FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

MEETING OF THE

ARTHRITIS ADVISORY COMMITTEE

8:10 a.m

Monday, September 29, 2003

Versailles Ballroom
Holiday Inn
8120 Wisconsin Avenue
Bethesda, Maryland

ATTENDEES

COMMITTEE MEMBERS:

GARY S. FIRESTEIN, M.D., Chair
Division of Rheumatology
Allergy and Immunology
Department of Medicine
University of California at San Diego
9500 Gilman Drive
Basic Science Building, Room 5098
La Jolla, California 92093

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

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

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

ATTENDEES (Continued)

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

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

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

ATTENDEES (Continued)

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

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

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

THOMAS A.J. LEHMAN, M.D.
Hospital for Special Surgery
Cornell University
535 East 70th Street
New York, New York 10021

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

ATTENDEES (Continued)

SPECIAL GOVERNMENT EMPLOYEES: (Voting)

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

GUEST SPEAKERS: (Non-voting)

BARBARA BOYTS
President
Alliance for Lupus Research
28 West 44th Street, Suite 1217
New York, New York 10036

MARGARET DOWD
President
Lupus Research Institute
149 Madison Avenue
New York, New York 10016

SANDRA RAYMOND
Executive Director
Lupus Foundation of America
2000 L Street, N.W., Suite 710
Washington, D.C. 20036

KATHERINE M. SNIDER
Executive Director
Rheuminations, Inc.
221 East 48th Street
New York, New York 10017

ATTENDEES (Continued)

GUEST SPEAKERS: (Non-voting) (Continued)

VIBEKE STRAND, M.D.
Biopharmaceutical Consultant
Adjunct Clinical Professor
Division of Immunology
Stanford University
306 Ramona Road
Portola Valley, California 94028

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

MARC WALTON, M.D.
Office of Drug Evaluation VI

JAMES P. WITTER, M.D., PH.D.

ALSO PRESENT:

LISA AMATO
PAUL BRUNETTA, M.D.
KELLY COOPER
BETTY ANN EXLER
SCOTT EXLER
LINDA L. NARDONE, PH.D., RAC
VENETIA THOMPSON

C O N T E N T S

Systemic Lupus Erythematous Concept Paper

* * *

C O N T E N T S (Continued)

P R O C E E D I N G S

(8:10 a.m.)

DR. FIRESTEIN: I thank everybody for coming today. I am Gary Firestein and this meeting is to talk about systemic lupus erythematous.

Now, unfortunately, we don't have the questions yet. They will be here very shortly, I'm told, and will be passed around so that we can give them intense scrutiny during Lee's opening statement.

In any case, do you want me to go around with introductions first? Okay. There are a number of new people at the table, so why don't we go around before have the meeting statement, starting at that end.

DR. SIMON: So I'm Lee Simon. I'm a rheumatologist. I'm the Division Director of Analgesic, Anti-Inflammatory and Ophthalmologic Drug Products at the FDA.

DR. SCHIFFENBAUER: Joel Schiffenbauer. I'm a medical officer in the Division of Analgesic, Anti-Inflammatory and Ophthalmologic Drug Products.

DR. SIEGEL: Jeffrey Siegel, Acting Branch Chief, Division of Clinical Trials, Office of Therapeutics at the FDA.

DR. WEISMAN: Michael Weisman, Director of Rheumatology, Cedars-Sinai Medical Center.

DR. WALLACE: Dan Wallace, Professor of Medicine, UCLA, member of the division at Cedars-Sinai, Los Angeles.

DR. BUYON: Jill Buyon, New York University Hospital for Joint Diseases.

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

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

DR. FINLEY: Michael Finley, Western University, Southern California.

DR. ILOWITE: Norm Ilowite, Schneider Children's Hospital, New York.

DR. MANZI: Susan Manzi, University of Pittsburgh.

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

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

DR. FIRESTEIN: And again, I'm Gary Firestein, a rheumatologist from UCSD.

DR. WILLIAMS: Jim Williams from the University of Utah.

DR. ANDERSON: Jennifer Anderson, statistician, recently retired from Boston University.

DR. CUSH: I'm Jack Cush. I'm a rheumatologist from Presbyterian Hospital, Dallas.

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

DR. PISETSKY: David Pisetsky, rheumatology, Duke University Medical Center.

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

DR. DOOLEY: Mary Ann Dooley, rheumatologist, University of North Carolina, Chapel Hill.

DR. HAHN: Bevra Hahn, rheumatology, UCLA.

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

DR. LOONEY: John Looney, rheumatologist, University of Rochester.

DR. LEHMAN: Tom Lehman. I'm Chief of the Division of Pediatric Rheumatology at the Hospital for Special Surgery in New York.

DR. FIRESTEIN: We have one person who has a telephone connection. Dr. Liang, are you there?

DR. LIANG: Yes, I am.

DR. FIRESTEIN: Would you like to introduce yourself?

DR. LIANG: I'm Matt Liang, Harvard.

DR. FIRESTEIN: Thank you very much.

DR. STRAND: Gary, am I allowed to introduce myself?

DR. FIRESTEIN: I don't know.

(Laughter.)

DR. FIRESTEIN: But that's Vibeke Strand.

Why don't we go ahead with the opening statement, please?

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 the meeting.

The committee will discuss the proposed systemic lupus erythematosus 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 proposed sections regarding the current state of the art, the claims for treatment, and clinical markers.

The topic of today's meeting is an issue of a particular matter of broad applicability. Unlike other issues coming before the 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 had 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 many firms and institutions, it is not prudent to recite all potential conflicts of interest, but because of the general nature of today's discussions, these potential conflicts are mitigated.

In the event that any discussions involve any other products or firms not already on the agenda for which FDA participants have a financial involvement, 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 firm whose products they may wish to comment upon.

Thank you.

DR. FIRESTEIN: Thank you very much. So we actually do have a very busy schedule today, and I'm going to ask, please, if people can try to stay on time with regard to their talks.

So as a way of introduction, I'm going to ask Dr. Simon to give his overview and he also is asked to stay on time.

(Laughter.)

DR. SIMON: First, I'd like to welcome everybody here, thank the committee and all the people that have volunteered -- well, not volunteered but have donated some of their time to come to this meeting. The committee is much larger than normal which is partly related to the importance of the discussion we're going to have.

This is really an interesting time. This is the second iteration, I presume, of a meeting that was held about six to eight years ago with this committee discussing very similar topics, but I believe that not only has the science evolved but also our thinking has evolved.

The division and the agency has been very hard at work since I arrived two years ago in looking at the question of systemic lupus and the lack of a guidance associated with that field. Some of you may be quite aware that we in the division, as well as what has now become

part of CDER, the CBER folk -- and I do welcome as of tomorrow our CBER colleagues into CDER as part of the CDER/CBER merger -- we have been trying very hard to get clarity over very many important issues that will allow us to create a forward-thinking, flexible, and appropriate document that will lead us to a new understanding of how to design clinical trials and study patients with this disease.

Therefore, we put this meeting together to discuss an ongoing document development, now presently called a concept paper, due to issues regarding getting draft clearance of a document as a true draft guidance. So we're discussing at this meeting a concept paper which has been developed by multiple groups within the agency to determine the basis for guidance for the development of therapies in systemic lupus.

We are going to spend time today and tomorrow discussing aspects of claims, discussing aspects of trial designs, and the issue of application of possible accelerated approval in the context of subpart H and E, depending if it's a drug versus a biologic.

The idea is that we are doing two things. One is getting clarity about how to study this disease, and secondly, how to foster interest in the field, particularly from sponsors and others that have the wherewithal to be

able to allow clinical trials to ensue.

I thought it would be useful to take a minute for everyone to be up to snuff on what is a guidance document. So what is a guidance document? Well, many of you do not know that the CFR is the Code of Federal Regulations. This is actually not your most entertaining reading, but nonetheless it is reading that's imperative to understand how the government and the agency works. You'll hear more about this from others, but basically the 21 CFR 10.115 defines a guidance document as those prepared for FDA staff, applicants and sponsors and the public that describe the agency's interpretation of or policy on a regulatory issue. A very key important issue associated with the development of drugs or biologics or devices.

Guidance documents could include description of the design, production, labeling, promotion, manufacturing and testing of regulated products, the processing content and evaluation and/or approval of submissions, and inspection and enforcement policies.

It is important for everyone to understand and recognize that guidance documents do not establish legally-enforceable rights or responsibilities. They do not legally bind the public or FDA. Anyone can choose any other approach than one that's set forth in the document, especially in that science is constantly evolving.

I want to point out, however, that the alternative approach must comply with relevant statutes and regulations and the FDA is willing to discuss alternative approaches that will make sure that they comply with these requirements, and although not legally binding, everyone should be aware that it's important to note that such a document represents the agency's current thinking.

It always seems to be frustrating to people to recognize that the document itself changes on a regular basis as well through these kinds of meetings as people begin to discuss aspects in the real world once the document is created. If the FDA departs from the document, it does require open public discussion about it and it requires justification and supervisory concurrence in a hierarchical way, in a hierarchical structure.

Now, what are the procedures for developing such a document? I've taken great advantage of that as Division Director at 550, or Analgesic, Anti-Inflammatory and Ophthalmologic Drug Products. Before a working draft is developed, we can as the agency seek or accept early input from individuals or groups outside the agency and it can be done through participation in or holding public meetings and/or workshops, and for anyone in the room who actually has been in Antarctica, we've actually had a lot of these meetings already. I have to thank in public the

four interested voluntary agencies, the four foundations that have spent a lot of time and effort with us in helping us understand some of these questions and providing great support for the community to be able to determine some of these important issues.

Once these kinds of early discussions take place, then a formal process takes place, and the formal process is that a document that's created is then reviewed internally by the people that know about documents and know about words and ensure all the words are appropriately created to be able to represent a very safe document from a government point of view, and then we publish a notice in the Federal Register, and then we post a draft on the internet. We make hard copy available. We invite comment. We hold further public meetings perhaps, if necessary, and then once decided and finalized, again notice is published in the Federal Register and the document is placed on the internet and made available as hard copy.

Typically, the agency functions, once it is in draft format, as if it exists and we work with it as such. In the context of this particular arena, because we'd like to foster development as rapidly as possible and given the importance of this meeting, we will go forward in thinking about the approach as we discuss today and as I'm sure you've been discussing with the various different divisions

in the past to ensure that we have continuity and that we will actually live up to our commitments that we've had before as we're beginning to evolve into this new realm of an accepted guidance.

I want to remind everybody the importance of why we're doing this seemingly interminable work. We want to get clarity and public buy-in of what we'd like to achieve, and one of the reasons for that is that for the last 30 years, we've not had a lot of drug development in this field. These are the three agents that are presently approved for the treatment of various different aspects of lupus within the United States: hydroxychloroquine, glucocorticoids, and more recently low-dose acetylsalicylic acid to prevent cardiovascular complications.

These are the drugs that are used off-label in the United States for the treatment of systemic lupus and its manifestations, and I won't go through the entire list. Many of you know what these things are.

It's important to recognize that to the agency's point of view -- and it's an important component of comparative trials that we will discuss as the day goes on today and tomorrow -- is that the prospective proof of some of the utility of these agents is quite lacking, and one of the reasons it's lacking is because of a lack of clear understanding about how to get therapies approved for

this particular disease, only one of the reasons, not all of the reasons.

One of the major issues that we all need to remember is for agency approval in comparative studies, one has to compare against a drug that's already approved in the field for the new drug to be approved or the standard of care or drug that's already in the field serves as a "placebo" compared to the new drug, but you'll hear much more about that as we go through. The important thing to remember is what we're trying to accomplish, which is to develop understood therapies for the treatment of this disease.

There are issues about lupus that make it difficult to do this, and it only should be something that we have to grapple with and get over rather than to leave it as an obstruction. The unique characteristics of the studies in lupus are highlighted by the heterogeneity of the disease, its unpredictability, and the heterogeneity of the patients and their manifestations.

We've noticed over the years that morbidity and mortality has spontaneously improved in the last three decades. One can argue it's just better therapeutic approaches and using unproven therapies, not that we don't believe they may work, it's just from a regulatory perspective, it's hard to find the data that proves that

they work. Remember, the agency regulates interstate commerce. It does not regulate medical care.

There's a lack of clear outcome measures and there's an issue of fixed damage and how one analyzes that, damage either due to therapy or damage due to the disease.

What about the length of time observed that can lead to a desired response? Sometimes some people believe that the ability to understand lupus nephritis might take three to five years from looking a real clinical outcome being either saving renal parenchyma or preventing end-stage renal disease, the lack of clear guidance so far.

And the disease course is extremely difficult to predict a priori, typically observed flares and remissions, multiple organ system involvement, and progression is quite variable and recurrences are very hard to predict.

One of the major issues that we need to discuss today and we have several questions that are actually devoted to this particular area and you will hear a talk or components of a talk about this particular issue -- so we believe that this is very important, and we have major discussions going on within the agency -- is the area of the applicability of disease activity indices and what they really mean and how to use them in the context of trial design.

What are the issues regarding them, including reliability and validity? What are the usual levels in active disease, and what is their responsiveness to treatment defined in corroborating prospective studies?

We at the agency in thinking about the utility of outcomes think about them in the context of their being proven within prospective analyses, not just retrospective studies, not to suggest that that is not important, but it's the prospective proof that allows us to understand their applicability.

So what we're doing here is thinking about the issues for regulatory approval, and again I would like to point out, as we have discussed before in this venue and in other venues, this is not to suggest that other trial designs are not important. There are many, many, many questions that all of us think about that are important from an academic clinical point of view but are not applicable to the regulatory process. So this is a parallel issue that goes on and drives the entire field.

So what are the issues regarding the effects of the systemic inflammatory disease on the whole person in the context of a regulatory approach? Well, we're interested in disease activity or measurements of disease activity. We're interested in measuring replicate data that describes response to therapy, amount of damage

prevented which would have been caused by the disease versus the fixed damage not able to evidence improvement, and maybe an important observation, if not worsening, and damage caused by the treatment. All of these things are very important to the regulatory environment.

We're also interested in thinking about not only the whole person but individual organs, the disease activity within that organ as to how it's measured, how one measures response to therapy in a time that's applicable to a clinical trial, but also in a time that may allow us to understand that the long-term nature of the changes within that organ may be predictable based on what we're using to measure the change, how much damage would be prevented which would have been caused by the disease, and how much damage that would be caused by therapy.

Then we're also interested in improvement in target area, such as in the organ we just talked about, but that the disease does not worsen elsewhere. Obviously a therapy that makes patients sick is not a particularly useful therapy, and one way patients may be sickened is in fact if you make something better in one area and you get worse in other areas.

In the context of trying to predict improvement in a relatively short period of time of a clinical trial that actually has importance as it relates to clinical

outcomes has been a discussion that's been going on for quite some time, both within and without the agency. This relates to surrogates and biomarkers.

Now, surrogate endpoints are candidate criteria for drug approval, but a broader term, biomarker or early marker, is commonly used. We actually had a meeting about this issue to discuss this related to lupus per se, but there are many other areas within the agency that are grappling with these particular problems.

It is important to remember that an early marker or a biomarker does not have the same regulatory implication as something that is labeled surrogate, and a surrogate may be a biomarker or early marker but not all biomarkers are surrogates. You will hear more about this discussion in our section talking about surrogate markers.

The importance of this is to gain clear, absolute understanding if we identify some marker as a useful way to predict response, that there is a clinical link to that marker that is well understood, well accepted, and has prospectively been proven. So in the scenario of a complex disease, it is important to be creative in trial design. We need to get much clearer clinical endpoints so that we understand what we're measuring and what the outcome will mean in the long run as well as in the short run.

It's important to remember that determining safety requires a robust data set. That has been a major debate both in the field of rheumatoid arthritis and in drug development, as well as in this one. We are looking at enough patients to understand the implications of intervention.

We will support the idea of studying therapeutic effects regarding the state of disease as well as specific organ involvement, yet the overall state of disease cannot worsen. That's a very critical part of what kind of measures that you build into a clinical trial.

Given the heterogeneity of the disease, we might consider the development of a responder index. Perhaps that may be one way to go to get away from multiplicity issues of measure which are always a bugaboo to our statistician colleagues.

Early and active dialogue with members of the agency is strongly recommended. Whether you are an academic investigator looking for an IND approval or a sponsor sitting out in the audience, it's critical for you all to talk to us as much as possible, whether we're in ODE 6, in the old CBER group that's now in CDER as of tomorrow, or whether in ODE 5 or even in ODE 2 with cardio-renal. Basically, early discussion is critical.

So our agenda that we're going to be following

today is going to be somewhat unique, and I would ask everyone here to bear with us as we work through this process. I ask the chair to do the same. So we're going to review the state of the art and have some very specific talks, and I'm incredibly grateful to the people that have volunteered their time to be able to do this, either those on the committee or those that are not on the committee.

Two is we're going to talk about potential claims. We're going to talk about the potential for accelerated approvals and application of early markers or biomarkers, and then we're going to talk tomorrow about trial designs.

What we've constructed here are three opportunities for people from the floor to actually make comment spontaneously, in addition to the open public forum that is traditionally held that requires you to submit an a priori application to be able to talk, and in between each major topic, we will have these new unusual discussions. There are microphones on the floor available for you for this.

At the discretion of the chair -- and I reiterate at the discretion of the chair -- you can get called on by standing at the microphone when we open up this section. Our exec sec, Kimberly Topper, will make an announcement before each of these sessions identifying what

you need to do when you come up to the podium and get called on. You need to disclose the issues about why you're here, support, conflicts of interest, for the community to understand in the context of your comments.

So that's all we're going to do. We're going to hopefully get major discussion going. I remind the people who actually come up to the microphone that one of the no-nos in this process, you cannot actually ask the committee members anything. You cannot engage in a discussion with the committee members. The committee members have their own discussion.

Once we actually finish the discussion regarding each of those sections, we will then move on to have a committee discussion and answer the questions. Again, the questions will be handed out as soon as we get them in an organized fashion.

The only other departure we're going to do this time as opposed to other meetings that you've experienced, we've invited the four major advocacy groups to actually spend some time sharing some of their thoughts with us about the importance of this document and the importance of the process at the beginning of the meeting, near the beginning of the meeting.

So I'm giving you heads up about some of the differences and uniqueness of this meeting, and I

appreciate everybody's up-front cooperation with the various different things that we're trying to accomplish here, and I wish you the best of luck and thank you very much, Mr. Chairman.

DR. FIRESTEIN: Thank you, Dr. Simon. We're a little bit out of order here actually because the discussion of the state of the art is going to begin with Jill discussing objective laboratory measures.

DR. BUYON: I thank you for inviting me. I'm afraid I've caused some AV disruption this morning.

And I've also done something I've never done before: I've taken the liberty of changing the topic. So instead of talking about anti-DNA and creatinine, I've chosen to consider that a typo and I'm actually going to talk about anti-DNA and complement because that seemed to make the most sense to me. So, Dr. Simon, my apologies.

So at any rate, I start out with basically taking us back probably to the '80s and this was the paradigm that we were all taught in medical school. The longitudinal, clinical, and autoantibody profile in the patient with lupus nephritis. It was very obvious and it was simple. Following along the blue line, patient's anti-double-stranded DNA antibody would go up. Concomitantly, the complement would drop. We'd all predict there would be active renal disease. The patient would be treated with

prednisone, anti-DNA antibodies would fall, complement would come up, and everybody would be happy. Voila. We have our biomarkers.

What was also interesting is that certain autoantibodies, the ENA, SMITH, RNP, RO, LA, did not track disease. What I want to evaluate really is now moving forward, evaluating anti-DNA and complement proteins in diagnostic testing and that will just take us back to the basics.

Well, what is the scientific reason for using these as biomarkers? Anti-DNA antibodies are specific to lupus, anti-DNA antibodies can deposit in the glomerulus. They're generally of high avidity, IgG, cationic and fixed complement.

Well, what about complement? There's evidence that complement consumption indicates immune complex-driven inflammation. Genetic alterations in the early components of complement, the classical pathway, are associated with lupus, and there's clearly an association between genetic polymorphisms of FC receptors clearly in these immune complexes and renal disease.

Just to take us back to biology for one second, we can see that anti-DNA antibodies and their antigen, the immune complex, clearly activate in the classical pathway, and the consequences are generation of C3a and C5a, both of

which are strong chemotactic factors and anaphylotoxins and most likely play a major role biologically in glomerular nephritis and fetal loss.

Just to remind you a little more of the science of why is it important to consider these as biomarkers is C3a and C5a do something. We all understand a little bit about the vascular disease of SLE and what I'm showing you in the slide is that this is the endothelium. The endothelium lining in the normal case has some of the adhesion molecules, a little bit of ICAM-1, and these are our neutrophils. When we generate C3a and C5a, perhaps through anti-DNA antibody complexes, we cause increased CR3 expression on the leukocytes, we increase the adhesion molecules, and we get leukoaggregation which may be relevant to the vascular disease of SLE. So complement plays a key role biologically.

So from the tissue perspective, this is a cerebral infarction, subtended by a vessel occluded by an aggregate of polys.

Well, what are the playing rules then, having looked at what the pathology might be and substantiating that anti-DNA antibodies and complement play a role in the pathogenesis of this disease? Let's look at the playing rules for evaluation of the biomarker.

It depends on how you test these biomarkers,

and I submit that is part of the playing rules. The first assay over here is the Crithidia assay, which picks up high and low affinity antibodies. It could be IgM or IgG, and it clearly favors and does pick up double-stranded DNA antibodies, but if you're measuring the antibodies by FARR -- that would be FARR out there -- high-affinity antibodies, IgM and IgG, picks up single-stranded and double-stranded DNA antibodies. The ELISA high- and low-affinity antibodies, you can choose IgM or IgG, you can choose single-stranded or double-stranded, and you can see there are differences in sensitivity versus specificity.

What about complement? There are immunochemical assays picking up native C3, C4. The specimen used is serum, generally done by nephelometry. You can measure functional integrity, CH50, EDTA plasma, measuring red cell lysis, and then you can measure the catabolic state, for example, looking at activation products, such as C3a, again EDTA plasma, measurement ELISA.

So it matters what rules, what instruments you use to measure the biomarkers and not only that, we obviously have to define the parameters of change for these candidate biomarkers.

So the question we'll ask is: does the candidate biomarker predict flare? Does it associate with

flare? Does it respond to therapy in parallel with favorable clinical outcome? An association between a factor and the risk of a disease does not guarantee that drug-induced changes in that factor will produce a corresponding change in risk.

So now let's examine several of the studies, and this is going to be very hard for me to go through without you being able to read this. But what we're looking at is a study by Michelle Petri and Audrey Ho, and she evaluated the percent of visits with flares categorized by prior and concomitant changes in the levels of anti-DNA antibodies and that's very important to keep in mind. So she defined "prior" as between visits 2 months and 1 month before the visit with flare, and she defined "concurrent" as between the previous visit and the current visit.

What she showed in this paper is that a prior increase of DNA antibodies to just 10 percent, when she compared, there were 70 visits that met that criteria, 30 percent associated with flare, compared to 19 percent in the overall group with a significant p value. But oddly and unexplained, when she increased the bar to greater than 25 percent, in fact, there was no significant difference.

However, if we look at a prior increase of DNA antibodies by doubling of the Crithidia as the criteria of change, there was a highly-significant difference in

detecting flare or rather predicting flare.

But one of the other points of her paper, somewhat unexpectedly, is the concurrent decrease in DNA antibodies by either ELISA or Crithidia at the time of flare was also a very strong association. So her message was that yes, some of these tests may be useful for prediction, but it is the decrease of the anti-DNA antibody by ELISA or Crithidia that went along with the disease flare.

Now, interestingly, Arthur Kavanaugh did a re-analysis of these data -- and I found this interesting -- looking at the likelihood ratio. And the LR for a positive test is the extent to which a positive test increases pre-test the likelihood of disease. So a high number is good and 10 would be good, and what you can't see here is that says sensitivity and 1 minus specificity. When both are up, obviously your number will be up, and if you recalculated Michelle's data, that turned out to be 2.7.

An LR for a negative test determines the post-test probability of disease after a negative result, and again what you can't see here is that the 1 minus sensitivity over specificity. If both are up, then the number is down. Hers was .081. So the conclusion would be that these tests have limited utility in predicting or excluding lupus flares.

Now, what about taking this the other way around and that is clinically-active, serologically-quiescent disease. This was a study out of Toronto. 514 patients were followed at the Toronto Lupus Clinic between 1991 and '95. 62 patients had clinically-active, serologically-quiescent disease and, interestingly, 43 with CNS renal or vasculitis. 58 patients had follow-up after the last defining visit, 9 remained that way for 3 years, 23 became active, 21 became clinically and serologically active, and 5 became serologically active but clinically stable.

That brings us to the last two studies I'd like to highlight, both unpublished as opposed to the others I just presented, and this is really data from NYU looking at the evaluation of the sensitivity and specificity of C3, C4, CH50, anti-DNA, and C3a for detection of lupus flares within 3 months. The cohort were actually patients enrolled in the SELENA trial. This was a randomized, double-blind, placebo-controlled trial, 496 females were enrolled, and SLE patients were treated with either HRT/placebo or OCP/placebo.

The analytes measured, as you can see, were the complements and complement-split products in anti-DNA done at baseline every month, monthly times 3, then every 3 months over a 12-month period, and the outcomes looked at

were severe flares and mild/moderate flare. Disease activity defined by SELENA SLEDAI and PGA.

Well, the approach taken here was to define the change in analyte prior to beginning the study. So the measurements are shown on the side, as you can see. C3a, a greater than 50 percent increase from the previous visit and an absolute level greater than or equal to 500 nanograms per ml. The CH50, greater than or equal to 25 percent decrease from previous visit. C3, same; C4, same; and anti-DNA antibodies, greater than 25 percent increase from the previous visit. And the previous visit by definition had to occur within 3 months from the date of measurement.

Our definition of flares is shown here, mild or moderate flare, a change in SLEDAI greater than 3, new or worse lupus rash, nasopharyngeal ulcers, pleuritis, pericarditis, arthritis and fever, any increase in prednisone up to .5 milligram per kilogram per day for treatment of lupus, added nonsteroidals or Plaquenil for disease activity, or a physician global assessment with an increase greater than 1 but less than 2.5. Severe flare was very specifically defined as a change in SLEDAI to greater than 12, new or worse CNS lupus, vasculitis, nephritis, myositis, thrombocytopenia, hemolytic anemia, requiring at least a doubling of prednisone greater than .5

milligram per kilo, and hospitalization, or the institution newly of Cytoxan, azathioprine, or methotrexate, and increase in PGA to greater than 2.5.

So these patients were available, as you can see, 496 total patients: 328 on HRT and a 168 OCP. 428 patients had levels that were available, and these are the differences. And flares, including multiple flares, there were 491 mild/moderate flares, and 39 severe flares.

And these are the data looking at the sensitivity and specificity of analytes to predict flares. What I want to point out is that every one of these measurements were highly specific for both mild/moderate flare and severe flare, but only the C3a to a level of greater than or equal to 500 nanograms per ml conferred a somewhat decent sensitivity. So to go over the limitations and implications, if the utility of analytes improved, perhaps a definition of positive test would be less stringent as in Dr. Petri's study. Perhaps analytes every 3 months is insufficient and we must consider monthly, and the absence of abnormal analytes does not equate with clinical stability but the presence may be predictive of flares, and then finally, a priori treatment with abnormal analytes may be appropriate since few patients will be unnecessarily exposed.

Just to finish up with this study,

serologically-active, clinically-stable patients -- and here the objective was to evaluate steroid treatment in averting flares when elevations of plasma C3a are accompanied by rising anti-DNA titers. The inclusion criteria that anti-DNA antibodies had to be present within 2 years -- and that's an important point. Are we studying patients who've never had anti-DNA or are we studying patients who have had DNA? It's different perhaps in a level that's rising versus de novo. Prednisone had to be less than 15 milligrams, no active infection, and stability of disease and medications for 2 months prior to study.

The study design was that patients were followed monthly for 12 to 18 months. They had history and physical, analytes, and SLEDAI. The randomization criteria was a rise of C3a greater than 50 percent, an absolute level greater than or equal to 500, rise of DNA 25 percent, as in the other study, and the absence of clinical activity. Meeting those criteria, the patient would be randomized to prednisone on the schedule that I've shown you, 30 milligrams for 2 weeks, 20 and 10, or placebo.

This is the flow chart, which again you really can't see very well, which is looking at patients followed in observational study for up to 18 months, and we had a 180 patients enrolled and I'll just point out the green side of the interest of time. 41 patients met

randomization criteria. There were 11 who wound up having a clinical flare, 30 had no clinical flare, 5 were mild/moderate, 6 were severe. This is the ethnic breakdown.

As it turned out when we analyzed the severe flare rates, the flares within 90 days, what you can see is prednisone and placebo. None of the individuals who received a priori prednisone had a severe flare, 21 no flare, and placebo 6 and 14, with a Fisher's exact of .009.

What was the nature of these flares? Timing and clinical features of the 6 flares, pre-C3a and DNA, placebo or randomized to the steroids, within 1 month, 3 renal, 1 CNS. And one of those renal was de novo; the other two were recurrences. Within 2 months, 1 pyodermic gangrenosum and pancytopenia, and 1 pleural effusion, hospitalization, and high fevers, non-infectious.

Well, the other question to ask is, that's fine, but does the biomarker respond in parallel with the clinical response? This is a summary of results of the outcome variables by treatment group, and what you can see is that the SLEDAI after 1 month appropriately decreased in the prednisone group as did the double-stranded DNA, as did the C4, and there was certainly a trend of to decrease in the C3a. So again, the clinical markers went in parallel with the response.

This basically shows you patients who received placebo and the C3a continues to rise, as does the anti-DNA, and we saw the different effect with the prednisone, that the marker also responds in parallel with the clinical effect.

So I leave you with anti-DNA antibodies and complement as candidate biomarkers for clinical trials in lupus.

Clearly, clinical laboratory correlation in lupus is a heterogeneous relationship, and these are the unanswered questions. Are these serologic parameters useful as predictors of flare and/or an assessment of flare in response to therapy? Which tests are best and are combinations superior? What is the optimal time interval in which to study a patient? Finally, what is the outcome being measured? In other words, defining a flare, what organ, and could renal be the most relevant?

We started out with a slide that probably was from the '80s that we were taught in medical school. But this is a table from Dubois current textbook and it's a chapter written by Schur and Glickstein, and this does project, which is sort of interesting, and this is basically very interesting in that it very simply tells us that when complement falls and anti-DNA rises, this should reflect active nephritis. But I leave with you a quotation

from a Roman dramatist Terrence. "One easily believes what one earnestly hopes for."

Thank you.

(Applause.)

DR. FIRESTEIN: Thank you. Please hold your applause.

The people on the committee should have received copies of the questions that were handed out during the talk, and then for those in the audience that have not received them, they are apparently available on the table outside.

