FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

ARTHRITIS ADVISORY COMMITTEE (AAC) MEETING

Thursday, July 25, 2019

8:29 a.m. to 4:30 p.m.

FDA White Oak Campus

White Oak Conference Center

Building 31, The Great Room

10903 New Hampshire Avenue

Silver Spring, Maryland

Meeting Roster

DESIGNATED FEDERAL OFFICER (Non-Voting)

Yinghua Wang, PharmD, MPH, RAC

Division of Advisory Committee and Consultant Management

Office of Executive Programs, CDER, FDA

ARTHRITIS ADVISORY COMMITTEE MEMBERS (Voting)

Mara L. Becker, MD, MSCE

Associate Professor of Pediatrics

Division of Rheumatology

Duke Children's Hospital and Health Center

Duke Clinical Research Institute

Durham, North Carolina

Jeffrey Curtis, MD, MS, MPH

Marguerite Jones Harbert – Gene Ball Endowed Professor of Medicine

Division of Clinical Immunology & Rheumatology

University of Alabama at Birmingham

Birmingham, Alabama

Jennifer Horonjeff, PhD

(Consumer Representative)

Founder and Patient Advocate

Savvy Cooperative

Patient Outcome and Quality Consultant

Division of Rheumatology

Department of Medicine

Columbia University Medical Center

Sunnyside, New York

Martha C. Nason, PhD

Mathematical Statistician

Division of Clinical Research

National Institute of Allergy and Infectious Diseases

National Institutes of Health

Rockville, Maryland

Alyce M. Oliver, MD, PhD

Joseph P. Bailey MD Chair in Rheumatology

Professor of Medicine

Medical College of Georgia at Augusta University

Augusta, Georgia

J. Steuart Richards, MD

Chief

Division of Rheumatology

Veterans Affairs Pittsburgh Healthcare System

Clinical Associate Professor of Medicine

University of Pittsburgh

Pittsburgh, Pennsylvania

Daniel H. Solomon, MD, MPH

(Chairperson)

Professor of Medicine

Matthew H. Liang Distinguished Chair

Harvard Medical School

Chief, Section of Clinical Sciences

Division of Rheumatology

Division of Pharmacoepidemiology

Brigham and Women's Hospital

Boston, Massachusetts

TEMPORARY MEMBERS (Voting)

William Calhoun, MD, FACP, FCCP, FAAAAI, FACAAI

Nelda C and HJ Stark Distinguished Chair in

Internal Medicine

Professor and Vice Chair for Research

Department of Internal Medicine

University of Texas Medical Branch

Galveston, Texas

Brian Garibaldi, MD, MEHP

Director, Johns Hopkins Biocontainment Unit

Associate Professor of Medicine and Physiology

Johns Hopkins University School of Medicine

Baltimore, Maryland

Nancy Geller, PhD

Director, Office of Biostatistics Research

National Heart, Lung and Blood Institute

National Institutes of Health (NIH)

Bethesda, Maryland

Todd Gilligan, MA

(Patient Representative)

Dubuque, Iowa

James Katz, MD

Senior Research Physician

Director, Rheumatology Fellowship and Training Branch

National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH

Bethesda, Maryland

Gail Kerr, MD, FRCP(Edin)

Chief, Rheumatology

Washington DC Veterans Affairs Medical Center and Howard University Hospital

Professor of Medicine

Georgetown ad Howard University Hospitals

Washington, District of Columbia

Susanne May, PhD

Professor of Biostatistics

Director, University of Washington Clinical Trials Center

University of Washington

Seattle, Washington

Carrie Redlich, MD, MPH

Professor of Medicine, Pulmonary Section &
Occupational and Environmental Medicine
Director, Yale Occupational and Environmental
Medicine Program
Yale University School of Medicine
Professor of Epidemiology, Department of
Environmental Health Sciences
Yale University School of Public Health
New Haven, Connecticut

James Stoller, MD

Professor and Chairman, Education Institute
Staff, Respiratory Institute
Jean Wall Bennett Professor of Medicine and
Samson Global Leadership Endowed Chair
Cleveland Clinic
Cleveland, Ohio

Michael Weisman, MD

Distinguished Professor of Medicine Emeritus

David Geffen School of Medicine

University of California Los Angeles

Professor of Medicine Emeritus

Cedars-Sinai Medical Center

Los Angeles, California

ACTING INDUSTRY REPRESENTATIVE TO THE COMMITTEE

(Non-Voting)

Sean Curtis, MD

(Acting Industry Representative)

Vice President, Clinical Research

Vice President, Clinical Therapeutic Area Head

Respiratory and Immunology

Merck Research Laboratories

Rahway, New Jersey

FDA PARTICIPANTS (Non-Voting)

Sally Seymour, MD

Acting Division Director

Division of Pulmonary, Allergy, and

Rheumatology Products (DPARP)

Office of Drug Evaluation II (ODE-II)

Office of New Drugs (OND), CDER, FDA

Nikolay Nikolov, MD

Associate Director for Rheumatology

DPARP, ODE-II, OND, CDER, FDA

Rachel Glaser, MD

Clinical Team Leader

DPARP, ODE-II, OND, CDER, FDA

Nadia Habel, MD

Medical Officer

DPARP, ODE-II, OND, CDER, FDA

Yu Wang, PhD

Statistical Reviewer

Division of Biometrics II

Office of Biostatistics

Office of Translational Sciences

CDER, FDA

C O N T E N T S

C O N T E N T S (continued)

C O N T E N T S (continued)

P R O C E E D I N G S

(8:29 a.m.)

Call to Order

Introduction of Committee

DR. SOLOMON: Good morning. I would first like to remind everyone to please silence your cell phones, smartphones, and any other devices if you have not already done so. The FDA press contact for today's meeting is Nathan Arnold, who is not present. His contact information is available on the press handout at the check-in table.

My name is Dan Solomon, and I'm the chairperson of the Arthritis Advisory Committee and for this meeting. I will now call today's meeting of the Arthritis Advisory Committee to order, and we'll start by going around the table and introducing ourselves. We'll start with the FDA to my left and then continue around the table.

DR. SEYMOUR: My name is Sally Seymour. I'm the director of the Division of Pulmonary, Allergy, and Rheumatology Products at the FDA.

DR. NIKOLOV: Good morning, everyone. My

name is Nikolay Nikolov. I'm an associate director for rheumatology in the Division of Pulmonary, Allergy, and Rheumatology Products.

DR. GLASER: Good morning. I'm Rachel Glaser. I'm a clinical team leader in the Division of Pulmonary, Allergy, and Rheumatology Products.

DR. HABAL: Good morning. My name is Nadia Habal. I'm a medical officer in the Division of Pulmonary, Allergy, and Rheumatology Products.

DR. YU WANG: Good morning. I'm Yu Wang. I'm a statistical reviewer in the Office of Biostatistics.

DR. BECKER: Good morning. I'm Mara Becker. I'm a pediatric rheumatologist at Duke University Medical Center.

DR. RICHARDS: Good morning. I'm John Richards. I'm a rheumatologist at the VA Healthcare System in Pittsburgh.

DR. OLIVER: Good morning. I'm Alyce Oliver. I'm an adult rheumatologist at the Medical College of Georgia.

DR. NASON: Good morning. My name is Martha

Mason. I'm a biostatistician at the National Institutes of Health, NIAID specifically.

DR. CURTIS: Good morning. I'm Jeff Curtis. I'm a rheumatologist at the University of Alabama at Birmingham.

DR. REDLICH: I'm Carrie Redlich. I'm a pulmonologist at Yale School of Medicine.

DR. WANG: Yinghua Wang, designated federal officer.

DR. SOLOMON: I'm Dan Solomon. I'm a rheumatologist and clinical scientist at Brigham and Women's Hospital in Boston.

DR. KATZ: Good morning. I'm James Katz. I'm a rheumatologist at the National Institutes of Health.

DR. CALHOUN: Good morning. My name is Bill Calhoun. I'm a pulmonologist and allergist in the adult world at University of Texas Medical Branch in Galveston.

DR. HORONJEFF: Good morning. I'm Jennifer Horonjeff. I am serving as the consumer representative, as I am a rheumatology patient.

I'm also a patient-centered outcomes researcher at Columbia University Medical Center and run the Savvy Cooperative, which is a patient organization.

MR. GILLIGAN: Good morning. I'm Todd Gilligan. I'm a patient representative, and I'm an SSc-ILD patient.

DR. MAY: Good morning. I'm Suzanne May. I'm a professor of biostatistics at the University of Washington in Seattle.

DR. GARIBALDI: Good morning. I'm Brian Garibaldi. I'm a pulmonologist at Johns Hopkins.

DR. KERR: Good morning. Gail Kerr, rheumatologist, D.C. VA Medical Center.

DR. WEISMAN: Good morning. I'm Michael Weisman, a rheumatologist in Los Angeles, California.

DR. STOLLER: Good morning. My name is Jamie Stoller. I'm a lung doctor at the Cleveland Clinic.

DR. GELLER: Hi. I'm Nancy Geller. I'm the director of the Office of Biostatistics Research at the National Heart, Lung, and Blood Institute.

DR. CURTIS: Good morning. My name is Sean Curtis. I'll be serving as the industry representative today. I work in clinical development at Merck.

DR. SOLOMON: For topics such as those being discussed at today's meeting, there are often a variety of opinions, some of which are quite strongly held. Our goal is that today's meeting will be a fair and open forum for discussion of these issues, and that individuals can express their views without interruption. Thus, as a gentle reminder, individuals will be allowed to speak into the record only if recognized by the chairperson. We look forward to a productive meeting.

In the spirit of the Federal Advisory Committee Act and the Government in the Sunshine Act, we ask that the advisory committee members take care that their conversations about the topic at hand take place in the open forum of the meeting. We are aware that members of the media are anxious to speak with the FDA about these

proceedings. However, FDA will refrain from discussing the details of this meeting with the media until its conclusion.

Also, the committee is reminded to please refrain from discussing the meeting topic during breaks or lunch. Thank you.

Now I will pass it to Yinghua Wang, who will read the Conflict of Interest Statement.

Conflict of Interest Statement

DR. WANG: The Food and Drug Administration is convening today's meeting of the Arthritis Advisory Committee under the authority of the Federal Advisory Committee Act of 1972. With the exception of the industry representative, all members and temporary voting members of the committee are special government employees or regular federal employees from other agencies and are subject to federal conflict of interest laws and regulations.

The following information on the status of this committee's compliance with federal ethics and conflict of interest laws, covered by but not

limited to those found at 18 U.S.C. Section 208, is being provided to participants in today's meeting and to the public.

FDA has determined that members and temporary voting members of this committee are in compliance with federal ethics and conflict of interest laws. Under 18 U.S.C. Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have potential financial conflicts when it is determined that the agency's need for a special government employee's services outweighs his or her potential financial conflict of interest or when the interest of a regular federal employee is not so substantial as to be deemed likely to affect the integrity of the services which the government may expect from the employee.

Related to the discussions of today's meeting, members and temporary voting members of this committee have been screened for potential financial conflicts of interest of their own, as well as those imputed to them, including those of

their spouses or minor children and, for purposes of 18 U.S.C. Section 208, their employers. These interests may include investments; consulting; expert witness testimony; contracts, grants, CRADAs; teaching, speaking, writing; patents and royalties; and primary employment.

Today's agenda involves discussion of the supplemental new drug application 205832 for nintedanib capsules, drug name Ofev, sponsored by Boehringer Ingelheim, for the treatment of systemic sclerosis-associated interstitial lung disease. The focus of the discussion will be whether the application provides substantial evidence of efficacy for the proposed indication.

This is a particular matters meeting during which specific matters related to the Boehringer Ingelheim's sNDA will be discussed. Based on the agenda for today's meeting and all financial interests reported by the committee members and temporary voting members, no conflict of interest waivers have been issued in connection with this meeting.

To ensure transparency, we encourage all standing committee members and temporary voting members to disclose any public statements that they have made concerning the product at issue.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Sean Curtis is participating in this meeting as a nonvoting industry representative acting on behalf of regulated industry. Dr. Curtis' role at this meeting is to represent industry in general and not any particular company. Dr. Curtis is employed by Merck Research Laboratories.

We would like to remind members and temporary voting members that if the discussions involved any other products or firms not already on the agenda for which an FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement, and their exclusion will be noted for the record. FDA encourages all other participants to advise the committee of any financial relationships that they may have with the firm at

issue. Thank you.

DR. SOLOMON: Thanks, Yinghua.

We'll now proceed with the FDA's opening remarks from Dr. Rachel Glaser.

FDA Opening Remarks - Rachel Glaser

DR. GLASER: Good morning. I'd like to welcome you to the Arthritis Advisory Committee meeting for the new drug application, or NDA, 205832 supplement 12, nintedanib for systemic sclerosis interstitial lung disease. My name is Rachel Glaser. I'm a clinical team leader in the Division of Pulmonary, Allergy, and Rheumatology Products, and I'm also an adult rheumatologist.

