people who have been on dialysis who have come off dialysis who have had acute renal failure.

I'm not suggesting that type of patient be included, but certainly people with high activity indices and increases in creatinine can be randomized under this scheme, and I think within a period of time that we think is reasonable -- reasonable as judged by experts -- we could have a bailout within even a period as short as 4 to 8 weeks before taking people out of a steroid-only arm and randomizing them into a standard of care versus test agent arm.

DR. WILLIAMS: I think we've gone as far as we can on this.

The next question is please discuss the importance of blinding in pivotal trials. In the context of phase I to IV trials, which trials can be performed unblinded and what is the justification? Bevra?

DR. HAHN: I've done some thinking about this one because with the new biologics, the difficulty of administering them, many of them are IV and it gets pretty complicated to do a placebo IV, and the IRB has some difficulty with the ethics of doing an IV in someone that

is getting a placebo through the IV. So it seems to me that if the assessors are blinded, the assessors of outcome, then it might not be even desirable to blind a study. I'd kind of like to see what the rest of the committee thinks about that.

DR. WILLIAMS: Norm?

DR. ILOWITE: I think if the parameters that are being looked at are very objective and not subjective, which would perhaps not include some of the domains and activity indices, then that would be legitimate. But an IV itself has a powerful placebo effect, so only half the patients would be getting that, and if there were subjective parameters measured, it might introduce bias. But if it were very objective parameters measured, I think it would be fine.

DR. WILLIAMS: Jill?

DR. BUYON: I would have a tremendous problem if it were a health assessment and that were unblinded. I can certainly tell you that in the SELENA trial, one of the points that was brought up and makes it difficult is if you think you know what someone is on, you're going to push harder for them to stay in a study. And I do want to emphasize how important that is, even though we're not talking about outcome measures, just having compliance and coming to visits, there's a push on the part of the

investigator who knows. We had a lot of issues with unblinding, and I actually would say blinding, as best as you could, would be important.

DR. WILLIAMS: Joan?

DR. MERRILL: I'm 100 percent for blinding in everything. There are so many little subtle things that happen. Even for a patient, it's depressing to find out you're not getting the treatment. As long as you're doing okay, it's reasonable to stay in the trial now knowing, and patients understand what they're doing when they go into a blinded trial.

You can't do an SF-36 unblinded. It's going to be useless information. Even though it might be a nephritis trial where I'd be very comfortable with the nephritis outcomes, you're going to want to be doing the other instruments. They're going to give you valuable information about your drug, and you really can't do the other instruments unblinded.

DR. WILLIAMS: The comments seem to be unanimous, and we're going to move on.

Question number 7. What would be the recommended duration of trials for non-major organ system studies? Could a therapy which treats constitutional manifestations be approved with a 3-month trial? What is the appropriate duration of trials to evaluate major organ

system involvement?

We'll take non-major organ system involvement first. Is a 3-month trial adequate? Jack?

DR. CUSH: Again, I won't discuss "constitutional." We voted that off the island yesterday.

(Laughter.)

DR. CUSH: I would stick with 6 months. I don't know why we have to go with 3 months. I think you can achieve maybe quick outcomes but then show maintenance or sustaining outcomes. So I think whether we're talking major organ involvement or signs and symptoms through disease activity measures, 6 months would be my minimum trial duration.

DR. WILLIAMS: Gabor?

DR. ILLEI: I agree. I would say 6 months is the minimum.

DR. WILLIAMS: Joan?

DR. MERRILL: I could imagine circumstances under which a primary outcome measurement might be much earlier but I'd still want the trial to go 6 months to see if it's maintained.

DR. WILLIAMS: Gary?

DR. HOFFMAN: I think 6 months is essential, or longer, because part of what you want to build into a study that looks at major organ or even minor organ involvement

is the ability of the test agent to allow that patient to have a very meaningful reduction in steroids or get off steroids, and I don't think you'll be able to say anything about the durability of that therapy in terms of its steroid-sparing effects within 3 months.

DR. WILLIAMS: Jack?

DR. CUSH: I'll stick to 6 months, but I will speak to issues as it relates to placebo-controlled trials and when you can exit them out, especially if it's life-threatening organ involvement. There should be earlier exit points with rules for that built into the system to allow for appropriate analysis maybe at an earlier point, but the desired outcome still should be 6 months.

DR. WILLIAMS: Again, there doesn't seem to be a lot of controversy on that particular one.

How about for major organ involvement? How long should the trials be?

DR. HOFFMAN: I think that would need longer, 1 to 2 years.