So the next discussion will be from Dr. Matt Liang who will be transported here magically through the wonders of modern technology. Dr. Liang, are you there?

DR. LIANG: Yes, sir. I'm also here with Professor (inaudible) from Sweden, and I really want to discuss my chart.

DR. FIRESTEIN: Matt, you're breaking up.

DR. LIANG: I wanted to say that I hope (inaudidble) 3-plus years of work by many people, I'm sort of at a disadvantage because we also had technical problems. So what I'm going to do in the presentation is to read the title of the slides, as I think they are in sequence, and then I think everybody has a copy of the individual slides as well as the two source manuscripts

that have been recently submitted to Arthritis and Rheumatism, which really details the work.

What I'll do today -- and my colleagues will cringe -- is give you clinic in bad slides, something inscrutable, something unreadable, and many are too busy, but it's really just a map for those two documents and the thing that's in your hand, and if I don't project well or my voice gives out, please let me know because I'm not as full-throated as I usually am.

In any case, the first slide should be "ACR SLE Response Criteria Initiative," and this is just to remind me that four years ago roughly, the ACR saw three or four different groups, sometimes with overlapping membership, trying to develop response criteria. And the ACR, having a tradition in providing some guidance to both nomenclature, taxonomy, and case definitions, thought it could play a very important role by convening a consensus-building process toward three initiatives which are on the slide and you can read them.

One was to define a priori minimally clinically important differences in the metrics of overall disease activity, which are usually a combination signs and symptoms and laboratory manifestations of the existing SLE activity measured. We didn't want to play favorites and we wanted to make sure that everyone could play no matter what

they measured.

The second subgoal was to do that in combination with selected target organ systems, and then tomorrow, I will go into what we have tried to do to develop criteria for steroid-sparing agents that are tested.

So the next slide should be the support provided in kind or with dollars by multiple groups which are on this slide. This took more time to do actually than some science because we were constrained by ACR rules to only get funds from certain kinds of sources, but this project wouldn't have been possible without these donations from these groups and we're really grateful for that.

The next slide should be the committee. I'm sure you can't read this. It includes some people in the room today, and like a lot of big projects, it involves an international village. It was represented on this committee as well as by the invited consultants and also the people who volunteered their time to do a web survey, which I'll describe in a second. Those are the experts. This is in your handout. You can't read this.

Then the next slide should be "Methods," and this is just an overview of the first paper and I'll get into the details. But basically we tried to do an empirically-based exercise but using real patients. We got

300 patients in trials, observational cohorts, from three or four different countries, and we asked a rheumatologist who wasn't involved in the care of these patients to abstract the clinical data into standardized vignettes. All these patients had at one point disease activity measured based on one of the six available measures in real time. There were a couple instruments where we had no data on specific disease activity measures, and these were done post hoc by Jill Buyon in one instance and by David Isenberg's group in another instance, so that we could actually have data in the SELENA SLEDAI and also in the BILAG.

From these 300 vignettes, we created a very complex but I think rigorous sampling frame that recognized that these activities are not normally distributed even in observational cohort or trials, and we wanted to ensure that we covered the range of activity.

We then got a tremendous donation in kind from the University at Dusseldorf who maintained a web site, and we were able to use their information sciences to create a secure web site where we basically pulled SLE experts. These experts were gotten from the editorial boards of our distinguished publications, presenters at ACR meetings, attendees of the lupus international meetings.

Then we asked the experts certain questions and

I'll detail that in a second, and then we had another group which is my committee, meet and to examine the data but blinded to the instrument and because there aren't any statistical or any other standards for what level of agreement should determine a significant agreement, we asked the participants to vote on this.

Then basically we now had a data set where we had experts rating whether a patient had changed in an important way and we had independently -- and this information was blinded to the survey respondents -- the actual disease activity measures as assessed by the clinician in real time. Therefore, we were able to create the data to establish the relationship between the clinically meaningful, important difference and a change in any given disease activity measure that we looked at. This operationally was corresponding to 70 percent or higher agreement between the experts of either improvement or worsening.

So the next slide, I think, is an "Example of Baseline Vignette," and again it's unreadable at a distance. But basically it has a piece of the history -- and this is a real patient -- the laboratory information, and then we asked the respondent to rate or ask specific questions and also to indicate what kinds of changes they would make in various classes of medications that are used

in lupus. We used that as sort of the functional operational definition of whether we thought that someone was getting worse because we argued that a clinician sensing something was important and different and worse would elect to go beyond symptomatic treatment to more toxic and possibly more effective treatment.

So the next slide, I think, says the "Same patient, two month follow-up." The survey respondents answered what they could after the first vignette and could not get back to change those answers and then was presented information from the same patients two months later, and again the format is as I've described, history, laboratory, and then the same questions at this new time point.

So the next one is "Assigning Vignettes," and what we did was with the bank of vignettes was to take from the bank five standard vignettes as sort of our internal control that all the respondents got, so that we could see what the reliability of those assessments were, and then we also had the rest of the vignettes stratified by the disease activity so that we could sample and cover the range of disease activity.

So as you work your way from the egg, two eggs after that, the box in the middle, you get down to the fact that the respondents got 5 standard vignettes and then 30 vignettes over-sampled for higher activity because, as in

most data sets, most lupus patients lump toward the mild to moderate level of activity. And these were given in randomized order to eliminate order effects to the respondents.

So I think the next slide should be the results of what the experts said or responded to in terms of the 5 standard vignettes, and the vignette numbers are down the left-hand column. They just correspond to the number of subjects, and the M0, M2, M6, are the months after the initial month. I think the key thing here, which I think is not a new finding for people who are involved with lupus, is that if you take the same patient, i.e. vignette 54, giving the experts the same information, you have roughly 6 percent saying they're worse and 80 percent saying they're improvement, and you can see that there's variation across all the vignettes.

The panel in Germany looked at the vignettes themselves and you could make up all kinds of explanations in terms of bad wording of vignettes, et cetera, but I think that this basically underscores the fact that given five lupus experts, we get six opinions.

So the next slide is the kind of data that we got on each instrument. And I have to single out the singular creative contributions of Dr. Michal Abrahamowicz, who's a professor of statistics at McGill, who's had a lot

of experience developing performance curves for various kinds of measures, but we were able to develop these kinds of data and curves for each of the instruments over the M0, M2, M0 2 to 6 period. And if I could, I'd just like to walk you through this because this is a data-driven exercise.

So if you look at the 3 colored curves, the blue corresponds to the probability of the experts saying that the patient was better, and if you look at the dotted box that goes from point A-2 on the vertical axis and minus 4, you'll see that there's a relationship of the probability of the experts saying that they were better but it's not the same. It follows the trajectory over the differences in disease activity scores. So if you looked at that rectangle, the dotted rectangle, you can see that the cross section should add up to 1, but on any given instrument of a minus 4 decrease measure, you have these probabilities with those confidence intervals. What the committee was asked to do was to pick out which probability would be the one that they would use as a consensus probability and that turned out to be 70 percent.

Now, the next slide is basically the bottom line in a sense. Here's where we gave the clinically meaningful differences for specific instruments for both improvement and worsening, and you can see down the left-

hand column the instruments that we evaluated and what differences in the metrics of that instrument corresponded to the clinician's assessment of improvement and worsening. This is very important information to drive sample size calculations. It also, I think, in the data is information about sensitivity.

Then finally, I just want to conclude. We had a number of recommendations about the conduct of trials, which is summarized in the paper, but I think that the data and our exercise concluded that for X rheumatologists, there are always X plus 1 different judgments in the assessment of the disease activity.

We felt and this is also based on the experience of Professor Abrahamowicz that the performance curves on disease activity measures, albeit not perfect, have more than adequate psychometric properties to distinguish different categories of response.

And then it was the feeling of the committee and also, I think, part of the assumption of the exercise that activity measures are summary measures. Some of them are implicitly weighted, others are explicitly weighted, but we felt that in a clinical trial and certainly, I think, for most individual patients, a change of therapy is usually driven by some key target organ which we're trying to control, and we thought that a priori target organ

response criteria, which we are currently doing, should be used with these measures.

There are two other reasons we think that this is true. In our data set, 10 percent of the subjects had some organs getting better and some getting worse, so that an overall summary would not capture that necessarily and that individual organ responses should at least be presented and documented.

Then finally, when you look at these measures, what they have in terms of breadth they lack in depth, and so when you look within an organ scale on any of these measures, the scales are most likely insensitive to change. And we thought that that's the other reason that these should be considered ancillary metrics to target organ response criteria.

That's the end of my presentation.

DR. FIRESTEIN: Thank you very much, and we appreciate you taking the effort to make the presentation from your home office.

All right. The next talk will be from Dr. Vibeke Strand, who will discuss fatigue and function in lupus.

DR. STRAND: Thank you, Dr. Firestein and members of the committee. As a nonvoting member who wasn't introduced, I will show you why, but this is my effort, of

course, to disclose the fact that I do a lot of consulting, and I do also teach at Stanford and I am a rheumatologist.

I've been very interested in lupus for a long time, both as a treating physician and in my role as a consultant in trying to develop new trial designs and hopefully facilitate the approval, one of these days, of a new product in lupus. We've had a lot of false attempts or, shall we say, a lot of hard work that so far hasn't been successful.

I think we know why. We just discussed disease activity indices, and I think part of that is because they were not designed as outcome measurements. The majority of them really have been used to determine when therapy should be changed, and that is a good means of using them and they can function that way in a clinical trial.

Perhaps one of the more important issues is that patients often say what they think of how they're doing and it's not very concordant with what the physicians think of how the patient is doing and, of course, progressive renal insufficiency is a good example. Until one is fully symptomatic with renal insufficiency, it's very hard to explain to a patient why we worry about their BUN and creatinine.

We have not so far been very successful in using responder analyses and presumably that's because

we've actually proposed them in advance of actually getting the data in a clinical trial. We now have some trials from which we've learned.

And, of course, change in medical practice occurs all the time and it may well confound outcomes.

But I've been asked to talk about fatigue and function and health-related quality of life, and I just want to remind you that back in 1998, at the lupus module at the OMERACT meeting, Outcome Measures in Rheumatology Clinical Trials, we developed consensus on the required domains to be assessed in either clinical trials or longitudinal observational studies. Those domains were very important and the one, of course, that's highlighted is health-related quality of life.

What is health-related quality of life? I think most of the people in the audience know, but it's certainly not the economy, it's not the geographical situation or the politics or the recall election in California, and it's not the status or access to resources. In other words, it's really in all the ways that your disease affects you, how are you doing today, and it has a great deal to do with how it's asked.

Lupus does affect all domains of health-related quality of life, but specifically patients in comparison to other rheumatic diseases complain of fatigue, complain of

the inability to plan ahead, and complain of changes in their appearance.

The SF-36 is one of the instruments that's been mostly widely used to measure health-related quality of life. It is a generic measure. It has 8 domains and 2 summary scores which sum the domain scores. It's been used in a variety of diseases. Lupus has been one of the newest ones it's been applied to. It's been validated in rheumatoid arthritis, osteoarthritis, and a variety of cardiovascular diseases, as well as diabetes and other chronic illnesses. In many ways, it's a useful instrument for us in rheumatology because we can then show other organizations how the diseases that we treat impact our patients and that can be compared to what happens with chronic renal disease or coronary artery disease or diabetes.

Now, the four domains that are positively summed in the physical component score and negatively summed in that one include: physical function, role physical, bodily pain, and general health perceptions. Vitality, social function, role emotional and mental health, are positively scored in the mental component summary score and these are negatively weighted.

There's been some question about whether using these component scores is useful or whether it's better, in

fact, to look at the individual domain scores. But just to remind you of the nice editorial that Michael Ward wrote, while he was still at Stanford, basically the coping mechanisms are most consistently associated with health-related quality of life in lupus patients but not necessarily their morbidity. We do know that ethnicity and socioeconomic status are important and all of these are very variable in trying to assess it.

It's pretty clear that social support mechanisms are fairly consistently associated with how patients report the mental health aspects of their health-related quality of life, and organ damage, as in the example of renal disease, is less associated with how a patient reports their health-related quality of life than disease activity in terms of how they perceive that. And this has been derived from cohort studies, as well as randomized controlled trials.

Fatigue and depression are quite important. Disease activity and damage do not equal health-related quality of life. Interestingly enough, disability does not necessarily equal impairment in physical function in lupus patients. For instance, if we look at a varied series of lupus patients here, Baltimore, published by Hochberg, Cleveland and Canada, published by Milligan and Dafra Gladman, we see that patients basically have relatively low

HAQ disability index scores, much lower than we would expect in patients with longstanding rheumatoid arthritis. Many of them will have a disability index of 0 indicating they have actually no impairment in physical function, and very few of them actually report requiring help to perform the physical activities queried in the HAQ.

So basically, in comparison with RA, lupus patients complain of loss of energy, unpredictable course of disease not different from RA, but they complain of fatigue much more prominently. They have much more dissatisfaction with their perceived control of their bodies, and more importantly, they report a lot more dissatisfaction with understanding of their disease on the part of other individuals, including their physicians, and specifically their handicap is invisible to others.

So if we look at prospective study of 82 patients with RA, 82 with lupus and match gender and age controls, we see that the diseases impact all dimensions of health status, but there's actually less disability in RA and lower visual analog pain scores, although in fact both groups of patients, RA and lupus patients, complain of bodily pain in their domain scores which indicate a significant impact of their disease. In fact, the SF-36 correlated best with patient global assessment and accumulated damage.

Now, Dr. Gladman was the first to actually show the SF-36 was sensitive to change in a longitudinal observational series in lupus and that the baseline domain scores were very, very low in all 8 of those domains, and that a variety of series have now shown in cohort studies, as well as some limited clinical trial data, that basically decreases in disease activity do translate into improvement in physical function, bodily pain and general health perceptions. Worsening disease activity actually shows worsening in all the domain scores, especially physical function, and more damage eventually translates into poorer physical function and poorer general health perception.

So SF-36 has been demonstrated valid and sensitive to change. The decrements in the multiple domains do, in fact, correlate with increases in disease activity and damage, but these are generally weak correlations. They also correlate with use of immunosuppressives, and they reflect end-stage renal disease where, once patients go on dialysis, they show very significant improvement.

So one of the things about looking at the SF-36 specifically in lupus was a series of observational studies with Thumboo, et al., in Singapore who was able to show that the SF-36 was sensitive to change, but in fact the greatest variability in reporting was in role physical and

role emotional domains. And he did not agree with how the PCS and the MCS, the summary scores, were put together and, instead, took the 4 physical domains and meaned them into what he called the PHS and the 4 mental into the MHS, saying that they better reflected the individual domains.

Interestingly -- this should again be arrows but e-mail always changes the symbols -- the PHS was negatively correlated with increased steroid doses and worsening BILAG score and the MHS was negatively correlated with increased steroids, use of cytotoxics, and also worsening BILAG scores.

Now, we talk about minimum clinically important differences and Matt Liang just discussed the exercise that we did with the disease activity scores at Dusseldorf almost two years ago now. Originally, MCID was defined by patient query and Delphi technique, but for instance, with both the HAQ and the SF-36, it's now been looked at in comparison to global visual analog scores on the part of patients in randomized controlled trials so that it's a statistical definition as well. For the HAQ disability index, it's an improvement of minus .22, and for the SF-36, it's generally considered to be improvements in domain scores of about 5 to 10 points across many different disease states, including cardiovascular and pulmonary disease, and about 2.5 to 5 points for the component

scores.

Another point is the confounding issue of fatigue. We talk about fibromyalgia being an important part of what lupus patients complain of. In various series, there may be as few as 10 percent or as many as 30 percent, and it does, in fact, significantly impact the SF-36 because fatigue is measured in several of the questions and is associated in several domain scores.

Fatigue is also directly assessed in the SLAM and the SLAM-R, and the Krupp Fatigue Severity Scale, which was developed for use in MS patients, has also shown that the fatigue that's reported by patients with lupus is different and involves different domains in fatigue. Either way, the fatigue can be assessed and is included in the assessment of the SF-36, and whether fibromyalgia is impacting the patient with lupus or not, one understands, I think most of us clinically, that if their lupus is improved to at least some degree, their fibromyalgia is as well.

Now, anti-double-stranded DNA antibodies do predict disease flares and Jill Buyon gave a nice summary of the data in the clinically-quiescent but serologically-active patients. You'll see here -- and I'm sorry they don't go through one at a time, which is how it was supposed to work -- that there are several series where

prospective treatment of patients who have elevated double-stranded DNA antibodies actually improves either their ability not to have a flare or actually decreases the number of relapses, and most recently, this was also published by Bijl in terms of using mycophenolate mofetil based on an increase -- again this should be two arrows up -- in double-stranded DNA antibodies.

Now, this is the LJP394 study, a phase II/III study that's been published and shown previously, to simply show that using this particular agent, active treatment resulted in improvement in double-stranded DNA antibody levels and increases in complement 3 levels, and such a relationship was not seen in the placebo group.

This was analyzed in terms of looking at SF-36 data, and this was specifically looked at in a longitudinal analysis of the first 18 weeks of patient treatment because that was when they all received the same 100 milligram dose weekly. This was a group of patients, a 179 intent-to-treat and a 157 who were defined as having high affinity double-stranded DNA antibodies; in other words, antibodies that had high affinity binding to the LJP394 epitope. As well, we looked at patients in 37 who had had a flare to see whether there was a difference between their reported HRQOL before and after the flare.

This is in fact the baseline for all treated

patients in green versus the age and gender matched norms, showing you that, with the exception of the mental health index, HRQOL was significantly impaired in the patients with lupus. These patients were required to have elevated double-stranded DNA antibodies at enrollment and to have had a history of renal flare but were clinically stable at the time of enrollment.

These are the changes in the domain scores over the first 16 weeks of treatment. One can see here now that the active agent is in green and placebo is in blue, and there are more improvements relatively in the active group with diametrically-opposed change in role emotional.

If one looks now at the pre- and post-changes with renal flare, one can see that the patients receiving active treatment do not show or report the deterioration that's seen with the placebo. And this is true in all domain scores, and it's also true if those patients who are receiving high-dose corticosteroids or cyclophosphamide are removed from the analysis, as they would be expected to report more deterioration.

So the conclusions from that particular study are that patients with clinically stable lupus reported impaired health-related quality of life, and even during the induction time when they had not had a flare, one could see improvement with active treatment which was associated

with a decrease in double-stranded DNA antibodies. The differences pre- and post-flare appear to be related, at least in some part, to those reported changes associated with active treatment.

Now, I want to quickly show you one more thing which is longitudinal changes in two randomized controlled trials and this is now looking at changes in double-stranded DNA antibodies, regardless of treatment groups; so both active and placebo are combined. The definition here is actually a greater than or equal to 10 percent reduction in anti-dsDNA antibodies in more than two-thirds of all the determinations. This definition was derived based on the standard deviation of the assay and the fact that patients were required to have a baseline of 15 on the FARR assay. One can use another definition, such as 20 percent, and see similar findings.

So one could see the responders are defined here and they show a sustained reduction in double-stranded DNA antibodies, and as you can also see here, even using the definition of 10 percent, the majority of the active responders actually also will come to a definition of either 20 or 30 percent decrease and such a magnitude of change or increase is not seen in the non-responders.

These are the health-related quality of life scores in both groups, now looking at responders in blue

versus non-responders, and this is at month 4. So one can see -- this is approximately the week 16 time point -- that despite clinically stable disease, patients report improvement in all domains of health-related quality of life, if their double-stranded DNA antibodies have gone down.

If one looks at a second series at 6 months and again at 12 months -- and these are in your handout -- you can see very similar types of findings, with in fact some deterioration in those patients who are not defined as responders.

Now, these analyses excluded even the patients with the renal flares who might have been attributed as reporting the worsening and in fact showed very little change in the analyses.

Now, are these changes clinically meaningful? I mentioned to you before that we think MCID is an improvement of about 5 to 10 points in domains. Well, you have in your handout global assessments which actually show improvement over time in the responders, and what we can also see here is that in one of these two series, the 15-point scale by Guyatt, et al., was used, asking patients in the past 3 months, has there been any change in your overall quality of life related to your lupus. We looked at those patients who said they were a little bit better,

which was 6 on the scale of 15, or those patients who said they were a little worse, which was 10 on the scale of 15. What we found was that the improvements indicated mean change scores improvement of 6.7 to 11 in all of the domains and about 3.4 to 3.9 in the two component summary scores.

Worsening, which was interesting, might have been expected had patients actually determined worsening a little bit sooner than they determined improvement or no change. We can see that the domain decreases or increases in fact range from 1.7 to a worsening of about 15 points, and in physical component and mental component summary scores, the worsening was approximately 1 to 2 points. So this is quite consistent with the published data suggesting 5 to 10 points for domain scores.

So it is difficult to assess outcomes in lupus, and the data derived from randomized controlled trials are very limited and have yet to result in approved therapy. However, I think it is important to look at patient-reported health-related quality of life. It is different from what we assess in RA. It means that we need to be looking at a generic measure that looks at all domains, such as the SF-36, that physical function is only one of those components and a limited one.

Health-related quality of life is improved.

Patients do report improvement when their disease activity scores go down. They do respond and prove that they feel worse when they're getting high-dose glucocorticoids or immunosuppressives. This kind of data has correlated with longer-term outcome, and I think that this data has preliminarily shown you in two series of patients that there appears to be reported improvement in health-related quality of life in many of those domains with sustained reductions in double-stranded DNA antibodies which are clinically meaningful.

Now, this data will then, of course, need to be confirmed in other clinical trials but suggests again that there is a relationship between a biomarker, a marker of disease activity, and a patient-reported outcome.

Thank you.

DR. FIRESTEIN: Thank you very much, and the next discussion is from Dr. Tom Lehman on pediatric lupus.

DR. LEHMAN: I'm going to take a slightly different approach this morning because I think there's a number of difficulties that are inherently obvious in how we analyze SLE disease activity. I think it's clear that the pathogenesis of disease in multiple different organisms is not the same and that when we use a generalized marker like the SLEDAI or the SLAM or the BILAG, which demonstrate evidence of SLE activity overall, if we try to use a single

marker when we're measuring differences in brain disease, skin disease, lung disease, or kidney disease, has inherent problems that assume a common pathogenesis which probably doesn't exist.

In order to deal with that in pediatrics, I'm going to show you one study we've done where we've deliberately restricted ourselves to children with biopsy-proven diffuse glomerular nephritis. By restricting ourselves to a specific organ system and a specific pathogenesis, I think we have a much better chance of showing a specific role of different antibodies, et cetera.

I'm not going to prolong the rationale of immunosuppressive therapy. I think everybody here is aware of this.

What I'm going to show you is data from using "our standard cyclophosphamide" therapy of a gram per meter squared per dose, given routinely in a prospective manner, 7 doses at monthly intervals, followed by 10 doses at 3-month intervals for a total of 36 months of therapy.

If you then go to look at responsiveness and say what measures of response can we show had a clearly meaningful effect, you can see initial SED rate comes right down over the time of treatment and persists in remission as the patient's persistent remission. Serum creatinine. We started off with people with basically normal

creatinines and you can see over a 5-year period, despite the fact they had biopsy-proven diffuse glomerular nephritis, there is no increase in serum creatinine levels. Creatinine clearance improves.

One of the things we're going to have to watch for is illustrated by this data point. All through this period, the creatinine clearances are being done on children while they're in the hospital during 24-hours receiving IV cyclophosphamide therapy and during which they're receiving 2 liters per meter squared of hydration. This point is attempting to follow up these patients who no longer require hospitalization with out-patient creatinine clearances.

So the data here probably is nowhere near as reliable as the data here. Because we've changed the timing and circumstances of the collection, we've introduced a degree of unreliability that if we're going to have meaningful results needs to be excluded.

The same is true here. These are C3 levels. Again when therapy stops, they drop down a little bit but remain at normal range. 24-hour urine protein. Again, these are all easily measured, easily quantifiable outcome measures.

Prednisone dosage goes down over time very clearly. One of the things that will need to be considered

is, is there a minimum prednisone dose on which we want patients to remain, and therefore if we're looking at changes in prednisone dose over time, if there's a floor which we've created that will need to be remembered in doing all the calculations.

In renal disease, we can do activity and chronicity indexes on renal biopsies. Obviously, that's not going to be possible when we're discussing renal flares as a whole.

Perhaps most importantly as we're talking about patient subjective sense of well-being, socioeconomic status, sense of depression, psychosocial factors, concurrent fibromyalgia.

A long time ago, one of first studies of outcome factors in adults with lupus showed that one of the best predictors of disease activity, when you got past C3, C4, et cetera, was plain old simple hemoglobin. If you want to know whether or not your patient is doing better and you want to avoid psychosocial factors, you want to avoid socioeconomic factors, you might want everybody to be on a vitamin pill that contains iron to minimize dietary issues, but hemoglobin coming up and normalizing clearly is associated with improvement in disease overall status without being organ-specific.

Indeed, here in our children treated with SLE

with Cytoxan, you can see the hemoglobin is normalized promptly over the course of therapy and remained normal at 5 years, 2 years after the last dose. I have further data now. We're 10 years out with the same exact results. I'm just showing you -- and we're going to be presenting at the ACR meeting -- 10 years.

When we do this, we still have to represent the fact that there are going to be problems and there are going to be failures, but I think the most important thing for everybody here to remember is that when we talk about lupus, we're talking about a very heterogenous disease. We're talking about the fact that we know there are racial differences in the incidence of lupus. Are we including in our studies the fact that there seems to be racial differences in the severity of lupus, not to mention the confounding socioeconomic factors, et cetera?

All of these things are not being accurately represented in the current measurement and outcome statistics that we're doing. I don't think anyone here would like to say that the average oriental patient has the same general disease activity level as the same white patient who lives on the Upper East side of Manhattan. There's different genetics. There are different confounding social variables. There are different confounding psychological variables, which we're going to

have to take into account as we do these things and as we measure the outcome of our patients. Clearly Minneapolis, New York, London, Singapore are not the same in patient population, in patient background genetics, and until we standardize treatment, in treatment, or even in the way they're going to interpret reading the SLEDAI to do the scoring.

Our major needs at present. We need standardized criteria for the initiation of therapy. Those are going to be present in drug trials. What we really need for our children right now is an early intervention that can prevent both corticosteroid and disease-related complications.

Thank you.

DR. FIRESTEIN: Thank you very much and that is the end of the state of the art discussion, or at least the presentations, and I want to thank all of the speakers for doing a wonderful job of summarizing the data.

The next section is a series of short presentations from a number of the groups that have been intimately involved in supporting the research, as well as our patients with lupus. The first is Barbara Boyts representing the Alliance for Lupus Research.

MS. BOYTS: Good morning, Dr. Firestein and members of the Arthritis Advisory Committee, Dr. Simon,

members of the FDA, other individuals here from industry, from academia, members of the public, and my lupus colleagues, and the other lupus organizations. It is indeed a great pleasure to be here with you today. My name is Barbara Boyts. I am the President of the Alliance for Lupus Research.

Last March, at the biomarker assessment meeting, many of you here embarked on the challenging process that lay the groundwork for a document that would have major impact on the lupus research community and on the individuals with lupus. I want to thank the FDA and particularly Dr. Lee Simon and the other members of the Arthritis Advisory Committee for organizing this meeting and showing such a strong commitment to making the lupus guidance document a reality.

I would also like to congratulate the many researchers whose collaborative efforts to identify targets for treatments served as a catalyst for this critical phase. Your rapid advances in scientific discoveries have brought new opportunities in many areas of lupus research and fueled the urgency to move forward. Examples include progress in genetics, molecular biology, molecular immunology, and complement biology. Each have yielded important knowledge for potential targets for new treatments.

I see evidence of progress throughout the field and can cite examples within the Alliance for Lupus Research Target Identification and Lupus Grants Program. There's huge potential for new treatment development, yet industry and research supporters have been reluctant to move forward because of the uncertainty surrounding the drug approval process.

Your efforts today provide hope that we can finally move past the handful of drugs that have been used for over 50 years to help manage lupus. They provide hope that we can find drugs that will do more good than harm. By providing clear ground rules for drug development and drug approval in lupus, we will expedite the process by which insights and discoveries in science translate into effective treatments for lupus. This in turn will stimulate industry and the Alliance for Lupus Research and other lupus research philanthropies to achieve their goals of new treatments to improve the lives of the more than 1 million people who live with this really terrible disease.

Working together with these new guidelines, I am confident that we can discover better treatments and even one day a cure for lupus. Thank you very much.

DR. FIRESTEIN: Thank you very much. The next speaker is Margaret Dowd, representing Lupus Research Institute.

MS. DOWD: Good morning. I'm Peggy Dowd. I represent the Lupus Research Institute, and I'm very grateful to be here this morning with you. This is indeed, I think, a pivotal point in the lupus world, and it is significant that we are here.

I bring you greetings from the LRI and our affiliates all over the country who are members of the board and members of the organization and that is the organizations of families and patients who comprise the Lupus Research Institute. They're the people who serve on our board. They're the policymakers. They're the decisionmakers. They're the funders of this organization that is seeking to bring new science to lupus. They are passionately devoted to the cause that addresses us today, and I am proud to bring you their greetings, their commitment, their thanks and their hope.

I think everyone has talked about significant meetings that we have sponsored and held over the past few years, but I think two years ago we co-sponsored a meeting at the NIH that had a session that many of you were at and I've noted this before in previous remarks on Friday night of that weekend in January 2002 that surfaced all the frustrations and almost the despair of clinicians, scientists and patients in that room. And it was a low point and it was a high point. I think it was a turning

point, and I think that one of the reasons that we are here today is that people left that meeting charged up and said we're going to do something about this.

One of the people who was there and who has done something about it and I'd like to personally thank today is Dr. Lee Simon. He was new to the FDA at that time and he told us that he was here and he was going to try to get things moving, and I think through the sheer power and force of his commitment and determination to move forward -- and this is not to slight anyone in the FDA who's done tremendous work on this disease for so many years, but I do personally believe that Dr. Simon has made an enormous contribution.

At the risk of political suicide perhaps, there's one other person that I would like to thank in my experience over the last 10 years with the SLE Foundation and the LRI and that is Dr. Matt Liang, who is not here with us today, but as I look back -- and as Matt just said, five experts, six opinions. As I look back at the studies that Matt has come to us with and we have helped and we have funded and we have worked with him over the past years, studies in nomenclature, studies in response criteria and finally bringing together and working to achieve consensus and agreement at difficult places, like Dusseldorf, I just think that that is about team. It's not

about turf. It's about agreement and it's about trying to get consensus. I applaud his work and I really, really hope that his leadership will take us forward to go the path that we need to do to get agreement quickly as we're going on.

The LRI continues to endorse clinical trials in lupus with the greatest of enthusiasm, and we are delighted that the issues of drug development are finally getting the attention they deserve. We have made a major commitment to advance clinical trial methodologies in lupus and we have an RFA on the market right now on biomarkers which we are working very hard to publicize among you all. We are seeking new projects to develop and to validate early markers, so you have our commitment on proceeding in that regard.