Before I begin, I would like to thank the members of the panel for your participation in this Arthritis Advisory Committee meeting. We consider your expert scientific advice and recommendations very important to our regulatory decision-making processes. In the next few slides, I will provide an overview of the nintedanib development program with an emphasis on efficacy, safety, and overall risk-benefit considerations.

We are here this morning to discuss NDA 205832 supplement 12, submitted by Boehringer Ingelheim, or BI, for nintedanib for systemic sclerosis interstitial lung disease. Nintedanib is an oral small molecule inhibitor of receptor and non-receptor tyrosine kinases. It is currently approved for the treatment of idiopathic pulmonary fibrosis or IPF.

The applicant has proposed nintedanib for a novel indication, the treatment of systemic sclerosis-associated interstitial lung disease. Systemic sclerosis interstitial lung disease will be referred to as systemic sclerosis ILD, or SSc-ILD, through the FDA presentations today.

Systemic sclerosis is a rare systemic autoimmune connective tissue disease involving the skin, underlying tissues, blood vessels, and major organs that affects approximately 100,000 people in the United States. It is characterized by microvascular damage and fibrosis of the skin and internal organs, including the lung, heart, kidneys, and gastrointestinal tract. Cardiac and

pulmonary manifestations are the most common cause of systemic sclerosis related death.

Interstitial lung disease occurs in approximately 55 to 65 percent of patients with systemic sclerosis. There are no currently approved therapies for systemic sclerosis or systemic sclerosis ILD. In clinical practice, treatment is based on expert-derived guidelines for the management of organ-specific manifestations.

Current guidelines from the European League Against Rheumatism, or EULAR, and the British Society for Rheumatology recommend consideration of immunosuppressive agents such as cyclophosphamide and mycophenolate for the treatment of SSc-ILD. These therapies are associated with significant toxicities, including cytopenias, infections, and malignancies. There's a high unmet medical need for therapies for these patients.

BI has proposed nintedanib for the treatment of systemic sclerosis-associated interstitial lung disease. The proposed dosing regimen is the same as the approved dosing regimen for IPF.

Specifically, BI has proposed the recommended dose of nintedanib is 150 milligrams twice daily, approximately 12 hours apart, taken with food. The recommended dose in patients with mild hepatic impairment is 100 milligrams twice daily. In addition, temporary dose reduction to 100-milligram treatment interruption or discontinuation may be considered for management of adverse reactions.

This slide summarizes the known safety profiles of nintedanib the IPF development program. Listed on the slide are the labeled warnings and precautions associated with nintedanib treatment. These include hepatic impairment; elevated liver enzymes and drug-induced liver injury; gastrointestinal disorders; embryo-fetal toxicity; arterial thromboembolic events; bleeding events; and GI perforation.

The applicant conducted a single clinical study, study 1199.214, to evaluate the efficacy and safety of nintedanib in SSc-ILD. To be concise, I will refer to the study as study 214. Study 214 a double-blind, randomized, placebo-controlled,

parallel-group study in which 576 patients were randomized one-to-one to receive nintedanib 150 milligrams twice daily or matching placebo.

The primary endpoint was the annual rate of decline in FVC in milliliters over 52 weeks. Key secondary endpoints included absolute change in modified Rodnan skin score, or mRSS, at week 52 and absolute change in St. George's Respiratory Questionnaire, or SGRQ, at week 52.

Additional secondary endpoints included time to death, the Health Assessment Questionnaire Disability Index, or HAQ-DI, and the Functional Assessment of Chronic Illness Therapy, or FACIT dyspnea scale. The study design will be discussed in greater detail by the clinical reviewer, Dr. Habal, later in the FDA presentation.

Forced vital capacity is a pulmonary function test that measures lung volume. It is the amount of air that can be forcibly exhaled from the lungs after the deepest breath possible. Clinically, it has been used to assess restrictive lung diseases such as IPF and SSc-ILD. In these

types of diseases, FVC decreases over time.

FVC was the primary efficacy endpoint using the IPF programs for nintedanib and pirfenidone. FVC had not previously been used as an endpoint in IPF, however, it was considered an appropriate endpoint to assess response in a disease that is marked by a progressive decline in lung function. Both nintedanib and pirfenidone reduced the decline in FVC over 52 weeks.

Other clinically meaningful endpoints were supportive, for example, exacerbations. For IPF, baseline FVC and decline in FVC greater than 10 percent have been shown to correlate with mortality. It was through review of these clinical development programs that we have accepted the use of FVC as a primary efficacy endpoint in IPF clinical trials.

Given what we know about FVC in IPF and its use clinically in the assessment of restrictive lung diseases, FVC is a reasonable primary efficacy endpoint in an SSc-ILD program. Both IPF and SSc-ILD are chronic progressive fibrosing diseases,

although there's less information about the magnitude of treatment effect that is meaningful in correlation with other meaningful endpoints in SSc-ILD.

To provide further context for the discussion, I will summarize the pertinent regulatory history of the submission. Nintedanib was approved for the treatment of IPF on October 15, 2014. The first communication regarding the clinical development in SSc-ILD took place in February 2015. At that time, the agency acknowledged that SSc-ILD is a slowly progressive disease manifestation, and it may take years to show benefit on disease progression.

In the absence of preliminary information on the effects of nintedanib on SSc-ILD, it was unclear if treatment could alter a natural decline in forced vital capacity in a one-year study in this patient population. However, the agency also acknowledged that a longer study may be challenging in this rare disease.

The agency noted that it may be difficult to

determine if a small improvement in FVC is meaningful without supportive efficacy endpoints that more directly measure how patients function and feel. Therefore, the applicant was advised to continue to follow the patients to the conclusion of the study. In addition, the applicant was advised to include all-cause mortality as an endpoint and to include secondary endpoints that measure how patients feel and function.

Whether a single well-controlled study would be sufficient to provide substantial evidence of safety and efficacy of nintedanib in SSc-ILD to meet the regulatory standard would depend on the persuasiveness of the treatment effect. An IND was opened with the proposed study in September 2015.

On July 6, 2016, nintedanib was granted orphan designation for the treatment of systemic sclerosis, including the associated interstitial lung disease. I will talk about this further on the next slide. Nintedanib was also granted fast-track designation for SSc-ILD in March 2018. The applicant submitted this supplement for the

treatment of SSc-ILD on March 7, 2019, and the application was granted priority review.

The orphan drug designation program provides orphan status to drugs and biologics, which are defined as those intended for the treatment, prevention, or diagnosis of a rare disease or condition, which is one that affects less than 200,000 persons in the U.S. or meets cost recovery provisions of the Act. Orphan drug designation qualifies the sponsor of the drug for various development incentives of the Orphan Drug Act.

Orphan designation does not alter the standard regulatory requirements and process for obtaining marketing approval. Safety and effectiveness of a drug must be established through adequate and well-controlled studies. However, for rare diseases, additional considerations to the design of a clinical program include the amount of clinical data that balance providing an adequate assessment of efficacy and safety and the feasibility of conducting clinical studies. In

that context, sometimes a single clinical study may be acceptable.

On this slide is an excerpt from FDA guidance for industry providing clinical evidence of effectiveness for human drug and biological products. This guidance indicates situations where a single study of a new treatment may be sufficient to support a marketing application; in particular, when there's independent substantiation from related supportive study data and/or when evidence from the single study is both clinically and statistically very persuasive.

The considerations of the single-study approach for nintedanib for SSc-ILD included that SSc-ILD is a rare disease, IPF and SSc-ILD are both chronic progressive fibrosing lung diseases, and while there are differences in gender ratio and age of onset, with SSc-ILD affecting middle-aged females and IPF affecting older males, both result in pulmonary fibrosis. In addition, the IPF studies with nintedanib had a similar design to the study in SSc-ILD.

Based on these considerations, the agency agreed to consider a single study to provide substantial evidence of safety and efficacy of nintedanib in SSc-ILD if the observed treatment effect was robust.

Study 214 showed a statistically significant lower annual rate of decline of FVC with nintedanib treatment compared with placebo over 52 weeks. The treatment difference was 41 milliliters per year. The observed decrease in FVC decline was not supported by improvement in other measures of pulmonary function, disease activity, or physical function, including endpoints that directly assess how a patient feels, functions, or survives. In addition, the treatment effect was less robust in subgroups, including patients from the U.S. and Canada, as well as the subgroup on mycophenolate at baseline.

The clinical significance of the treatment effect of 41 milliliters per year in the absence of supportive efficacy from other secondary endpoints is for your consideration and discussion today. In

study 214, the safety profile was generally consistent with the known safety profile of nintedanib in IPF. Death and serious adverse events were balanced between the treatment groups.

Differences between treatment and placebo were primarily related to the gastrointestinal and hepatic events, which is consistent with the labeled adverse reactions. In addition, there was a numerical increase in pneumonia in the nintedanib treatment group, however, overall infections were similar between treatment groups.

SSc-ILD is a rare and serious disease associated with high morbidity and mortality. It is also a disease with high unmet need for new therapies. Study 214 demonstrated a statistically significant decrease in the annual rate of decline of FVC with nintedanib treatment compared with placebo. As previously noted, the observed decrease in FVC decline was not supported by improvement in other measures of pulmonary function such as SGRQ or FACIT dyspnea scale and other measures of disease activity such as mRSS or in

differences in mortality.

FVC is an endpoint that does not directly measure how a patient feels, functions, or survives. In IPF, a decrease in decline in FVC has been demonstrated to result in clinical response. Of note, the treatment difference in an nintedanib IPF program ranged from 94 to 131 milliliters per year as compared to 41 milliliters per year in the SSc-ILD clinical study. However, the relative difference in FVC decline in nintedanib treatment arms versus placebo were similar between the IPF and SSc-ILD programs.

To what extent the treatment effect in IPF can be relied upon to support the modest effect observed in the SSc-ILD population is for the committee's consideration today. The safety of nintedanib in SSc-ILD is generally consistent with the established safety profile of nintedanib in IPF, which includes risk of gastrointestinal disorders and liver toxicity.

The warnings and precautions for nintedanib are listed on the right side of the slide. In

addition to the established safety risks in the SSc-ILD population, there were increased number of serious infections driven by an increase in pneumonia in the nintedanib treatment group.

In summary, while the efficacy data are consistent with the treatment effect of nintedanib versus placebo, the committee is asked to address whether the observed treatment effect on FVC is clinically meaningful in patients with SSc-ILD.

I will now introduce the discussion and voting questions that the committee will consider today. The first discussion point refers to the efficacy data for nintedanib for the treatment of systemic sclerosis interstitial lung disease. We would like to obtain the committee's input on the clinical meaningfulness of the changes in FVC observed with nintedanib treatment in the SSc-ILD population.

We also request the committee's input on the efficacy in the subgroups of patients from the U.S. and Canada, and the patients who received background mycophenolate treatment at baseline

versus those who did not receive background mycophenolate at baseline. Discuss the implications, if any, of the results of these subgroups for use of nintedanib in patients in the U.S.

Then the committee will be asked to vote whether the data provides substantial evidence of the efficacy of nintedanib for the treatment of systemic sclerosis interstitial lung disease. This will be followed by a voting question on whether the safety data are adequate to support approve of nintedanib for the treatment of systemic sclerosis interstitial lung disease. We will conclude with a separate voting on the overall benefit-risk profile to support approval of nintedanib in the proposed indication.

Thank you for your attention, and I'll turn the podium back to you, Dr. Solomon.

DR. SOLOMON: Thanks. That was a great overview.

We're going to now move to the applicant's presentation, and I want to make a couple comments

before that.

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 advisory committee meeting, FDA believes that it is important to understand the context of an individual's presentation.

For this reason, FDA encourages all participants, including the applicant's non-employee presenters, to advise the committee of any financial relationships that they may have with the applicant, such as consulting fees, travel expenses, honoraria, and interest in a sponsor, including equity interest and those based upon the outcome of the meeting.

Likewise, FDA encourages you, at the beginning of your presentation, 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 presentation, it will not preclude you from

speaking.

We will now proceed with presentations from Boehringer Ingelheim.

Applicant Presentation - Kay Tetzlaff

DR. TETZLAFF: Good morning, members of the Arthritis Advisory Committee, FDA representatives, and members of the audience. My name is Kay Tetzlaff. I'm the head of medicine, therapeutic area respiratory at Boehringer Ingelheim. I'd like to thank you for the opportunity to discuss our development program for nintedanib capsules in treating systemic sclerosis-associated interstitial lung disease.

Systemic sclerosis-associated interstitial lung disease, or SSc-ILD, is a serious and life-threatening disease with very limited treatment options and no approved therapies in the United States. Nintedanib is a small molecule tyrosine kinase inhibitor that blocks numerous profibrotic pathways implicated in pulmonary fibrosis.