DR. WILLIAMS: Joan?

DR. MERRILL: I'd like to get back to the idea of induction and maintenance. I don't think you're going to approve a drug without knowing its long-term effects, but it might be that a trial could be an induction trial and be very helpful to collect extra data or peripheral

data. And maybe that's a more complicated trial, but then another maintenance trial. That's one concept I hope we could leave on the table.

DR. WILLIAMS: Lee?

DR. SIMON: In relationship to both Joan and David's comments, I'd like to point out that we've learned a lot about 2-year trials. They can't be done in the context of a pre-approval, real trial design. Patient dropouts are too dramatic. Rescue therapy intervenes. The interpretation of the data becomes very difficult.

So our experience basically is a 1-year trial is about the limit that you can get from the point of view of a trial trial, and then you can do extensions in that patient population that have built within them the caveats of dealing with the extensive dropouts in patient attendance and a zillion different reasons, moving away, getting married, not really safety issues, but the normal everyday things that we all have problems with. Coercion of patients into participating longer based on rewards and whatever, as everyone here knows, is a very big no-no according to IRBs.

So in the context of that, I think the induction idea is a wonderful one because you can get an early response, maybe even as Joel on the side here suggested, an escape at 1 month and then go on looking at

some other issues.

But what about extension trials? Because one of the big issues is durability of response. So in your comfort zone, you see a response at 6 months. Do you expect to be able to -- let's say, it's a significant improvement, a la Matt's definition of that in lupus nephritis. Would you like to see that that response is maintained for another 6 months? 18 months total? What would be your comfort zone there?

DR. PISETSKY: Some of this depends I think on the mechanism of the agent and what you were doing. Obviously, anti-TNF drugs you'd keep on forever and you want to see them sustained. You stop, things get worse. But it doesn't mean they're not useful. On the other hand, Cytoxan, interestingly enough, is a drug you stop and then you observe.

DR. WILLIAMS: Mike?

DR. WEISMAN: Lee, I think that you have to define whether or not you're talking about disease activity or maintenance of a disease-free state. I think if you're just looking at disease activity, I don't have a problem even with a 3-month trial, if that's all you're looking at. But when you're talking about taking patients from one state to another, which is the issue we had with ankylosing spondylitis, you remember, how long do you need to observe

that patient or that trial for the durability of continuing the patient from one state to another? I think that at least here, 1 year has got to be the maximum, according to you, and probably 6 months would be the minimum.

So if our goal here is to provide impetus to companies to push drugs further -- and I think that is one of our goals -- I would move the threshold for disease activity to 3 months and maintenance of a disease state, whatever you want to define, load, activity, remission, whatever, between 6 and 12 months. That's how I would vote it.

DR. WILLIAMS: Jack?

DR. CUSH: I don't think that lupus, for most people, especially the problematic patients we may be talking about here, is a disease where we're on and off therapy, much like the gastroenterologist may be for Crohn's or the dermatologist may be for psoriasis where they think of the interventions they do as being short-term. I think that when we step up our therapies based on disease activity, we do so for a sustained period of time because lupus doesn't quickly remit.

I think that if you show efficacy, whether it be for signs and symptoms or for organ-specific indications, for 6 months, I think you've met the bar. I think beyond that you're only showing durability, A, and B,

safety. You still only need to meet the bar at 6 months. I think the 6-month extension should be strongly recommended for those other two caveats, but for purpose of approval, I don't know we need to go beyond 6 months.

DR. WILLIAMS: John?

DR. LOONEY: I guess I'd agree with Michael. For a disease activity index where you're trying to show that the drug is sort of globally effective for signs and symptoms, 3 months seems to be fine to establish that. I think that I wouldn't want to set the bar higher in lupus than it has been in the past for rheumatoid arthritis.

DR. WILLIAMS: Jill?

DR. BUYON: I would just actually disagree. It depends on what you're looking at, and for renal disease, anything less than a year to me would be inadequate, despite the fact that I understand 2-year trials have their problems. We are looking at an undulating disease, and we could easily get caught in a capsule of time, wind up with an indication and be severely slapped in the face within a year afterward. I would personally find that an embarrassment of the FDA to approve such a drug. If it were renal disease, I think 1 year would be the absolute minimum, and I'd be worried about that too.

DR. WILLIAMS: I think we have to move on.

Should pediatric patients be incorporated into

trials of adult SLE or studied separately? Norm?