First and foremost and before all else, we petition the FDA to have and to maintain a deep and serious concern for the safety of lupus patients as we proceed with drug development. We ask you, above all, and members of the committee to, of course, protect our patients.

But we also ask you that once that safety is determined to please work to let these projects go forward. We don't need perfection. We don't need total and complete consensus on which markers or which measures or the details, and this should not be a stumbling block to going

ahead with the trials we need so badly for our patients. Let's agree on and let's get a document that's usable and doable and that works, that can be the basis for moving ahead as long as safety is not an issue.

We ask you on behalf of our patients to take some risks. At the LRI, we take risks. We fund people whose work wouldn't be funded immediately at the NIH because there isn't enough data for it. It's just a good hypothesis. It probably would be called kind of a lame business plan, I guess, but boy, is it turning out to be very productive. The people that we take risks on on good ideas are going on to the NIH for funding and their work is making a significant difference in this disease.

So we ask you to do what we preach about the LRI, to think outside the box and to take some risks with the details and not get bogged down in a process that doesn't give us the product that we need.

In closing, I would just like to cite two of the people who are very important to the LRI. They co-chair our board of directors. Robert Ravitz and Jack Lavery. They are parents of daughters with lupus, and some of you in this room know them intimately because you've treated their kids, Annie Ravitz and Dena Lavery. They're two young women that I would like you to keep in mind as you proceed with your deliberations. Now in their 30s,

they have suffered the ravages of this disease since childhood. They have lost eyes and fingers and toes. They have lost the ability to conceive and bear a child. They have lost the experience of quality of life or nothing. They have suffered heart attacks. They have suffered open heart surgery and now Annie, looking for a kidney and on dialysis.

Their fathers co-chair the Lupus Research Institute board and for 20 years, their parents have been contributing millions personally and raising millions more to get some relief for their children and we don't have it yet.

So I ask you in the process of deliberation, as you go forward, to consider these young women who suffer with this disease. We can't let problems delay the process, and I implore you to let trials go ahead as soon as possible.

Thank you.

DR. FIRESTEIN: Thank you very much. Representing Rheuminations, Incorporated, is Katherine Snider.

MS. SNIDER: Good morning. My name is Kit Snider, and I'm the President of Rheuminations. I want to thank the FDA for providing the opportunity for Rheuminations to speak this morning.

Lupus is a disease that my family has lived with for a long time. My mother was diagnosed with lupus in 1973 and my own diagnosis followed 7 years later. In 2001, our wish to respond to our personal experiences with lupus gave birth to a private charitable foundation, Rheuminations. Our goals have been to fund excellence in scientific research leading to better understanding of and treatments for lupus and to offer education, empowerment, and support to patients in fresh and innovative ways.

Our first project was to establish the Mary Kirkland Center for Lupus Research at the Hospital for Special Surgery in New York. Our second project has been to design a comprehensive and ongoing web site that will adapt itself to the changing needs of people with lupus over time. Our most recent project has been to create a separate public charitable organization known as the Lupus Clinical Trials Consortium.

One of LCTC's current goals is to give grants to over 25 academic institutions to support their infrastructure for clinical research activities focused on bringing new, safer, and better treatments for lupus to market. Most of the current treatments for lupus are off-label, borrowed from other diseases, very powerful and pockmarked with side effects. Some of these drugs may cause infertility, cancer, bone and joint damage, and

infections that can lead to death. Many of these were approved to save lives and not to treat chronic illness.

I would like to quote a few patients who discussed some of the worst side effects of these treatments.

Debbie, now 43, was 21 when she developed lupus. She said, "The worst part was the physical change in my appearance. I blew up. My face changed and people I have known all of my life walked right by me and did not even recognize me. All my joints were hurting pretty badly. My hips collapsed and I could not walk."

Tiombe, older sister of Kai, who was first diagnosed with lupus at the age of 13, describes the way treatment affected her sister. "Along with the medication, the doctor said she would become very moody, gain weight, and her hair might never grow back. It was so painful to look at my sister and not see her as the same person."

Ellen, who has lived with lupus for many years, describes her flares as a series of "little deaths," referring to losses of health, independence, self-esteem, and quality of life. The reality is that those little deaths are due not only to lupus but also the treatments currently being prescribed for lupus.

Advocacy groups have worked hard to bring lupus into the public eye. Foundations have been diligent in

their support of research. Lupus researchers remain passionate, dedicated and tireless in their pursuit of new discoveries leading to new therapies. We must now move forward to identify biomarkers and innovative drugs that can pass through clinical trials and on to market, but it will take the commitment of all areas of the lupus community, including government, academic centers, advocacy groups, the public, and, of course, industry to support this effort and defeat this devastating disease.

Thank you.

DR. FIRESTEIN: Thank you. The final presentation, representing the Lupus Foundation of America, is Sandra Raymond.

MS. RAYMOND: Good morning and thank you. I'm very pleased as the President and CEO of the Lupus Foundation of America to share the podium with our sister lupus organizations.

The Lupus Foundation of America is dedicated to improving the diagnosis and treatment of lupus, educating health professionals about lupus and supporting individuals and their families while educating the public and hopefully finding a cure. We vigorously pursue this mission through programs of research, public and professional and patient education and advocacy.

I'm here today representing almost a million

individuals, women, men, children of all races and ethnicities, who implore you to issue a guidance document for industry that will offer the absolute stimulus necessary for major pharmaceutical and biotechnology companies to invest the hundreds of millions of dollars it will take to bring a lupus drug to market.

You will hear all day today and tomorrow that there has not been a new lupus drug in the last 30 to 40 years, and you will hear the reasons why this is so. There is no question but that the disease is complex and that there are many gaps in the science of this autoimmune disease, but there are other factors that are equally true. Quoting Dr. Dan Wallace, "In the year 1948, half of those with lupus died within 2 years. By the year 1960, 60 percent of people with lupus were living 10 years, and by the '90s, 90 percent were living 10 years or more."

This improvement in mortality from 60 to 90 percent took place during a time when no new lupus drugs were introduced to the market. It was the skill of clinicians in learning how to use a variety of existing drugs and dialysis and interpreting markers that made the difference, but gains in survivorship, however, have come at a very high price since the morbidity associated with existing treatments may be worse than the original lupus symptoms.

While the document lays out the gaps in science and in doing so puts forward a robust research agenda, its purpose is to provide a road map for industry, to encourage investment in lupus research. We believe that clinical experience in lupus should not be ignored.

Recently, in preparation for this meeting, we conducted what I would call a convenience survey by e-mail of 341 clinicians who provide treatment to people with lupus. These individuals were randomly selected from a list of 1,000 clinicians who we know treat many lupus patients because their names appear on the physician referral list maintained by our 50 chapters nationwide.

While we recognize that the results cannot be projected to represent the practice of all clinicians who treat people with lupus, with only one exception, every one of the 132 clinicians who responded answered yes when asked if they used complement levels and antibodies to double-stranded DNA to evaluate disease activity in lupus patients.

The almost unanimous agreement by those who responded indicates to us that these markers are used widely in this country by clinicians and represent a so-called standard of care in the management of lupus patients.

If we do not find a way to broaden out this

document, children and women of childbearing age diagnosed today may well experience the same future as those who have lived with lupus for the past three to four decades. I respectfully ask you to recognize the dire circumstances in which these patients find themselves as they continue to take toxic drugs to control their lupus and suffer the side effects that can be worse than lupus itself.

We have very brilliant people in the field of lupus here today, and I ask them to find a way to open up this document beyond the subpart H or E to allow for full development of safe and effective therapies for very, very brave people who have waited much too long for your help.

Thank you.

DR. FIRESTEIN: Thank you, and that brings us to the end of the first section today, and we're going to take a break, a 15-minute break. So according to my watch, it's 10:05. So we'll start at 10:20.

(Recess.)

DR. FIRESTEIN: Why don't we go ahead and get started then with some of the questions that have been asked by the agency regarding the state of the art section? We didn't have much time to contemplate this in advance because we just received the questions this morning, and I have actually a slightly altered form from the time that the questions were passed out an hour and a half ago.

So I'm going to read the first question. Disease activity indices may be useful in assessing overall disease activity in lupus. Please discuss the utility and potential limitations of disease activity indices. Please discuss the acceptability of a single DAI applicable as a stand-alone primary measure of disease state in response to therapy versus the use of several DAIs. Please discuss the use of DAIs in the context of treatment of specific organs as an outcome. For example, nephritis improves at 1 year, but SLEDAI must also improve or cannot worsen.

So I'm going to open this up now to the panel and hopefully get a lively discussion. Certainly significant aspects of this were discussed in a number of the talks that we've heard today, but with regard to the utility and potential limitations of disease activity indices, does anybody want to begin with a comment? Certainly, again, there are multiple indices that have been discussed today, the SLEDAI, SLAM, BILAG, et cetera.

DR. ILOWITE: As a pediatrician, I just want to mention that although the SLEDAI and the SLICC have been validated in children, there are limitations with regards to the sensitivity of the instruments. In children, for instance, they don't assess growth, school performance, and sexual development, things like that, that would be important to include in a pediatric trial.

DR. FIRESTEIN: Thank you. That's very true.

What about some of the limitations of the SLEDAI, for instance, where only changes in activity are necessarily monitored as opposed to some of the other indices? Jill. I'm sorry. Bevra first and then Jill.

DR. HAHN: I want to comment on that. I saw that paragraph in the draft document we have, and I don't think it's quite correct. I think that refers to maybe older versions of SLEDAI, but the SELENA SLEDAI, you get points if you still have activity in arthritis or you still have oral ulcers or whatever, you still have malar rash. You get points for that. So I think that might be a misconception about SLEDAI. I'd like to hear what other people think, that it doesn't measure ongoing disease activity, only new things.

DR. BUYON: I was going to echo that sentiment exactly and point out that several of the parameters have been changed so that it reflects ongoing activity, not just new.

The other is to recognize that SLEDAI actually misses some organ systems. So, for example, you could have hemolytic anemia, which I think we'd all agree would be very serious, and that would not even be captured in the SLEDAI.

So one of the problems about using SLEDAI as

disease activity or even a flare index is recognizing that it's not all-encompassing and that in using it, you'd have to mandate equally that there be guidelines for using it because interpretation of SLEDAI, given the descriptors being rather perfunctory really in what it states, you have to have not only that document but then there would have to be a compendium or what we would call a glossary of terms.

So I would submit that, number one, it's not all-inclusive and that number two, it needs definite education for uniformity, and what it doesn't really encompass at all is the intention to treat which obviously the BILAG incorporates in a different way.

DR. FIRESTEIN: Joan.

DR. MERRILL: I really want to second what Jill is saying, but I do want to point out that the SLEDAI that is being used in clinical trials today is mostly the SELENA SLEDAI which does address some of the problems that Jill brought up. There are ways for the SLEDAI, with the flare index and with the global assessment put into it, to reflect things that may not be on the list of categories.

Having said that, I think it is really important to recognize that the SLAM, the SLEDAI, the BILAG are three very different instruments that are useful for different purposes. The SLEDAI is the instrument that is probably least susceptible to the placebo effect from the

point of view that it is measuring mostly objective criteria.

The SLAM has a weighting that has to do with whether you're getting better or getting worse, but if you had CNS disease and it was at its top worst point, you'd be getting the same score as if you had fatigue and it was at its top worst point. So that weighting doesn't really factor in that some organ system disease is much worse than other organ system disease.

The SLEDAI does the opposite of that. The SLEDAI weights by organ, so that if you have very severe thrombocytopenia and you have a platelet count of 5, you get 1 point. If you have fairly mild-to-moderate arthritis and have two or more swollen joints, you get 4 points. So sometimes these instruments just don't reflect what's really going on with the patient and aren't optimal to compare drug versus placebo. The instrument that solves this problem is the BILAG because the BILAG enables you to look at both of those qualities at once.

DR. FIRESTEIN: Yes.

DR. DOOLEY: I think the other difficulty with the SLEDAI is it's got a threshold effect, so that if you have 2 or more swollen joints, you get the same points as if you had 20 joints. Moreover, if you go from 20 joints to 3, your score doesn't change. So it has a disadvantage

of a threshold effect and then also reflecting a fairly short period of time of 10 days prior.

DR. FIRESTEIN: Graciela, did you have a comment?

DR. ALARCON: Yes. The only comment is that regardless of the instrument, not only do you need the glossary, you really need training. Unless the training is accomplished, then you really are going to guess how to score an instrument, and I think that's quite important when we are talking about multi-center clinical trials.

DR. FIRESTEIN: Dr. Hahn, did you have a question before?

DR. HAHN: No.

DR. FIRESTEIN: Okay.

Dr. Simon, yes.

DR. SIMON: If I just may ask a question for a little bit more clarification there? We grapple at the agency with the idea of a memory score. The idea that you're asking a question about a patient to remember how they were beforehand, and in the VAS scale for pain, in other circumstances, other kinds of interventions from an outcome point of view, we grapple with this all the time.

Could you comment on the utility of an instrument that looks at a 10-day window to week window, whatever the window is, and how accurate it might be, one;

two, and how one would validate that for the today versus the 2 weeks before; and how important is it to validate it for the today versus 2 weeks before in a disease such as this?

DR. FIRESTEIN: Yes, Joan.

DR. MERRILL: I'd like to make a comment about that. I think one of the most important things to do when there is a long window is not to depend too much on subjective parameters because those are almost impossible to talk about a month ago, or even 10 days ago, and people's emotional baggage does get involved in these things.

However, having said that, I've observed in doing clinical trials for many years that when you ask people to fill out one of these analog scores, if you don't let them look at their last score, the data jumps all over the place and has nothing to do with your assessment of how their disease is doing. If you let people look at how they were last time and then you say now move that line this time, I think you get beautiful data.

Now, I haven't proven it. I haven't published it. I think I'd love to hear the comments of some of the other people, Mary Anne and Bevra and Jill, but I think that you can depend on people to know if they're better or worse than they were before. Having them just simply

subjectively tell you how they are is a little bit more difficult.

DR. BUYON: I actually agree, and from the physician's perspective, just to echo that, one of the things we did in the SELENA trial is that everything had to be documented and you were encouraged always to go back and look at your note from the month before or three months before, and in our educational sessions, everything that we scored on an instrument had to be in the source document. So you were describing the joints, you were describing the skin, and then with that document in hand, you could "remember" and make an assessment the next time and move forward.

DR. DOOLEY: I think that also reflects in fact how we care for patients with lupus, and I think patients are very good at letting you know if they feel that they're better or not, and moreover, you can tell them your labs look great and they can tell you quite explicitly that they don't feel as well, and typically they are good predictors of what their clinical status will be.

So I think that in fact we do make treatment decisions and we alter therapy based on what our patients tell us their status is at the visit compared to their prior visit which may be as much as three or four months ago, depending on their activity.

DR. FIRESTEIN: Yes.

DR. MANZI: I was just going to comment to Leigh that I think unlike perhaps other diseases, just by the nature of lupus, we actually have to be able to do that because of the up and down course of the disease as opposed to a progressive course and that actually, I think, brings up an issue. Is it valid to take a pre- and post-snapshot and think you have captured what's gone on for the course of the trial? So, for example, pre- and post-SLEDAI. Does that really tell you how the patient's done over the course of the trial? I would venture to say that it may not, and so I think there's imperfections in that, but I think we really have to do that to reflect this particular disease.

DR. FIRESTEIN: Bevra.

DR. HAHN: It's an interesting idea that has come up here. Personally, I think that any of the scales are okay and they can't stand alone as the only measure of outcome in a trial, and if you want to have less argument about people who will be reviewing results as the trial goes on, you might be smart to use two of them so that those that favor one over the other, at least they'll have something they like.

But what I'd like to talk about is the question that's arisen here. I'd know from the experts that do this kind of study is there a precedent for -- let me say I'm

not so sure I agree that most patients can tell you if they're better or worse. I come out of an examining room many times having no idea what somebody thinks about that, what a patient thinks.

So is there any precedent for doing the global assessment scales or quality of life measures or something like that with looking at what prior scores have been, the person is looking at what prior scores have been, so either the patient or the physician? Is there any precedent for doing the science that way?

DR. MERRILL: SELENA.

DR. LIANG: I have a comment, Mr. Chairman.

DR. FIRESTEIN: The chair recognizes Dr. Liang.

DR. LIANG: I just want to clarify one thing and that is, all the measures are meant to be done by experienced people who in their clinical wisdom will filter out these kinds of issues because obviously in an individual patient, their anchor point, their cognitive function, all that stuff, whether it's being done after a steroid dose or whatever, these all play into it.

I think clinical judgment is meant to be interpolated in the completion of these instruments, and I also want to say that I don't think the issue is coverage of subjective versus objective. The patient owns their feelings and we use them, as Mary Anne Dooley pointed out,

in our assessment. Of course, we try to incorporate in that our assessment of their previous state, the worst lupus patients we've ever seen, whether they are people with low symptom-reporting thresholds, et cetera, et cetera. It's complicated, but I think that we still resolve this after every office visit. We come down to some assessment.

And I think that it's more important that we try not to reduce this ad absurdum and recognize that life is much more complicated than we can ever measure, but the key thing is that we capture it, that we do it in an accurate reliable way, and that we also report it. I mean, we don't think of a baseball player just by their batting average. We like to hear about other contributions they make, and I think that's what I believe is important, is that we recognize that you can't describe an individual by gender alone. You need to get ideas of their vital signs, et cetera, and I think all of these things are actually a description and they should be reported so that the results are transparent.

And Bevra's question is, I think, it depends. If you were to show a patient their subjective rating from a baseline, for instance, Guyatt has done that with his (inaudible) scale. It actually improves the sensitivity of the measure, but I think it depends on the state that

you're trying to measure, and I think that's actually a testable question for some of the subjective symptoms that lupus patients have.

DR. FIRESTEIN: Dr. Cush?

DR. CUSH: I would offer a contrasting view that I don't believe that global visual analog assessments should be relative to that which went before. I think that in doing different trials in different areas, you do the assessments based on what an ideal outcome is, maybe no disease, and what the worst outcome is and that's the span of disease one is looking at. And whether the VAS that you use or the patient uses are maybe not descriptive enough, that might go into it, but you have these in line. I think they're also somewhat dependent upon the tools that you either use as a clinician or the patient is using to make these assessments.

In RA trials, we know that patient assessments are very, very valid and are done without prior information of what they were doing. They know how they're doing since last week and 2 months and when they entered the trial. They don't have to look at their scores to actually come up with an independent assessment for today, and I think the same should be true here, that this should be a slice in time of what's going on today and how the patient is doing at this point in time, based on all the things that affect

them from their disease. To introduce past assessments into that equation, I think, muddles it up and doesn't really clarify the issue at all.

Maybe the problem then is the assessments themselves are not fine enough that they can distinguish true changes in disease activity.

DR. FIRESTEIN: To come back to the question that's asked, there are a number of instruments that have been suggested as being useful in these clinical trials. Do any of them rise above the others as individual endpoints for potentially drug approval, or is some sort of composite of composites going to end up being the gold standard?

Michael.

DR. WEISMAN: Gary, that's a really good question, and I was thinking about that as I was listening to the discussion because in my mind, it's not clear whether there really is a significant difference between the instruments or, really, are the issues mostly how the instruments are used; that is, the standardization, the training, the glossary, and all the very important scientific aspects of doing any instrument in the disease. And that's true for rheumatoid arthritis or anything else as well.

So from what I hear, there are some differences

between the instruments or among the instruments. Some are a little bit more subjective than others, but really those differences aren't as great as the differences in how they're implemented; that is, if they're implemented with the proper training, with the proper glossary, with the proper standardization and all the other scientific methods. So I come down more on that side than answering your question if there's one better instrument than the other.

I wonder, also, the BILAG instrument has always been cited in this group as well as being the best. I don't know what exactly that means, but it's been cited several times in the previous discussion, and yet it's not widely used at all and what the difficulties are with that may, in fact, be that it's hard to standardize, it's hard to train, it's very difficult to use. So that almost answers my question.

So in summary, then I think it's more the training and the scientific methodology that's important than the choice among instruments.

DR. FIRESTEIN: Yes, Joan.

DR. MERRILL: I have to agree that all of the instruments are fine, if properly-trained people are using them, and I'm sure that Matt knows these data better than I do, but they've all been shown to be sensitive to change

and reliable with different observers who know how to use them.

I would like to say, however, that I think the BILAG is by far and above the most flexible instrument. It can be used for so many different kinds of assessments. It is not true that it's hard to use. It's quite easy to use. What's hard about it is to do the statistical analyses, but in fact, in a clinical trial, that's not up to the individual investigators, and so a person developing the drug can work with the BILAG people and there's now computerized support for it. So it's actually quite easy to use.

DR. FIRESTEIN: Are there any other comments specifically on that question with regard to which instrument? Well, first, Jill, did you want to say something?

DR. BUYON: I don't know if we're going to readdress the patient assessment versus the physician because I feel a little dissenting about that, and the other actually slips my mind at the moment. But I don't necessarily want to leave that issue because I think that attribution is extremely important and we didn't mention that yet. But unlike seeing other patients, lupus patients require a lot of time and sometimes how they feel is reflected by other things going on in their lives that are

very exaggerated by this disease, and I do think we need objective anchors and I certainly don't want to leave that point.

I also do recall, I disagree a bit with Dr. Cush, at least with regard to lupus. If I couldn't go back and the patient couldn't go back and look at how they were even a month ago, I think those scores would be useless.

DR. FIRESTEIN: Dr. Simon, did you want to say something?

DR. SIMON: I just want to remind you all that what we're asking about is not clinical care, and we really need you to focus on -- not that clinical care is unimportant. We really need you to focus on what kind of instruments will be useful in a clinical trial setting for regulatory approval? What do these instruments tell us?

So my question to you, Jill, in particular, was that you inferred that to be used. The question I have to ask you is to be used for what? Not to follow the patient over time from the point of view of a clinical practice is very important.

DR. MERRILL: For determining the difference between an effective treatment and placebo, and the reason for that is because you can break it down by organs and your final assessment actually solves the problem that is a problem in RA trials as well, which is the all-or-nothing

problem.

Now, the SLAM also solves this problem and can be used effectively in this situation as long as there's some way to differentiate between the really tough organs and the really not-so-important organs, which I'm sure could be incorporated into the SLAM.

But the point is that a lupus patient can be a whole lot better, but unless everything is gone, the SLEDAI is not going to react, and so you're going to have a narrowing of gap between drug and placebo because of that. So the BILAG actually solves all the problems. You can look at this organ versus that organ which, as Matt pointed out, is a very important thing to do.

Now, you can do this with any of the instruments. It's just that there's actually a composite score that you can get that's already been sort of built in to the BILAG which factors in the importance of the organ and whether or not the patient is somewhat improved or totally improved, and you get points for all of that.

DR. DOOLEY: I think the other unique aspect of the BILAG is that one of the things that you want to be sure is not happening during a trial is not only is the patient's active symptoms getting better, but that you aren't acquiring de novo or new manifestations of disease that might be an inadvertent effect of the medication.

DR. DAVIS: I would like to disagree with some of the comments that have been said, too. I don't think all the instruments are equal or can give us just as good of information. I think the ones that have more subjective scales in them and use longer periods of time are more susceptible to the placebo effect, and I think that's maybe one of the reasons why in past clinical trials, we haven't seen a difference.

I also think that patients with lupus have a lot of cognitive difficulties which would be another reason that I want to use more objective outcome measures and for shorter periods of time to help them recall those things.

I really think that we don't have good weighting scores, both on the SLEDAI and on the SLAM.

DR. FIRESTEIN: We still have a couple of other questions to get to in this section.

DR. ALARCON: I think another thing to consider when you're talking about overall assessment and asking the patient -- I think you have no way not to do that -- is that the visual analog scales have a floor effect and a ceiling effect. So if the patient has scored herself to be really at the high end of the score and today is worse, there's no way really to get worse with visual analog scale. So I think that's something that should be considered.

I think that you have no way to exclude the patient. You have to ask the patient how the patient is, and I disagree with the fact that you really can only be reliable if you use a measurement that goes for a very short time because in a disease that is so variable as lupus, the value of asking what happened over the last month as opposed to today is that your patient happened to be very sick, being in the trial or not being in the trial, in the first 2 weeks of the month and then comes back to you in the last 2 weeks and she's fully recovered or over the flare, you're not going to capture that.

DR. DAVIS: And I'm not going to put her in a trial.

DR. ALARCON: She is on the trial already.

DR. FIRESTEIN: In terms of the specific question that was asked, Lee, if I can just try to summarize, none of these instruments are perfect and there wasn't hue and cry for a composite of composites that I was able to discern. It sounded like, among the many that are available, the BILAG seemed to have some advantages compared with the others, but again it also had some issues associated with it.

Can we just comment briefly on organ-specific outcome measures? That was one of the things that was asked by the agency. Joan.

DR. MERRILL: Yes. I think that's a very important aspect of what we need to be able to do. One of the things we need to do is not just prove that something works globally for lupus, which may be impossible, but if we can prove that it works for some aspect of lupus, it will become available and then time will tell.

So I think this is very important work, and I believe Matt is doing a lot of work trying to begin to sort some of those things out to support the fundamental research into how that ought to be done, is that correct? Is he here?

DR. LIANG: We're bringing it to the village.

DR. FIRESTEIN: I guess the question that arose was if you have something that prevents renal flares, for instance, but exacerbates CNS disease, how does one evaluate that?

DR. MERRILL: That's something Matt has been working on, isn't it, Matt?

DR. LIANG: I and others.

DR. MERRILL: Yes. You want to comment?

DR. LIANG: What we're trying to do -- and this was also supported by the ACR -- was to pick some major organ systems where we would likely need new agents and, again, because it's really difficult to amass any significant numbers or to, obviously, get information from

an ongoing clinical trial to, again, do the exercise of reviewing measures for specific organ systems, borrowing, stealing other people's work, but having a committee come up with what they sensed was a clinically meaningful difference and a recommendation for an appropriate scale.

The first of these is nearing completion and that is the renal criteria. And then after that, we had done background work and had actually some position summary papers on other organ systems, and we were hoping that in the absence of data, it's more important to be consistent than to be right and that that would ensure a level playing field and perhaps avoid what is commonly done in trials and that is to do post hoc data dredging for statistically significant differences.

This is our sort of effort, and obviously we would want to test these out, but I think realistically, it would be very difficult, for instance, to amass enough patients with specific neuropsychiatric phenotypes to actually test these out. So I think we're left with trying to do a sensible but not perfect nor completely evidence-based job, but I think we need to do it, otherwise we will be having these meetings endlessly about how difficult it is.

DR. FIRESTEIN: Let's move on now to the quality of life questions, and these discussions are, of

course, restricted to the panel members only.

One of the questions is whether or not this should or could be a primary outcome measure in evaluation of lupus. Jack.

DR. CUSH: Well, my comments actually also relate to organ. I don't know that you can have an organ-specific indication or a quality of life indication without actually having a disease-improving indication. So meeting a criteria for a SLEDAI or SLAM or BILAG, along with an organ-specific like renal or musculoskeletal or heme or a quality of life, that makes more sense to me.

I don't know that you'd want a therapy to be approved for something that in trials might, for instance, improve quality of life but not actually improve SLEDAI, SLAM, or more global measures. I don't know if you've really gained anything in the treatment of lupus.

DR. FIRESTEIN: Yes.

DR. LOONEY: Just a query, I suppose. One aspect of lupus you could consider would be antibodies against phospholipids, and Coumadin would be a pretty good drug to test as an effective treatment for that but would have no beneficial effect on most of these parameters at all.

DR. FIRESTEIN: Mike.

DR. WEISMAN: Go ahead, Dan.

DR. WALLACE: I think Vibeke presented some data that was published on lupus this month on the quality of life indices with the LJP394 where she showed dramatic improvements in quality of life with just lowering anti-DNA and nothing else.

On the other hand, I think in the Gene Lab's DHEA studies, you have improvements in quality of life without improvements in those parameters.

So quality of life can be improved with or without other instruments necessarily improving, and I think it's a very, very important component, and I think every study that's been done with quality of life has really validated the current indices' use in lupus.

DR. FIRESTEIN: Michael.

DR. WEISMAN: I think Jack raised the important question and that is, that if you can improve quality of life and make no change in any of the other parameters or instruments, is that sufficient for drug approval in this disease.

I'd like to hear Lee's comment on that because it's very clear in the draft guidance documents submitted to us that the agency will not tolerate, if you will, worsening in the disease with improvement in something else and that's throughout the document. But what if everything is stable and there is improvement in quality of life or

even improvement in lupus nephritis, if a drug is very specific for lupus nephritis, and everything else is the same? Is that sufficient in the agency's view for approval of a drug in this disease? That's a question, not a comment.

DR. FIRESTEIN: Well, Michael, that's the question that's being asked of the committee.

(Laughter.)

DR. FIRESTEIN: So what do you think?

DR. WEISMAN: I think it is and I think that we've reached the point now where we understand the value of each one of these individual organ-specific measures. We know that. We've had 20 some years of experience in understanding what the predictors are for mortality. That's again in the draft guidance document. The agency understands that as well, and I think the experience with quality of life and its meaning for lupus patients is also well understood. So my opinion is, yes, I think that's sufficient, as long as the rest of the disease doesn't get worse.

DR. BUYON: I would say that I agree, as long as the other doesn't get worse, but that's too vague for me, and I personally would vote down quality of life as a single outcome. I would also point out we have a very heterogeneous group of patients from an educational point

of view, and I know you don't want to talk about care of patients, but we're still dealt with Belleview Clinic, clinics in inner cities versus private practices, that we're drawing the group of patients here and certainly we've seen many individuals who feel very fine and want to refuse all of our therapies when we see the creatinine rising. So I would be extremely against quality of life as being the single outcome measure.

DR. FIRESTEIN: The analogy in rheumatoid arthritis, by the way, is that there are composite indices for disease activity but there are quality of life indications as well, and I think at least from my perspective that is a reasonable approach to this disease as well.

Jim, do you have a comment?

DR. WILLIAMS: Yes. I agree that quality of life is very important as an outcome measure, but I would agree that I do not see it as the primary outcome measure. It see it as an adjunctive outcome measure.

DR. CALLAHAN: I just wanted to clarify. When we talk about quality of life, I think it's very important to have it as adjunctive, and lupus is not my main area, but I would think you'd have to have other primary outcomes.

Are we having a discussion, too, about whether

it should be a generic quality of life measure versus more disease-specific or just any one? As long as they have any measure, it's fine?

DR. FIRESTEIN: I don't know.

Yes. Were you going to address that question?

DR. WALLACE: I think the quality of life indices are hampered a bit in that they don't take into account disability and they don't take into account fatigue, and I think they have to really be improved before we can use it as a single parameter.

DR. FIRESTEIN: Joan, and then Richard, and then we're going to move to the next question.