Nintedanib has established safety and

efficacy in idiopathic pulmonary fibrosis, or IPF, and has been approved for that indication in more than 70 countries. SSc-ILD is another target in fibrosing interstitial lung disease treatment that shares clinical and pathologic features with IPF.

The clinical development program for nintedanib spans the whole spectrum of fibrosing interstitial lung diseases, which include IPF, SSc-ILD, and other progressive fibrosing interstitial lung diseases. Nintedanib was first approved for IPF in 2014, based on the data from the SENSCIS and INPULSIS trials.

The application for SSc-ILD, based on the SENSCIS trial is currently under review, and today we are presenting the data supporting this application. The program for progressive fibrosing ILDs for the INBUILD trial is ongoing.

Nintedanib's approval as a treatment for pulmonary fibrosis in IPF was based on the efficacy established to replicate 52-week phase 3 trials, INPULSIS-1 and INPULSIS-2. These studies were designed to investigate the effects of nintedanib

on lung function decline, a hallmark clinical feature of pulmonary fibrosis.

In these trials, nintedanib significantly reduced the annual rate of decline in forced vital capacity, or FVC, in patients with IPF, consistent with slowing disease progression. FVC has been established as a preferred outcome in trials investigating new treatments in fibrosing interstitial lung disease. FVC is accepted as a surrogate for mortality in IPF.

The pooled analysis of the two INPULSIS trials showed that the treatment difference between nintedanib and placebo corresponded to a relative reduction in the rate of decline in FVC of nearly 50 percent. The safety and tolerability of nintedanib are supported by the clinical development program in IPF, long-term experience in phase 4 trials, and postmarketing exposure of more than 80,000 patient-years.

Systemic sclerosis, also known as scleroderma, and we will use both terms throughout our presentation today, is the chronic connective

tissue disease characterized by progressive fibrosis, which has a high disease burden and high rate of mortality. Interstitial lung disease is a common manifestation of systemic sclerosis and is the leading cause of death.

Pulmonary fibrosis in systemic sclerosis is progressive over time with a variable clinical course. The accelerated loss in lung function, however, is irreversible. It has been shown that short-term changes in FVC as a surrogate for progression of pulmonary fibrosis may predict mortality in SSc-ILD.

The efficacy and safety of nintedanib in patients with SSc-ILD was evaluated in a large phase 3 trial, SENSCIS study. In fact, this was the largest randomized placebo-controlled trial conducted in SSc-ILD to date.

The trial design of SENSCIS mirrored the design of the INPULSIS trials with a 52-week treatment duration and evaluation of lung function decline using FVC as the primary endpoint. The trial population included patients with diffuse and

limited SSc, allowed concomitant immunosuppressants such as mycophenolate, and included a wide range of baseline lung function with no upper limit on the FVC. Therefore, the SENSCIS population is representative of patients seen clinical practice.

Similar to IPF, patients were offered to roll over into a long-term, open-label extension trial, the SENSCIS-I study. Ninety-four percent of SENSCIS patients chose to do so, indicating the high unmet medical need in this population.

In January 2015, BI submitted the pre-IND meeting package, seeking FDA's feedback on the proposed design of SENSCIS and the overall clinical development program in SSc-ILD. FDA advised that for the primary endpoint of the annual rate of decline in FVC, observed values instead of percent predicted values should be used.

FDA recommended that patients are followed up for longer than the initially planned 52 weeks. FDA also stressed the importance of minimizing missing data and of all-cause mortality as an additional endpoint. FDA's advice was implemented

in the clinical trial protocol, and subsequently the IND for nintedanib in SSc-ILD was submitted and went into effect in September 2015.

FDA granted orphan drug status to nintedanib for the treatment of SSc-ILD July 2016 and fast-track designation in March 2018. The sNDA was submitted in March, and in May 2019, BI was informed that the sNDA is under priority review with the FDA.

We will provide information to support the following proposed extension of the current indication for the use of nintedanib soft capsules to include treatment of SSc-ILD. Nintedanib is indicated for the treatment of systemic sclerosis-associated interstitial lung disease. The dosing regimen we propose is the same as has been approved for IPF.

Nintedanib has been formulated as a capsule and will be available in 100 milligram and 150 milligrams strengths. 150 milligram twice daily will be the recommended starting dose, and there will be recommendations in the labeling to reduce

dosing to 100 milligrams twice daily in specific cases to have patients manage certain adverse events.

Today you will hear that the SENSCIS trial met its primary efficacy endpoint. It showed a significant reduction of FVC decline over a 52-week treatment, and in a large population of patients with SSc-ILD, SENSCIS is the first positive placebo-controlled phase 3 study in SSc-ILD. The degree of relative reduction in lung function decline versus placebo was consistent with the experience from IPF. Safety and tolerability were similar to IPF and no new safety signals were detected.

Nintedanib has an antifibrotic treatment option with a target of slowing down loss of lung function in SSc-ILD. The data we present today support a positive benefit-risk assessment and improved nintedanib soft capsules for the treatment of SSc-ILD.

This morning, Dr. Seibold will set the scene on the disease background and unmet need of SSc-

ILD. Subsequently, Dr. Stowasser will present the clinical development rationale for nintedanib in SSc-ILD and will summarize the clinical evidence available from treatment of patients with IPF. Dr. Clerisme-Beaty will review the efficacy data of the SENSCIS trial, and Dr. Kohlbrenner will review the safety data. Then I will come back and briefly summarize our conclusion on the benefit-risk of nintedanib in SSc-ILD. Finally, Dr. Brown will summarize and provide his clinical perspective on the data reviewed today.

The advisors identified on this slide will be available to address specific questions or clarifications requested by the advisory committee during the meeting today. Now, I'd like to invite Dr. Seibold to the podium to discuss important background information about systemic sclerosis-associated interstitial lung disease.

Applicant Presentation - James Seibold

DR. SEIBOLD: Thank you, Dr. Tetzlaff. My name is Jim Seibold from Scleroderma Research Consultants. My task today is to orient you to the

clinical aspects of scleroderma, also known as systemic sclerosis, and the clinical implications of interstitial lung disease, which continues to represent a significant unmet medical need.

I see my role here as an advocate for the community of scleroderma patients and caregivers. That said, I am a paid consultant to the sponsor, although I have no financial interest in the outcome of this meeting.

Systemic sclerosis is a rare disease with the best estimates of annual incidence in the U.S. population ranging from 20 to 24 per million per population per year. The U.S. prevalence is thus estimated at approximately 300 million individuals, which translates to between 70[000] and 100,000 U.S. adults with the disease. This permits it to be classified as an orphan disease.

Interstitial lung disease occurs in the majority of these patients with the estimates ranging from 52 percent to 79 percent, depending on the methodology employed. Scleroderma is dominantly a disease of women, occurring 4 times

more frequently in women as in men, with a peak age of onset between 40 and 50 years of age. There's some data to suggest that scleroderma is more severe in African Americans.

Systemic sclerosis generally segregates into two patterns of clinical behavior. This graph shows the extent of skin involvement over time. Diffuse scleroderma is characterized by widespread skin involvement that increases very rapidly early on, and then plateaus at around 18 to 24 months. Thereafter, it is stable and typically spontaneously improves.

In contrast, limited scleroderma has a much slower onset and an indolent progression, with skin involvement even after many years being restricted to distal areas of the body. Diffuse and limited scleroderma differ considerably in terms of the extent of skin involvement in the pace of disease, but they share many common clinical features, including Raynaud phenomenon, digital ulcers, esophageal involvement, and pulmonary hypertension. But the other notable shared clinical feature is

interstitial lung disease, which tends to begin within the first one to two years after diagnosis, and which continues to worsen over time.

Thus, while patients are seeing improvement in their skin, their lung fibrosis continues to worsen. Studies that address both skin and lung simultaneously are extremely difficult to perform.

In my 41 years as a caregiver for patients with scleroderma, I've become all too familiar with the extreme impact of this disease on patients' lives. SSc often begins in the prime of life in women who serve as caregivers and the family anchor for children and aging parents. We also have to consider the impact of this life-changing illness on their careers and social activities.

As physicians, we offer little to address the uncertainty regarding the likely clinical course of the disease. Patients face a multitude of clinical issues, from cosmetic effects, Raynaud phenomenon, hand dysfunction, and fatigue, in addition to the issues stemming from internal organ involvement.

In the face of this onslaught of day-to-day issues impacting function, quality of life, and survival, evidence-based treatment options are limited. Currently, the only approved therapies are for pulmonary arterial hypertension.

SSc-ILD has emerged as the leading cause of death. The clinical presentation is described here. It's present in the majority of patients with the most common form of involvement as fibrotic, nonspecific interstitial pneumonia or NSIP. It is clinically progressive, although the pace of decline is variable.

Around one-third of patients experienced rapid progression. Lung function decline begins early in the disease course, but continues to decline over time, and median survival is somewhere in the range of 5 to 8 years after diagnosis.

Importantly, over the years, we've identified several putative risk factors for progressivity of SSc-ILD. The most rapidly progressive form is typically seen in patients with diffuse scleroderma and those with disease duration

less than 5 years. Extent of lung involvement by high-resolution CAT scan has a very strong predictive value for continued progression, as does a forced vital capacity less than 70 percent predicted. There are sure serologic tests such as the antitopoisomerase 1 antibody that correlate with progressivity of the ILD.

The SENSCIS trial was large enough and inclusive enough to permit assessment of all of these factors. These recent data reported by Guler, et al. showed that the rate of decline in FVC is clearly separating the population into distinct cohorts with different survival prognosis. The SSc-ILD has a heterogeneous disease course in terms of the pace of FVC decline. The group with rapid progression has particularly poor survival, with the most aggressive form associated with the survival of less than 4 years.

These findings have also been recently borne out in the study of SSc patients in Norway, which showed that the ILD is associated with mortality in patients with SSc, even amongst those with

preserved lung volumes at baseline when compared to matched healthy individuals.

This analysis further showed that approximately 30 percent of patients with scleroderma in Norway died from ILD. Amongst patients with scleroderma in ILD, the risk of death from ILD increased from 50 to 70 percent. If a patient with SSc-ILD from Norway met the inclusion criteria for SENSCIS with a greater than 10 percent disease extent on HRCT, the prognosis was even worse, with a dismal 10-year survival rate, considering this a relatively young population.

Recent data from the EUSTAR cohort study, A collaborative effort in Europe involving over 11,000 patients, has confirmed that pulmonary fibrosis is the leading cause of disease related death in the scleroderma population. This is clearly a clinical problem that deserves our close attention.

The pathogenesis of scleroderma is complex. It apparently begins with vascular injury, but immune activation, including disease-specific auto

antibodies, are present at the earliest recognizable stages of disease. These events are relatively short-lived, but they initiate a cascade of cellular responses and cytokine release, many of which are tightly linked to the genesis of fibrosis.

More specifically, cytokines and growth factors stimulate the migration and differentiation of the activated myofibroblasts, the primary source of overproduction of extracellular matrix. While the initiating events may differ, the diverse interstitial lung diseases share this final common pathway of fibroblasts activation and fibrosis.

Dr. Stowasser will show you how the pleiotropic effects of nintedanib interfere with this fibrotic process at multiple levels, which applies across a broad branch of fibrotic bronchomo-lung diseases.

We mentioned IPF a lot in our presentation because it is the foundation for the sponsor's development program in the scleroderma ILD. While IPF and SSc-ILD are fibrosing interstitial lung

diseases, they share many pathophysiologic features, but they also have some important clinical differences.

IPF is a disease of males over the age of 70, where scleroderma is a disease predominantly of women ages 45 to 55. The pathologic findings in IPF are typically UIP, usual interstitial pneumonia, whereas in SSc-ILD, nonspecific interstitial pneumonia dominates. This generally implies that scleroderma ILD has more of an inflammatory component, at least at early stages.

The key point of difference lies in the pace of progression. The pace of decline and forced vital capacity is rather rapid in IPF, whereas it is generally much slower in scleroderma ILD. Consequently, median survival is 3 to 5 years in the IPF population compared with 5 to 8 years in scleroderma.

Dr. Brown and I had the privilege of serving as co-chairs on a 270-physician panel convened under the egis of OMERACT to develop consensus criteria for outcome assessment in connective

tissue disease-associated ILD. This slide illustrates what we developed as the core set measures.

This doesn't define how to do lung studies now, but rather defines the agenda for research; what's validated, and what's not, and what's needed. There was unanimous agreement, though, that measures of lung physiology, most notably forced vital capacity, was a core measure and is a robust surrogate for all-cause mortality.

We also included lung imaging, recognizing the quantitative HRCT required further study before it could be considered completely validated. Then we wrestled with the importance of patient-reported outcomes, recognizing that an ideal outcome in an ILD trial would be to prevent deterioration. Therefore, we would not expect measures of shortness of breath in other patient-reported outcomes to improve, but rather simply not worsen. Unfortunately, there are no independently validated patient-reported outcome instruments in scleroderma, and there are many challenges in

developing such given the multi-system features of the disease.