DR. ILOWITE: It depends what you mean by "incorporated." Certainly issues to consider would be that it's likely that the centers would be different. Different data would have to be collected, including things like, depending on the length of the study, growth, sexual development, cognitive development. The children wouldn't be static in any of those areas. The SLEDAI might have to be modified to include things like school performance, school attendance. Pharmacokinetic data may need to be obtained differently because most children won't submit to sampling over a course of a day, and population PK methods would have to be used or likely to be used. So, sure, they could be incorporated, but it would almost be a separate study that was ongoing with the adult study.

I think most pediatric rheumatologists agree that lupus in children as a disease is very similar to lupus in adults, and it's just the children that are different than adults that makes it different.

DR. WILLIAMS: Joel?

DR. SCHIFFENBAUER: Can I just clarify? If the primary outcome was some measure of renal disease, could you mix that outcome, forgetting for the moment that you would need to look at growth and sexual development in the kids, but if the primary outcome were development of renal

disease of some shape or improvement in renal disease, could that all be mixed together in a single trial?

DR. ILOWITE: Yes, I believe that the measures would be very similar, and especially if it were noninvasive, that shouldn't be a big problem.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: I would agree that pediatric patients should be considered to participate, but I think there are a couple of issues, as Dr. Ilowite has suggested. I think it would be folly to have a trial at an institution where you didn't have close collaboration with pediatric colleagues.

And then I think the other issue is about corticosteroids during the trial because the younger lupus patients are oftentimes given twice the dose that adult patients are given and may be tapered more slowly than our adult patients. At our institution, our pediatric nephrologists define children as up to age 21. So there's a slippery area in there.

So I think that there would be unique considerations but that we certainly should make every effort to include pediatric patients.

DR. WILLIAMS: Jeff?

DR. SIEGEL: One part of this question that maybe wasn't made explicit is that studies in children are

often delayed until after approval of the agent for adults. So one question I would like to get some feedback on is whether this model should be practiced in lupus or whether there's a sense that children should be included in clinical trials before approval in adults.

DR. WILLIAMS: Jack?

DR. CUSH: There's sort of a practicality behind that, but I think there's also a safety issue behind that, and I think that safety should reign and to allow this to be tested in adults first, and to look at the most common or major toxicities that may arise and how that's going to impact the pediatric population would be prudent before going forward in at least a few studies, maybe a large phase II or at least have a reasonable amount of information before proceeding to an initial phase II in children.

DR. WILLIAMS: Norm?

DR. ILOWITE: I agree with that. Especially if there's animal data to suggest a unique toxicity in young or developing organisms, then it would be more ethical to test it in adults first.

DR. WILLIAMS: Gabor?

DR. ILLEI: I agree with Jack.

DR. WILLIAMS: Bevra?

DR. HAHN: Could we find a compromise age? I

mean, there are so many people who start with lupus when they're 15 or something like that. Is there an age at which we worry less about effects on growth, effects on sexual development? After people have passed puberty, is it okay to include them in these studies? Because it's a tremendous delay for people in that age group to have to wait for 2 or 3 years, if they have bad lupus, to get something experimental.

DR. WILLIAMS: Mary Anne?

DR. DOOLEY: If you have lupus as a child, you're much more likely to have renal disease and much more likely to have frequent relapses. So in some respects, they have a more severe disease. So if we could define a group that could be included earlier.

DR. WILLIAMS: Joan?

DR. MERRILL: I fear that what ends up happening is, for well-intentioned reasons, people are a little scared to test things on kids, and what ends up happening is we never do find out how things work in kids. And the trials, after drugs are approved, really aren't funded very well or much. So I have a lot of teenage lupus patients in my practice, not because I wouldn't want them to see a pediatrician, but because of the availability of doctors. I know that these people and their parents would like them to have access to the opportunities that other

patients have on a case-by-case basis, and I'd like to make it available to them.

DR. WILLIAMS: Norm?

DR. ILOWITE: Well, certainly I agree that it's important to study these medications in children as soon as possible.

Bevra, in answer to your question, if we make the entry criteria for older children, essentially we're studying them in young adults and it's a advantage/disadvantage continuum, whereas we like to get the data in young children also because they're the ones who are going to differ from adults the most, and that's where we get the new information. So, yes, there is probably a cutoff where adolescents could be included in an adult trial without much modification, but it would give limited information.

DR. WILLIAMS: We are overtime on this open session. There are still three more questions. Can we delay those?

This will end the open session. The closed session will begin in 10 minutes. We need to have everyone but the FDA and the committee leave the room in that time. So we'll reconvene here at 11:15.

(Whereupon, at 11:05 a.m., the committee was recessed, to reconvene in closed session at 11:15 a.m.,

this same day.)