DR. MERRILL: I thought that Vibeke's data were compelling and my own instinct agrees with her data which is that if patients get better, their quality of life improves. I would rather see the focus of what we're doing here, which is a very serious intent, which is to try to finally figure out a way to develop drugs for lupus be on improving lupus. I have to agree with Jack on that.

I don't need an approval for lupus for Prozac. I can give Prozac anyway.

DR. FIRESTEIN: I forgot about Jennifer.

DR. ANDERSON: Well, I would submit that Prozac may affect some aspects of quality of life, but the way that quality of life is measured, it actually helps status,

and it has physical and mental components, and it actually reflects some aspects of disease activity and some aspects of damage as well. They're all interrelated. So I don't think that's a reason to reject quality of life.

The thing that I wanted to say was really following the item that Vibeke Strand presented relating to outcome domains recommended by OMERACT. The three, disease activity, damage, and health-related quality of life, also adverse events and economic costs, but those are in a different realm. It's important that all three, disease activity, damage, and health-related quality of life, be included, I think, in an outcome measure for use in clinical trials because they're all important.

I'd like to understand more about specific measures for health-related quality of life. I don't know that they're going to be discussed today, but given that fatigue and some other aspects of disability that may not be covered in the SF-36 are important in lupus, I don't know whether any attention has been given to developing lupus-specific quality of life measures, but that should be examined further.

DR. FIRESTEIN: Just a quick comment from Richard and then Jeff, and then we're going to move on to the last question.

DR. LOONEY: As I understand it, the question

was could it be used as a primary outcome and can I envision a group of lupus patients in which that would be a reasonable outcome to be the measure as opposed to it always being or being in people with organ damage. I think in people who have a specific organ which is the target of your therapy, then no, it wouldn't be an appropriate outcome, but people who don't have that, it would seem like that would be actually a very good outcome to use.

DR. SIEGEL: In the previous discussion, we've heard a couple of different points of view with respect to the importance of improvement in one specific organ system, and since this is a major part of the concept paper, I wonder if at some point we could hear a little bit more from the rest of the committee.

Some people have said that they don't think improvement in one organ system would be enough. The disease as a whole should improve. Other people have suggested that they thought improvement in one organ system would be enough, so long as the other organ systems don't worsen. So at some point, it would be helpful to get feedback from the other members of the committee.

DR. FIRESTEIN: I didn't get the sense that people were opposed to a single organ indication, as long as the other aspects of the disease didn't worsen.

DR. BUYON: A very brief comment. As I was

reading through the document that you handed us, I'm not sure I even agree with that. If renal disease was made better and a malar rash might be made slightly worse, I think we have to be very cognizant of the fact that perhaps that would be okay, and I would not want to close the barn door as an absolute, and I actually very much disagreed with the idea that everything had to be okay.

If we find a medication that literally stops diffuse prolific glomerular nephritis in its track and there just might be a little more hair loss, let the patient decide that, not us.

DR. FIRESTEIN: I don't think anybody suggested that actually. I think the notion is that if there is significant worsening that is of the same order of the original disease and then the patient is no worse off, but worsening of malar rash or alopecia.

What direction would you like to go at this point? Because we can go on longer on this, if you'd like.

DR. SIMON: It would be very helpful to hear the other two points from people have already raised their hands, or three.

But also, I'd like to end off a question. This may seem self-evident, but to us it's not. I'll go back to lupus nephritis yet again and their lupus nephritis is being treated with some specific agent and it improves. Do

you all expect that, in addition, you would like some other disease activity index to be measured as well as the indicator of the overall disease lupus improving, one, two, or three different measures? Would one be enough or do we need more than that?

Then furthermore, as Gary had alluded to before, a la the rheumatoid arthritis guidance document, in a tiered nature of the indications, moving from signs and symptoms to x-ray to physical function, would you all see that HRQOL would serve in the realm of a health-related quality of life indication further enhancing the investment into trial development and indications to allow more studies to be done as we did with the rheumatoid arthritis guidance document?

DR. FIRESTEIN: So who were the two people with their hands up?

DR. WILLIAMS: I just wanted to comment that I agreed with Jill. I read the document the same way, that any worsening would make it unacceptable.

DR. FIRESTEIN: Was that the intent?

DR. SIEGEL: No.

DR. FIRESTEIN: No.

DR. WILLIAMS: I wasn't speaking to intent. I was speaking to our interpretation.

DR. FIRESTEIN: I understand. I'm just trying

to sort through if that's what they wanted it to say.

Then Jack, and we'll just go around the table.

DR. CUSH: I was raising my hand in response to the organ-specific question. Why do we step up our therapy in patients? Usually it's because we deem lupus to be active based on several parameters or we deem one organ specific to be out of control that we have to treat specifically, thrombocytopenia or renal failure or CNS disease. So I would be in favor of a single, sole organ-specific indication, as long as those are well defined, based on some study.

But I still think the better way to go in trial design to get a drug approved, to answer Lee's question, is that more than one measure must be done for disease activity, to get that indication and then to get one of these other indications, whether quality of life or organ-specific. I'd like to see at least one of those improving with these organ-specific measures improving or quality of life improving. That would be what I'd like to see.

DR. MERRILL: I'd like to give an example of a situation where you could have a drug that would be wonderful for one organ in lupus and really might make lupus worse and that's thalidomide which is highly effective for discoid lupus and is a tumor necrosis factor alpha blocker. I think widespread use of thalidomide in

lupus might cause some pretty bad flares. I think the question is still kind of open, and I think it's a drug worthy of study. I think the approach of tumor necrosis factor alpha blockade for certain manifestations of lupus might be worthy of study, but we would be going in understanding that there's the possibility you could in certain situations cause flares.

When we're more sophisticated, we may be able to do this better and be able to tread lightly and know what to measure and not get the patients into any trouble doing it, but I have to vote that we keep a little bit of an open mind about worsening in other organs because I agree with Jill. She gave examples that were straw men that were easily knocked down, but it may be that for some people with devastating, disfiguring discoid lupus, it would be worth it to them to risk an arthritis flare.

DR. FIRESTEIN: I'd just add one small point to that and that is that the mechanism of action of thalidomide is not certain.

DR. MERRILL: Fair enough.

DR. FIRESTEIN: It is not at the therapeutic doses necessarily a TNF blocker.

DR. MERRILL: Fair enough. But there are actually other theoretical reasons to consider TNF blockers for discoid lupus.

DR. FIRESTEIN: I understand that.

Any other comments? Graciela.

DR. ALARCON: The question of whether or not you need to use a general instrument in addition to measure organ-specific, the answer is absolutely yes. I will favor to use more than one activity index. If you really examine your patients and ask your patients all the right question, you can score the SLAM, the SLEDAI, and the BILAG with your source document, as Jill mentioned, without any difficulties. So I think that really and truly you will be better off examining and scoring your patient completely and then scoring everything.

Then as part of the trial, I think that you have to include the quality of life, and I think that therefore it's not a lot more effort to do it if you are really spending millions of dollars in developing your medication or getting your medication to the market. So I'll go for all of them.

The fact that the SF-36 doesn't cover fatigue I don't think is true because one of the scales of the SF-36 is vitality and if you actually correlate that, very well, you can see that it correlates very well with the degree of fatigue the patients experience. So I think that's a very good instrument.

DR. DAVIS: I have a comment too. I'd agree

with the single organ --

DR. FIRESTEIN: Excuse me one second. We're going to have to, I think, move on at this point. One last very quick comment.

DR. DAVIS: Okay. It was just hearing other people's opinions, and my opinion of it is single organ system would be fine, and I wouldn't require multiple other markers to change if I had a very, very effective drug, for instance, for lupus nephritis.

DR. FIRESTEIN: Question number 3 regarded responder indices, and I think essentially most of the instruments that we're using or suggesting to use are responder indices. So I think that question is probably answered.

The last question relates to clinical trials with regard to irreversible damage, and I would ask Dr. Simon whether he wants to go into that discussion now or maybe come back to that later because we have another presentation at this point.

DR. SIMON: Well, actually, I would like to return to one other thing because it has to do with the responder indices issue and it actually does relate to 4. If one is to measure SLAM, SLEDAI, and BILAG in the same trial, there is an issue of multiplicity. You're doing multiple measures, and a responder index which actually

might/should possibly be developed would theoretically provide a bar of response that you'd be looking for that would take into consideration MCID that had been defined in some fashion and then would be able to then provide a bar that the SLAM should change by X, the SLEDAI by Y, and BILAG by Z, achieves that bar in A percentage of patients and that that A percentage of patients is acceptable to the community as a substantial response compared to placebo or standard of care.

That is the kind of thing we were looking for in the context of a responder index, but you're correct, any one of these things is a responder index. But if we're really going to go the route of multiple different measures within a trial, then I was wondering what the community thought about then inventing an ACR, WHO, ILAR, blank something for response in lupus that would take all of these measures into consideration.

DR. FIRESTEIN: I didn't sense a lot of enthusiasm for that in the previous discussion. Does anybody else want to comment on that? Bevra?

DR. HAHN: Personally, I think it's a good idea. I've been thinking about it for years and those of you who work in RA could maybe guide us better. That seemed to be a big breakthrough in RA.

I think the reason I'm reluctant to get into it

is that we're talking about years of study to validate it while potentially nothing in the pipeline gets released for use in SLE because we're all waiting for the development and validation of a response instrument.

So I guess I would like a sense from around the table, was it the breakthrough in RA therapy that it seemed to be to me as an outsider in the RA clinical trials? Are our measures in lupus so inadequate that we need it pretty badly before something new should really be considered approvable?

I guess I'd like those answers mostly from everybody.

DR. FIRESTEIN: Well, the ACR criteria for rheumatoid arthritis aren't exactly analogous to these other instruments because they actually are more analogous to individual components of the lupus instruments. So in a sense, you already have those for better or worse.

Tom.

DR. LEHMAN: Yes. I'm right here. If I could just make one point. I think we would be doing ourselves a disservice if we both slow down production of drugs while we try to determine the optimal index or assume that we can determine in a committee like this what the optimal index is.

If we go ahead and encourage in fact the

companies to include multiple indices when they do these studies, then although we are going to introduce some confounding in terms of multiple measures, the ability to use that data in the end by somebody like Matt in a computerized analysis to determine what in fact are the best factors to be included in a global index will present itself automatically.

DR. FIRESTEIN: David.

DR. PISETSKY: I was going to say to a certain extent for rheumatoid arthritis that these were retrospective data and they were based on clinical trials of existing agents, both the Paulus criteria and ACR. So there was a data set that allowed you to distinguish what worked and what didn't work, and this is, I think, a very different situation where you don't have the background of clinical trials to go forward.

DR. FIRESTEIN: Jennifer, Jack, Joan, and then Jeff.

DR. ANDERSON: Yes. A lot of the impetus for the development of improvement criteria in rheumatoid arthritis was the existence of multiple measures and the multiplicity of answers that you could get and it was very unsatisfactory for deciding whether a drug had really improved in comparing from one drug to another and so on.

Yes, enough of the measures had been used in

the past that it was possible to use preexisting data to develop the measure, and I would think in a few years' time, if trials have been done with using the various measures that are being suggested here, there would be enough data to develop a response criterion that would be a single outcome measure for use in SLE clinical trials, but I believe that it can't be done just yet. But if the data is gathered properly from all the trials and made available to somebody to do analysis, then it's not a difficult matter to come up with that within only a few years' time, I would expect.

DR. FIRESTEIN: Jack.

DR. CUSH: I would be against the combination of these tools as a responder index. I think you have the tools right now to give you the indications for control of signs and symptoms and control of quality of life and control of an organ-specific thing, and I would stick with those individually.

DR. FIRESTEIN: Joan.

DR. MERRILL: I agree with Jack actually. I think that for different drugs that you're developing, one of these instruments might be better than another, and I think they should be able to choose what their primary measurement is from some of the options that exist.

I do want to point out, however, that I think

the data are out there that could be used to develop really much better assessments of our tools and figuring out what the margins ought to be from the FDA's point of view, what kind of improvement they would want to see.

We've had completed clinical trials by Gene Labs, LJP, IDEC, Biogen. There's a lot of data out there that really hasn't been mined for what it could tell us about how to develop drugs.

DR. FIRESTEIN: Okay. As Dr. Siegel gets ready, one last comment from Dr. Diamond.

DR. DIAMOND: I think what Bevra said is very important, that we have enough tools that we shouldn't wait on anything before going forward with clinical trials, and I think the other thing that's true is what's been said many times. When we have a good therapy, we'll be able to assess which of these tools is best. And while I think it's very important to use a multiplicity of tools in clinical trials, I think it's also very important that when we do organ-specific clinical trials, which I certainly think we ought to be able to do in lupus, that we not require that one meet any standard on these global assessments, that we have them, but that part of the efficacy of the drug not be determined by that.

So we need to do them to learn, but not necessarily to approve the agent.

DR. FIRESTEIN: Thank you. Well, we're all in complete agreement, as usual.

(Laughter.)

DR. FIRESTEIN: Now Dr. Siegel will talk about SLE claims.

DR. SIEGEL: Thank you. In my talk this morning, what I'd like to do is to discuss some considerations with respect to deciding what types of claims, what types of benefits should be recognized for agents undergoing clinical trials in systemic lupus erythematosus.

As background for my talk, I just want to review a few points that I think many people in the audience were already aware of. We have not had any new products approved for lupus in recent years, and while products can be developed without guidance, formal guidance from the FDA can be helpful. Guidance on what can represent adequate evidence of efficacy can have an important role in facilitating drug development. In a disease like lupus, ideally guidance should recognize a broad range of potential benefits that therapeutic products could achieve in this disease.

In formulating a claim structure, as I mentioned, it's desirable to include a wide range of potential clinical benefits, but there are a number of

challenges in reaching this goal. For one, as has been mentioned many times, lupus has very widely different manifestations from patient to patient and over time. Disease activity has a tendency to wax and wane over time, making assessment difficult and complex, and there's a paucity of randomized clinical trial data to be used to characterize the clinical benefits of many of the currently used agents.

On this slide, I'm showing some of the potential claims that are under consideration in the agency. The first would be perhaps the most straightforward in some ways, which is that a new therapeutic product would improve disease activity in a specific organ.

The next, which is anything but straightforward, would be reducing signs and symptoms, and a claim of this type would be based on a trial that showed improvement in a disease activity index -- and I'll call this DAI in the rest of my talk -- compared to a control arm.

But there's one very difficult problem that we've been grappling with, which is that if improvement in a trial like this concerns non-internal organ system manifestations, perhaps the benefit that such a trial would show would be better described as improvement in

constitutional symptoms and constitutional aspects of the disease rather than improvement in overall disease activity, and I'll talk about this in more detail as I go on.

The other claims that are under consideration is prevention of lupus flares, complete response or remission, and improvement in health-related quality of life. So turning first to organ-specific disease activity, the evidence that we're talking about here would be based on a study that enrolled patients with active disease in a specific organ system. So, for example, patients could be enrolled who have disease in renal aspects of disease, hematologic, pulmonary, or central nervous system disease.

In addition, such a study could enroll patients who have disease in more than one organ system, but you'd use stratification, so patients with disease in each organ system would be balanced across study arms to be able to reach conclusions about each organ system manifestation in the trial, and a successful trial would demonstrate better control of disease in the involved organ system with study drug compared to control.

In many cases, however, outcome measures are not yet well-defined for organ-specific manifestations and this presents really an enormous challenge for optimal design of clinical trials. One possibility is to use

portions of disease activity indices that assess specific organs to explore those for their suitability as outcome measures in clinical trials and this would need to be done on a case-by-case basis and validated.

The definition of success could be restricted to complete remission in that organ system, or it could allow partial responses in control of disease activity to also be recognized as a clinical benefit.

One specific example of improvement in an organ-specific manifestation would be lupus nephritis. Lupus nephritis has, of course, represented a major cause of morbidity and mortality in the past. However, modern management has been associated with improved outcomes compared to earlier eras. Nonetheless, current treatment modalities are associated with considerable toxicity in many cases.

Possible outcome measures for lupus nephritis are shown here. Survival and progression to end-stage renal disease represent clear clinical benefit but may occur too infrequently to serve as sensitive indicators of treatment effect. Other potential outcome measures include doubling of serum creatinine, and this has been reported to predict progression to end-stage renal disease, at least in certain populations. Others include smaller increases in serum creatinine, such as an increase of 50 percent, or

sustained attainment of renal remission using accepted criteria, such as normalization of an active urine sediment, improvement in glomerular filtration rate, or improvement in proteinuria.

Turning next to the next potential claim, reduction in signs and symptoms, this claim would represent success in a clinical trial that showed benefit in signs of disease activity and the associated symptoms, but as I mentioned before, there's been considerable internal agency discussion about the relative merits of calling such an improvement a signs and symptoms benefit versus improvement in constitutional symptoms. And I'll tell you more about exactly what I mean by constitutional symptoms in a minute.

But such a clinical trial showing improvement in signs and symptoms would assess overall control of disease activity, so in contrast to an organ-specific, overall control of disease activity using a disease activity index, such as the SLEDAI, the SLAM, the BILAG, or another validated index.

Since disease activity indices measure a wide range of disease manifestations, defining the clinical benefits demonstrated in such a successful trial may be quite complex, and I'm going to illustrate that with two extreme examples in the next two slides. This isn't in your handouts. I apologize.

Consider trial number 1. Such a trial would enroll patients with active lupus stratified for the type of internal organ system involvement. Let's imagine trial 1 at the end of the trial showed scores on the disease activity indices that were statistically significantly reduced with study drug compared to control. And further imagine that the percent of patients with renal, pulmonary, CNS, hematologic manifestations, each represented about 25 percent of the overall study population; namely, there were enough patients to assess that there was improvement in each organ system with study drug compared to the control. Based on such a study, you might conclude that the study drug showed efficacy on a variety of major internal organ system manifestations of lupus.

Now consider another study, trial 2. This study would also enroll patients with active lupus, perhaps a similar size study, perhaps patients with similar overall baseline scores in their disease activity index, but this trial does not stratify for the type of internal organ system involvement. Imagine that trial 2, at the end of the trial, shows scores on the disease activity index that again are statistically significantly reduced compared to control.

However, imagine that the percent of subjects with each of these individual organ system involvements,

renal, pulmonary, CNS, hematologic, in this case only represent, say, 10 percent of the overall study population, and with such small numbers, there's no clear evidence of improvement with study drug compared to control.

But imagine that the improvement in disease activity can be largely attributed to improvement in arthritis, skin, fatigue, and other non-internal organ system manifestations. Here you might conclude that there is a drug effect but you cannot conclude that there's clear evidence of efficacy on internal organ system manifestations.

So these are the two extremes. For a signs and symptoms claim, to attain such a claim, a product would need to show benefits in control of the common and serious manifestations of lupus. Therefore, a trial showing efficacy would need to enroll subjects with disease affecting the major target organs in lupus, and it would need to demonstrate that the efficacy is general and not restricted to specific organ systems, otherwise perhaps you would have a claim for those specific organ systems but not for signs and symptoms of lupus in general.

Now, let's talk in contrast about what reduction in constitutional symptoms might mean. Here, the idea is that some products may demonstrate an effect on disease activity indices without affecting disease activity

in major internal organ systems by affecting what you might call the constitutional aspects of disease, for example, effects on arthritis, rash, fever, fatigue, serositis. So perhaps reduction in constitutional aspects of disease should be recognized as a distinct claim.

Such a claim would represent improvement in constitutional symptoms as a clinical benefit of products that don't affect the internal organ system manifestations. Now, one of the challenges here is that currently there are no validated instruments for assessing constitutional symptoms.

The next claim I'd like to discuss is prevention of lupus flares. Here, the idea would be that a product showing this benefit would have demonstration of efficacy in preventing lupus flares in trials of adequate length that showed one or more of the following potential benefits: reduced frequency of flares over an adequate time frame, increased time to flare compared to a control arm, or reduced severity of flare. And for a trial like this, use of a validated definition of flare would really be essential.

Now, you could argue that efficacy in prevention of flares is really similar to the benefit of control of disease activity. However, some products may be effective at preventing flares but could not be used in

treating acute disease. So, for example, high-dose corticosteroids may treat acute disease but may be too toxic to use to prevent flares in long-term use. Other products, in contrast, may be better tolerated long-term and could have utility in preventing flares, and for an example, I would give a study from John Esdaille, et al., in the New England Journal that assessed the use of hydroxychloroquine in preventing flares. So the idea here would be that prevention of lupus flares may represent a distinct benefit in some circumstances.

Next, turning to complete clinical response and remission, this claim would be defined by analogy with a similar claim for rheumatoid arthritis as a prolonged absence of disease activity in patients who previously had active disease. The clinical trial evidence would involve absence of disease activity, for example, for 6 consecutive months. The study could represent a 12-month clinical trial with disease activity score achieving 0. For a complete response, the outcome would be achieved while patients were also receiving other lupus-directed therapies, whereas for remission, the outcome would be achieved in patients receiving no other lupus therapies. And furthermore, the claim could pertain to one single organ system or could be for treatment of lupus generally, depending on the patient population studied.

I want to present a few caveats in these claims. The proposed claim structure I've talked about allows for approval of products affecting targeted organ systems. However, products to be approved must show an overall favorable risk-benefit ratio. So for example, there shouldn't be any worsening of other aspects of lupus that would counterbalance the benefit seen in the particular organ system under study and no unacceptable adverse event profile again that would counteract the benefit seen.

Of necessity, regardless of the claim that was being sought, clinical trials should assess all relevant disease domains, including disease activity, irreversible damage, and health-related quality of life.

Now, let me turn to heath-related quality of life as a claim for a minute. OMERACT recognized health-related quality of life in lupus as a key domain in assessment of lupus in clinical trials. Recognizing a claim of improvement in health-related quality of life is under consideration. A product that attains this claim would previously be shown to also reduce disease activity or in the same trial, and evidence should include a validated health-related quality of life measure in lupus and a patient global assessment. Assessment of health-related quality of life outcomes in clinical trials should

include the same statistical rigor as other endpoints under assessment.

In conclusion, the proposed claim structure would recognize a variety of potential clinical benefits. Some of the challenges are: one, designing clinical trials that clearly demonstrate which patients would benefit from the therapeutic product and what benefits they would attain; and, two, describing the benefits seen in the studies in a useful and accurate manner for patients and clinicians. Finally, clinical trials should assess the effects of the therapeutic product on all domains of disease in order to fully characterize risks and benefits.

Thank you.

DR. FIRESTEIN: Thank you very much.

Yes, Dr. Diamond?

DR. DIAMOND: Can I just make one comment? I mean, I'm sure there are lots of comments that we all want to make about this, but can you tell us why a claim cannot be for reducing morbidity from therapy? Because I think all of us who work in lupus know that we want drugs that don't have the side effects of Cytoxan and steroids, and it would seem to me that that would be a major advance and it's not included in any of the claims.

DR. SIEGEL: I think we are interested in hearing additional claims that the committee might think

are worthwhile and important that we did not include in our talk paper or in my talk today, but I can discuss a little bit about why we didn't include that.

I think some people might think that it would be straightforward to show that a product reduces toxicity due to another therapy, but it actually can be quite difficult. For one thing, suppose the product reduced the toxicity from another product short-term but it had to be used long-term, and when it was used long-term, there were additional toxicities that came out after only 6 months or a year of treatment. Then what we'd have to do, to assess whether there's a favorable risk-benefit of the new product, is not just to show that you don't get the toxicities of the old agent but to characterize the toxicities of the new agent and somehow weigh one against the other.

I think this would depend on what product you're talking about, but there could be some cases where the new toxicities would clearly counterbalance the benefits of not having the toxicities of another product. I don't want to close off discussion. We'd very much like to hear your thoughts, but that's one of the concerns.

DR. FIRESTEIN: Yes, sir.

DR. SIMON: Thank you. Without dragging this on longer, I think it's important for the committee to

remember that the FDA is charged with approving drugs that demonstrate efficacy and safety, and under those circumstances, such a putative agent would have to demonstrate efficacy and that alone in the context of safety, the safety issue would be highlighted significantly within the label in clinical trial descriptors, as well as potentially ways that the agency might choose, to capitalize and emphasize the safety issue.

So the idea of an indication for safety is a difficult one because in fact it's an efficacy indication that is associated with an improved safety profile and that would be heavily described if that exists.

DR. DIAMOND: So can you just clarify something for me? So that means something could be approved if it performed as well as Cytoxan and had a better safety profile or if it performed as well as Cytoxan and had the same safety profile. Both? Neither? That's what I'm not understanding.

DR. SIMON: I'll make it even clearer. A, no, because Cytoxan is not approved and has no indication and furthermore has had no real good clinical trials that show efficacy.

But let's say it did and let's say Cytoxan was an approved therapy. A new product that came along that was efficacious that was not inferior to cyclophosphamide

in the context of lupus nephritis, let's say, and had improved safety and was proven -- and important improved safety -- would receive a label that would describe such a thing and would give an approval.

If in fact it was superior to cyclophosphamide, it would not require cyclophosphamide to be approved at all anyway and therefore would get that, and again if it had an improved safety signal, that would be highly described within the label.

DR. FIRESTEIN: And with that, we'll have lunch. So we will reconvene at 12:40.

(Whereupon, at 11:38 a.m., the committee was recessed, to reconvene at 12:40 p.m., this same day.)

AFTERNOON SESSION

(12:40 p.m.)

DR. FIRESTEIN: If everybody is ready, we'll go ahead and get started, and if people aren't ready, we'll go ahead and get started anyway.

So this is an interesting change in terms of how these meetings are held with an open public hearing and some statements from a whole list of folks over the next 30 minutes. So we have six people that are scheduled and potentially others that may not be on the schedule who will have the opportunity to speak. So if you will each come up and introduce yourselves, and I believe each individual gets 5 minutes to make their comments, unless otherwise indicated, yes.

So the first is Dr. Paul Brunetta from Genentech.

(No response.)

DR. FIRESTEIN: Well, that was shorter than I anticipated.

(Laughter.)

DR. FIRESTEIN: The next is Dr. Dan Wallace.

DR. WALLACE: Thank you. My name is Dan Wallace. I'm a member of the Division of Rheumatology at Cedars-Sinai Medical Center and a Clinical Professor of Medicine at UCLA.

DR. FIRESTEIN: Hang on one second. I'm sorry for interrupting. I'm supposed to read something in advance of this which I thought was going to be read by somebody else.

In any case, 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 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.

Now, I apologize for interrupting you.

DR. WALLACE: I understand the other speaker came in. I don't know if you want to have him first.

DR. FIRESTEIN: I think that's very reasonable. So the first speaker is Dr. Brunetta.

DR. BRUNETTA: I just wanted to read a very brief statement related to claims for treatment, and this is a point that's been touched on by this committee previously.

Cyclophosphamide and prednisone remain the standard of care for treatment of severe lupus nephritis. The concept paper comments on improved renal survival with the use of these medications, and Cytoxan, as mentioned by Dr. Diamond, in particular is known to have significant treatment-related morbidity. The concept paper does not assert that a Cytoxan-sparing program would be a claim for treatment in a complete clinical response or induction of remission trial in lupus nephritis. So that is part of a question whether or not a Cytoxan-sparing claim would be acceptable to a committee such as this.

If a Cytoxan-sparing claim is not acceptable to the FDA and a placebo-controlled trial in lupus nephritis is considered, we then have to determine what would be an ethically acceptable period wherein a patient with severe

lupus nephritis would be treated off of Cytoxan and how we would then determine rescue for that patient.

So that's the main point that I want to make, that Cytoxan-sparing is quite important to patient care, very important to investigators and to clinicians and how we would consider Cytoxan-sparing in a program.

DR. FIRESTEIN: Thank you. Did you state your financial interest?

DR. BRUNETTA: Yes. I'm Assistant Medical Director at Genentech in the Biotherapeutics Division.

DR. FIRESTEIN: Thank you very much. Now, back to Dr. Wallace.

DR. WALLACE: Even though I'm not running for Governor of California, I appreciate you giving me 5 minutes of your valuable time.

The FDA needs to generate a guidance document giving industry a crystal clear road map which will lead to the burgeoning and not discouragement of clinical trials. The nearly 1 million Americans with SLE demand no less.

I've been in the trenches, so to speak, seeing 20 lupus patients a day for the last 20 years, and it's discouraging to see industry implement their best ideas and initiatives in other disorders. I am fully cognizant of the weaknesses, confounding factors, and biases of every statistically validated inflammation quality of life damage

index and biomarker evolved for the disease.

I am thrilled with Lee Simon's leadership and I am thrilled that the FDA has made several seminal suggestions in their draft document that improve our current methodology, such as looking at area under the curve for SLEDAI, and I think it makes sense to rely on at least two activity indices in a clinical trial.

Let me embark on a historical perspective for a minute. In 1948, the year LE preps and steroids became available in Marian Ropes' Lupus Clinic at the Mass General, half with lupus died in 2 years. This observation divided those with organ-threatening from non-organ-threatening disease.

As Sandra Raymond pointed out earlier, by the mid-1960s, 60 percent with lupus were living 10 years and by 1990, 90 percent in the United States. This improvement in survival rates took place during an interval when no lupus drugs were introduced into the market. We had Cytoxan in 1963.

It was largely due to the skills of practicing clinicians in learning how to manipulate steroids, alkylators, antibiotics, antihypertensives, and the availability of dialysis. This improvement can be attributed to the clinical skills of rheumatologists in interpreting SED rates, complements, anti-DNAs, 24-hour

urines, and urine sediments and managing patients accordingly. Improving the survival rate was due to physicians being able to identify a flare and managing it accordingly.

I don't know of a single rheumatologist in private practice, other than myself, who has BILAG software or calculates damage indices, but my point is that assessing improvement in lupus is not rocket science and that finely-honed clinical acumen is all that is needed to ascertain if a treatment regimen is effective or not.

Lupus should be an easier disease to quantitate than RA because there are fewer subjective factors, such as morning stiffness, that are used in clinical trials. The weaknesses of the ACR-20, 50, 70, the DAS28 score and Sharp scores are no more or less serious than those of the SLEDAI, SLAM, or BILAG, and the FDA is actively promoting RA clinical trials in spite of these deficiencies.

Over the last weekend, I read a transcript of the hearing relating to the advisory committee recommendations for fibromyalgia drug trials which took place on June 23rd of this year. In my opinion, those recommendations seem clearer than what we have for lupus, and the irony is that in fibromyalgia, just about everything followed is subjective.

Matt Liang and his committee have explored

markers to evaluate organ-threatening disease and assess steroid-sparing regimens. The drafts of his paper endorsed by the ACR and submitted to ANR for publication are positive, cogent, and constructive. They provide hard evidence that the FDA should include in its road map.