So let's look at forced vital capacity. It's emerged as the primary outcome measure in all trials of interstitial lung disease. It's defined as the maximum amount of air in milliliters exhaled after maximum inhalation. It's reproducible, and it offers real-time quality assurance of the inspection of the flow-volume loop. It's a measure of lung elasticity, although it's also affected by age, gender, ethnicity, and height.

Most of the literature describes FVC as percent predicted, but it can also be expressed as absolute volume in milliliters. It is important to mention that this normal physiologic decline in FVC as we age is approximately 25 milliliters per year. In contrast, patients with ILD can lose somewhere in the range of 100 to 200 milliliters per year, and the rate of loss is dependent on the specific disease entity.

How to assess dyspnea scleroderma? There are many challenges in establishing a validated

patient-reported outcome. They're exemplified when we considered dyspnea. In scleroderma, dyspnea and exercise capacity are influenced by many domains of the disease, including muscular skeletal involvement, skeletal muscle perfusion, deconditioning, and concomitant heart and pulmonary vascular disease.

One good example is the U.S. trial of cyclophosphamide, whereas the modern or transitional dyspnea index improved on drug, whereas the FVC actually decreased. This reflects the impact of cyclophosphamide on musculoskeletal features rather than an effect on the lung. Currently, there is no validated dyspnea PRO in scleroderma ILD.

There are no currently approved therapies for scleroderma ILD, and the goal of treatment, again, is to prevent worsening of lung function. Based on our current knowledge, regeneration of functional alveolar surface is not biologically plausible. Nonetheless, there are a number of therapies employed in the community for the

management of ILD. These include cyclophosphamide, both oral and intravenous, and mycophenolate mofetil. There are also limited data available with rituximab.

Cyclophosphamide use is limited by adverse events, including gastrointestinal upset, bladder irritation, depressed white blood cell counts, premature infertility, and carcinogenesis after prolonged exposure. Mycophenolate is better tolerated with its use mainly limited by GI upset and fatigue. Importantly, mycophenolate has evolved to become the standard of care in the United States and is used in around 80 percent of patients. I will show you the trials that led to the use of these agents.

The use of cyclophosphamide is supported by a randomized trial known as Scleroderma Lung Study I, an NIH-supported trial that compare one year of oral cyclophosphamide with placebo. The study was designed to test whether immunosuppressive therapy could slow the progression of disease. Cyclophosphamide was discontinued at the end of one

year because of its high carcinogenic potential, and patients were followed for an additional year to assess the durability of the response.

What one can see here is that cyclophosphamide therapy led to less loss of forced vital capacity over one year of treatment. A 1 percent decline shown here at 12 months is roughly equivalent to a 27-milliliter annual rate of decline, which approximates the expected annual age related decline in FVC, and loss of FVC was greater on placebo. Unfortunately, at the end of two years, there were no differences between the active treatment arm and placebo.

This led to a follow-up study with slightly more optimistic results. Scleroderma Lung Study II was a small randomized study that compared one year of cyclophosphamide followed by one year of observation, with two years of continued mycophenolate mofetil in patients with early-stage SSc-ILD. There was a trend towards improvement in FVC or at least stability over the 24-month observation period.

This study is the major source of support for the prevalent use of mycophenolate in the scleroderma community, particularly in the United States. These studies suggest that in some patients, particularly in the earlier phases of disease, there's an inflammatory component of SSc-ILD that benefits, at least in the short term, from immunosuppressive therapy.

Unfortunately, the total story is not as positive as we would like. The group that led both SLS I and SLS II have recently published a long-term analysis of survival outcomes in each cohort. In the cyclophosphamide study, SLS I, they were able to follow patients out to 12 years. Long-term survival was dismal regardless of initial treatment or follow-up.

We can also see that survival was influenced by the rate up FVC loss. A greater than 10 percent loss in FVC percent predicted was significantly associated with mortality as early as two years.

A similar analysis looked at SLS II with follow-up here restricted to the 6 to 7-year range,

and showed no difference in survival between the two treatment groups. Patients continued to die of their lung disease at an accelerated rate despite immunosuppressive therapy. Again, loss of forced vital capacity over 1 to 2 years was significantly associated with mortality. These relationships were true for each individual clinical trial but also for analysis of the combined trial populations.

In summary, patients with scleroderma have a remarkably high disease burden with multi-organ involvement, but with lung and fibrosis as the leading cause of death. A reasonable therapeutic goal in SSc-ILD would be to prevent or slow the worsening of lung function based on evidence that decline in FVC is associated with increased mortality.

There are no approved therapies in the United States. Available immunosuppressive therapies may provide short-term benefit and selected subsets, but they do not appear to provide a long-term survival benefit.

The holy grail of drug development in this space is a treatment that can prevent progressive fibrotic destruction of the lung. One can hypothesize that an antifibrotic therapy would change the natural history SSc-ILD. I thank you for your attention, and I would now like to invite Dr. Stowasser to the podium.

Applicant Presentation - Susanne Stowasser

DR. STOWASSER: Thank you, Dr. Seibold.

Good morning. My name is Susanne Stowasser. I'm the associate head medicine for interstitial lung diseases at Boehringer Ingelheim. Before we get into the data from the SENSCIS trial, I will briefly review the rationale and the broader context for the development of nintedanib in systemic sclerosis-associated ILD.

First, as mentioned by Dr. Seibold, patients with SSc-ILD have a high unmet need for treatments targeting ILD, a potentially devastating organ manifestation. Second, we know nintedanib works in IPF, the most progressive and devastating form of ILD. In addition, fibrosing interstitial lung

diseases such as IPF and SSc-ILD share pathophysiologic similarities in fibrotic remodeling despite differences in initiating or amplifying events.

Finally, nintedanib has demonstrated antifibrotic activity in several in vitro models with human fibroblasts and in various models replicating different features and triggers of pulmonary human pathology.

Nintedanib has a very distinct inhibitory profile. Shown here in the simple graphic are several receptor and non-receptor kinase targets of nintedanib with potential relevance in fibrosing interstitial lung disease, including SSc-ILD.

For simplicity, not all growth factors and pathways involved in fibrosis are depicted. By binding to tyrosine kinases such as VEGF, PDGF, FGF, CSF1 receptors or Lck, nintedanib inhibits downstream signaling pathways implicated in fibrotic remodeling in pulmonary fibrosis.

Specifically, nintedanib inhibits the differentiation and migration of fibrocytes, the

migration and proliferation of fibroblasts, and their transformation to active myofibroblasts, and consequently extracellular matrix deposition.

Furthermore, nintedanib blocks the differentiation of alternatively activated macrophages and the release of profibrotic mediators from T cells involved in the initiation of fibrosis. The similarity of these final pathways in fibrosis, irrespective of initiating events, has provided a strong preclinical rationale to develop nintedanib as an antifibrotic treatment for pulmonary fibrosis.

Therefore, it is important to understand and interpret the SENSCIS trial and the context of the broader clinical development of nintedanib as a treatment for fibrosing interstitial lung diseases.

IPF as a prototypical fibrosing interstitial lung disease is the most progressive form of pulmonary fibrosis and was the initial indication. Nintedanib has proven efficacy in this indication based on the two replicate phase 3 INPULSIS trials and was approved in October 2014 in the United

States.

SSc-ILD progresses more gradually compared to IPF, but it shares also common features with IPF, as well as with other progressive fibrosing ILDs, which are currently included in the ongoing INBUILD trial.

Now, I will describe the key clinical trials that we conducted in IPF that serve as a foundation for the development of nintedanib in SSc-ILD. More than 1500 patients have been exposed to nintedanib across the global development program, and the main body of evidence stems from the pivotal phase 2 and 3 studies or 52 weeks duration.

The phase 2 TOMORROW trial has clearly shown that 150 milligram bid is the most efficacious dose, which was taken forward into the phase 3 INPULSIS trials. Equally important are long-term data from two open-label extension that provide a robust safety database. In those studies, patients were exposed to nintedanib for up to 68 months, and they confirmed the safety profile observed in the parent trials with no new safety signals.

There are several common features across the pivotal phase 3 studies in IPF and SSc-ILD. The dosing regimen used in INPULSIS and SENSCIS is the same. It's 150 milligram bid with the option to reduce the dose or interrupt treatment to manage adverse events. The randomized treatment period for the assessment of benefit-risk is the same, 52 weeks in all trials. And last but not least, the primary outcome measure is the same. It's a measure of lung volume, specifically forced vital capacity or FVC.

One may ask why FVC was chosen as the primary outcome. First, it reflects the underlying pathophysiology of the scarring process in the lung, and it is a simple and reproducible measure that is central to monitoring patients with interstitial lung disease in clinical practice.

In IPF, the most fatal interstitial lung disease, FVC, is accepted as a surrogate for clinically meaningful benefit based on its association with mortality in 6 interventional trials. In SSc-ILD, despite a more gradual

trajectory compared to IPF, numerous longitudinal studies have demonstrated that FVC decline is associated with increased mortality. Therefore, slowing the loss of lung function should ultimately prolong survival in these patients similar to IPF.

Given the lack of other validated outcomes, FVC is the preferred outcome in SSc trials that assess ILD progression, especially following Scleroderma Lung Study I. And finally, FVC has also been proposed as a core outcome measure by the connective tissue disease ILD working group under the direction of the OMERACT Initiative, a consensus process which included both physicians and patients.

In IPF, nintedanib significantly reduced the annual rate of decline in FVC by approximately 50 percent in both replicate INPULSIS trials, which was a breakthrough after two decades of failed clinical studies in IPF and has changed patient management.

Similar to other studies, the data suggests that the average loss of lung function in IPF, as

reflected in both placebo arms, is around 200 to 240 milliliters per year, which is a multiple of the physiologic decline in healthy adults of about 25 to 30 milliliters per year.

That is why we studied nintedanib in IPF first, because it's the most rapidly progressive form of pulmonary fibrosis . In comparison, and that's already alluded to by Dr. Seibold, the pace of decline in FVC is less rapid in SSc-ILD, although the underlying pathophysiology of fibrosis is similar.

These curves over time illustrate the mean change from baseline in FVC for both INPULSIS trials pooled. As you can see, the nintedanib and placebo groups separated early, and the curves continued to diverge over the 52-week treatment period.

In summary, nintedanib addresses the same underlying pathophysiology in IPF and SSc-ILD, which allowed us to build on the IPF experience. We set up SENSCIS as the largest randomized placebo-controlled trial in SSc-ILD, and we have

included a broad patient population, which is representative of the patients likely to be treated.

We used the same dosing regimen that was established in IPF, which is further supported by the fact that the PK properties of nintedanib are comparable across populations. And finally, the SENSCIS trial used the same primary endpoint used in the IPF program, the annual rate of declining FVC, which is a physiologic surrogate that reflects the pathologic fibrotic process in the lungs and is associated with mortality in both diseases.

Thank you for your attention, and now I will hand off to my colleague Dr. Clerisme-Beaty, who will present the efficacy results from the SENSCIS study.

Applicant Presentation - Emmanuelle Clerisme-Beaty

DR. CLERISME-BEATY: Good morning. My name is Emmanuelle Clerisme-Beaty, senior clinical program leader at Boehringer Ingelheim. This morning, I will be presenting an overview of the trial design, along with a summary of the efficacy

results from the SENSCIS study.

As previously mentioned, SENSCIS is the largest trial in patients with SSc-ILD to date. The trial was conducted in more than 30 countries, including North and South America, Europe, Asia, and Australia. Patients were recruited from over 190 sites of which approximately 25 percent were in the U.S. and Canada.

SENSCIS was a randomized, double-blind, placebo-controlled trial, evaluating the efficacy and safety of nintedanib in patients with SSc-ILD. Following screening, patients were randomized to either nintedanib or placebo, and were followed for a minimum of 52 weeks. Based on data from the literature, suggesting that antitopoisomerase antibody, or ATA, may be a prognostic indicator of viral de-progression, randomization was stratified based on ATA status at baseline.

Similar to the IPF program, the primary efficacy assessment is based on data collected over 52 weeks. However, as requested by the FDA, patients were continued on blinded treatment for up

to 100 weeks until the last randomized patient completed 52 weeks treatment.

To minimize missing data, all patients, including those who prematurely discontinued study drug, were asked to complete all study visits as scheduled. After completing the blinded treatment period, patients were followed for an additional 28 days off treatment primarily for collection of safety data. All patients who completed the trial on study medication were then eligible to roll over into the open-label study SENSCIS arm.

Consistent with the study design, treatment period beyond 52 weeks varied among patients, depending on when they entered the trial. As illustrated by the blue arrows, all treated patients were to complete a minimum treatment period of 52 weeks with the trial ending after the last patient end reached 52 weeks.