The experience in clinical trials conducted thus far validates the use of the ACR classification, the use of quality of life indices, the use of damage indices. When plugged into Matt's specific organ markers, a combination of BILAG with SLEDAI in a response index, I feel confident that investigators now have more than enough of an armamentarium to conduct an honest, rigorous lupus clinical trial. Adding a few biomarkers or surrogate markers, such as anti-DNA or C3 complements, is icing on the cake.

The draft document wants more documentation that anti-DNA or C3 can be biomarkers. We've already heard about the LFA poll where a 131 of a 132 rheumatologists polled feel that they are lupus markers. You already heard what Dr. Buyon talked about this morning.

If you look at Frank Quismorio's chapter in the 2002 edition of the Dubois textbook, he reviews 6 perspective and 6 retrospective trials validating the use of anti-DNA and 11 prospective studies validating the use of C3 in over 1,500 patients. There is even a paper that's

going to be presented with myself as a co-author with LJP at the ACR validating the use of anti-DNA further.

I cannot prove it, but in my opinion, the ability to follow anti-DNA and C3 are one of the major reasons why mortality rates plummeted between 1960 and 1990.

I had the privilege of serving on Matt Liang's ACR Nephritis Guidelines Committee and it addresses the concerns relating to confounding variables with renal function, validation of doubling of the creatinine, induction of renal remission, surrogate renal markers, the issue of cellular casts in a very comprehensive manner, and I'm sure that this would be acceptable when incorporated in the final document.

Finally, the issue of measuring flare which, in my opinion, is only one of six major categories of outcome measures in conducting a trial, is the weakest link we have right now in validating a lupus trial, but this should not delay trials. The flare indices were premeditated and plugged into these trials and we're just analyzing the data now. The FDA should propose provisional parameters for measuring flare that can be changed and adapted as current trials are analyzed.

In my opinion, the menu of ascertainments we have now, while flawed, are as good as what the FDA has

endorsed for RA, fibromyalgia, osteoarthritis and osteoporosis. My lupus patients deserve the same considerations as their rheumatic disease compatriots and nothing less.

Thank you.

DR. FIRESTEIN: Thank you. The next speaker is Kelly Jean Cooper.

MS. COOPER: Hello. My name is Kelly Cooper. I live in Chicago, and the Lupus Foundation of America is underwriting my trip here.

I have quite a dramatic story. I was diagnosed five long years ago and have not experienced remission once since then. My usual symptoms include inflammation, high fevers, facial rash, chronic fatigue, hair loss, general malaise, cognitive difficulties, painful joints, and chronic severe chest pain.

When I first became ill, my anti-DNA numbers were so high, I was immediately put in the hospital and pumped with a three-day megadose of steroids. This treatment was effective briefly but then my numbers slowly started to climb. In the past five years, I've had that same treatment three times and suffered some of steroid's awful side effects.

I also began taking oral steroids which I'm still on. I couldn't even attempt to count how many times

I've raised and lowered my oral doses, guessing on my own the appropriate amount to take in relation to the symptoms I was having.

As most lupus patients are, I was put on Plaquenil but this drug had no effect on me. My doctor then proceeded to put me on Imuran which my body just couldn't tolerate, then CellCept which made me feel worse than Imuran. Next in the line-up was methotrexate, started at 10 milligrams, then went to 15, and still my body just didn't respond.

I've traveled to the Mayo Clinic and the NIH to see if there was any reason why my body was not responding to anything at all, but left their care with only the suggestion to up my dose of methotrexate yet again. I have finally found some stability on 20 milligrams of methotrexate, but let me reiterate it has only stabilized me. I am not getting better.

In the past five years, I have had and still do have pleurisy, pericarditis, fluid in the lungs, dry skin, dry eyes, and I carry the anti-RO, anti-la, anticardiolipin and antiphospholipid antibodies which will make carrying children a very dangerous endeavor for me. And I am literally in Northwestern Memorial's ER no less than twice a month for pain.

I've been on pharmaceuticals for malaria, anti-

organ rejection, cancer, and arthritis, but not one of these drugs has been approved for use in lupus patients, and I cannot say strongly enough how important it is for the FDA to stimulate private companies to do research specific to lupus.

What you are doing here today is very important to me and many others like me trying to live with this awful disease. I asked for the opportunity to testify during this particular part of the meeting because, as I understand it, the document as it now stands includes three claims that a drug company can make on behalf of a new drug for lupus. However, for each of these claims, there appears to be problems that may be seen as impossible to overcome by the very drug companies that we want to attract to work on lupus.

To overcome these problems, the FDA must be willing to invest its funds in helping to solve problems, such as coming up with an accepted definition of a flare or validating biomarkers for lupus.

Will the National Institutes of Health step up to the plate and help address the gaps in science that the draft document cites? Will the NIH provide research money to find answers to these questions?

Your decisions on these matters have life and death consequences for me and many other people with lupus.

I ask you on behalf of myself and all others who suffer from this disease to please make your decisions now rather than later.

Thank you.

DR. FIRESTEIN: Thank you very much. The next speaker is Betty Ann Exler.

May I ask the speakers not only now but also throughout the day and tomorrow to try to stay within the 5-minute guidelines? Thank you.

MS. EXLER: Hi. My name is Betty Ann Exler and this is my son Scott, and we were asked to come by the Lupus Foundation of America. They are underwriting our trip, and I want to thank you for the opportunity to speak here today.

I am here as the mother of a child with lupus. One of the most important issues for parents of children with lupus is the long-term effects of the drugs that the children must take to treat this disease. Safety and effectiveness of drugs for lupus is very much on our minds and we are very concerned about the toxicity of the current treatments.

One day when Scott was in second grade, he came home from school. After leaving to go to school just fine, he dropped to the floor and said, I don't feel good. I hurt everywhere. He had a fever, did not want me to touch

him anywhere, not even to help him get up. I immediately took him to the doctor who found that his spleen was enlarged and tested him for strep. The test was positive. He was treated with antibiotics. He seemed to feel better, but within a few days, he was feeling too ill to go to school and the doctor treated him two more times with antibiotics, until it became clear that they were not helping. His spleen was still enlarged. Mono tests were coming back negative.

Our doctor then sent us to an infectious disease doctor who ran a number of blood tests and discovered Scott was ANA-positive and IgA-deficient. His blood cell counts were very low. He sent us to a hematologist who found Scott's immune system was destroying healthy blood cells. He sent us to an immunologist who found other immune system problems.

At this point, Scott was having severe knee and ankle pain and was having a great deal of difficulty even walking. We went to a rheumatologist who tested his urine and found blood and protein present. He sent us to a nephrologist.

This all started in February, just a week after Scott's 8th birthday. At the end of July, almost 6 months later and 40 blood tests later, the doctors informed us that our son had systemic lupus with kidney involvement.

His blood count by this time was so low, he was near needing a transfusion. His enlarged spleen and achy joints were keeping him from going to school and playing the sports that he so enjoyed.

He was immediately put on 40 milligrams of prednisone daily which helped alleviate the most severe symptoms quickly. It also made him gain weight and changed his appearance so much so that when school started a month later, some of the children didn't even recognize Scott. The medication made him very edgy and unable to concentrate in the classroom. Noise gave him headaches and playing sports was difficult.

Scott responded well to the prednisone but is very sensitive to changes. When the doctor tried to lower his dosage, his symptoms would immediately worsen. He spent his third grade year going through the misery of withdrawal from prednisone only to have the levels raised again. This scenario repeated itself constantly for the next two years.

The next summer, Scott developed a skin rash on his face, arms, chest, back and legs, more severe than anything I've ever seen in my life. His skin was red, swollen, hurt, itched and was so fragile, that if I tried to play with him, his skin would break and bleed. To combat this, the doctors raised his prednisone to 60

milligrams per day. By the time Scott started fourth grade, his skin was doing better, but he continues to have a much milder rash than that today.

In the meantime, the doctors started him on CellCept to suppress his immune system. He was put on Plaquenil to help his skin. He was put on Prinivil to hopefully decrease the amount of protein leaked by his kidneys. All of these medications cause harmful side effects, cataracts, deposits in the eyes, inability to fight infection, increase in the risk of cancer, sterility.

As a parent, I was shocked, heartbroken and devastated. The doctors told us we really had no choice but to put our young child on these toxic medications. Without them, he would continue to deteriorate with no chance for improvement, and if these medications don't work, our only choice would be to treat him with even more toxic medications. He has not been able to discontinue any of his medications since he was diagnosed over three years ago.

The only word I can say to describe my feeling when the doctor informed me of his diagnosis is devastation. I never really truly knew the meaning of this word until then. The feeling never leaves me. I am heartbroken that this child who is so full of life, kind, funny, generous and well-loved, so talented and fun, must

live a life confined by his disease and the side effects of his medications.

Scott is very difficult to manage medically and the doctors have few choices in medications or therapies to even try. The doctor tells me I am so very sorry. I wish I could do something, but we must have hope. Hope is really all we have, and the hope is in this room.

I cringe and want to jump up and down and scream whenever I hear that this is a manageable disease. You can live a normal life and the prognosis is so positive. I know this is not a graceful or complimentary mental image, but Scott's disease is very complicated and difficult to treat. He does not live a normal life and he does not have a good prognosis. Our family lives with the pain of this disease every day as we watch Scott deal with this life-altering, life-threatening disease.

As I looked over the concept paper, it struck me that there are both positives and negatives. For example, one of the ways the document suggests to prove a drug is working is to show a decrease in the frequency or severity of flares. However, in the same paragraph, the document states that no measure of a flare has been validated.

In a different section, the document suggests using disease activity indexes, but the document then goes

on to say these indexes have not been validated in clinical trials. The document calls for making claims based on knowledge that presently does not exist.

These contradictions are frustrating for people with lupus, for parents as well as for sponsors of clinical trials. I hope additional efforts will be made to address these gaps in knowledge and contradictions so safer and more effective therapies can be developed for children and adults who suffer from this disease and also suffer from the side effects of the present toxic medications.

I came here today as a member of the Lupus Foundation to ask you and as a mother to beg you to give research clear-cut guidelines so that we may bridge the gap between the devastation and the hope.

Just real briefly, this is a book about pediatric lupus that Scott's third grade class worked on and did, and I left a few of the copies with Kimberly if anyone would like a copy of this. And this is Scott, and if you don't mind, he'd like to say a word.

MR. EXLER: Hi. I'm Scott Exler. I'm 11 years old, and I have lupus. I would really like you to try and find a cure for lupus because it is really not fun to have lupus. Thank you.

MS. EXLER: Thank you so much for your time.

DR. FIRESTEIN: Thank you very much, Scott, for

sharing your story.

The next speaker is Lisa Amato. Again, if you can comment on potential conflicts of interest, I would appreciate it.

MS. AMATO: Hi. My name is Lisa Amato, and the Lupus Foundation of America has underwritten my trip here to speak to you today.

Thank you for the opportunity to address the committee. I want to comment on the document, but first I want to take a few minutes to tell you how I got here.

It all began with low-grade fevers, joint pain, and loss of appetite. For years, we could not find the cause. After going to numerous doctors and undergoing a battery of tests, including an invasive biopsy, I learned my kidneys were slowly deteriorating. I was 21 years old and diagnosed with lupus.

My story is not unique. Many suffer from the same initial symptoms I had but many have yet to be diagnosed. When young people begin treatment to fight lupus, they are susceptible to complications from long-term use of the medications, such as diabetes, high blood pressure, high cholesterol, osteoporosis, and obesity, which can lead to heart attacks by the age of 40.

Despite high doses of steroids to fight lupus, my health worsened, with high blood pressure, anemia, fluid

retention, and protein in my urine. After nine years of undergoing several more biopsies, weekly blood tests, and an angiogram, my doctor told me I had end-stage renal disease. With only 10 percent kidney function, I was expecting to die, until we learned that my sister was able to donate a kidney. Our transplant team told me I was going to feel better with a new kidney.

Remarkably, I was a totally new person the day after the operation. To prevent rejection, I was given prednisone, Imuran, and cyclosporine. Three years after the transplant, I was diagnosed with lymphoma caused by the drugs. My life with lupus has not been easy.

For the last 18 years, I have been fighting for my life, taking an arsenal of powerful medications each day to prevent my immune system from destroying the vital organs of my body. These medications often are worse than the disease itself. After a few years on prednisone, I developed avascular necrosis, forcing joint replacement surgeries for both knees and twice on both hips. Clearly, we need safer, more effective treatments without the severe side effects.

As I read this document -- and clearly it was difficult to read -- I felt as though the document was laying out an agenda for research on lupus. It points out that there exists many gaps in scientific knowledge for

lupus. I feel that the contradiction will discourage the development and testing of safer and more effective drugs for lupus.

As a lupus patient who had end-stage renal disease, I am concerned that the document creates an impossible hurdle to overcome. It appears that lupus is being held to a higher standard that makes it difficult to prove that prospective new drugs are effective. This will cause drug companies to avoid working on lupus.

For me and others who have had major organ involvement, this would be very disappointing. We need safer drugs now. If not now, when? Thank you.

DR. FIRESTEIN: Thank you very much. The last scheduled speaker is Venetia Thompson.

MS. THOMPSON: Good afternoon. My name is Venetia Thompson. My trip here has been underwritten by the LFA, and I am the wife of a retired worker from Monsanto. I'm not sure if that has any bearing. We are still holding some stock.

(Laughter.)

MS. THOMPSON: I want to thank you very much for the opportunity to provide comments regarding the proposed draft document. I am particularly interested in this effort because of the higher prevalence of lupus among African Americans as well as the higher numbers of African

Americans that suffer with serious lupus nephritis and the importance of minorities being well represented in any type of a study that's being done. For these reasons, I wanted to present some brief comments.

At the age of 9, I was hospitalized for chronic fatigue, abnormal blood tests, unexplained fevers and headaches. Suspecting that I was suffering from juvenile diabetes, when those tests found out that I was not, then I was sent home. Symptoms continued to plague me as my body went through growth cycles, so did my symptoms. Pregnancy proved to be the most difficult for me.

When a butterfly rash appeared on my face at the age of 35, my obstetrician suggested that I take a lupus test. I was ignorant of what lupus was and was not aware that there were no known tests, and he told me that the test came back negative.

At the age of 40, a co-worker watched my fingers as they turned white and were numb. Believing that I had a connective tissue problem, she encouraged me to see a doctor immediately. I did and I tested positive for ANA connective tissue problems, Raynaud's phenomenon, and rheumatoid arthritis, but it wasn't until my boss had to drive me home twice in one week and my family members had to pick me up off the bathroom floor that I realized I really could not work any longer.

Fatigue and chronic pain from pleurisy has taken over my body. Six years later and three specialists, who all diagnosed me differently, finally have brought my symptoms under bearable control.

The draft document points out the disagreements among researchers about the ability to measure disease activity and how to weigh these measures, and there are situations where changes in scores do not necessarily reflect or relate the changes in disease activity. This makes it difficult to know if the drugs are having any effect.

I have been disabled for the last 4 of my 46 years. My doctor draws blood three to four times a year and I'm on six nonsteroidal medications a day. I watch my diet very carefully and I exercise. It was difficult to resist using steroids to help resolve pain issues, but I did not choose to use them because I was too concerned about the side effects from using them.

My treatment is somewhat costly, time-consuming, and labor-intense, but I do believe that over time, the costs will prove to be far less substantial and ensure greater quality to my life. I've lived 30 years with my husband. We have a daughter who is pursuing her Ph.D. in biochemistry. I have one son who is an Army officer that's serving in Iraq right now, and my other son

is an Army officer who will be joining him in March. I would like to live long enough to see grandchildren.

Thank you very much.

DR. FIRESTEIN: Thank you, and thank you to all the speakers.

Now, this is the time when I'm supposed to open this up to the audience. It's an interesting conundrum because there are 30 minutes allocated to this section and there are six speakers each allocated 5 minutes. So the time remaining will be available. So if people would like to make a comment, they can be entertained now, but they should be, please, brief.

(No response.)

DR. FIRESTEIN: In that case, we will move on. So for the next hour or so, the goal is to go over the next series of questions posed to the committee from the agency, and actually some of them can be done together here. So why don't I just read them?

The first one is: would a claim for "treats constitutional manifestations" that include such manifestations as arthritis, skin involvement, fatigue, fever, weight loss, be acceptable? Then the next question actually is: for an individual without specific major organ involvement, should a claim for "treats constitutional manifestations" be considered as an

indication? What outcome measures are appropriate to support this claim, such as a DAI?

I guess I'm not sure I like the term "constitutional manifestations," and we'll find out in a minute if there's much disagreement that this would be a reasonable indication.

DR. HAHN: I'm stimulated to talk about this one. Jeff was talking about it. I think if we're going to talk about constitutional symptoms, we should maybe back off into items like fatigue and pain and maybe disability or something like that.

I was a little bothered by skin and arthritis being rolled in here because in some people, they can be very bad and require so much treatment that it's quite dangerous. So I was a little bothered to separate out skin and joints from kidney and brain and hematologic. Although maybe not as consistently life-threatening, they certainly can be.

So I was thinking if we talk about constitutional, maybe we should back off or make a little more global description of people's constitutional disability and not imply that some organs are less important than others.

DR. WALLACE: Yes. I agree. Constitutional is simply fatigue, fever, and weight loss, nothing else. Pain

falls into quality of life and disability. So when you're talking about constitutional, it's something that doesn't apply to a specific organ system but all organ systems and it's only those three entities.

DR. FIRESTEIN: Well, I think that was one of the reasons why I wasn't thrilled with the notion of constitutional although, again, many of the therapeutics that are being entertained are generally divided into those that might be useful for major organ system disease and then all others, and I think what you're driving at here is all others.

Jack?

DR. CUSH: I would take "constitutional" off the table. I think that while it is a major problem for many patients, I think it's also hard to define, and it's really analogous, I believe, to RA. We don't treat the pain of RA with narcotics. We try to control the inflammation of RA and pain will take care of itself in the vast majority of individuals.

The same is true here. If you control the disease, you control many of these hard-to-describe aspects, fatigue and poor sleep and weight loss or not feeling well. You'll control those aspects or constitutional aspects of the disease as well by meeting the signs and symptoms indication alone.

DR. FIRESTEIN: So what you're suggesting is that those are not independent variables, that they're dependent on the activity of disease as measured by other parameters?

DR. CUSH: Yes.

DR. FIRESTEIN: Yes.

DR. MANZI: I would generally agree with that, but the only component that I think may stand alone might be the fatigue component. I mean, I liked the idea when I saw this that there might be an indication for lupus fatigue, even in the setting of relatively inactive organ system involvement. So I might be a proponent of fatigue in a constitutional symptom claim, but I do agree with taking out the other organs in constitutional symptoms.

DR. FIRESTEIN: I don't know that we should get hung up on whether they should be called constitutional or not. Lee, is the question related to true constitutional symptoms or non-major organ systems? Not that skin is not a major organ. I know that it is.

DR. SIMON: It's actually really driven by the fundamental question, do we approve a drug to treat lupus, whatever that might be, and, of course, it's open to debate? We have a criteria for diagnosis of 11 different things. We put together for clinical trials to look at 4 of 11. Is that what we consider lupus, and thus any of

those things that go into our ingredient list to make the diagnosis could be those things that we look at for improvement as it relates to an indication for "lupus."

We thought that that was a difficult approach. So we were thinking that we could ask a different question, and a la the RA guidance document again as has already been alluded to, in a signs and symptoms way, could we look at that question and identify those things that might be amenable to some kind of therapy that may not be the same kind of therapy to treat a specific organ?

Because we've only had those therapies that seem to treat many different things at the same time, it doesn't mean that in the future, we will not be able to have something that might just treat the fatigue and fever and weight loss of the constitutional components. So that's what we were trying to get at.

DR. FIRESTEIN: It seems to me that it would be inadvisable to set the bar so high that only therapeutics that will treat major organ system disease that have significant mortality associated with them, for instance, would be approvable and a good example would be antimalarials, which clearly have benefit in patients but would not fare well in a trial head-to-head, say, to cyclophosphamide in renal disease.

Is there any disagreement on that?

DR. HAHN: But they would in arthritis and skin. You know, if you picked your organs.

DR. FIRESTEIN: Right. I agree. I think the line is being drawn here between major organ systems and the rest as opposed to constitutional which is fevers, weight loss, fatigue.

Joan.

DR. MERRILL: I think the issue really is do we really have that patient who only has fevers, weight loss, and fatigue and doesn't develop a flare in any organ. In my clinical experience, that's fairly rare and probably would, if it happened, respond to a short course of steroids. So it's not one of our major needs, and I guess that's why you're not getting a lot of enthusiasm.

DR. FIRESTEIN: Well, based on prescribing patterns for antimalarials, I think that there's a lot of individuals where that would have value.

DR. MERRILL: I'm thinking of my own clinical practice which is lots and lots of lupus patients, and I'm prescribing antimalarials for arthritis and fatigue and weight loss and fevers, but there's always something else going on when you see those.

DR. ILOWITE: So couldn't arthritis and skin involvement qualify as organ system disease for organ-specific claims in one paradigm, and then I would be in

favor of leaving fatigue, fever and weight loss for a constitutional indication.

DR. MERRILL: Yes, and I think Bevra is making a very important point. You can have someone very sick from arthritis or have disfiguring discoid rash and then you can have mild rash and mild arthritis, and it may not be the same drug for both.

DR. LOONEY: Would it be possible to let whoever is proposing the trial select the manifestations that they want to say the drug is going to be good for that they want to have constitutional skin, serositis, and arthritis and say that that's what they want to get an indication for? Would that be acceptable to people?

DR. FIRESTEIN: Yes.

DR. ILLEI: Leaving the label of this aside, I thought that the screen would be to include patients with different manifestations into the same study and show that a drug works or to be able to compare patients who have arthritis to a patients who has skin disease and then label the drug that it can treat lupus in general. And I think that it, to a certain degree, brings us back to the question of a responder index where it could be individualized. So the patient who comes with arthritis should respond in the arthritis, and if they meet a certain response, then they are responders, and the same can be

brought down for any manifestation.

DR. FIRESTEIN: Dr. Hoffman.

DR. HOFFMAN: I'd just be concerned that if someone were to design a trial to diminish malaise and fatigue, that that is likely to be confounded by concurrent therapies that patients are on that may, in fact, cause malaise and fatigue. While I think it's important to track, in the context of a trial that deals with a broader concept, I think just malaise and fatigue, if patients are on background therapy of cyclophosphamide or methotrexate and perhaps other agents, CellCept, Imuran, that the effects of those drugs in causing malaise, fatigue, even perhaps weight loss, may not be easily sorted out from the underlying disease.

DR. WILLIAMS: I would just like to agree with Dr. Cush. I would take skin and arthritis and move them to organ-specific or even signs or symptoms but maybe organ-specific, and I would not consider constitutional symptoms as a primary indication for the drug.

DR. HAHN: I'd like to speak to rescue it. I see what the FDA is trying to do, and I think that it's a good idea and maybe it's the wording hanging us up. I'd still like to have pain in the definition because I don't think it is like RA. I think pain and fatigue often persist after everything else looks like it's better in SLE

patients.

It's a little hard to quantitate, so you'd have to have people presenting something for that indication, would certainly have to have very rigid measures and inclusion and exclusion criteria and this sort of thing. Certainly if we had something, in addition to Plaquenil, that had these effects, it would be something that we all wanted to use in subsets of patients and patients after they improve.

So I'd like to think of a way to rescue it. Taking those two organ systems back, I agree, putting them back into the organ-specific part, and I don't know. I'm thinking about how to do that. So Matt, are you still on? He's not on?

So I was interested in whether in anybody's work, can you pick out a symptom complex that would define this subset? We all know what we're talking about, right? So how would you define it?

DR. FIRESTEIN: David.

DR. PISETSKY: To a certain extent, this is a matter of severity. I mean, this is what could be called milder lupus of skin, joint, that would respond to antimalarials, nonsteroidals, and other medicines like that, and I could certainly see a value for developing alternatives for that group of medicines. But here, it's

constitutional. It's really this complex of less non-severe organ system disease. But I don't know that I would pull arthritis out of it nor do I know I would pull all skin disease out of it because there's a value to other drugs --

DR. FIRESTEIN: This doesn't exclude having organ-specific indications for those, but as a more global non-major organ system, and I tend to agree with that.

Joan?

DR. MERRILL: Actually, this work has been done. There is a definition for this. It's a BILAG C score, and I think that would probably address Bevra's concern. When you have a BILAG of an A and you treat it or a B and it gets down to a C, you've pretty much got what Bevra described, and so you could consider that as an indication for a certain kind of therapy that might be a little safer and might be not quite as powerful as some of our others.

DR. WILLIAMS: However, if you're getting down to definitions by BILAG score, isn't that coming under signs and symptoms rather than constitutional?

DR. MERRILL: I mean, it can be either one. BILAG is a very flexible instrument.

DR. FIRESTEIN: Can we take the word "constitutional" out so we don't argue about it anymore?

DR. SIMON: Yes, but if you're going to do that, let me just step back for one second. Internally, we were thinking about the issue that constitutional was consuming signs and symptoms, but yet I've heard several people separate that out. I would just like to have some understanding of what then is signs and symptoms.

DR. PISETSKY: To a certain extent, some people would put serositis in this, too, which I think would be reasonable. It's not listed here, but I think it would be included. So I think it's that constellation of problems that you would like to have a term for that is not strictly constitutional since it does involve organ systems. It's just, I think again, degree of severity.

DR. SIMON: I don't mean to be facetious or to drive a drug in turn or deal with semantics, but in that context of the non-organ system-based symptoms and signs, then we're talking about the disease lupus unrelated to organ involvement specifically.

DR. FIRESTEIN: Essentially, yes. I would hate to see the bar raised so high for a therapeutic that it would discourage drugs that would address that.

Jack was next.

DR. CUSH: What are signs and symptoms? You define that in your protocol, based on a certain level of SLEDAI or SLAM or BILAG and that's your threshold. Now, a

BILAG C, if that's the trial you want to do, then that's where you start. You're starting at a lower level, but at least you set the bar at whatever level you want for signs and symptoms and that's your indication.

I think to rely on single variable poorly-defined outcomes like the ones we mentioned -- now, fatigue has poor tools for measuring now, and those can be measured and they get some sort of secondary credit or secondary outcome measures. But as a primary outcome or as a primary indication, I still don't think that that's wise.

DR. FIRESTEIN: Jill.

DR. BUYON: I would just make some disagreement. There's specificity, and I think serositis, arthritis, and skin disease may be specific for lupus. I'm not sure about fever and weight loss and fatigue. So I personally would separate those, but this is one's person opinion. Not to say that I wouldn't use those three if that is what industry wanted. Perhaps they could make that very specific. But I would never lump what is specific to lupus with what isn't specific to lupus.

DR. FIRESTEIN: Jeff.

DR. SIEGEL: I think the discussion from the committee has been very helpful, but there's one aspect that I want to focus on. We imagine that there will be clinical trials not just of patients who have one specific

manifestation, like renal, or a group of patients with renal manifestations, a group with skin, a group with joint, but clinical trials that would take all-comers who have active disease based on a disease activity index.

Our concern is suppose you do such a trial and you show a reduction in the disease activity index, but you don't have enough patients and enough data to say that it affects each major internal organ system. Would that be a basis for a claim and, if so, how would you describe it? The specific concern was you may not have evidence in that trial that it improves CNS lupus, that it improves renal lupus, but yet the disease activity indices come down.

The term "constitutional symptoms," was brought up in part to help describe something that hadn't been shown to improve all the internal organ system manifestations, but nonetheless did decrease the disease activity index.

So I think it would be helpful to learn from the committee how it would see a trial like that that showed a decrease in disease activity index in an all-comers trial without necessarily showing that it specifically reduced renal, CNS, so on.

DR. FIRESTEIN: David.

DR. PISETSKY: You used two terms. One is "symptoms" and the other is "manifestations," and I think

they do really have different meanings. The problem here is things like serositis and arthritis may not be very objective, but they are subjective and patients will know the difference. I think the term "symptom" may have some value here.

DR. FIRESTEIN: Jack was next.

DR. CUSH: I think the way you sort of set it up in your cases is the way you set up your protocol. So if the company that's sponsoring a product wants to go after an organ-specific indication, they have to structure their protocol to answer that question. But I would argue that again everyone should meet at least some measure of a disease activity improvement by, hopefully, more than one measure and that's your indication for lupus activity, signs and symptoms, and then if you want to go even further, much as like in RA, we go for quality of life or x-ray improvement, so it affects renal or hematologic or articular outcomes, you have to structure your protocol and power it appropriately to go for that indication as well. Then for quality of life, maybe it has to be longer, more than 6 months' duration, maybe as we do in RA, a year or 2 years.

DR. FIRESTEIN: Mary Anne.

DR. DOOLEY: I was just going to say that in many respects, we oftentimes concentrate on organ damage

from the disease, but I think if you survey patients, the most troubling symptoms oftentimes can be these so-called constitutional symptoms, particularly relentless fatigue. So certainly there are many people who cannot take Plaquenil or there may be other drugs that might do a better job than the existing drug. I don't think we should ignore the less organ-threatening spectrum of disease and favor only organ-damaging disease in terms of allowing people to develop drugs.

DR. FIRESTEIN: Joan, Jim, and then Michael.

DR. MERRILL: I think that most of us who treat lupus patients -- I want to agree with everyone. I think there really are sort of three levels of lupus. One is exactly what Mary Anne just described, but it could be almost anything. It could be a little mild thrombocytopenia that you're not worried about but it's there. It could be fatigue. It could be arthralgia, myalgia, fevers, but not high fevers, not toxically-ill. These people may be going to work. They just feel lousy, and we'd like to have one kind of drug for that and we'd hope that would be a very safe drug, too.

And then there's sort of the moderate people and then there are the really severe people where right now we're going to cyclophosphamide or high-dose bolus steroids.

I don't think clinicians are all that confused about how to stratify those patients, but I don't think it's kind of organ so much as it is severity of disease.

So again, I don't mean to sound like a broken record, but the BILAG took care of that for you already. A, B, C.

DR. WILLIAMS: Actually that was my same comment, that I think for what Dr. Siegel is referring to, if you're going to use the disease activity index, you're really talking about signs and symptoms and that would cover it.

DR. FIRESTEIN: Dr. Weisman?

DR. WEISMAN: I think there are some circular reasoning here. The disease activity measures were developed to capture this panoply of non-organ-threatening signs and symptoms, and that's what we have. That's lupus. Then there's some specificity for what historically has been felt to be organ-threatening/life-threatening lupus which is renal disease, and that's where we've been.

I don't think we're going to change that. I agree totally with you, Gary, that we just have to get rid of the term "constitutional." We understand where the meaning comes from. We understand where Jeff is coming from. It means the non-life-threatening/organ-threatening signs and symptoms, and just get on with it. So I agree

with you.

DR. FIRESTEIN: So we've reached consensus. I'm just kidding.

(Laughter.)

DR. FIRESTEIN: Jill.