Due to staggered entry, treatment duration beyond 52 weeks varied with only 146 patients, or 25 percent, of the study population completing 100 weeks of treatment. As a result, the evaluation of

efficacy beyond 52 weeks was prespecified as exploratory and is considered only as supportive evidence.

Key inclusion criteria are shown here. Adults with SSc diagnosed within 7 years were eligible if they had active lung involvement as demonstrated on HRCT with greater than 10 percent fibrosis as confirmed by a central reviewer. In addition, eligible patients also had to have evidence of functional lung impairment based on an FVC percent predicted greater or equal to 40 percent and a diffusion capacity, DLCO or percent, between 30 to 89 percent of predicted.

Based on the known safety profile of nintedanib, patients with an ALT, AST, or bilirubin more than 1.5 times the upper limit of normal, as well as patients with bleeding cardiovascular or thromboembolic risk factors were excluded. In addition, due to potential effects of nintedanib related to its anti-VEGF activity, patients with significant vascular involvement related to systemic sclerosis, based on the following

criteria, were also excluded, as were patients with an FEV1 FVC ratio less than 70 percent in order to minimize potential confounding related to concomitant obstructive airway disease.

This slide summarizes the key stipulations in the protocol regarding use of concomitant medications At baseline, patients were allowed to be on prednisone at doses less than or equal to 10 milligrams per day, or on stable treatment with methotrexate or my mycophenolate for at least 6 months prior to randomization.

To minimize potential confounding, use of other immunosuppressive therapies, including the following, was not allowed at baseline. However, it is important to note that during the course of the study, initiation of any immunotherapy was allowed to manage disease worsening at the discretion of the investigator.

In line with the primary objective and experience in IPF, the annual rate of decline in FVC over 52 weeks was selected as a primary endpoint. In addition, the study was also powered

to evaluate the following two endpoints referred to as key secondary endpoints at 52 weeks. Based on preclinical data suggesting potential effect of nintedanib on skin fibrosis, the modified Rodnan skin score, or mRSS, a subjective measurement of skin thickness was used.

As there is no disease-specific quality-of-life instruments developed for SSc-ILD, the St. George's Respiratory Questionnaire, or SGRQ, was used to measure quality of life. The SGRQ was originally developed for use in COPD, with a total score ranging from 0 to 100, higher score representing worst quality of life.

To protect against type 1 error, the primary endpoint and key secondary endpoints were analyzed using a hierarchal testing procedure. For the analysis of the primary endpoint, all available data over 52 weeks were used to calculate the FVC slope, including data from patients who prematurely discontinued treatment. Adjustment for ATA status, the stratification factor used for randomization, as well as selected patient characteristics known

to impact rate of FVC declined were prespecified.

Out of 819 patients screened, 580 patients were randomized. The primary reasons for screen failure were due to patients not meeting the imaging or lung function criteria. Overall, 288 patients were treated with nintedanib and placebo, respectively, over 52 weeks.

There was a higher rate of premature study drug discontinuation noted in the nintedanib arm, 19 versus 11 percent, with a primary reason for discontinuation in both arms being due to adverse events. However, since patients were expected to attend all scheduled visits, even if they discontinued study drug, more than 90 percent of patients in both treatment arms completed observation up to 52 weeks.

It should be noted that 94 percent of eligible patients from the trial elected to roll over into the ongoing open-label extension trial underlying the high unmet need in this population.

While all patients with at least one host baseline FVC measurement contributed to the

efficacy evaluation, only FVC measurements collected within a predefined time window of the week 52 study visit were included in the prespecified analysis over 52 weeks. This resulted in 78 patients being considered as having missing FVC measurements at week 52. However, of these 78 patients, 28 were still and the study and had an FVC measurement with a median of 9 days after the week 52 time window.

Therefore, when I present the sensitivity analyses a bit later, I will also show revised analyses using all available data at 52 weeks, including 52-week data for those 28 patients. We believe this is important, as it minimizes missing data.

The SENSCIS trial included a spectrum of patients who are representative of those who would be treated in clinical practice. The study population consisted primarily of women in their early to mid-50s. Most were white with approximately 25 percent Asian, and 6 percent black or African American. Overall, baseline

characteristics were well balanced between the two treatment groups.

Disease characteristics with regard to FVC were representative of the diverse patient population routinely seen in clinics and were balanced between treatment arms. Slightly more than 50 percent of the study population had diffuse cutaneous SSc, with a mean mRSS of approximately 11 points, indicating mild extent of skin involvement.

The median time since first non-Raynaud symptom was approximately 3 and a half years, and 60 percent were ATA status positive. A significant portion of patients were on background immunosuppressive therapy with approximately 50 percent being on stable dose of mycophenolate or corticosteroid at baseline.

Baseline prominent characteristics were also balanced across two treatment groups and were consistent with that of a population with mild to moderate lung function impairment. The mean extent of fibrotic ILD was approximately 35 percent with most patients having evidence of reticulation with

or without ground glass opacity, and 15 percent were findings of honeycombing in HRCT. There was a slightly higher mean FVC at baseline in the placebo group. However, the mean FVC and DLCO percent predicted were balanced.

Now for the study results. As previously mentioned, the SENSCIS trial is the first positive registration trial in SSc-ILD having met the prespecified primary endpoint. Treatment with nintedanib led to preservation of lung function by significantly reducing the annual rate of FVC decline over 52 weeks.

Looking at the results of the primary endpoint compared to a placebo decline of 93 milliliters per year, treatment with nintedanib was associated with an annual rate of FVC decline of 52 mL per year. This corresponds to a significant difference of 41 milliliters compared to placebo, which is equivalent to a relative reduction of 44 percent.

Although the overall rate of FVC declined and absolute reduction in this population is less

than that observed in IPF, the relative treatment effect is consistent with what we've seen in the IPF trials. Furthermore, this treatment effect is considered meaningful, as ILD progression is associated with increased mortality.

Similar findings were also seen when we look at other FVC endpoints. Looking now at the rate of FVC as present predicted rather than in milliliters, we see that compared to a placebo decline of 2.6 percent, treatment with nintedanib was associated with an annual rate of FVC decline of 1.4 percent.

This translates into a 46 percent relative reduction, which is consistent with the primary findings. And in addition, the magnitude of effect is in line with that reported for the SLS I trial with cyclophosphamide, as presented earlier by Dr. Seibold. Similarly, nintedanib treatment also was associated with 46 percent relative reduction in the absolute change from baseline in FVC compared to placebo at 52 weeks, as shown here.

Looking now at the trend in FVC over time,

for those treated with nintedanib in the top curve compared to placebo, you can see that the curves separate early and continue to diverge up to 52 weeks, indicating sustained benefit over time.

To get a better picture of the impact of treatment on loss of lung function, we also looked at the proportion of patients meeting various cutoffs of FVC change at 52 weeks. The following analysis uses worst observation carried forward for those without an FVC measurement recorded at week 52, which is slightly different from the analysis in the FDA briefing book.

Bars on the left indicate every see worsening, while the bars on the right show FVC improvement. Overall, patients on placebo were more likely to worsen while patients treated with nintedanib were more likely to improve.

Although there is no established MCID, or minimal clinically important difference, for change in FVC for SSc-ILD, a recent publication by Kafaja and colleagues, using data from SLS I and II, identified a range of FVC cutoffs potentially

correlated with several patient-reported outcome measures.

We conducted a post hoc analysis using the lower cutoff of FVC change proposed in this article, which is shown here. The graph shows the proportion of patients having 3.3 percent or greater decline in FVC classified as disease deterioration on the left, and those with a 3 percent or more improvement in FVC on the right. Consistent with the primary findings, these results support the treatment that treatment with nintedanib is associated with meaningful slowing of disease progression.

Although the study was not powered to look at individual subgroups, as is normally done, the primary endpoint was investigated in several prespecified subgroups to confirm consistency of the observed treatment effect. This shows subgroup analyses based on prognostic factors that have been associated with ILD progression in the literature, including ATA status, SSc subtype, baseline FVC percent predicted, and extent of fibrotic ILD on

HRCT.

The findings from the various subgroups are consistent with the overall analysis as evidenced by the broadly overlapping confidence intervals and high interaction p-values far from 0.05. This provides reassurance that the treatment effect is consistent across subgroups, including ATA status, which was used to stratify randomization in the trial.

We also conducted subgroup analyses based on relevant patient demographics as required for regulatory submission, and this included age, gender, race and region. Mycophenolate used at baseline was also prespecified as a subgroup given its prevalent use in SSc. Here again, you see that the findings from all subgroups are consistent with the primary analysis with broadly overlapping confidence intervals and high interactions p-value, again, providing reassurance that the treatment effect is consistent across these subgroups.

While we should not interpret findings from any individual subgroup, since you are being asked

to consider the FVC data in the North American region and mycophenolate subgroups, I would like to briefly share some additional data on this subgroup to assist with your assessment.

Obviously, the study was not designed to assess the effect of mycophenolate on lung function. Based on the prevalent use mycophenolate in the U.S., we conducted a prespecified analysis to look at whether treatment effect of nintedanib differed based on mycophenolate use at baseline.

This graph shows the annual rate of decline in FVC by mycophenolate use in patients who had been stable on mycophenolate for at least 6 months on the left compared to those not taking mycophenolate at baseline on the right.

Looking first at the placebo groups, we see that the rate of FVC decline in patients on stable mycophenolate treatment is less than that observed in those not taking mycophenolate. However, while the lower rate of FVC decline in the placebo group on mycophenolate led to a lower absolute treatment difference, the relative treatment effect was

comparable between both subgroups and is consistent with the relative treatment effect of 44 percent seen in the overall trial.

Based on this, we conclude that the benefit of nintedanib is independent of mycophenolate use. Of note, in patients treated with both nintedanib and mycophenolate, the rate of FVC decline was close to the expected age related decline for healthy population.

With regard to the U.S. and Canada, in general, patients from U.S. and Canada were comparable to the overall study population with the exception of a high proportion of African Americans at 15 percent and a higher proportion of mycophenolate use in 80 percent.

With regard to the FVC results, in addition to the prespecified subgroup analysis, we also conducted categorical analysis, which shows significant consistency with the overall results. Looking again at the MCID threshold previously presented, with long function deterioration on the left and improvement on the right, we see

proportionally fewer patients on nintedanib had deterioration in lung function and more patients had improvement compared to placebo, thus, further supporting that the benefit of nintedanib treatment in patients with SSc-ILD also applies to the U.S. and Canada.

Now for the key secondary endpoints. As mentioned, in addition to its impact on lung function, we also wanted to evaluate the potential systemic effect of nintedanib on skin fibrosis and overall quality of life. First looking at the modified Rodnan skin score, there was approximately a 2-point or 4 percent decrease in mRSS at 52 weeks compared to baseline in both treatment groups with no significant difference between the groups.

With regards to quality of life, there was minimal to no change in SGRQ in both treatment groups at week 52 compared to baseline. The minimal change in SGRQ of less than 1 percent at 52 weeks is considered to be within the measurement error for the tool.

Similarly, as detailed in the briefing

document, no meaningful between-group differences were observed in the other efficacy endpoints. However, when interpreting the results of the patient-reported outcome measures, it is important to understand the challenges in demonstrating the benefit of a treatment that stabilizes or slows lung function decline in a chronic disease such as SSc-ILD.

Due to the low number of events during the study, we cannot draw any definitive conclusions related to the impact of nintedanib on mortality. However, there was no difference in mortality observed between the two treatment groups.

Before concluding, I would like to present the following additional analyses that were done to further explore the data and confirm our conclusions.

As detailed in our briefing document, we conducted two separate sensitivity analyses of the primary endpoint. This shows the results from the prespecified sensitivity analyses using three different imputation approaches for missing data.

Despite making conservative assumption regarding patients with missing data on nintedanib, it is reassuring to see consistency of these sensitivity analyses with the primary analysis shown at the top and the broadly overlapping confidence intervals.

When we revised the insensitivity analyses to include all available observed data, including data from the 28 patients who had FVC measurements just outside the week 52 window, you can see the results remain consistent.

In addition to the prespecified sensitivity analyses, we also conducted a post hoc tipping-point analysis at the request of the FDA. This was done to assess how robust the data are across various missing data assumptions. The intent of this analysis was to determine what magnitude of FVC decline would be required in patients with missing data on the nintedanib, assuming no change in the placebo group, in order to lose statistical significance of the primary endpoint.

Both our and the FDA's analyses show that a

penalty of at least 30 mL per year for patients with missing data in the nintedanib group would be required for the trial to lose statistical significance. When we revised this analysis to include all available data, imputing data only for the 50 patients who truly did not have FVC measurements at week 52, we see that the penalty required to lose significance is 120 milliliter per year for patients with missing data on nintedanib. We believe this estimate is the most appropriate, as it minimizes imputation for missing data.

Taking under consideration the totality of the data, we therefore conclude that the results are robust. The magnitude of the penalty required to lose statistical significance seems implausible given the results from the primary analysis.

Lastly, even though complete 100-week follow-up data are available for only 25 percent of patients, we conducted exploratory analyses to evaluate the effective of nintedanib over the entire trial. Here you see the estimated treatment difference is 65 mL approximately at 100 weeks

using an intention-to-treat approach, which provides a conservative estimate. Although the estimated treatment effect out to 100 weeks varied depending on the statistical approach used, as described in our briefing book, the effect of nintedanib on lung function decline appears to be sustained beyond 52 weeks.

In summary, SENSCIS is the largest SSc-ILD study to date and the first positive registration trial for this indication. Similar to its demonstrated effect in IPF, nintedanib was shown to reduce ILD progression in patients with SSc-ILD with a 44 percent relative reduction in the annual rates of FVC decline over 52 weeks in a clinically representative population of patients with SSc-ILD.

The benefit was consistent across all subgroups, including patients on background mycophenolate, different SSc subtypes, varying severity of lung disease, and across all regions. In addition, the findings support the robustness of the results, and exploratory analysis beyond 52 weeks suggests that the benefit is sustained.

There was no effect of nintedanib on secondary en points, including MRSS and the SGRQ. Nevertheless, the observed benefit on lung function is considered clinically meaningful, as ILD progression is the leading cause of death in this patient population, and since slowing FVC decline has been associated with improved outcomes in IPF, we expect similar benefit for patients with SSc-ILD.

I will now like to invite Dr. Kohlbrenner to the podium.

Applicant Presentation - Veronika Kohlbrenner

DR. KOHLBRENNER: Good morning. My Name is Veronika Kohlbrenner, and I'm a physician in the global pharmacovigilance department at Boehringer Ingelheim. I will provide an overview of the safety data from the SENSCIS trial. As you will see, there is a high level of confidence that the safety data in the new patient population with systemic sclerosis is consistent with the known safety profile of nintedanib in patients with IPF, as established in the INPULSIS trials.

These are the topics I will cover in my presentation. First, I will review exposure in the SSc-ILD population in the SENSCIS trial, then I will provide a summary of overall adverse events reported in SENSCIS compared to the INPULSIS trials to demonstrate the consistency of the safety data across both the SSc-ILD in IPF patient populations. Further, I will provide details of safety topics of special interest in the SENSCIS trial.

Exposure to nintedanib in the SENSCIS trial was substantial. As you have heard, patients were followed for a minimum of 52 weeks and some for up to 100 weeks. Based on the time of randomization into the trial, the duration of the study varied for each patient. Mean exposure for the 52-week treatment period was about 11 months in both treatment groups, and mean exposure for the entire trial was about 15 months. Importantly, approximately 40 percent of patients had exposure greater than 18 months.

To demonstrate the consistency of safety in the SSc-ILD population compared to the IPF

population, I will show data for both SENSCIS and the INPULSIS trials. As you can see, there was an overall similar frequency of any adverse event reported in both populations. There was also comparable rate of discontinuation due to adverse events, with discontinuation mainly due to gastrointestinal adverse events.

Serious adverse event reports were less frequent in the SSc-ILD patients than in IPF patients. Overall, the safety profile in patients with SSc-ILD was reassuring and consistent with what has been observed in the IPF population.

This slide shows the most common adverse events in both SENSCIS and the INPULSIS trials. Gastrointestinal side effects of diarrhea, nausea, vomiting, and abdominal pain were the most frequently reported events in both populations. Of note, gastrointestinal events were more common in SSc-ILD patients compared with IPF patients. However, this holds true for both the placebo and the nintedanib group and likely reflects the underlying GI symptoms of patients with systemic

sclerosis.

Weight loss is of concern in patients with systemic sclerosis. Weight decreased and decreased appetite are known side effects of nintedanib and occurred with similar frequency in the SENSCIS trial as compared in the INPULSIS trial. Notably, no serious weight loss was reported in SENSCIS.

As expected, skin ulcers were reported only in SENSCIS, but were reported with similar frequency in the nintedanib and placebo groups. Other common events reported were in the respiratory system and were in line with the underlying interstitial lung disease. In general, the data in patients with SSc-ILD showed consistency with the IPF population.

Now, I'd like to turn your attention to safety topics of special interests, which were defined by the safety experience in the INPULSIS trial. Diarrhea and hepatic events are presented due to their frequency of occurrence and their importance in patient management. Bleeding and cardiovascular events are presented because they

have been associated with the mechanism of action of VEGF inhibitors.

As you can see, hepatic events, bleeding events, and cardiovascular safety were comparable amongst SSc-ILD in IPF patients. Now I will describe in more detail each of these safety topics specific to patients in the SENSCIS trial.

Recognizing the concerns regarding diarrhea that are particular to patients with systemic sclerosis, this provides more details around this commonly reported event. As I have mentioned, diarrhea is the most frequently reported adverse event with use of nintedanib reported in 75 percent of patients in SENSCIS. In the majority of patients, diarrhea was characterized as mild or moderate. Four percent of patients reported severe diarrhea.

Diarrhea was initially managed with antidiarrheal medication and assurance of adequate hydration. As needed, dose interruption followed by dose reduction were employed. Of the 75 percent of patients on nintedanib who experienced diarrhea,

dose reduction was instituted in 20 percent.

With these mitigation measures, most patients were able to continue in the trial. Seven percent prematurely discontinued nintedanib due to diarrhea. Diarrhea was manageable, and the mitigation measures enabled continuation of nintedanib in the majority of patients with recovery reported in over 90 percent.

Shown here are all hepatic events that were reported as adverse events, which included predominantly mild transient liver enzyme elevations. Investigators were instructed to report liver laboratory abnormalities as adverse events, regardless of the level, if they were considered clinically relevant.

Ninety-four percent of hepatic events were non-serious. For treatment management guidelines, 4 percent of nintedanib-treated patients were dose reduced and 2 percent discontinued treatment due to hepatic events. There were no cases of liver failure and there were no liver related deaths.

This shows the laboratory test results for

liver enzyme elevations. Most occurred early after the start of treatment. Therefore, laboratory testing is recommended routinely in the first 3 months and periodically thereafter. The majority of liver transaminase elevations were less than 3 times the upper limit of normal. Five percent of nintedanib- treated patients experienced and/or AST elevations at or above the 3 times the upper-limit-of-normal threshold. Of those, 3 patients had elevations above 5 times the upper limit of normal.

There were no cases that met Hy's law constellation predictive of liver failure, and for patients who dose reduced or discontinued nintedanib due to liver enzyme elevations, liver laboratory abnormalities completely resolved.

Bleeding events where predominantly mild and non-serious, and they occurred with similar frequency in both nintedanib and placebo groups. The most frequent bleeding events were epistaxis, skin contusions, or rectal hemorrhage. There were 2 bleeding events involving the central nervous

system in the nintedanib group. Both events had clear precipitating factors, and study treatment was able to be continued uninterrupted in both patients. All patients with bleeding events were able to continue treatment uninterrupted.

Overall, cardiovascular events were rare in SENSCIS, and there was no imbalance amongst treatment groups. MACE events, as reported by the investigator, were balanced between the two treatment groups. An independent committee reviewed these and adjudicated 3 events in the placebo group as MACE versus one event in the nintedanib group.

There was no imbalance of cardiac failure or venous thromboembolism. Pulmonary embolism was not reported in the nintedanib group. Pulmonary arterial hypertension was noted with low frequency in both treatment groups. Hypertension is a known side effect of nintedanib. Although the numbers are small, these data are reassuring for patients with SSc-ILD.

In summary, the safety and tolerability

profile of nintedanib in patients with SSc-ILD in the SENSCIS trial was consistent with that observed in patients with IPF. There were no new safety findings for nintedanib in SENSCIS. The common adverse events associated with nintedanib were manageable with existing strategies as outlined in the product label. The safety results support the treatment of patients with systemic sclerosis-associated interstitial lung disease.

Now, I will turn the podium back to Dr. Tetzlaff to briefly summarize benefit-risk.

Applicant Presentation - Kay Tetzlaff

DR. TETZLAFF: Thank you, Dr. Kohlbrenner.

Now, I'd like to share our perspective regarding the overall benefit-risk profile of nintedanib in SSc-ILD. Fibrosing interstitial lung disease is a common manifestation of SSc that is associated with early mortality. Progression of pulmonary fibrosis is irreversible. No approved treatment exists to slow down the accelerated loss of lung function associated with pulmonary fibrosis.

As presented today, the SENSCIS trial showed that nintedanib significantly reduced the annual rate of decline in FVC by 44 percent relative to placebo in a patient population with SSc-ILD that was representative for patients seen in clinical practice.

The relative treatment effect was in the same range as has been observed in patients with IPF in the INPULSIS trials. This figure is illustrating the consistency of the relative effect obtained in the 3 placebo-controlled 52-wee phase 3 studies of nintedanib in fibrosing interstitial lung disease, namely the two INPULSIS trials in IPF and the SENSCIS trial in SSc-ILD.

Why we did not see significant changes in skin fibrosis symptoms and health related quality of life over one year of treatment, the effect of nintedanib on slowing lung function decline in SSc-ILD patients is clinically important, given the typical age of onset of SSc-ILD and the natural progression, with gradual and irreversible loss of lung function accumulating over years.

As we've just heard, nintedanib was safe and well tolerated in the SSc population, and the safety profile was comparable to the experience in IPF. Hence, we conclude that nintedanib has a positive benefit-risk profile in patients with systemic sclerosis-associated interstitial lung disease.

Now, I'd asked Dr. Brown to provide his perspective on what nintedanib may add to the physicians' armamentarium in treating SSc-ILD.

Applicant Presentation - Kevin Brown

DR. BROWN: Good morning. My name is Kevin Brown. I'm a lung doctor in Denver, Colorado, where I'm also a professor of medicine at National Jewish health. I'm a paid consultant for the sponsor. I have no financial interest in the outcome of this trial.

More importantly, I want to point out a couple of things, number one, being much younger than Dr. Seibold, I've only been thinking about lung fibrosis for the past 30 years, both what causes it, and more importantly how to control it;

really, the latter I've been pretty poor at. Secondly, I want to thank the committee for their service. I realize that this is an imposition on some of you, but it's greatly appreciated.

When I think about scleroderma, I think about the following. Two scleroderma patients of mine, one on your left without evidence of significant lung disease, and no one would call you a liar if you said there's a problem with the one on the right. You do not need to be a thoracic radiologist to understand that the radiograph on the right is abnormal. It is fibrotic, and it is likely associated with significant impairment in someone's quality of life, their functional status, and potentially their long-term outcome.

If you were to do a surgical lung biopsy on those two CT scans, you might look at the one on your left and say that looks pretty normal, and you would be correct. But then the one on the right is what we call non-specific interstitial pneumonia, a fibrosing interstitial lung disease known to occur in scleroderma and associated with shortened

survival.

I'll show you the natural history of one of my patient's loss of forced vital capacity over time. And you could ask me why in the world did you watch this happen? You must have tried something, and the answer is, yes, we did try something.

We tried the following. We tried 18 months of cyclophosphamide, and while there was an initial sense that probably we had attenuated the loss of lung function after I'd created early menopause and the oncogenic potential, we switched to an alternative, azathioprine, initially, prior to the mycophenolate data.

Recognizing that there was a sense of maybe stability during that time, the mycophenolate data became available, and we said let's switch to that. But now with the benefit of hindsight, one could see that there's probably no effect on what we did with her rate of FVC decline. Rituximab had early data, and again, no obvious benefit to that therapy.

So we know this is an issue. Data from 20 years ago, if you looked at patients with scleroderma without heart, lung, or kidney disease and saw a 15 percent mortality at 10 years, the simple presence of interstitial lung disease, absent any of the other end-organ damage, you saw a third of patients dead after 10 years.

More recent data from the Norwegian study, a national study, looking at patients who had normal lung function at baseline and a little bit of fibrosis, at 10 years, their mortality was the same despite half of these patients being treated with immunosuppressive therapy; three-quarters of these patients dying of their underlying lung disease.

Even in the Scleroderma Lung Study, one population, recognizing that these patients got treated with cyclophosphamide for at least a year and likely got treated with additional immunosuppressive therapy, after 10 years, the mortality has not changed. Now, there's always a risk of comparing studies that were not performed identically over time, but this is not the progress

that we would like to see.

In this most recent American Thoracic Society pro-con debate, when all other therapies have failed in scleroderma-associated interstitial lung disease, we'd like to be able to say at least we have lung transplantation to offer; that it was a legitimate question to ask should that therapy even be offered to patients with scleroderma-associated lung disease at the light stage of their disease? Because of the adverse effects associated with the therapy, the likelihood of long-term outcome, and whether or not transplant even should be a treatment option was a legitimate topic for debate.