DR. BUYON: I would make actually two points. One, I wouldn't want to a priori discourage anyone from making a claim about anything that a priori they made that claim for. So if industry wants to say these are the items that we would like to make better, who are we to say not to do that? Why can't we agree up front yes, you have something objective. You've written it out. Whatever level it is, severe or not, I don't think it's our job here to sit here and discourage that. So that bothers me.

The other is these composites are the sum of the parts. So if you look at an activity index and you see what gets better, it's not a number. It's what constitutes that number, whatever instrument you use. And I think we're moving away and I'm actually a little discouraged by these comments.

DR. FIRESTEIN: Joan, Bevra, and then we're going to go on to the next question.

DR. MERRILL: Don't be discouraged, Jill. No rheumatologist is going to go to hear a presentation about a trial and not ask, well, what percentage had arthritis.

We want to know what was being treated and the rheumatologists will demand to know that.

I think that when you develop a drug that has a certain mechanism of action and you go to the FDA and say look, now I can tell you I can predict this is only for nephritis, the FDA's going to listen and there are already medications like that. Their medication is aimed at nephritis like a bullet and they're not going to fix other things. Those are legitimate medications to develop, and I don't think anyone's discouraging that.

DR. FIRESTEIN: Dr. Hahn.

DR. HAHN: I want to go back to Jeff's question which has been bothering me. In your scenario, you said that drug X would have been shown to reduce activity indices, but there weren't enough individuals in each cell of organ involvement to say it reduces arthritis or it reduces nephritis.

I guess I'm bothered about although that would be nice, would that mean that the FDA couldn't approve a drug if it reduced disease activity in SLE? Isn't that a good starting place?

DR. SIEGEL: We're asking you how you would see such a trial, how you would describe the benefits that trial showed to help us advise sponsors on how to design their trials and so on.

DR. HAHN: I wouldn't have any trouble with that outcome being worthy if reducing disease activity.

DR. FIRESTEIN: That really leads into the next series of questions, which I'm going to lump together, which again is related to using organ-specific endpoints versus activity indices.

It says, for an individual with major organ system involvement, is this best studied utilizing a DAI or organ-specific endpoint or both? That's number 3. Number 4 is, is the claim for treatment of signs and symptoms which includes individuals with major organ system involvement acceptable, and would a DAI be an appropriate outcome measure? Last, should a claim of "treats lupus" be considered? If so, how many organs should be studied, et cetera?

So, again, we're circling back around to the same type of question. Can there be organ-specific endpoints? Would that stand alone, or should we only be looking at DAIs, or should there be a combination of them?

Yes, go ahead, Mary Anne.

DR. DOOLEY: I think that we should use a combination. The organ-specific endpoint should establish the efficacy of the drug and the disease activity indices in that setting would just ensure that in fact you're not inducing new manifestations of lupus with that agent. So I

would see them not as primary endpoints but rather as secondary or safety points.

DR. MERRILL: I really think again Jill is looking very grim, and I agree with you. I think you've got to have some flexibility here because we don't even understand the biology of lupus well enough yet to know whether some of these medications that are being developed out here might treat arthritis and skin and not nephritis whereas another one will treat nephritis. So we have to let them find out. So I think there has to be flexibility here.

Someone has got to go into a trial with a primary outcome measurement but that will be arrived at based on drug mechanism and what's predicted.

DR. FIRESTEIN: But that's what phase II is for. By the time you get to phase III, you should have a specific indication.

DR. MERRILL: Correct. By the time you get to phase III, I think that some medications, it might be best to be organ-specific, and some, it might be perfectly acceptable to treat lupus because that might be what they do best.

DR. FIRESTEIN: David.

DR. PISETSKY: But some of this would even be more than organ-specific, it would be manifestation of

specific organs. So what you may do for diffuse proliferative nephritis would not be what you did for membranous and you'd need something separate, or diffuse CNS disease as opposed to other CNS disease would also be separate. So I think you're going to have to be encompassing enough to allow people to focus on particular organ-specific manifestations.

DR. FIRESTEIN: Lee.

DR. SIMON: So I have two questions that have grown out of this, and I've now learned to ask one at a time.

The first is, is there something called "indicated to treat lupus" because it has three organs involved versus four organs involved that you're addressing versus two organs? Can I get some clarity about treats lupus and what would be those things that would get you that? Is there a minimum number of things that would allow you to say that's your indication?

DR. FIRESTEIN: Isn't that what the DAIs are designed to do?

DR. SIMON: Well, some of us don't believe that. That's the problem about DAIs.

DR. FIRESTEIN: Joan.

DR. MERRILL: Well, I think that this was to some extent addressed by the SELENA SLEDAI document which

defined severe flare, mild to moderate flare. I think going through these trials was very instructive because I thought I was a very astute clinician, but when you start tallying things up, you do learn a few things.

I think most of us who see patients who have what we would call by a SLEDAI a severe flare, which is a score of 12, those are very sick patients and those are not the same people as a SLEDAI of 4 or 6. So really what that ends up being is more organ systems because you can arrive at that number various ways. So unless you have a really severe organ system like CNS involved, you don't get 8 points at once and most of the patients who get to 12 are accumulating organs.

So you have a good point. I am not sure whether or not the science is out there to tell you what that number is yet, but it's probably worth doing some reviewing about that.

DR. SIMON: And my second question is -- and it's more of am I correct in hearing this, yes or no -- I'm hearing that we should be flexible enough so that our indications are really what the community defines in designing their trials that they're trying to prove, whatever that indication would be, meaning if you think you have a therapy that would treat arthritis, that should be enough, if you're successful in treating arthritis, and

thus the label should reflect this is for the treatment of lupus arthritis, and that we in fact are trying to lump too much too early when it should be split more until we have further information.

DR. FIRESTEIN: I think the corollary to that, Lee, was the use of a DAI, in addition to that, to show that there's not something untoward that has happened.

Jill.

DR. BUYON: I would say 100 percent second the motion. I'd like to hear that as being the directive of the FDA. I think the activity index is your safety net, but to not encourage industry to go for either specific manifestations or global, I think that would be very unfair.

I also would say "to treat lupus" is a very broad term. It could be the arthritis of lupus, the renal disease of lupus. I don't see anything wrong with that, and I'm not sure we ever could sit here and say, well, we have a drug to treat lupus. If we have a drug to treat discoid lupus, we have a drug to treat lupus.

DR. SIMON: Could I ask your indulgence then, Mr. Chairman? This is really a critical question. Could we determine that everyone on the committee does or does not agree with this kind of approach? It changes entirely the way we traditionally think about approaching such a

document. It's not that we're against it. We'd just like to know that everybody agrees with it.

DR. FIRESTEIN: Are you asking for a vote on a specific question?

DR. SIMON: The question I would ask for the vote on would be to split rather than lump the individual manifestations and allow the sponsors to define what they are targeting and whatever they are targeting, if they win, with the appropriate provisos and built-ins and the issue that disease activity indices do not worsen concomitantly, et cetera, and that statistics that are associated with that, is that an acceptable way to go?

DR. MERRILL: Clarification. Acceptable, in other words, they could lump? They have the choice. Correct?

DR. SIMON: No, no. That would be the choice.

DR. MERRILL: They can choose to lump or split.

DR. FIRESTEIN: So we're going to go around the table, but because there are 25 of us, we can't have any testimonials. We just have to basically say yes or no with one or two sentences at most. So the question was as you just defined it. Well, you're still allowed to lump.

DR. MERRILL: It's both or lump.

DR. SIMON: Lump and/or split or just lump.

DR. FIRESTEIN: Sounds like a menu.

DR. LOONEY: I guess I go for the lump and/or split.

DR. FIRESTEIN: I've had a motion here to raise hands so we don't have testimonials. I've been advised that there have to be individual votes. So we will again endeavor to go around the room.

I just don't exactly understand what we're talking about.

(Laughter.)

DR. FIRESTEIN: You're allowed to split not in a vacuum but with some lumping in addition to that. Okay. Now I understand.

DR. CUSH: I am not sure that lumping and splitting is raising it right. Are you saying that by lumping a disease activity indication, that's global, and by splitting, you're saying splitting aside from organ-specific indications, we're actually going to allow now symptom-specific indications? That's what you're allowing when you say splitting. So not just organ-specific but symptom-specific, including itching and fatigue and things that might --

DR. FIRESTEIN: Well, which ones are being split off needs to be defined separately, but I think we're probably talking about some of the things that have been discussed, such as renal disease.

DR. CUSH: That's actually very different than fatigue. I mean there are millions and millions --

DR. FIRESTEIN: I understand.

DR. CUSH: -- of people with fatigue and ANAs. Are we going to call them lupus or are we going to now allow this be an indication that's going to be treated with a new drug?

DR. FIRESTEIN: No, but this is just as a matter of philosophy, not for that specific indication that you're talking about. Would it be reasonable to have an indication for lupus nephritis?

We're going to take a vote on it, so everybody now will be able to be heard, and then we'll just tally up the votes. So why don't we just start off?

DR. LOONEY: What are the A and the B here? Could you give us the statement of what A is and what B is?

DR. FIRESTEIN: Okay. A is an organ-specific or a manifestation-specific indication along with a global index as a safety net, as was pointed out, or just a global indication. That's the B.

DR. DIAMOND: It's whether the claim is this treats lupus or this treats lupus nephritis or this treats thrombocytopenia or this treats lupus skin disease. It's whether you can have a restricted claim or whether you have to say that your drug treats all of lupus, so you have to

show that. That's essentially what it is.

DR. WALLACE: Essentially, you have four things. You have symptoms, signs, laboratory abnormalities, and organ disease.

DR. DIAMOND: But we're not deciding now --

DR. FIRESTEIN: Hold on just a second, please. All right. So the vote is -- there are only two choices -- for either a global lupus indication, that's B. I'm doing this in reverse order to confuse people even more. Or A, you can have an organ-specific claim.

DR. DIAMOND: A restricted claim. We're not deciding now whether fatigue is in or out.

DR. MERRILL: But it's not one or the other. You could have either one. It's do you give people a choice or not.

DR. FIRESTEIN: Please. Folks, we're just going to go around the table now.

DR. SIEGEL: I'm going to try to again. A is an organ-specific claim. B is only a lupus claim.

DR. LOONEY: A.

DR. ILLEI: A.

DR. HARDIN: A.

DR. HAHN: A.

DR. DOOLEY: A.

DR. ALARCON: A.

DR. PISETSKY: A.
DR. MERRILL: A.
DR. GIBOFSKY: I think A. I'm a lawyer here, and I'm confused by all this parsing that we're doing.
DR. FIRESTEIN: Dr. Gibofsky, you've used your time allotment.
(Laughter.)
DR. GIBOFSKY: A lupus general claim.
DR. HOFFMAN: B.
DR. CUSH: A, only organ-specific, not symptom.
DR. ANDERSON: A, either organ or symptom.
DR. WILLIAMS: A.
DR. FIRESTEIN: A.
DR. CALLAHAN: A.
MS. McBRIAR: A.
DR. MANZI: A.
DR. ILOWITE: A.
DR. FINLEY: A.
DR. DAVIS: Organ-specific or symptom.
DR. DIAMOND: A.
DR. BUYON: A, whatever the claim may be.
DR. WALLACE: A.
DR. WEISMAN: A.
DR. HOFFMAN: We can make that actually unanimous because of the confusions of A and B. It

actually was A.

DR. FIRESTEIN: Oh, all right. So it's unanimous. Therefore it must be correct.

(Laughter.)

DR. FIRESTEIN: All right. So we have a few more minutes in this time allotment in order to talk about the last couple of questions. Number 6 is what criteria should be used to define remission, and how long should a remission be observed? Dr. Williams.

DR. WILLIAMS: Many years ago, the CSSRD did a study on early undifferentiated disease and we did a 5-year follow-up. Then they asked us to do a 10-year follow-up. One of the interesting aspects of that was patients who were in remission at 5 years with lupus were not in remission at 10 years. Over half of them went back into disease.

I'm uncomfortable with calling remission in lupus. I think you can say they're no longer active, but I think these patients went in and out, whereas rheumatoid patients tended to stay in remission.

DR. FIRESTEIN: But that's not a problem in a number of other diseases where people can be in remission and then the disease can come back or recur at a later time.

So how long does somebody have to be in

remission before you would feel comfortable calling it a remission? Joan?

DR. MERRILL: The remission has to be permanent. I mean if you're allowed to get sick again and still you called it a remission at that time, then I would say 1 year.

DR. WALLACE: What do we mean by remission? Clinical? Laboratory?

DR. FIRESTEIN: Well, that was defined by Jeff, wasn't it? The actual definition of remission was --

DR. WILLIAMS: Jeff's definition, I think, was no disease off medication.

DR. FIRESTEIN: No disease off medication.

DR. WALLACE: Well, is that a positive ANA in somebody off medication who feels fine?

DR. FIRESTEIN: No, I don't think so. A positive ANA is not lupus.

DR. LOONEY: You're saying that you would not say somebody is in remission if they had a positive ANA?

DR. FIRESTEIN: My personal opinion? If they were clinically in remission and were not on therapy and their ANA were persistently positive, I would still be comfortable calling that a remission.

DR. LOONEY: Sure. I think everybody would agree with that.

DR. MERRILL: This has been done before. It's called a BILAG D.

DR. FIRESTEIN: Are there 26 BILAG --

(Laughter.)

DR. PISETSKY: Depending on the agent and the manifestation, being off all medication is actually fairly stringent because there are patients with lupus who are going to be on a variety of things that are not necessarily immune-related. Anticoagulants, you know the list. And I don't think you really expect people to be off of those in the entirety.

The other thing is whether there should be separate consideration for being on nonsteroidals as opposed to steroids or antimalarials.

So I think all medication is really a little stringent and not realistic.

DR. DIAMOND: The definition was lupus-directed medication.

DR. PISETSKY: Right. But I think you can argue where some of these should be fit in. Does that include ACE inhibitors or aspirin?

DR. FIRESTEIN: Do you have a comment, Dan? Jack, and then Gary.

DR. CUSH: I think this is putting the cart way before the horse. I think that since we're struggling with

BILAG C's and D's, that I think it's a lofty goal but one that must be, I think, based on some study and a guidance document or the result of some consensus on what defines remission. It would be a very rigid definition. There would be a time element. I don't know that that's up to the committee to give guidance on that right now.

It's such a high bar, I don't know that it needs to be developed right here and now. I think in developing your guidance document, you can talk about it and take suggestions for it and wait for developments in that field, but I don't think we have any firm definition of it at this time.

DR. LOONEY: I think we're already faced with the possibility that treatments are actually going for remission. I think that's the goal of the Hopkins study with ablative therapy for Cytoxan. They're looking for treatment-free remission of disease.

DR. FIRESTEIN: Gary.

DR. HOFFMAN: I think the quibbling is going to be about the suggestion that we consider remission off all drugs because there are other disease activity tools, and in fact in oncology as well. Remission is not defined based upon being off or on any drugs. It's the absence of any signs of active disease.

So I think we can define that remission in fact

takes place if someone is on Cytoxan and prednisone, but then we have to subqualify great remissions when people are off steroids entirely and the greatest remissions when they're off steroids and adjunctive therapy, whether that be cytotoxic drugs or some of the newer biologics.

But I think this runs a little counter to some of the other guidelines and even dictionary definitions of what constitutes remission, where in fact if you look it up in Dorland's, it's only disease improvement. It doesn't even mean absence of all signs and symptoms of disease.

DR. FIRESTEIN: Mary Anne, and then Bevra.

DR. DOOLEY: I think speaking about remission, too, may not be able to speak about remission in lupus in general, but also be organ-based. So, for example, there are guidelines being presented for remission in lupus nephritis and those might be in some respects more easy to define than remission in other organ systems. So I suspect it will end up needing to be more organ-based.

DR. DIAMOND: Can I ask a question? I'm sorry. Go ahead.

DR. HAHN: I want to pick up on Gary's theme, and I think that these are different levels and maybe it's too hard to recognize them because of numbers, but if we call remission treatment-free absence of symptoms and signs, leaving the lab out, and then we call complete

response absence of symptoms and signs and activity but you could still be on some medication, and partial response, there's some level of medication that's required and there are still symptoms and signs, I think we could follow what Gary was saying and we could make some words or definitions that would capture all three of those.

Now, my question for the statisticians was if we do that, will we never get enough n? Is that too much? That's reality. Is it too much subsetting to get statistical validity, do you think?

DR. FIRESTEIN: Did you want to comment on the statistics? Then we'll go over to this side.

DR. ANDERSON: I would just say that some years ago, it was thought that there would never be complete response in rheumatoid arthritis and so there wasn't much point in defining it, but now, they're getting more and more. So it really depends on what proportion of the patients are going to end up in that state and maybe now you don't have it but in 5 years or 10 years, you may. So that isn't a reason not to do it because you don't have it right now.

DR. FIRESTEIN: John, and then Norman.

DR. DAVIS: I think we are putting the cart before the horse. We've been dealing with this with spondylitis for the past couple of years as well, and it's

a totally separate disease. I agree with what Mary Anne has been saying, which is for each organ-specific claim, we can have a definition of a remission, but for lupus overall, I can't think of a good-enough definition.

So I'd like to throw another word into the pot of low disease state, being maybe 20 percent on a number of different scales. I know it's a BILAG whatever, right?

(Laughter.)

DR. MERRILL: This has been published. This has been validated. You guys are reinventing the wheel. I think we should all go read the literature.

DR. DAVIS: But I think we don't know enough about the disease pathology to be able to define really what a remission is, and if we go back even to looking at these sort of constitutional things that we were talking about before, a patient could totally look very well but feel very unhealthy but be off of all medicines. Are they in remission at that point? I have nothing to measure it.

DR. ILOWITE: It seems to me that this is a relatively semantic argument because there are descriptive terms to describe things that we can define any way we want. But the real question is how long are they going to stay in this state and not relapse, and we're not there yet. So we don't know how long you have to be in remission or complete clinical response or off medication or not

before you're less likely to relapse, at least not to my knowledge.

DR. DIAMOND: I completely agree with Norman. I think that when you need this discrimination is when you have two good drugs and you want to know if one is better than another. But we're not there yet, and at the moment, we can use our words to describe that it improves renal disease, that it improves hematologic disease, that it improves total disease global assessment, whatever, and we would not do a service to ourselves and patients by getting hung up on these definitions when we don't have a reason to need them yet. Hopefully we will.

DR. FIRESTEIN: So there you have it. We don't know what a remission is and so we don't know how long it should last either, I guess.

Is there another comment? Joan.

DR. MERRILL: I thought our charge was to say how long a patient would have to be disease-free to consider it a remission, and I thought the definition of remission was not cure, it was just absence of disease. So I would be convinced that someone had absence of disease if they really had it for a year.

DR. FIRESTEIN: Since we can't agree on what a remission is, then --

DR. MERRILL: It's a BILAG D.

(Laughter.)

DR. FIRESTEIN: I understand that and we've been admonished to read the literature which we shall do

So did you have another question about that?

DR. SIMON: So after this long discussion and the various different questions, I wanted to pose another question to the committee. You don't have to take a vote, but I'd like to hear your answer, which is, so if we identify that a sponsor decides that arthritis of lupus is important to them and they want to develop a therapeutic for that, that might be a nonsteroidal anti-inflammatory drug. And that would be acceptable to this panel that the sponsor would go ahead and design and implement a discovery program to demonstrate that a specific nonsteroidal anti-inflammatory drug would be approved for the treatment of the arthritis of lupus.

DR. FIRESTEIN: I can't speak for everybody else, but from my perspective, if they wanted to invest their resources in that manner, there are certainly broader target audiences for that particular indication in terms of arthritis that can also work in lupus as well, but if that's what they choose to do.

So the last couple of questions here are related to the use of flares and how we assess them and how we use them. So should a new therapy also study the

treatment of active disease if a prevention of flares claim is sought? Is it acceptable to consider all flares of equal importance as an outcome measure?

DR. CUSH: I think flare trials are probably best if you're looking to do a placebo-controlled trial, meaning that you have stable disease and you're looking to prevent flares. Then that's the way you can place your placebo-controlled trials and maybe that's during phase II. But then when you want to treat nephritis or active cerebritis or active severe hematologic disease, really sort of threatening sorts of organ-specific manifestations, that's where maybe an active control randomized trial might make more sense, comparing your new intervention with whatever the standard is at that time.

DR. FIRESTEIN: Jill.

DR. BUYON: I couldn't emphasize more strongly that to lump all flares together I think would be a major mistake and that we should clearly recognize the difference and implication of a mild/moderate flare versus a severe flare. I think our experience with the SELENA trial unequivocally dealt with that.

And patients, by the way, also want to know. When they say is something going to happen to me, there's a very big difference to them between going on dialysis and again we were discussing hair fall, for example. Or an

oral or nasal ulcer is an area which we didn't even touch upon. So I would make a great push that we differentiated mild/moderate from severe flare.

DR. FIRESTEIN: Jim, did you have something to say, and then Joel.

DR. WILLIAMS: Well, I was just going to say that based on our previous vote, that if they wanted to go in for an indication of flare, they define it and do the study and then we base it on that study.

DR. SCHIFFENBAUER: The reason behind the question was if a drug is approved that prevents flares, it's likely at some point to be used off-label to treat active disease. So the implication was when and if it should be studied to treat active disease because that's how it would potentially be used. That was the reason behind the question.

DR. HOFFMAN: I think while it's important to track flares and types of flares, the proposed document gives us lots of options about ways that we can perhaps more accurately get a feel for the disease. For example, I think on page 8 and 9, looking at disease activity, there's the suggestion of also looking at area under the curve for disease activity throughout the trial with frequent intervals of assessment. So I would suggest to our partners in industry, should they be interested in

utilizing such a tool, that they also take to heart the AUC determinations as an important endpoint.

DR. DOOLEY: In terms of saying that there may be an agent that would prevent flares but that would not be substantial enough to treat active disease, I certainly think that exists. There have been some small trials of Plaquenil, for example, that suggest that it may prevent serious flares of lupus, not just mild flares, but no one would presume to treat active nephritis with Plaquenil alone.

So I don't think that because there may be physicians who choose to use drugs off-label that we should make regulatory requirement that these drugs that might be good enough for maintenance but not strong enough for active disease would have to prove efficacy against it.

DR. FIRESTEIN: Gabor.

DR. ILLEI: I just wanted to say the same. I do think that a drug that has the claim for prevention of flare has to be proven to treat active disease.

DR. FIRESTEIN: Bevra.

DR. HAHN: I agree with that, and I wanted to be educated about this. So what is the FDA's responsibility to try to prevent using drugs off-label? What are the guidelines?

DR. SIMON: That is exactly what I was going to

address. We don't regulate medical care and the community should do what is standard of care. The key issue is understanding the safety of such a thing.

The use of it might be slightly different. It might be used for a different period of time. The analogy is in the pain field, an acute pain drug, we know if people believe it works for acute pain, will be used for chronic pain. We've had any number of examples of that, and thus it's used for a longer exposure time and a different kind of exposure time, thus potentially opening up safety risks that could not be seen in a 2- or 3-day trial for acute pain.

So under those circumstances in creating this analogy, one would think that it is possible that entrepreneurial and aggressive and interesting physicians might do interesting studies to demonstrate that a drug to prevent a flare would then be used to treat active ongoing disease in a way that might be slightly different or for a longer period of time. So it allows us to understand a little bit more about its safety issues than just its efficacy.

DR. PISETSKY: I was going to say that there's plenty of precedent of drugs used preventively to prevent a bad outcome. You lower lipids to prevent MIs, but there's no assumption you're going to treat MIs with lipid-lowering

agents. So I think it's really quite fair, if you had some idea of mechanisms, to have agents that prevent lupus from getting worse, and I wouldn't discourage their development.

DR. SIMON: Don't get us wrong. It was not an understanding to discourage development. It was an attempt to understand how best to understand its use. We're not wedded to it one way or the other. We're just asking questions to get a clarity point of view about what you all are thinking about.

DR. FIRESTEIN: Two more comments and then we're going to move on to Dr. Witter's presentation. Graciela and Mary Anne.

DR. ALARCON: Yes. I think when you're talking about prevention of flare, you really should expand and say prevention of the damage caused by the disease or the treatment. So really and truly, if you get a drug that not only does prevent you from having an acute exacerbation but also prevents damage, then you really have a winner.

DR. FIRESTEIN: All right. Well, that answers that question, I guess.

Dr. Witter is now going to talk to us about clinical markers.

DR. WITTER: I have in my hand here a form 356H. It's entitled Application to Market a New Drug, Biologic, or An Antibiotic Drug for Human Use. If we

wanted to start filling this out, there are some things that I think we need to understand and that's part of the reason that I'm giving this presentation today.

I might be dating myself, but my mother used to give me cod liver oil and told me it was good, so I held my nose and I took it. So if I see some of you holding your nose, I won't be offended. But I do think we need to get on the same page literally with some of these concepts so that we understand where we are and, more importantly as we've been discussing, where we need to go.

So as you've heard already from Dr. Simon, the FDA approves drugs and biologics -- and I didn't leave devices out to be spiteful -- therapeutics for interstate commerce. The FDA does not regulate medical care, although I think one could certainly argue that if a drug is withdrawn or withheld from the market, that may in fact be regulating medical care. Therefore, we come to this issue we've been talking about as standard of care or off-label use and Dr. Simon again talked about what we're all familiar with which are "approved" drugs and then "off-label" use.

So I'd just like to take a little bit of time to go back to explain how it is that we arrived -- the FDA that is -- where we are with our current thinking, so that we understand the rules and I think we'll better understand

the exceptions to those rules.

So the FDA really started in about 1906 and it was really established really just to respond to problems. There was no specific requirement for testing or approval, but then some things happened in the '30s which woke a lot of people up. Dinitrophenol was being utilized at that time for weight loss and if you go on the internet today, there's still an active amount of discussion for that particular usage for that compound. But it was discovered that about 1 percent of people -- and this was mostly women -- developed cataracts and there were also some deaths associated with its use.

Then in 1937, there was the elixir sulfanilamide disaster. If you read the label -- and actually you can go on the FDA's web site and they have a picture of a bottle-- the label says, "For all conditions in which the hemolytic streptococci appear." That's when you should use it. There were a 107 deaths, many of whom were children, and this was come to understand that diethylene glycol, or essentially the component in antifreeze, was being utilized as a solvent. It seems kind of ridiculous today, but this was how we came to understand that things can have an unwanted effect.

So the Food, Drug and Cosmetic Act was written in 1938, and really what that act did was establish the

requirement for safe therapeutics. Marketing required an NDA, New Drug Application, but it was really a passive process in the sense that if FDA did not object, then it was okay. So for example, FDA at that time could refuse an application if the investigations did not establish safety under the proposed label, the tests showed that they were unsafe or not safe, there was insufficient information to establish safety, or the label was false or misleading.

Jumping ahead now to 1962, the Act was amended to add the requirement -- and the word is "requirement -- for efficacy, and it laid out some mechanisms to conduct clinical studies. The goal was to predict safety and efficacy when the product was to be marketed, and this was accomplished through carrying out, and the importance of words here, "adequate and well-controlled trials."

Jumping ahead to current times now, if we look at the Food, Drug and Cosmetic Act, section 505, which I'll talk about in just a bit more in a second, again we have now arrived. Now I think you understand why we say for a traditional approval, what we're after is that there needs to be substantial evidence of safety and efficacy as the basis for approval. This time now, the approval is on a positive approval. It's not negative. We actively have to be involved in it. And of course, as you've been discussing, this then gives FDA the right to grant

exemptions from this Act to allow IND studies to be conducted for both drugs and biologics.

So we'll come back to my application here. We're going to fill it out later, I hope, in some way or another at some point in time. One of the questions on there says you have to fill in whether this is a section 505(b)(1) and that is what we call a traditional pathway for approval of a new drug, for example. In there, it states the application has to have full reports of investigations to show whether a drug is safe and effective and it has details about components, composition, methods, and controls.

On this same form, you also have to fill out some areas that talk about the Code of Federal Regulations. They are both laws but the Code of Federal Regulations is a way to implement the act, and so that's what we tend to talk about more at these kinds of meetings. So really what the Code of Federal Regulations is, it's a codification of rules published in the Federal Register by the executive department of the Federal Government. It's divided into 50 titles and these titles generally represent broad areas that are subject to federal regulation. The titles are divided into chapters which often beara the name of the issuing agency, and then the chapter are divided into parts and subparts. So we've now come to at least half of my

talk. Subparts.

Title 21 then is really composed of nine volumes with parts. Parts 1 to 1299 comprises the first chapter and that really describes what we do at FDA. Part 1300 to the end which is only a single volume, has really two chapters in it. One describes the Drug Enforcement Agency and how it works and chapter 3 talks about the Office of National Drug Policy. So we'll be focusing only on some of those.

So if we went then to part 314 and we looked at that -- this is an application to market a new drug -- we would see subparts A, B , and all the way down to H, which is what I'll be talking about today.

So now we should understand when we say 21 CFR subparts H and E, how we got there. So subpart H is 314.500, as I've indicated here, and it reads, "Accelerated Approval of New Drugs for Serious or Life-Threatening Illnesses." Just to remind you again, 314 regulations are really new drug regulations. Subpart E is in a different area. It's under what are the IND regulations, so it's 312.80, and it is entitled "Drugs Intended to Treat Life-Threatening and Severely-Debilitating Illnesses."

So I think we should take a second to make sure that we understand, at least in terms of the Code of Federal Regulations, what these definitions are meant to

mean.

Life-threatening is two things. You can substitute lupus, as you see fit, into these definitions as I move forward. Diseases or conditions where the likelihood of death is high, unless the course of the disease is interrupted, and diseases or conditions with potentially fatal outcomes where the endpoint of clinical trial analysis is survival.

Severely debilitating, on the other hand, are diseases or conditions that cause major irreversible morbidity.

So now let's talk about surrogate approval. I'd like to not use the term "accelerated approval" because I think it's a little confusing. So a surrogate approval, subpart H -- now you know where the citation comes from -- reads as follows: "FDA may grant marketing approval for a new drug on the basis of adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit or on the basis of an effect on a clinical endpoint other than survival or irreversible mortality."

There are caveats to subpart H. For example, there is a requirement that the applicant must study the

drug further to verify and describe its clinical benefit where there is uncertainty. So in the one instance where we're utilizing the surrogate to a clinical benefit or the observed clinical benefit to the ultimate outcome of survivability, for example.

These studies that are done post-marketing are expected to be underway and they also are expected to be adequate and well controlled and they must be carried out with due diligence.

Other caveats to pay attention to in subpart H. The FDA may withdraw approval following a hearing if any of the following apply: post-marketing clinical studies that are underway fail to verify the clinical benefit; the applicant fails to perform the required post-marketing study with due diligence. I find this one particularly interesting. The promotional materials are false and misleading. Even in that instance, a part of the agency is looking at this, called DDMAC. Other evidence demonstrates that the drug product is not shown to be safe or effective under its conditions of use. These types of caveats don't apply to traditional approvals.

Subpart E also has its caveats and some of these are really quite interesting, I think. It says that FDA can exercise flexibility in applying standards while preserving safety and effectiveness, much of what we've

been talking about today so far, and it states that these procedures reflect the recognition that physicians and patients are generally willing to accept greater risks of side effects from products that treat life-threatening and severely-debilitating illnesses than they would accept from products to treat less serious illnesses.

Another caveat I think important to bear in mind is that, for example, when the agency is looking at the risk-benefit analysis in the review of a marketing application under subpart E, that it's not necessarily a done deal, that you can get, for example, a non-approvable letter, if it's a drug, or a deficiency letter, if it's a biologic, that may be issued after the review. In other words, there is a decision that has to be made here. And phase IV studies seem to be very likely because the FDA may seek agreement from the sponsor to conduct certain phase IV studies to delineate additional information about the drug's risks, benefits, and optimal use. So that sounds pretty much like subpart H.

So we've come to a part of the talk then that deals with the subparts. Now let's talk about surrogates. Maybe many of you were at the meeting four or so years ago. It was an NIH/FDA-sponsored meeting that really talked about biomarkers and surrogate endpoints. That was very much coming into a lot of people's radar screens at

that point in time, and this meeting was a very interesting meeting. We talked about at that point definitions, conceptual models and possible relationships. So I thought I would just go over some of those for a bit because it might be useful for today's and tomorrow's discussions.

The conceptual models that were really talked about at that time were that biomarkers included measurements considered directly related to clinical outcomes but are not the outcomes themselves. We've heard some of that discussion already today. Biomarkers can evaluate the safety or efficacy or potentially both of therapeutic interventions, and some biomarkers may achieve the status of a surrogate endpoint in a clinical trial, but at that time, it was thought to be difficult because diseases are generally very complex and single markers have their limitations.

Some of the relationships that were discussed at that point were that a biomarker, for example, may be of no value as a surrogate marker and, for example, the intervention may affect the disease and not the marker at all. It was talked about that biomarkers may measure an unfavorable outcome, and I'll talk about an example of that in a bit. It may be that a biomarker has the partial value and that the intervention's positives and negatives are not fully measured, and this may be where most current

surrogate endpoints are today, or it may be that the biomarker is in fact an ideal surrogate endpoint which would be what would be desired.

So biomarkers in an SLE may, for example, be utilized then in exploratory studies. They may help identify or prioritize new therapies. They may help to assess safety. They may help to compare therapies. They may help patients and doctors to select and monitor therapies, and if they're good, they may then function to help assess efficacy, particularly as a surrogate.

So let's talk about surrogates for a second and make sure that we are again understanding what the definition is. A surrogate endpoint of a clinical trial -- this has been described -- is the laboratory measurement or physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions, or survives. And I'd like to stress that again. When you're looking to be approved without any caveats, when you're looking for a clinically meaningful endpoint, that's what this is, is that it has to describe how a patient feels, functions, or survives.

Changes induced by therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint. The surrogate endpoint concept is only valid if the effect on the surrogate leads to a

clinical benefit.

So as Lee talked about a bit earlier, then the distinction then between surrogates versus biomarkers is that surrogate endpoints are candidates for drug approval and biomarkers do not have the same regulatory implication and some surrogates may be biomarkers but not all biomarkers are surrogates.

I just have a little slide here kind of showing this in a picture form, cartoon form, and I think it's re-illustrating the fact that there are a variety of ways for a biomarker to become a surrogate marker and a surrogate marker to become a clinically meaningful endpoint.

So what is the current status of surrogates? Blood pressure, for example, is one that's utilized. Lipid-lowering we heard just before. Blood sugar. Bone mineral density and HIV load. If you were, for example, to go to the FDA's web site and look under subpart H, there have been since 1992 49 approvals under subpart H, and 50 percent of those, about half, have been for HIV. Another 25 percent have been for oncologic-type indications. So there's not a lot of experience in terms of looking at surrogate endpoints in situations outside of this, so hopefully we'll be able to have some of that discussion today.

Well, what are some of the problems with

surrogates? One of the most worrisome is that they do not always account for adverse event effects which may cancel out part or all of the apparent treatment effect. So, for example, one that's often talked about is the Cardiac Arrhythmia Suppression Trial, or the CAST trial, which was published in the New England Journal back in 1991. The idea going in there, which was agreed to and made sense to everybody, was that it was good to suppress arrhythmias, but in fact what came out of the trial was exactly the opposite in the sense that I have listed here, for example, deaths and cardiac arrests in the placebo group, which was 3.5 percent, and in the active treatment groups, which was 8.3 percent. So in this regard, the problem is that the surrogate marker was -- no pun intended -- dead wrong.

So subparts H and E, then hopefully I've explained, they have some potential advantages. They can do this that we've been describing as an accelerated approval, but they have a potential disadvantage in that you can also have an accelerated withdrawal because there are certain requirements put on a compound if it takes this route.

So let's just finish up and talk about a few potential biomarkers or surrogate markers, starting off with uric acid as a potential example. We all know that serum uric acid is a laboratory measure and that in the

right patient, elevated levels can correlate with gout attacks or tophaceous disease or renal disease.

So the question then becomes, in terms of lowering serum uric acid, are we looking for then decreasing the incidence of what could be argued to be a robust clinical endpoint of end-stage renal disease or are we simply looking for the reduction of stone formation? Are we looking for then to decrease gouty arthritis or simply the size of the tophi? How much is enough? Do we have to come to a certain level there? Do we have to beat placebo, and does it have to be in everyone or just a subset of patients? These are issues which we may want to discuss as we proceed here in SLE.

So let's throw out a for instance. This is a hypothetical example of a surrogate approval. Say that it is proposed by a sponsor that double-stranded DNA, the antibodies against such, are proposed as a surrogate in a trial for lupus, for renal disease in this case, and that they are proposing a responder approach to analysis. Hopefully we've heard enough about responder analysis to understand that it's interesting because it's highly malleable and it can be adaptable to different situations, which makes it appealing.

So they would then have proposed this based upon certain endpoints in phase II and then look at it in

phase III, and this would be then addressing a short-term benefit. So we would propose or we would be discussing, as you've done today, that it seems obvious then that you have to have some kind of benefit from a renal perspective, but then what else do you need that shows that you have clinical benefit? Would one of the quality of life measures that we were discussing earlier today be sufficient to allow it to get on the market with a robust due diligence post-marketing commitment to verify long-term clinical benefit and what would that then be? Preservation of renal function? Some of these will be described after the break.

So I think really what we've been discussing all day and I'm pretty sure what we will be continuing to discuss is that when you talk about risk-benefit, there really are different levels that need to be considered. We at the agency look more at a population level. Those of you out here that are providers, you evaluate it for your patients and then those of you that are patients, you obviously evaluate this from your own terms and what makes it of importance to you. So hopefully we can keep all of these balancing acts in mind as we move forward with our discussions.

DR. FIRESTEIN: Thank you very much. We are now at our next break time. So we will break for 15

minutes, 15, 1-5, and then reconvene.

(Recess.)

DR. FIRESTEIN: The next portion of the meeting is going to be an open public hearing.

As before, we have several individuals who have asked for time and again just to remind those individuals to please state their potential conflicts of interest. I have to read it again? Can we just play it back?

(Laughter.)

DR. FIRESTEIN: I really have to read this again?

MS. TOPPER: You really do, yes.

DR. FIRESTEIN: 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, FDA believes that it is important to understand the context of an individual's presentation.

For this reason, 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 that may include a company or a group's payment of your travel, lodging or other expenses, in connection with your attendance at the meeting. Likewise, FDA encourages you at the beginning of your statement to advise the committee if you do not have 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.

The first speaker is Sandra Raymond.

MS. RAYMOND: Thank you. I don't believe that I have any financial or conflicts of interest.

Let me just say by virtue of the discussion this afternoon, I believe we're moving toward claims that are more realistic in terms of the state of the science, but I would say to you that language is very important and that this document frames the science in very negative terms and those terms in my view are very unattractive to potential sponsors. I understand that there are gaps in the science, but I think the way in which this is framed really is very negative.

In terms of markers, I recall in the original osteoporosis guideline document that bone mass was a surrogate marker, and certainly measuring bone mass was very, very important and was central to clinical trials.

The fact is that the technology at that time had not fully evolved and there was great controversy about the accuracy and the precision of DEXA and other forms of the technology, yet the FDA was very generous, I think, in allowing that technology to be included in the document. It really did allow for the evolving therapies that now exist today, and you know we have very good therapies in osteoporosis. So I would ask the FDA not to set the bar at an impossibly high level.

With respect to the claims, I think that the current claims -- and I know they may be changing -- but when you talk about language, for example, the document up front lays out the science. In fact, it lays it out in glaring detail, and it does raise uncertainty about the disease activity indices, and from what I've heard here and from what I'm told by the experts, we do have indices that are pretty good in this field. They may have some imperfections, but the fact is that they, through clinical experience, have proven to be pretty good indices.

We may have several markers, and I'm not here to tell you whether you have them or not, but certainly clinicians have been using complement and have been using ds-DNA and have been managing patients that way for quite some time. So is it rocket science?

In terms of definition of a flare, I suspect

that there is a definition that's been used out there. It may not be validated. And I don't know how fair it is to ask a sponsor to both validate definitions and indices while conducting their clinical trial. Maybe that is not the role of the sponsor. Maybe it's the role of other agencies.

In the second claim, I think the document lays out unrealistic outcomes, sustained doubling of creatinine, which might be too harsh for patients over that period of time, or progression to end-stage renal disease which, from what I hear, you have to conduct a pretty long trial in order to get the number of patients you need to power a study. So I think that those are pretty unrealistic outcomes. I believe there are five in claim 2 and three in claim 1, if you include quality of life.

The other issue I think that we need to think about is this whole idea of complete clinical response and clinical remission. I hear around the table that there may be some definitions that have been used that seem to work in clinical trials.

I would say about subpart H and E the following, that for companies working in the field today, this is a thin ray of light, but for others, this is a high-risk, high-return strategy unlikely to be so attractive in today's drug development environment. I

could be 100 percent wrong on that, and certainly it did wonders for HIV and if it can help in lupus, that's terrific, but I wonder whether in fact the drug industry will be attracted specifically to that. So that means that claims in the front part of the document need to be spelled out, I think, the way in which you discussed it this afternoon.

Thank you.

DR. FIRESTEIN: Thank you. The next speaker is Linda Nardone.

DR. NARDONE: Thank you for agreeing to allow me to speak. My name is Linda Nardone. I am the Vice President of Clinical and Regulatory Affairs at Elusys Therapeutics, Inc., and Elusys is a biotechnology company that is developing a drug for SLE. And certainly a clear road map -- and you mentioned that before -- is very, very important to a company like us.

As you have heard today, systemic lupus erythematosus is a complex disease and there are a variety of manifestations in different organ systems at different points in an individual's long-time battle, long life battle. It follows that SLE has been a very difficult disease in which to conduct clinical trials. The disease population is certainly one with unmet medical needs, and the paucity of drugs to have even come before the agency

for approval in the last 20 years attests to that.

These differences in patients in expression of disease and in complexity notwithstanding, there is a well-documented body of evidence regarding double-stranded DNA autoantibodies. In fact, you heard data presented today, particularly those by Dr. Buyon and Dr. Strand, and even Dr. Dan Wallace mentioned that in the last public session there is a wealth of data about that.

Three points among others can be made. The first is that double-stranded DNA autoantibodies are the diagnostic hallmark of this disease. The second is that a correlation of double-stranded DNA autoantibodies with kidney pathology and function and even longevity has been certainly demonstrated in animal models, and we haven't talked much about the animal models in this forum. Finally, in the clinical setting, lupus nephritis is established as a major sequelae of the disease and large percentages of patients exhibit double-stranded DNA autoantibodies at some point in that disease. The correlation between the autoantibodies and this debilitating kidney damage continues to be studied.

We therefore urge the agency and this committee to recognize double-stranded DNA autoantibodies as a surrogate marker for clinical benefit in lupus.

We applaud the agency for the current

initiative designed to look at many aspects of drug development for this disease, including biomarkers and surrogate endpoints. The use of surrogate markers will enable the development of important new therapies for the treatment of these patients.

Thank you.

DR. FIRESTEIN: Thank you. Are there any additional comments during this open session?

(No response.)

DR. FIRESTEIN: If not, we can move ahead to the questions. So there are three questions that have been posed, each of which has many parts.

So the first is: would a change in anti-double-stranded DNA antibody level associated with a change in hematuria or proteinuria be considered reasonably likely to predict clinical benefit in treatment of lupus nephritis? That's in combination with would the following outcome measures together be reasonably likely to predict clinical benefit: (a) a change in anti-double-stranded DNA antibody levels, (b) along with some other clinical outcome measures, such as SF-36, et cetera, and then (c) no worsening kidney function over 6 months, and then subsequently be required to show in post-marketing a 3-year study of improvement in renal function? So it's a rather complex question, but I think people get the idea.

So this raises the important question of whether or not there exist surrogate markers versus biomarkers in lupus from a regulatory perspective, and who would like to begin? Jennifer.

DR. ANDERSON: I'd like to ask a question of Dr. Buyon, because in your presentation, it seemed at one point you were saying that the complement was associated -- I'm not sure if it was the complement or the anti-dsDNA -- with both increase and decrease of activity of disease.

DR. BUYON: I'd like to clarify the fact that some of the slides you couldn't see.

So, first of all, I was representing different studies, and I'd like to say that the answer is heterogeneous. To address your particular question, that was in reference to one study by Michelle Petri in one cohort looking at global lupus and what she found in her paper was that increases of anti-DNA doubling by a Crithidia predicted flare, but at the time of flare, there was a concurrent decrease in anti-DNA antibodies. I brought that up because I think that paper is highly quoted, yet most of us in fact have not actually been able to corroborate that.

DR. PISETSKY: Some of the differences allegedly have to do with how frequently you assess anti-DNA.

DR. BUYON: Yes.

DR. PISETSKY: The idea that anti-DNA goes down during disease actually has been around for awhile, and the interpretation is you form immune complexes at that point and it deposits in the tissue and therefore is just not measurable. So they're actually not inconsistent. So you can imagine time when anti-DNA goes up and then there's a separate event. You form immune complexes and it goes down.

I think that when you look at measure of change, I think one question, beyond the issue of what methodology you use, is how frequently you do it because I think you'll get very different answers.

DR. FIRESTEIN: Dr. Cush.

DR. CUSH: I don't treat lab tests. I treat patients, and while biomarkers and surrogate markers may be things I worry about and things upon which I base some of my treatment decisions and how often I'll see the patient, how often I'll do double-stranded DNAs, I do not respond to lab tests alone.

To allow a biomarker or surrogate marker to be the primary endpoint for an indication I think would be wrong. I think to use a biomarker or a surrogate marker as the hallmark for an organ-specific indication might be appropriate, if it was uniformly agreed upon that the

surrogate marker that was being used, 24-hour creatinine clearances or whatever, was felt to be highly predictive of what would happen for that organ outcome. But I would not allow just a double-stranded DNA as my sole outcome and upon that I base approval or give some approval to that.

DR. FIRESTEIN: Dr. Looney.

DR. LOONEY: In the Goulay study of lupus nephritis from the NIH, one thing that was impressed me was that when they looked at people who eventually responded, when they looked at 1 year, where most of them had not responded, you could pick them out using serological markers compared to the people who didn't respond. The change in the anti-double-stranded DNA in the people who didn't respond from initiation of treatment to 1 year was from 320 to 160 units; whereas, in the people who did respond, it went from 160 down to 10, which is essentially normal.

So I think if in fact you're talking about losing your anti-double-stranded DNA completely, I suspect that that would be a pretty good surrogate marker, at least for proliferative lupus nephritis.

DR. FIRESTEIN: Graciela.

DR. ALARCON: To echo the fact that you just cannot look at a marker isolated, it can be a secondary outcome measure but not really the outcome measure of a

trial.

DR. FIRESTEIN: Joan?

DR. MERRILL: I don't think you can treat lupus nephritis by looking at the patient alone. I think that as part of the picture, these laboratory indices are all we have between the first biopsy and whenever your second biopsy is to tell you how you're doing, and it's a conglomeration of things. It's not one, it's not just the antibody to double-stranded DNA. It's the complement. It's the protein. It's the sediment, and it's the renal function. And if you see all of that going in the right direction, you're pretty comfortable as a clinician.

DR. FIRESTEIN: So are you suggesting that an individual one of those components wouldn't be appropriate as a surrogate marker but that a composite index would?

DR. MERRILL: I'm saying that you can't take one of those things. It's just not going to work in enough people, but if you put them all together and make rules -- and the precedent for this, I think, was the original LJP trial. Now, they weren't talking about treatment, they were talking about flare, but they had a nice put-together definition of flare.

Another good example is the CellCept trial. We were really following specific things. There was a crossover point. If we weren't doing well, we were going

to change to the other treatment, and we set rules, but it was a conglomerate rule.

DR. MANZI: I guess I have a question for Dr. Witter. I guess by definition, a surrogate should stand alone as an outcome and that's, I guess, the presumption with the surrogates that you showed us. But what you're suggesting to us is really coupling the double-stranded DNA with other measurements.

My question is, is there precedent for a "surrogate" to be coupled with something else and still be a surrogate?

DR. WITTER: I think that's the question that we're trying to ask with these surrogate questions, is should we be doing that? Whether there's precedent for it, there probably is. I can't think of it off the top of my head, but I think what we're after is getting as much comfort as we can pre-approval so that we don't have to worry about certain issues post-approval.

DR. MANZI: Our response, my guess would be, that we could certainly come up with what might be response in renal disease, but the question that seems to be posed to us is double-stranded DNA a surrogate and would it stand alone, I think is what you're asking us, and yet the way you've posed it here, it's really can you couple it with other factors and come up with a response.

DR. FIRESTEIN: The way the question is written, it almost sounds like you've already decided that anti-double-stranded DNA is not a surrogate marker but it might be in a composite with something like proteinuria and hematuria.

Lee, do you want to address that?

DR. SIMON: Yes. I think that it's hard to hide in these circumstances. Internally, we've had that debate, and there are many people who are uncomfortable within the agency that anti-DNA today can stand alone as a marker, that it would make people feel more comfortable that if you're following anti-DNA, which you measure based on specific therapy compared to your active control, would then be corroborated with a longer-term post-marketing, post-approval phase IV trial, that the way to make you feel more comfortable with that decision was to link it to some other event, one of which might be a health-related quality of life measure, maybe perhaps other disease activity scores/indices, and then also obviously what we've talked about over and over again, that there wasn't worsening in other things, and in particular in this context, that there was not worsening in renal disease at the same time your anti-DNA fell.

So that is why these were lumped together. I would love to hear if you all would be willing to do an

anti-DNA as a surrogate predictive of end-stage renal disease and that you'd have to look at that in a phase IV marketplace with a 3- to 4- year study for end-stage renal disease.

DR. FIRESTEIN: Bevra.

DR. HAHN: I'm actually quite comfortable with this. I think that either the anti-DNA or hypocomplementemia with a clinical marker of short-term benefit is fine. I think it's a fine place to start. We wouldn't want to take just the change in the urinalysis either, I don't think, as indicative of improvement, unless it was sustained for a long time.

So I agree with all the people who have said the combination is reasonably predictive.

When you look into the literature, in general, I did want to make the point that the studies that show the best correlation which is never perfect in humans are the studies where the anti-DNA is done frequently, at a regular interval, independent of what's going on clinically, and it's all done in the same lab by the same method or, better yet, two methods, and they correlate each method independently.

When you do it that way as opposed to taking what comes into the chart from 20 different labs that the HMO is paying to do the anti-DNA that month and you do it

only when you think the patient might be deteriorating, then those don't correlate very well. So I think the way it would be done in a trial regularly, same lab, same technique, that we could hang our hat on a combination like this.

DR. FIRESTEIN: Bevra, if you weren't comfortable with the urinalysis, for instance, in combination, what would be an example of a lab that you would then link to anti-double-stranded DNA?

DR. HAHN: Well, I suggest a hypocomplementemia or a creatinine or a protein-creatinine ratio, something functional as well as immediate.

DR. FIRESTEIN: So I think the creatinine is particularly interesting, at least to me. If we were to do that, then we're back to looking at essentially renal function as the endpoint and we lose the power of a surrogate endpoint to get us around having to do a longer-term study looking at renal function specifically in a disease-oriented or an organ-specific endpoint.

DR. HAHN: I think part of this depends on how fast you think it would change. So I think I could use either one. They both change pretty fast, right, clinically, both the creatinine and the sediment, and the creatinine is a little more reliable in terms of accuracy. That's all.

DR. FIRESTEIN: John, and then David.

DR. HARDIN: I suppose in some ways, anti-DNA is to lupus as cholesterol is to cardiovascular disease. If we were to bring a drug to lower the serum cholesterol to the FDA, would you require a clinical response or would you limit it just to lowering the cholesterol effectively?

DR. FIRESTEIN: Well, I'm not the FDA.

DR. HARDIN: Well, maybe Lee or someone could answer that?

DR. SIMON: Well, initially, before all the enthusiasm and hype and any number of different trials that get very confusing were done, in fact that was required, that lowering of serum cholesterol was a surrogate marker for outcome, and subsequently, there have been trials that have claimed in the right patients and the right circumstances that lowering serum cholesterol has made a difference in clinical outcomes.

Therefore, it is the same route, meaning we're asking for something being reasonably likely based on either epidemiologic studies, which is what happened with cholesterol, and then furthering the drug development, demonstrating in large post-marketing circumstances that that data was corroborated. So it's in fact incredibly analogous.

DR. PISETSKY: I was just going to say that I

think while in many patients, anti-DNA is associated with renal disease manifestations, it's by no means all patients and there are certainly exceptions in both directions of people serologically active, clinically quiescent, and the other way around, whether's that's assay or not.

So if it's to be used as a biomarker or surrogate marker, it has to be very defined in terms of which patient population it's used in, and I think, in addition, there are issues in terms of methodology as to how broad or narrow you wish in terms of which types of anti-DNA you want to include.

But the other question I would sort of bring up is what constitutes a clinically significant change in anti-DNA. I'm quite surprised by seeing these 10 percent changes. When this system was originally described, it was notable for the huge range in anti-DNA. This was an antibody that could see extraordinary levels and went away with therapy, and now we're dealing with 10 percent levels. So I think it's something in between that's going to turn out to be informative.

DR. ILLEI: I think that the combination of hematuria/proteinuria and the anti-double-stranded DNA or some other serologic markers is reasonably likely to predict a clinical response. I'm not sure about double-stranded DNA in itself.

In the last NIH trial, a positive outcome was used as the primary outcome. That was called a response and that included normalization of proteinuria, normal creatinine, and then normal urinary sediment. We did do a follow-up study on those patients and we looked on the long-term outcome between those who were responders or non-responders in that study, and the responders who fulfilled the criteria for the remission did have much better long-term renal outcome than those who were either partial responders or non-responders.

We also did a study on renal flares including these patients and the subset of patients who were treated during the period where double-stranded DNA antibodies were routinely tested. Those who did have positive double-stranded DNA antibodies at the end of the treatment had a significantly higher probability of flaring. So I think including serologic markers in a combination endpoint is useful and it is reasonably likely to predict response.

I think the risk for using double-stranded DNA antibody in itself is that there may be treatments that do have a biologic effect on double-stranded DNA but do not influence other aspects of the kidney disease, and there may be a mixture of patients in trials, some of which may have already had some chronic damage to their kidneys. So I would be cautious in using double-stranded DNA on its own

as a surrogate marker.

DR. FIRESTEIN: Joan, and then Jill.

DR. MERRILL: I would think that you would have to be careful about what sort of trial you're talking about, but if this were a trial where you were entering patients who had antibodies to double-stranded DNA, and if this were a trial where nephritis is what we're talking about, then I think there's plenty of justification for considering anti-double-stranded DNA, plus one or two other markers, and I would say C3 would be a key one as being reasonable beginning steps to show the possibility that this could be an effective medication.

The goal of that would be to shorten the time it would take to get things moving for a drug? I'm not sure I quite understand what the goal is.

DR. SIMON: The goal has been defined by what has been used in the past to use as an endpoint. Remember what Dr. Witter's slide said. Function survives. So organ survival of end-stage renal disease has been classically considered the important clinically oriented outcome. You and I and everyone in the room know to do a clinical trial is impossible for that. So what we've been searching for consistently is something to allow a much shorter trial time to allow then a change to be monitored and measured that would be importantly linked to end-stage renal disease

or perhaps even not as an extreme example, just a 50 percent change in creatinine clearance and maybe that would be good enough under those circumstances.

DR. MERRILL: What I would just suggest under those circumstances, though, is that now there's a concept evolving of induction and maintenance. So the definitions would have to be very clear. Is this induction? Is this maintenance?

See, a lot of what Jill was talking about were studies that were looking at different kinds of flares, and what's interesting is that even some of them seemed, to some extent, to follow with the antibodies to double-stranded DNA. But I think you're going to get more of a connection if you stick to nephritis.

DR. FIRESTEIN: Jill, and then Michael.

DR. BUYON: I would just make two points. One, we have to remember to define the players. We've heard that a lot of times. We have to define what the players are. Two, I think we should take pause really in the estrogen story and that is here is clearly a medication that changes a surrogate marker. It changes cholesterol levels and we know what the data show with regard to the actual clinical benefit.

So I would say we could take the open-minded approach that a drug could change a surrogate marker. For

example, DNA as a stand-alone, but that it's mandate that you have to couple that perhaps in post-marketing with a clinical response. So I really don't have any problem with the concept of accepting a drug that does something to a surrogate marker that has reasonable chance of being something. We've all heard here about DNA antibodies and the association with TPGN, but it would have to be coupled with a clinical improvement in post-marketing.

DR. FIRESTEIN: But has there ever been a study where a therapeutic has lowered anti-double-stranded DNA and not shown efficacy in terms of --

DR. BUYON: I think we don't know that, but the same question could have been raised about estrogen about six years ago. That's what we have to find out.

DR. WALLACE: There was one study and that was the case of apheresis. They developed columns that removed anti-double-stranded DNA but the nephritis did not get better.

DR. WEISMAN: Let's put this in some perspective. In a disease that we already have drugs approved for, rheumatoid arthritis, right now companies and investigators can construct trials enriched with seropositive-only patients with erosive disease and a drug can be shown to eliminate or change the rate of erosive disease and therefore get a claim. But we don't know

whether seronegative rheumatoid arthritis erosions are going to respond the same way.

We've already allowed ourselves to do that and talked about half or two-thirds of the whole rheumatoid population and we've all agreed as members of the advisory committee and the public and everyone else that that's fine. So we have a claim.

So what's happening here? The argument is, are anti-DNA antibodies a sufficient marker for outcome? I think the issue for me is I'm wrestling with that. Is this the same as the erosion? I still feel that if I saw anti-DNA antibodies go away and it was coupled with some other clinical indicator, whether it's proteinuria, as Bevra suggests, or red cells in the urine or a change in creatinine, I'm convinced at this point.

DR. FIRESTEIN: Gary, then John.

DR. HOFFMAN: I pass.

DR. FIRESTEIN: John.

DR. DAVIS: I'm having a hard time with it as well as a stand-alone surrogate marker for a number of reasons, because even in proliferative patients, even in the most severe ones, it doesn't always correlate, and I'm also wrestling with the idea of what titer would be pathologic in my mind and what percent change, as David said, would be significant? If we're going to have it as a

stand-alone marker, what percent change are we going to set as the threshold then? If we're going to approve it now, what percentage would we want, and in the future, if we're going to tie it to other things, we're going to have to make darn well sure that we tie it to things that are temporally related, like complement, because proteinuria is going to take at least 3 to 6 months really to change. So you've got to keep those things in mind.

And I would not use serum creatinine. You're going to have too much damage going on before you're able to detect anything there.

DR. FIRESTEIN: But does it have to be 100 percent predictive? For instance, bone mineral density does not always predict someone who will have a fracture and people with high cholesterol don't always have myocardial infarctions.

DR. DAVIS: I don't know.

DR. FIRESTEIN: Mary Anne.

DR. DOOLEY: I think the way it's written here and we're saying it's reasonably likely to predict clinical benefit from treatment of lupus nephritis, then I think you have to couple the double-stranded DNA antibody with some measure specific to the kidney, whether that be proteinuria or whether that be creatinine.

I think if you look at nephrology literature,

it's true that doubling of serum creatinine is serious, probably a 50 percent loss of kidney function, and we certainly don't want that as a goal or need to demonstrate that to demonstrate the drug is not doing well, but you can certainly look at the log of reciprocal of creatinine or look at kidney function measures in a number of different ways and tie this more directly to the specific organ.

DR. FIRESTEIN: David.

DR. PISETSKY: The only comment I was going to make is that while we always think of anti-DNA as related to nephritis, there is emerging data in other situations that DNA/anti-DNA immune complexes have more widespread activity. I think there was considerable interest in the study presented that quality of life went up, and I think in current evidence you can explain that by sort of cytokine effects and sort of some well-being if you get rid of the component that's leading to the cytokine. So it may be reasonable to tie it to other things when we know more.

DR. FIRESTEIN: Graciela.

DR. ALARCON: I think that if you're going to do a study, why do you have to wait for the post-marketing data to actually show the component? If you're going to actually measure anti-DNA, you should as well measure all the other things that go along with it and then you don't have to wait for 2 more years of data.

DR. FIRESTEIN: Lee.

DR. SIMON: That's a very cogent point, Ciela, and I think that it really raises two other issues that have been brought up to us from clinical investigators who are very interested in lupus nephritis trials.

One issue is how long it takes to see a change like that. So doubling of serum creatinine is obviously not something we want. Serum creatinine has its own problems, although it's easy to measure. So we've been looking for other measures that would predict, one of which would be GFR as determined by creatinine clearances, but then we're told by lupus clinical investigators we can't do those because that requires too much burden to the patient to be able to collect the urine appropriately.

This raises the question that I'm going to ask, which is we can't have it both ways. We want rigorous trial designs, yet we hear from the community that we can't get that, so we have to settle for less useful measures, such as serum creatinine.

So it's trial design length is the reason that we're trying to look at shorter trials, Ciela, and allow a post-marketing period for corroboration of what the predicted result might be, and then, secondly, what is the issue about how difficult it is to do these trials because of these kinds of interventions? Is this really true? Is

it really hard to do a creatinine clearance when in fact that is the best way to measure what we're trying to answer the question about?

DR. FIRESTEIN: Susan, and then Gabor.

DR. MANZI: I just wanted to make one point about this idea that we tend to be perfectionists and we want every individual to fit the profile, and I think you were alluding to this. I think if you teased apart the hypercholesterolemia trials and the lipid-lowering trials, there are many individuals that don't fit the profile, whose cholesterol levels stay high, don't have an event, whose levels go low and have events, but you're looking at population effects. You're not looking at individuals.