That brings us the SENSCIS. This was a giant trial when we think about scleroderma-associated interstitial lung disease, and most importantly, from my perspective, it embroiled a broad population of patients, the patients that I see, and those of us who see interstitial lung disease see in their clinics. It did not exclude patients who are already on

therapy, therapy that some believe and obviously works for some.

When I think about this as a clinical trial person or a researcher, I know that when you look at a population decline curve, what in fact is happening is that there is a huge population of individual patients who are changing over time; that their FVC is individually changing, and not all at the same rate.

We can recognize this and we can make up curves with patients who are relatively stable and patients who are more rapidly declining. Clearly, those patients who are more rapidly declining in terms of the decline in their FVC are those who are at the greatest risk of death.

So when I think about SENSCIS, I put it in the context of those patients with fibrosing lung disease that I see, particularly idiopathic pulmonary fibrosis. We know that benefit accrues in terms of saving FVC over time; that it accumulates over weeks; that the benefits seen in terms of the relative preservation of FVC is

similar to what we see in idiopathic pulmonary fibrosis.

When I think about, rather than the population, individual patients that I see, recognizing that it is always risky to look at individual subsets, particularly when they're not powered, I can see that regardless of serologic status if I'm seeing a patient, regardless of their scleroderma subtype, regardless of the severity of their physiologic impairment or the extent of fibrosis on their CT scan, that they are likely to receive the same benefit over time.

Most importantly for me, those patients who come to me who have the most rapidly worsening disease as measured by a decline in their FVC, that it appears that there is a 30 percent lower risk to fall into that category if you're on active therapy.

In the end, with scleroderma-associated interstitial lung disease, where are we? Patients with scleroderma are affected in the prime of their lives. They're parents, they're children, they're

siblings, they are employers, they're employees, and they are caregivers. All of the major personal relationships in their lives are affected by this disease.

Lung fibrosis without question is the leading cause of death, and there are no approved therapies for their disease. We recognize that unapproved immunosuppressive therapies may provide short-term benefit in selected subsets, and in some maybe provide some long-term benefit. But most importantly, as with IPF, prevention or slowing of disease progression measured by FVC is the therapeutic goal.

A patient can reasonably ask me, "If I have lung fibrosis, if I'm going to progress like I have IPF, if I'm going to die like I have IPF, why shouldn't I be treated like I have IPF?" Effective antifibrotic therapies is what's needed.

So in the end, the way I think about it is as follows. While I always hope for a cure, that progress is painfully slow, it is intermittent, and it is never perfect. But this is what progress

looks like. Thank you very much.

Clarifying Questions

DR. SOLOMON: Well, thank you very much to the applicant and the speakers. That was a great overview. We now have about 15 minutes for clarifying questions; not discussion, but clarifying questions. If you can remember to state your name for the record before you speak, and if you can, please direct questions to a specific presenter. Yinghua will be taking a list, so just raise your hand, and we'll try to get through as many as we can.

DR. BECKER: Hi. Good morning. I'm Mara Becker. I had a question in the efficacy presentation, specifically slide CE-7. Just to clarify, I'm not sure if I misunderstood, it looked as if it was presented that patients who were on cyclophosphamide, azathioprine, rituximab, or cyclosporine were excluded. However, I thought I heard that it was permitted for clinicians to use these agents if there was clinical deterioration during the course of the trial.

I first wanted to clarify is that correct. And if so, do you have any data on if these agents were used between the placebo group or the active drug group?

DR. TETZLAFF: Yes, we do have the data, and I'd asked Dr. Clerisme-Beaty to come up to the podium to respond to your question.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. Indeed, that is correct. While we did restrict the use of these medications at baseline, we did allow the introduction of these during the study to manage events.

Can we have the efficacy summary, please? We're putting up the slide. Overall, in regards to the added restriction, during the study, about 9 patients on placebo and 11 patients on nintedanib were introduced to one of these medications. And we're trying to pull up the slides for you to look at the numbers.

Maybe we'll bring it later, but 9 and 11, respectively.

DR. SOLOMON: Dr. Calhoun?

DR. CALHOUN: Thanks. Bill Calhoun. I'm interested in the sponsor's conceptualization of this disease in the rationale we talked about, you folks talked about, this being a fibrotic lung disease, and that fibrosis was important. And you used the IPF data as a rationale.

The pathology, and one might legitimately argue the pathogenesis of fibrosis, and certainly the pathology, is different in interstitial lung disease related to scleroderma. The temporal heterogeneity that's present in IPF is not present in scleroderma-associated fibrosis. There's a bunch of ground glass opacities present on your films. I think 88 or 90 percent of them had GGOs as well. Then in one of your therapeutic subsets, the statistically significant benefit of nintedanib was lost in those who were on mycophenolate, which principally is anti-inflammatory.

So again, I'm a little confused about how you think this is working. If you're arguing that this is an antifibrotic drug and yet the benefit is

lost when patients are on an anti-inflammatory drug, and the pathology of interstitial fibrosis related to systemic sclerosis is really fundamentally different, then that seen with IPF, I'm confused about what the rationale is.

DR. TETZLAFF: Yes. Thank you for your question. Let me just clarify that the study was not powered to look at particular subgroups, such as I'm looking at the effect of the drug on mycophenolate. The use of mycophenolate at baseline was allowed to include a broad population of patients and not withdraw patients from a drug that is used as standard of care.

In terms of your comments on the rationale as to what an antifibrotic may add to this armamentarium, I'd ask our expert, Dr. Maher, to come to the podium and provide his clinical insight. He also happened to be an investigator on the SENSCIS trial.

DR. MAHER: Hi. I'm Ted Maher. I'm a pulmonologist at the Royal Brompton Hospital in London, Imperial College, London. I'm a paid

advisor to the sponsor, but I have no financial interest in the outcome of this meeting.

So I think you ask a valid question about the comparison between IPF and scleroderma. I think increasingly we recognize that when we look at the spectrum of fibrosing lung diseases, there are three components that predispose an individual to develop fibrosis. One is injury to the lung, the second is genetics, and the third is aging. I think in different disorders, those components vary. And I think in systemic sclerosis, our assumption is that its immune-mediated injury is the primary driver of developing fibrosis. But at the end of the day, fibrosis is an injury response, and it manifests only in a certain number of ways.

You also point out the sort of distinction between NSIP and UIP, and I apologize to the rheumatologists for the excess of acronyms that we have in interstitial lung disease. But essentially, these are different patterns that lie on a spectrum. When you look at fibrotic lung under the microscope, you can see NSIP where

typically there's preservation of the alveolar spaces, all the way through to UIP, where you actually get destruction of the lung with honey comb change.

The reality is, if you look at the whole organ, you will tend to see patchy areas where some of it looks more like UIP and patchy areas where some looks like NSIP. At a molecular level, these conditions behave in the same way.

So I think we've been prone in the past to making slightly artificial distinctions between groups of patients with fibrosing lung disease, when increasingly we're realizing that in practice, they behave in much the same way over time; albeit, Uh, I think the pathology or histopathology does speak to disease trajectory.

So I think the more honeycombing in UIP, you've got the more rapid progression; more NSIP, the slower the progression. But at the end of the day, the molecular process occurring in the lung and the destruction of the lung tissue is almost identical.

DR. SOLOMON: Thank you. Dr. May?

DR. MAY: Susanne May. I have a question with regard to the data specifically for U.S. and Canada. That was slide CE-25. Could you clarify how the responders were defined, and do you have data on the patients that were not included in this analysis?

DR. TETZLAFF: I'd ask Dr. Clerisme-Beaty to come to the podium and provide some further explanation.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. This analysis shows the responder analysis based on the cutoffs recommended by the Kafaja publication. More than 3 percent, or more, 3.3 percent decline in FVC, was considered deterioration, based on this publication, and more than 3 percent improvement was shown for this in placebo. This uses worst observation carried forward, and that includes all patients.

DR. SOLOMON: Can I just follow up on this U.S. analysis? Was this prespecified?

DR. CLERISME-BEATY: No. This was not. It was a post hoc analysis, given that the publication was recently published.

DR. SOLOMON: One follow-up on that. Those differences are not statistically significant, correct?

DR. TETZLAFF: I'd ask our project statistician, Dr. Voss to --

DR. VOSS: Florian Voss from Boehringer Ingelheim. These analyses, these are subgroup analyses of a post hoc analysis, so they are only based on U.S. and Canada patients. There was no statistical test performed. This is an exploratory analysis, and the study was not designed or powered such that these analyses are tested in a confirmatory manner.

DR. SOLOMON: Alyce Oliver?

DR. OLIVER: Hi. Alyce Oliver. Dr. Seibold mentioned putative risk factors for ILD progression, and included HRCT, extent greater than 20 percent and an FVC of less than 70 percent predicted.

My clarifying question has to do with slide CE-14, where the mean FVC was 72.7 in the placebo and 72.4 in the study group. My question is the HRCT features, what was the HRCT extent? You mentioned the reticulation, the GGOs, and the honeycombing, but what was the overall extent?

DR. TETZLAFF: I'd like to ask Dr. Clerisme-Beaty to directly respond to that.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. The mean extent of fibrotic ILD is shown in the first line, which was 35 percent for placebo and 36 percent for those on nintedanib.

DR. SOLOMON: Dr. Katz?

DR. KATZ: James Katz. My clarifying question is for Dr. Kohlbrenner concerning slides CS-11, cardiovascular events. I want to ask about the hypertension patients. Were there any episodes of renal crisis, and how are these patients who developed hypertension treated, and what happened to them?

DR. TETZLAFF: Dr. Kohlbrenner?

DR. KOHLBRENNER: There was one patient in the nintedanib group that experienced scleroderma renal crisis. That patient was treated with ACE inhibitors, and after some initial improvement, due to multiple complications, over the clinical course, the patient ultimately died.

In terms of treatment for hypertension -- do we have a slide for the antihypertensives? This shows medication use for actually both hypertension and pulmonary hypertension. As you can see, antihypertensives used in this trial ranged across multiple different agents, but the antihypertensive use was comparable in the placebo and the nintedanib group.

DR. SOLOMON: Dr. Stoller?

DR. STOLLER: Yes. I have a clarifying question regarding CE-21, the responder analysis. Recognizing that this is secondary, as you're aware, Kafaja and colleagues described a spectrum of MCIDs, depending on the anchors, and you've picked some of them. For example, they described a range for decline of 3 to 3.3 percent and a range

of improvements from 3 to 5.3 percent. So I wonder whether you prepared the responder analysis using the other extremes of their MCID estimates.

DR. TETZLAFF: Dr. Clerisme-Beaty will respond to you.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. The Kafaja analysis was done post hoc because of publication came after the protocol was finalized. Prespecified in the protocol, we did look at responder analysis using the cutoff of 5 percent deterioration because that's what had been previously suggested for IPF.

We will try to pull up the slide. Basically, we show a similar pattern that, overall, fewer patients in nintedanib met the criteria for worsening. This is showing different cutoffs for responder analysis. Looking at the 5 percent predicted, more than 5 percent worsening or more than 10 percent worsening. Again, the numbers are fewer because fewer patients met those higher cutoffs, but in general, the trend was similar.

DR. STOLLER: I have a follow-on. Just for

clarity, I understand those thresholds for responder. That's not exactly the question I asked, which was the specific ranges for the MCID in the Kafaja paper.

DR. CLERISME-BEATY: No, we did not do that. What we do have are the categorical analyses that they showed as part of the core presentation, which I can show you, which is slightly different, but I can offer this to take a look at. This is shown here where we have cutoffs 0 to 5 percent, 5 to 10 percent, and 10 to 15 percent on the improvement side, and then the opposite for worsening; so slightly different but similar takeaway and summary.

DR. SOLOMON: The last question for clarifying at this stage is going to be Dr. Geller, then we're going to have a break. When we come back, we'll try to fit in some more clarifying questions later on, but I think it's just best to continue, so Dr. Geller.

DR. GELLER: Nancy Geller. I'm still confused by CE-21 and CE-25, and that's because all

of the patients in the middle are left out. You don't show the results for patients who didn't fall into either of those categories, the ones in between.

DR. TETZLAFF: Dr. Clerisme-Beaty?

DR. GELLER: And even though you have expectations that you'd get some results in the middle and not terribly different, it ain't there.

DR. TETZLAFF: We can offer you another view on the data, and Dr. Clerisme-Beaty will show that.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. This shows you data that you're requesting, which is another way of looking at the data, with the percent improvement, the 3 percent improvement on the top; the stable change, which is between minus 3 and 3, and then the proportion of patients with 3.3 percent of more decline. This adds up to the overall population.

DR. GELLER: And what about for the U.S. and Canada?

DR. CLERISME-BEATY: I do not believe we have this for the U.S. and Canada, per se, because,

again, the numbers got smaller. But we can take a look and provide this after the break.

DR. SOLOMON: Chair's prerogative, we're going to take a break. We'll come back in 15 minutes, and then if we have time after the FDA's presentation, we'll have some more clarifying questions. Thanks.