We're very much influenced by our individual patients and the variability, and I think if we can step back and say let's not be perfect, but as a population, would that surrogate predict a good outcome and would we be comfortable with trying it?

DR. FIRESTEIN: One other interesting side bar on that is that it may be that the effects on cholesterol, for instance, are totally independent of the long-term beneficial effects of statins and that we were all fooled into thinking that that was the surrogate marker, but that's a whole other discussion.

Gabor, and then Dan.

DR. ILLEI: I pass.

DR. FIRESTEIN: Dan.

DR. WALLACE: First of all, if rheumatoid patients volunteer to get endoscopies all the time with nonsteroidal trials, I don't think it would be that hard to do a creatinine clearance on a lupus patient. I just think that anybody that's motivated to be in a clinical trial would do that, and I just don't think that's a major problem.

I think, also, rather than collecting 24-hour urines and new protein/creatinine ratios are very, very well-validated.

But thirdly, I think in two or three years, we're going to see a new marker, something like one of the urinary cytokines, like urinary IL-6 or urinary MCPs, that is going to be coupled with the anti-DNA and I think we have to be poised to be flexible and jump into some sort of evaluation along those lines.

DR. MERRILL: I actually was going to say exactly what Dan was going to say, and I wish I going to say what Sue said because I agree with her 100 percent.

Having schlepped through so many clinical trials, I do not think getting 24-hour urines is at all a problem. Yes, we lose a few. Yes, a few people forget. But they'll do it. So I don't think that's an issue.

The protein/creatinine ratio, I've been getting them lately, and they don't quite correlate but they go in the same direction at the same time, and I think they're very useful and you wouldn't need the 24-hour urine. So I think all of this is really open to us, but again you could potentially predict who's getting better relatively quickly.

It would be perfect if you had antibodies to double-stranded DNA, sediment, urine protein/creatinine ratio, or 24-hour urine and antibodies to double-stranded DNA. I would be highly comfortable with that. I'd probably be comfortable with less.

DR. FIRESTEIN: But if the creatinine clearance improves, do you still need a surrogate marker? Is that an endpoint in and of itself for renal function?

DR. WALLACE: No. You can improve it just by adding an ACE inhibitor. You can improve it by diet. So that's no.

The other thing is I think we should take the Crithidia assay out of the equation because its levels do not necessarily correlate with true improvement. I think we have to either use the ELISA or the FARR.

DR. FIRESTEIN: Dr. Hahn, you had a comment?

DR. HAHN: I just had a comment about the creatinine clearance, and the issue isn't so much in my

experience in clinical trials with patients not being willing to collect it as that they're so inaccurate. So I once did a trial where we had patients on the CRC and we did two 2-hours and a 24-hour in the same period and the results were all over the map on what the creatinine clearance was. They varied as much as 60 to 80 percent in the same patient in the same 24-hour period under a supervised CRC condition.

So I don't think we should require creatinine clearances if people have a way to do it that is as easy because of their inaccuracy, not because of the inconvenience to patients.

DR. DOOLEY: At least in the clinical trials group that Matt Liang had convened that included both nephrologists, as well as a number of rheumatologists, Crockoft-Gault formula was accepted as a good estimate of creatinine clearance.

DR. FIRESTEIN: So to focus the discussion a little bit, I guess one question we might ask is could anti-double-stranded DNA in and of itself serve as a surrogate marker because that is something that has been discussed?

I don't know. Lee, do you want some sort of formal comment from us on that. No, you do not.

DR. SIMON: I think we've heard what we've

needed to hear about this.

DR. FIRESTEIN: Gary.

DR. HOFFMAN: The thing that I haven't heard, and not being one of the parties to the multi-center lupus trials or the disease activity exercises, is when it comes to surrogate markers, I'm not sure that a surrogate marker and a single clinical marker, say, in lupus nephritis is better in terms of predictive value than using a surrogate marker, whichever one you choose, and the composite scores from the disease activity indices.

Does the disease activity index, if one is looking at lupus nephritis complement the renal outcomes better than in fact antibodies to double-stranded DNA?

Sue, you've been involved.

DR. MANZI: I'm certainly not the nephritis person here, so I'll defer, but I mean I think our understanding was that the disease activity indices are not as good or as sensitive as measuring change in renal disease which is exactly why this conference was convened to look at renal outcomes specifically because I don't think the indices can tease out renal change as well as they can global effect, but please comment if that's not true.

DR. PISETSKY: BILAG can.

(Laughter.)

DR. FIRESTEIN: BILAG. Which one exactly?

DR. MERRILL: I was just going to say that one of the problems with the SLEDAI is that you get too many points for different parts of nephritis, I think. Would you agree with that, Jill?

DR. BUYON: There's no question, the SLEDAI is definitely a problem in that regard because the point scale has a lot of redundancies, and it's really not clear enough and you have to have a lot of guidance. For example, if red cells can stand alone, do they have to have concomitant proteinuria? That particular instrument needs major guidelines.

DR. MERRILL: And I think the SLAM is actually okay for nephritis and I hate to say this but the BILAG works.

DR. FIRESTEIN: Lee, are there other issues in question 1 that you want us to cover with regard to combinations or not? It seems to me we've covered most of this ground.

DR. SIMON: I think that you've covered most of the things that we are interested in.

DR. FIRESTEIN: Is there anything that you're not interested in that we should cover?

(Laughter.)

DR. FIRESTEIN: That appears in the transcript

and it'll look just foolish as will this. Strike that from the record, please.

Would time to resolution of hematuria and/or casts in the context of proteinuria be considered as evidence of efficacy for lupus nephritis? This is a variation on a theme from what we've already discussed.

Gary.

DR. HOFFMAN: My concerns about that come from other than lupus nephritis, although I've had some experience with that, but the different types of glomerular nephritis that you see with vasculitides. I can tell you that if there's been significant delay before intervention, there's enough glomerular basement membrane injury, so that you can continue to see significant proteinuria, red cells and red cell casts, even a year later with a stable creatinine once effective treatment has been implemented. But if there hasn't been significant delay and there hasn't been irreversible damage, you might in fact see reversibility within a matter of a few months. I think in part it depends on what your starting point is for intervention.

DR. WALLACE: As good as hematuria's disappearance is, it's still a very bad marker for a clinical trial. First of all, 90 percent with lupus nephritis are women and if menses interferes, that's a

major, major problem.

The second is to look for casts and hematuria, unless you have a trained observer, if you're going to send it to Indianapolis or something and it's going to be frozen, it's not going to be reliable, unless it's looked at fresh by somebody who's really good, and it's really not going to be overly practical.

DR. FIRESTEIN: Yes, I agree. Now that urinalyses are essentially no longer done by the house staff or the medical students or anybody else except by central labs evaluating for casts makes it extremely difficult.

DR. SCHIFFENBAUER: I think the thrust of the question, though, is if we can have a trial that's relatively short-term, maybe we can afford to hire someone to do that specifically and make it a feasible outcome. I think that's the question. That's really behind the question.

DR. FIRESTEIN: Although would that address some of the issues that were raised by Dan with regard to menses and other confounding factors?

Mary Anne had a comment.

DR. DOOLEY: I actually do spin and look at urines every week in clinic from the unusual position of seeing lupus patients with nephrologists. My concern is

actually in the opposite direction, which is that you'll see patients' urinary sediment improve with steroids alone, and yet if you go and look at the subsequent biopsy, you see quite active disease, and so a teaching point for many of our nephrology fellows is the patient who presents with a flare is treated initially with prednisone while being referred to the nephrology clinic. They get to the clinic. Much of their hematuria is resolved or they may no longer have casts. Half of the people that we biopsy have creatinines in a normal range, and yet you see very active diffuse proliferative nephritis on biopsy. So my concern is in the opposite direction which is you can mask urinary sediment activity with steroids and yet obviously, as the NIH has shown, not affect long-term renal function.

DR. FIRESTEIN: Lee.

DR. SIMON: So with this discussion, could we refocus back to the first part of the question and let's not just use hematuria and casts. Let's use response Y time to resolution. Is that an important way to design a clinical trial? Time to the event is one way to think about that. So although we'll talk about trial design tomorrow morning, could you comment about the first part which is could you use time to resolution and that that time to resolution, given a disease that waxes and wanes spontaneously, as an outcome, whatever the outcome is?

DR. DOOLEY: I think using time to resolution of abnormality in renal disease would be an excellent outcome because the longer the inflammation is occurring, the more risk you're taking of damage that won't be reparable, and particularly since what we're talking about is in most of our therapies, we're trying to suppress the immune system and the immune response to prevent scarring. So if you're looking at agents which will be immunosuppressive, then shortening the time to response ought to minimize the risk of scarring. So I think it would be an ideal endpoint.

DR. FIRESTEIN: Graciela.

DR. ALARCON: Time to resolution would be fine, but you have to actually be sure that the manifestation actually is on remission or is resolved over time. So you have to measure that several months after to be sure that you really have achieved it, that in a disease that waxes and wanes, it is not just one time point.

DR. HAHN: Yes. I'm pretty uncomfortable with this one actually in terms of how short-term it could be, and I see what's disappeared here is something that we discussed at the Biomarkers meeting which is repeat renal biopsy. So I don't even know if that's a better marker. It sounds good. If you at 6 months showed that group A had less renal tubular damage and scarring than group B, have

you achieved your endpoint, but frankly I'm more comfortable with that than I am with whether you've changed what's on spot urinalyses over a period of time. It's so variable.

DR. FIRESTEIN: David, and then Richard.

DR. PISETSKY: I was just going to say the other meaningful thing to me is prevent progression, and if all you're looking for is resolution that may prevent you from seeing an important benefit.

DR. FIRESTEIN: I was going to say this is analogous to again rheumatoid arthritis studies where you have a chronic disease and you're looking at a very short-term outcome, whether or not that's going to have an impact on the true natural history of the disease.

Jeff, did you have a comment? And then Richard.

DR. SIEGEL: We've had concern raised by a number of members of the committee about using casts or hematuria alone. What about using a more comprehensive guide to renal remission? I think the NIH definition uses an active sediment returning to inactive, plus a return of the creatinine to normal and loss of proteinuria.

Would something that measured multi-parameters be more reliable?

DR. PISETSKY: To a certain extent, there are

some data available because the original NIH trials reported results almost from weeks after the onset. I mean, they go back into the '60s. Unfortunately, it took a real long time to see a benefit, but if you go back to those numbers, you could see the 6-month follow-up, the 1-year follow-up, and a few-week follow-up. It takes awhile to see these changes.

DR. FIRESTEIN: Richard.

DR. LOONEY: The one practical matter in designing a trial is if you don't take time to resolution, you have to pick a time when you're going to look at your response and looking at the different nephritis trials, when you see resolution is so variable in those trials, I think it becomes very difficult to pick a single time that you're going to use for your primary outcome. So to be able to use time to resolution which would allow you to look at a number of different time points would be a big advantage.

I would like to second the idea that a renal biopsy as an outcome would be very useful and probably could be done as early as 6 months, but both of these things I think really fall in the area of surrogate markers and you would have to have some kind of long-term follow-up to document that they were actually accurate.

DR. FIRESTEIN: Jack.

DR. CUSH: I think the time to resolution trial answers the question of acute therapy. This would be an acute indication for active disease. I think that certainly might be a means of getting accelerated approval using H&E as Jim outlined for us, but I think as everybody said, we're more concerned about the long run. But again, for acute therapy, it might be the way to go against an active control.

To answer Jeff's question, I do think that the NIH definition of response might be fine, but again, reliance on RBC casts is fraught with difficulty because of the inaccuracies in their measurement, even in good labs. I agree with Mary Anne Dooley, at the time of clinic which is not done because of CLIA, then why do it?

DR. FIRESTEIN: Mary Anne, and then Joan.

DR. DOOLEY: I think it would also need to be hypothesis-driven. If the drug under consideration is to treat inflammation, then looking at a relatively short time period and looking at repeat renal biopsy at 6 months would be reasonable. But if what you're trying to do is prevent progression, then you're talking about a much longer trial and that would be either time at remission or looking at a biopsy 2 years down the line. So in some respects, it would need to be hypothesis-driven, based upon the proposed action of your drug.

DR. MERRILL: I would suggest that trying to imagine all the different possible mechanisms of action of some of these new biologics coming down the pike, that you would add to your renal standard marker a marker that the drug had its biological effect and that might be related to, somewhere down the line, getting rid of DNA antibodies. That would be the first thing.

The second thing I want to just cement back is this idea that there may be a difference between what's necessary for induction and what's necessary for maintenance. The Europeans certainly believe this, that you don't need to use quite as toxic a medication for maintenance as you do for induction, and it could end up being that we would want to switch drugs at some point so that you might have different requirements for a drug to induce and that might be a more short-term marker than you would have to give an approval for induction and maintenance which is where your long-term going back and again nephritis comes in.

DR. FIRESTEIN: Wendy.

MS. McBRIAR: Just from a consumer point of view, if we can figure out a way to measure by lab tests rather than biopsy, I think that would be a positive thing for patients, not only the costs involved in doing it, but just the possible difficulties that could happen with

biopsies.

DR. SIEGEL: In that regard, a number of committee members have asked for biopsies either at 6 months, if I understand it, or at 2 years to corroborate that the other findings are accurate. We've had some push-back from sponsors who have told us that their investigators were unable to get a repeat biopsy through their IRB if the urine sediment was normal and there was no proteinuria and so on.

Could those of you who thought that a repeat biopsy was necessary comment on whether you'd still recommend it in the presence of absolutely normal function and sediment?

DR. LOONEY: I think the studies are probably not going to be able to be done out in the real world if everyone is required to have a repeat biopsy, but I think it will be possible to do that on a subset of patients. I think it would be a corroborative evidence rather than a primary outcome. But I think that it would be good to get repeat biopsies on people with a range of different responses because you would like to verify that people who have had a complete renal response actually have the kind of biopsy that you would predict when they do that.

DR. HAHN: Yes. I brought it up, and I also brought it up at the biomarkers meeting, that I don't think

the IRB will permit a renal biopsy in somebody who's otherwise doing well. I think that is a problem.

I also think that they're getting safer and with the new intravenous approach to renal biopsies, I've been happy with that in terms of really low, low, low morbidity. So I think maybe we keep in mind that the technology for that is also advancing and we might want to leave it as an open question.

I think it might be the best primary outcome measure actually, the most predictive, but I don't think it's practical.

DR. FIRESTEIN: Jim.

DR. WILLIAMS: I vice chair an IRB and I think that the major reason is the education of the IRB. If you point out that renal function being normal doesn't necessarily imply that there's no active disease. A lot of the times the decisions are being made by non-rheumatologists and non-nephrologists, and it may take better explanation, but I think with explanation, you could get it through an IRB.

DR. FIRESTEIN: Well, patient recruitment also becomes an issue.

Mary Anne.

DR. DOOLEY: I think I would be a very strong proponent of rebiopsy, and I would also say that we've

actually looked and surveyed the group of nephrologists that we work with about that issue, about the willingness to adopt a study to include a rebiopsy even in folks who appear to be doing well. And over a 10-year period of time, that group has now decided that it is quite reasonable and ethical to rebiopsy. The reason is that when we looked at our group of patients -- I'm from North Carolina -- predominantly African American, and we have very active patients, such that although most of our patients enter with normal serum creatinines, by the end of 5 years, 40 percent of our African American patients were on dialysis. So they didn't double their serum creatinine, they required renal replacement therapy.

When we looked carefully and we identified all of the clinical, histopathologic, serologic, and medication data that was present at the time of the initial renal biopsy and then the patients received the standard Cytoxan therapy, we could not pick out those patients who went to dialysis in any of those aspects from those who did well. So there was no data available to us at the beginning of therapy as to who would progress to end-stage renal disease.

We included 8 patients who actually required dialysis at the time of institution of Cytoxan. 5 of those patients came off and remain off dialysis, but a suitable

number came in with normal creatinines and required dialysis within 6 months.

So I would suggest, also, in looking at the patients as they go from monthly IV Cytoxan to quarterly Cytoxan, we also see a significant portion who look like they are staying in remission but who rapidly flare upon completion of their quarterly doses of Cytoxan. When we come back to rebiopsy them, we see significant chronic change, suggesting that even though clinically they appear to be in remission, that they had grumbling, ongoing activity that was leading to further damage. So I think repeat renal biopsy study would certainly help us to understand better what's going on during that time period.

DR. FIRESTEIN: Dr. Simon.

DR. SIMON: So in that context, people were talking about biopsies, people were talking about using them as a surrogate marker. I would presume you're not talking about it in the context of the WHO classification. I presume that the changes that you're talking about are the clinical activity inflammatory changes versus sclerotic changes. That's my first question. I have a second question after that.

DR. DOOLEY: Well, the first question about the change in WHO class -- and we certainly know that patients do change among the classes. I think it's important to

describe that. We will see patients who go from proliferative to membranous during a course of therapy, and certainly you can see a major difference in long-term renal survival in patients who have predominantly membranous disease compared to those who have proliferative disease.

So if patients are continuing to have proteinuria but are predominantly membranous, I think your impetus to treat with increased cytotoxic therapy is not as great. You may want to maximize ACE or ARB therapy or choose other means to decrease proteinuria.

Looking at activity and chronicity indices are very important, and I think looking and seeing that somebody has little activity but high chronicity may cause you to think that perhaps the damage is done and you don't want to subject that patient to further immunosuppressive therapy.

So I think there's information in both descriptors.

DR. SIMON: Thank you.

And the second issue is although we want to be as flexible and as open in a document as possible as relates to induction versus maintenance therapy, at the same time, if it's written in too structured a way related to that, it might preclude the newest development of therapy that would not require induction and maintenance

therapy. So that's a little tension there that we have to be careful about, not to suggest that at the present state of the art, that is in fact what we're working with.

Tomorrow we'll discuss this somewhat more, about what we know or think we know about the utility of induction therapy with cyclophosphamide and what it really tells us, if anything, about how we should be approaching this.

So thank you.

DR. FIRESTEIN: David, and then we're going to move on to the third question.

DR. PISETSKY: I was just going to say in the experience of the other North Carolina institution, if you have high chronicity, the outcome with Cytoxan is not favorable. It's predictive of poor outcome. So I'd just clarify that. I think one should bear in mind when you talk about trials that therapies presumably can treat activity but they don't yet treat chronicity and there should be some consideration as to what kind of patients enter these trials because if they have too much burden of disease, you don't see benefits.

DR. FIRESTEIN: Well, that actually moves us into the last question. It seems to me that it's revisiting the question of using one of these laboratory biomarkers in combination with a non-traditional domain for

approval, such as quality of life indicators.

Does anybody want to comment on that? For instance, anti-double-stranded DNA plus quality of life as an approvable endpoint.

DR. BUYON: I don't see how that could be approved without having some type of biopsy or other objective evidence, and I would strongly say you could not do that without linking the other.

I would also comment that something Mary Anne said was very disquieting, that if the sediment alone is not predictive and you just told me you're at the level of teaching that to renal fellows, then I don't see in a way how we can almost get away without biopsying to really sit back on our laurels and say a medication works or not. That may not be the first thing you have to do to approve, but just as you were saying before, it would be coupled by you get the claim and then you have to follow it by a phase IV trial. I don't see how we can get around that.

DR. FIRESTEIN: Dr. Looney.

DR. LOONEY: I guess this sounds like it's in the setting of renal nephritis, and if that's true, then I don't see how you could just have -- I mean quality of life is important. I think it's more important in non-organ-threatening diseases, but I think for nephritis, it's not as important an endpoint.

DR. FIRESTEIN: Joan.

DR. MERRILL: Yes. Mary Anne, can you clarify that? That's in the setting of an acute flare, isn't it? So that they came to your clinic, they got maybe a week or 2 or 3 weeks of steroids and now the sediment is clear and then within another week or 2, they get a biopsy. I'm not that surprised to see that, and I don't think it doesn't mean that they would be getting better. I think probably what's going on deep in the kidney is going to lag a little behind what's coming pouring out.

So I'm not sure I'm as concerned about that data as I am about your other data with your patients that went on dialysis.

DR. DOOLEY: Yes, that's correct. But then, even not that long ago, I think that as rheumatologists, we were taught that the first step in treating a patient who looked to have a flare of nephritis was to put them on high-dose corticosteroids and then reassess within a 1-month period of time. Then we expected, if we saw improved renal sediment, that we will have made a therapeutic impact, the concern being that you may actually mask the activity of the urinary sediment but not necessarily have resolved underlying nephritis.

Now, if the patient's serum creatinine remains normal, proteinuria is resolving, then I think you're in a

much safer ball park, but institution of steroids as part of an acute flare doesn't mean that you've treated the nephritis just because you've changed the urinary sediment.

DR. FIRESTEIN: Dr. Hahn, did you have a comment?

DR. HAHN: Yes. I was responding to something you said, and that is, that we have to remember that the nephrologists are coming up with experimental molecules that will prevent fibrosis and scarring in kidneys. So we want to keep in mind that we aren't looking just at what we currently think of for suppressing active lupus, but I'm hoping they'll be coming into the lupus field with their strategies to prevent damage, whether or not they probably have to be added to ours, and they might be the maintenance. So you might induce with ours and maintain with anti-scarring and that unfortunately brings me back to the biopsy. I just wanted us to remember that.

DR. FIRESTEIN: Yes. I think I'm going to add my name to the list of people that are uncomfortable with a biomarker like anti-double-stranded DNA and quality of life type of an outcome.

With regard to biopsies, I think that that would be an excellent choice, except for two potential issues. One is the issue of sampling error that can arise and it depends on how many glomeruli you can get in your

sample in order to get an adequate representation.

Then I also have some concern that we would have difficulty recruiting into a study like that, especially for the second biopsy. I have no doubt that the first biopsy would be doable. It's the second one, even if the IRB approved. Our experience has been that people are not anxious to be biopsied again.

Gary, you had a comment.

DR. HOFFMAN: I think everybody is on the same page as Mary Anne in suggesting that the first biopsy is always illuminating and sometimes actually very surprising, but when you look at patients who have responded unequivocally to treatment, whose urine sediment appears to be improving, whose creatinine is going down, perhaps is normal, it's very difficult in the context of routine patient care to tell that patient you would like to get a renal biopsy.

So I think studies can be designed where patients other than that type, where there are several markers, clinical or otherwise, suggesting continuing active disease, markers that may influence a change in therapy are present, where you could have a branch point in your study design where you could say that patient in the context of even routine patient care might be recommended to have a biopsy, to then be able to change treatment, and

in that way, you can get the data that I think other people are interested in.

DR. MANZI: I would just caution that what we feel comfortable with in patient care may be very different as to what we think is appropriate for a clinical trial. I do agree it's about education and I would be curious to hear Wendy's response. If this were a surrogate marker that could accelerate drug approval and this was a 6-month rebiopsy, I think you may have a very different response from patients willing to participate. I think it just depends on how important drug approval is to them. But I'm sure we've never approached them with that particular surrogate outcome, and maybe Wendy is in a better position to answer that.

MS. McBRIAR: I feel uncomfortable speaking for all lupus patients here, but certainly I think if there's a clear, defined benefit that may be shown using the biopsy that would give us a potential drug approval, I think most patients probably would go along with it.

Clearly, there has not been much in the way of good therapy for new medications for people with lupus, and I think that's a real important piece and there certainly have been plenty of people here today that have said we need to do something and so if we can give them a clear idea of the benefit they might receive from participating

in that, that certainly should help.

DR. FIRESTEIN: Dr. Simon, are you going to summarize for us?

DR. SIMON: No. I'm going to ask a question, if you don't mind. I don't want to parse, but given your last comments, Gary, about the idea that you would be uncomfortable with the anti-DNA associated with perhaps an HRQOL or something as a sole outcome to predict longer-term effects, may I ask the question?

Alternatively, I heard earlier that it's possible that people would be more comfortable with an anti-DNA and some urinary marker of inflammation that had been followed which perhaps would be something related to creatinine clearance or iothalamate along with an active urinary sediment, and if that was then correlated along with an HRQOL, would that significant change be enough, where BILAG doesn't worsen, to warrant at 6 months an approval with a commitment to prove over 3 years a change in organ survival?

DR. WALLACE: I think it would, but I just want to caution that at least a third of my nephritis patients feel fine. How are you? I'm okay. Their HRQOL is not going to change.

DR. BUYON: I want to really second that motion because unlike the extra-renal parameters, at least I would

totally agree with you, that's our biggest difficulty, is trying to convince young women to take Cytoxan when they feel fine and we tell them their creatinines are deteriorating. This is very different than arthritis or skin disease which is apparent to them as serositis. Renal disease is very often a silent killer, except that your ankles are a little swollen. So I totally agree and would not want the health quality and anti-DNA alone without some follow-up.

DR. ILOWITE: It seems to me that when you look at the other surrogates that have been approved, they all seem to reflect long-term accumulated consequence, like hemoglobin A1c, bone mineral density, HIV load, and we're not really there yet with DNA antibodies, unless we're creative about area under the curve of DNA antibodies and over a long period of time show that that affects outcome.

So that, I think whatever biomarker we choose, it has to be linked or linkable to evidence of accumulated damage, either on a biopsy, or if the creatinine clearance nuclear medicine scan is sophisticated enough, perhaps that's sufficient.

DR. ANDERSON: I'd just like to make a comment about the patients feeling fine. I think that health-related quality of life measures like the SF-36 are more sophisticated than just asking patients how do you feel and

their saying fine or bad. They do cover more domains than that.

Also, hearing that statement about that kind of measure makes me think that perhaps there aren't any -- I actually want to ask a question. Are there any long-term observational studies in lupus where health-related quality of life has been measured fairly early on, along with some biomarkers, where you do have long-term outcomes on patients as a function of those things measured early on? Because if you do, then this will give you some help in deciding whether these things are really useful.

DR. MERRILL: There are some ongoing studies. The SLICC cohort for atherosclerosis is taking patients with a new diagnosis of lupus and they're getting these done. I think there have been correlations between that and some of these disease activity indices, but I can't remember how to quote them off the top of my head.

Lee, to answer your question, I think that I wouldn't want to require health-related quality of life to improve for a nephritis drug. I sure would like to see what it did because it looks like there might be some very interesting stuff there. I wouldn't want to require that and I wouldn't want to require that the BILAG not get worse. I mean, if this is a medication that's aimed at the kidney, I guess I'd only want to see kidney parameters,

whatever seems to be enough.

For me, I think at this point, antibodies to double-stranded DNA, some measure of creatinine clearance or urine creatinine ratio, something like that, and complement would be plenty for me.

DR. FIRESTEIN: Graciela, then Jack, and then David.

DR. ALARCON: Just a comment about the SF-36 over time in our cohort, which is now about 520 patients. Over time, the best predictor was actually the baseline SF-36. So how bad the patients were at the beginning is what predicts how bad they were at the end in terms of quality of life, and we have not been able to correlate the SF-36 with any of the serological markers.

DR. PISETSKY: I was just going to say as a cautionary note here, certainly from animal models, you can have interventions that help kidney disease that don't change anti-DNA. You just prevent their deposition or the inflammation secondary. So I really wouldn't link these too closely.

DR. CUSH: My question was to the FDA with regard to this post-marketing commitment to verify long-term clinical benefit. Do you have any idea of how you would actually define that? Would that just be an open-label follow-up of that 140-patient 6-month blinded trial

and then follow them over time or would you actually want that expanded in the post-marketing era to a registry? I mean, would there be mandatory data collection to look at these outcomes?

I'm sort of concerned. I don't have a problem giving expedited approval for a life-threatening organ-specific indication based on some of the things we talked about, but I do have concerns about how that would be followed up longitudinally and then acted upon.

DR. SIEGEL: I can't really comment on how it would be applied in this particular situation, but in terms of other instances of accelerated approval, I think there are a variety of different post-marketing studies that are done. In many cases, it requires a randomized, controlled trial showing a clinical benefit afterwards, but in other cases, I think in oncology trials, the idea is to show that the benefit in terms of remission has a benefit in terms of survival and that would not be a separate randomized trial.

DR. FIRESTEIN: There's another comment from Marc. Did you want to say something?

DR. WALTON: Marc Walton in Office of Drug Evaluation VI.

Only to follow up on what Jeff has said, that the verification studies, the design is not in any particular way mandated in a global sense. However, it is

meant that the verification studies do obtain rigorous evidence of clinical benefit, and what design might be necessary is certainly going to vary from disease entity to disease entity.

DR. FIRESTEIN: Betty.

DR. DIAMOND: I just want to say as we talk about anti-DNA antibodies or complement or whatever as one of the biomarkers or surrogate markers even used in composite with something else, I think we should be careful about making the tacit assumption that any degree of decrement in antibody titer or increment in complement is associated with improvement.

There may be real threshold effects and you have to reach a certain decrement in titer, and in fact, while that's not been studied quite that way, if you go back and look at what data there are, you really have to normalize your titer. You don't need to reduce it by 10 percent, 20 percent, 30 percent. You really need to normalize and so I think we need to be careful.

I would certainly agree that it can be used as a marker, but I don't think that it can be used just as a statistically significant difference between two populations because that doesn't have a predictive effect that we know of.

DR. FIRESTEIN: Dan, did you have one comment?

DR. WALLACE: Vibeke wanted to be recognized. She has a lot of experience with quality of life indices, and I know she wanted to make a comment, if you would allow it.

DR. FIRESTEIN: I'm sorry. This is for the panel members only.

DR. WALLACE: Oh, okay.

DR. FIRESTEIN: Dr. Simon, would you like to summarize? Because I don't want to.

DR. SIMON: Well, it seems that we have looked at this from multiple different directions, and it seems that I have heard and we have heard that the community at this table believes that there is utility in certain measures, that that composite measure of outcome in nephritis, which was the majority of the time we spent talking, could be several different measures that each look at different aspects of the clinical scenario, and that that might be a useful way to study a patient over time.

I discerned a lack of comfort in applying that in the context of a surrogate outcome, but that if something just came along that showed clear change and it would have to be going to 0 in activity, that it would not be just a statistically significant percentage alteration, that that might be very important.

It does seem that at this stage of the game,

early marker development or surrogate marker development is still in development and that many of the people around the table didn't feel comfortable with some of the proposals that we did as straw men. At the same time, people raised the question of the utility of kidney biopsy and that perhaps that might be revisited as something that is an important outcome that would predict renal survival.

I also heard things like changes in anti-DNA levels would not be great as a measure of other aspects of systemic lupus besides nephritis. Perhaps there was even less enthusiasm about that as a measure for other components, and perhaps there are other measures out there that we did not talk about that would be useful in the context of other manifestations of the disease.

Is that fair?

DR. FIRESTEIN: I think that is a reasonable facsimile of the discussion.

Are there any other questions or comments at this point?

(No response.)

DR. FIRESTEIN: In that case, today's session is officially adjourned. Thank you.

(Whereupon, at 4:07 p.m., the committee was recessed, to reconvene at 8:00 a.m., Thursday, September 30, 2003.)