(Whereupon, at 10:44 a.m., a recess was taken.)

DR. SOLOMON: Well, we're off schedule already, so we'll just do our best. It's now time for the FDA to present.

FDA Presentation - Nadia Habal

DR. HABAL: Good morning. My name is Nadia Habal, and I'm a clinical reviewer in the Division of Pulmonary, Allergy, and Rheumatology Products. I'm also a practicing adult rheumatologist. I would like to thank the panel members for coming to share their expertise with us. We have heard the applicant's discussion on nintedanib, and the agency will now present its perspective on the efficacy and safety of nintedanib for systemic

sclerosis- associated interstitial lung disease.

This is an outline for the FDA presentation this morning. I will first begin by giving an overview of the clinical program for nintedanib in patients with systemic sclerosis interstitial lung disease. My colleague, Dr. Wang will then provide the statistical review of efficacy and detail, and then I will return to summarize safety and provide the benefit-risk considerations for discussion.

I will begin with an overview of the clinical program. As has been discussed, systemic sclerosis is a serious disease with considerable morbidity and mortality. The primary causes of SSc related death are cardiac and pulmonary complications of the disease. The target of this program was one of these pulmonary complications.

Currently, there are no approved therapies for patients with systemic sclerosis or systemic sclerosis-associated interstitial lung disease. Expert guidelines recommend consideration of immune suppressives such as cyclophosphamide and mycophenolate for the treatment of SSc-ILD.

Nintedanib is approved for idiopathic pulmonary fibrosis. We acknowledge that systemic sclerosis-associated interstitial lung disease as a disease process has similarities and differences from idiopathic pulmonary fibrosis. The similarities between the two disease processes include that they are both chronic, progressive, fibrotic lung diseases resulting in loss of pulmonary function and associated morbidity.

The two conditions, however, differ in demographics. Idiopathic pulmonary fibrosis is mainly seen in older men, whereas we see SSc-ILD in middle-aged women. In addition, the findings on high-resolution computed tomography and histology are different. For IPF, the classic signs on high RCT for usual interstitial pneumonitis include traction bronchiectasis with peripheral basilar predominant opacities and honeycombing.

In contrast, for SSc-ILD, nonspecific interstitial pneumonitis is associated with peripheral ground glass opacities. Idiopathic pulmonary fibrosis can have exacerbations, whereas

SSc-ILD is usually characterized more by a gradual decline. Finally, progression of IPF is more rapid than that of SSc-ILD with a shorter median survival. Despite these differences, both result in pulmonary fibrosis and associated morbidity and mortality.

I will now move on to the applicant's clinical development program for SSc-ILD, The clinical development program in SSc-ILD consisted of a single phase 3, double-blind, multicenter, placebo-controlled study to evaluate the efficacy and safety of oral nintedanib in patients with SSc-ILD.

576 patients were randomized one-to-one to treatment with nintedanib, 150 milligrams twice daily or placebo. Randomization was stratified by antitopoisomerase antibody status. The primary endpoint was the annual rate of decline in forced vital capacity, or FVC, in mL over 52 weeks.

Key secondary endpoints included absolute change in modified Rodnan skin score and absolute change in St. George's Respiratory Questionnaire at

week 52. The primary and key secondary endpoints were at week 52, but patients could remain on treatment up to a maximum of 100 weeks to collect follow-up efficacy and safety information, including mortality.

This is an informational slide about protocol-specified dose reduction, interruption, discontinuation, and rescue. You will hear more about how these could impact missing data in a later part of the FDA presentation.

The table shows the different reasons for modification of treatment in each of the categories shown. In the event of adverse events or liver enzyme elevations, dose reduction was considered. For adverse events considered drug related, treatment could interrupted for up to 4 weeks, and if adverse events were not considered drug related, treatment could be interrupted for up to 8 weeks.

Treatment discontinuation was to be considered if adverse events persisted at the lower dose or for severe adverse events and repeat elevated liver enzymes. Patients who experienced

clinically significant worsening could receive rescue therapy. Permitted rescue medications included prednisone greater than 10 milligrams per day and other immune suppressants.

Next, I will discuss the efficacy endpoints. As stated, the primary endpoint used was annual rate of decline in forced vital capacity in mL over 52 weeks. The key secondary endpoints were absolute change in modified Rodnan skin score at week 52 and absolute change in St. George's Respiratory Questionnaire at week 52.

Other secondary endpoints included time to death. Secondary endpoints related to pulmonary function and symptoms included annual rate of decline and percent predicted forced vital capacity, forced vital capacity in milliliters, and absolute change in DLCO percent predicted and FACIT dyspnea scale.

Secondary endpoints related to other systemic sclerosis disease manifestations and physical function included relative percent change in modified Rodnan skinned score, HAQ-DI total

score, CRISS index score, and digital ulcer net burden.

Next, I will review the demographics of study patients. The demographic characteristics of the two treatment arms were generally balanced, as summarized in this table. This was a predominantly white female population with a median age in the 50s. Overall, the patient demographic characteristics were balanced and representative of the intended patient population of SSc-ILD.

Of note, in this study population, 25 percent were from Canada and the United States. I mentioned this because it will be relevant to the discussion of efficacy in a later part of our presentation.

Next, I will review the baseline disease characteristics of patients from study 214. Baseline disease characteristics were similar between treatment groups. Sixty percent of the patients had antitopoisomerase antibodies. The mean time since first onset of non-Raynaud symptoms was 3 and a half years. Approximately half of the

patients had diffuse cutaneous systemic sclerosis and approximately half of the patients had limited cutaneous disease.

Measures of lung function, including mean percent predicted FVC and percent predicted DLCO were generally balanced by treatment group. The percentage of patients with pulmonary hypertension at screening and the mean baseline mRSS were also balanced by treatment group.

Prior digital ulcers were reported by 42 percent in nintedanib-treated patients as compared to 35 percent of placebo-treated patients. Other disease manifestations, including Raynaud phenomenon, diarrhea, malabsorption, bacterial overgrowth, esophageal dysphasia and reflux, and synovitis were balanced by treatment group.

At baseline, 48 percent of patients received treatment with mycophenolate. The use of mycophenolate at baseline will be a consideration as we discuss the efficacy results.

Next, I will review the disposition for the study patients. As shown in this table, there were

288 patients randomized in each treatment group. The trial's statistical analysis plan defined the time window for week 52 time point from week 44 to week 53. According to this definition, at week 52, the placebo arm had a study completion rate of 95 percent. The nintedanib arm had a slightly lower study completion rate of 92 percent.

There were more early study withdrawals, treatment discontinuations, dose reductions, and treatment interruptions in the nintedanib than in the placebo group. The most common reason for study withdrawal, discontinuation, dose reduction, and interruption was diarrhea in the nintedanib group.

I will now turn the presentation over to Dr. Wang for discussion of the efficacy results.

FDA Presentation - Yu Wang

DR. YU WANG: Thank you Dr. Habal.

Good morning. I'm Yu Wang. I'm FDA's statistical reviewer for this submission. I will present to you today our investigations on the robustness of treatment effect demonstrated by the

primary efficacy analysis and the collective evidence provided by study 214.

This presentation includes a review of key elements of the statistical analysis plan, an overview of patient disposition, a review of the primary and the key secondary analysis results, and I will end this presentation with a summary of our statistical review findings. I will begin with some important aspects of SAP.

Trial 214 was a randomized, double-blind, placebo-controlled trial designed to investigate the efficacy and safety of nintedanib in patients with interstitial lung disease associated with systemic sclerosis. Key efficacy endpoints were assessed over a 52-week period.

The primary endpoint was the annual rate of decline in forced vital capacity, or FVC, in milliliter over 52 weeks. The applicant predefined the multiple testing hierarchy. Also included is the following two key secondary endpoints: mean change from baseline is the modified Rodnan skin score or mRSS at week 52 and mean change from

baseline in the St. George's Respiratory Questionnaire total score, or SGRQ, at week 52.

The primary analysis model for annual rate of decline over 52 weeks is a restricted maximum likelihood based random coefficient regression model. A mixed model with the repeated measures approach was used as a primary analysis model for all changes from baseline endpoints in this study, including the two key secondary endpoints.

While missing data may have resulted from one of several mechanisms of missingness, these two analytical models both assumed missing at random; that is missingness may depend on observed covariates and outcomes. But given this, not on unobserved outcomes.

Aside from the primary endpoint, additional efficacy and endpoints to gain better understanding regarding treatment effect on FVC included the annual rate of decline in FVC in percent predicted over 52 weeks; responder analysis, based on change from baseline in mL and FVC in percent predicted.

Estimand is a target of estimation to

address the scientific question of interest posed by the trial objective. Study 214 targeted the de facto or treatment policy estimand, which is the difference in annual rate of decline in FVC, comparing all patients assigned to nintedanib to all patients assigned to placebo regardless of adherence to treatment or use of rescue therapies.

To evaluate this estimand, both on-treatment data, and where available off-treatment data were to be included in the analysis. While, the applicant prespecified analysis plan to include both on- and off-treatment data is consistent with the treatment policy principle, in a real-life clinical trial, unavoidably, there will always be some missing data despite all the planning and effort to prevent them.

Missing-at-random assumption in the primary analysis model is considered a strong and unverifiable assumption to explore the robustness of this inference, from the primary analysis as estimator to deviations from the underlying missing-at-random assumption. The Applicant

planned a series of sensitivity analysis, including approach that utilized the pattern-mixture modeling, or PMM, with multiple imputation.

The preplanned PMM sensitivity analyses do not comprehensively explore the plausible space of missing data assumptions. Therefore, the FDA review team requested an additional tipping-point analysis that systematically and comprehensively explores the space of plausible missing data assumptions.

Consistent with the treatment policy principle, the primary efficacy analysis population is a treated set, which included all randomized patients who received at least one dose of study medications. This set was also used for analysis of other efficacy and safety endpoints.

To control the type 1 error, a sequential testing procedure was used, so if a result was found to be statistically significant, then the next endpoint in the sequence will be tested. If the result for any of these endpoints was not statistically significant, then no subsequent test

will be performed.

In the next two slides, I will summarize patients' disposition in terms of trial medication discontinuation and the primary efficacy follow-up at week 52.

This table summarizes trial medication discontinuation status at week 52. In general, in trials for pulmonary drug, patients on placebo are more likely to discontinue than patients on study drug. However, in this study, we instead see that patients were more likely to discontinue the study drug than the placebo. In particular, the discontinuation rate was 19 percent for the study drug compared to 11 percent for the placebo.

Adverse events caused most of the discontinuations. Other reasons included patient's refusal to continue taking medication as trial medication and noncompliant with protocol.

This table displays primary efficacy follow-up status at week 52, according to FVC data availability, trial medication discontinuation status, and vital status at week 52. Despite the

off-treatment data retrieval plan and effort across two arms, there were roughly 86 percent of patients with their week 52 FVC data available, and there were 14 percent of patients with their week 52 FVC data missing. In the nintedanib arm, this rate was 17 percent, which was higher than the placebo arm. These four patterns described here were used in the pattern-mixture modeling approach in sensitivity analysis that will be presented later.

Now, we are going to look at the primary endpoint results. This table displays the primary analysis results. In the treated set, the adjusted rate of decline in FVC in mL was 52 in the nintedanib group versus 93 in the placebo group.

Compared with placebo, patients treated with nintedanib showed a statistically significant reduction in rate of decline, with an estimated rate difference of 41 milliliter per year. The comparison test in p-value was 0.035. While efficacy finding was statistically significant, we will further explore its robustness to assumptions with the missing data.

This figure displays the mean change from baseline on FVC in mL over 52 weeks by treatment group. Data are observed, the values. Vertical bars represent 95 percent confidence intervals. As we see in the primary analysis, the rate of decline for treatment appears to be less than that for placebo.

Given the statistically significant but small effect size from the primary analysis, which was based on strong and unverifiable missing-at-random assumptions for missing data, the FDA review team conducted a sensitivity analysis to assess treatment effect and the alternative missing data assumptions.

For supportive analyses in assessing the treatment effect of the nintedanib FVC, we also evaluated the treatment effect size in terms of FVC in percent predicted and conducted responder analysis of FVC change from baseline. Three of the analyses were preplanned in the study protocol and SAP, and additional tipping-point analysis was performed at FDA's request.

These are the three preplanned pattern mixture models. Each pattern mixture model assumes a certain level of deviation from the missing-at-random assumption but adopting an imputation rule based on observed FVC data using either on treatment referred to as pattern 1 or retrieved dropouts for patients who were off treatment, referred to as pattern 2.

For example, in pattern mixture model approach number 1, missing data for patients who are alive were imputed using retrieved dropout data from the same treatment arm. Missing data for patients who were deceased were imputed using the worst half of the retrieved dropout data from the placebo arm. Details with imputation algorithms for pattern mixture model number 2 and 3 are described in the briefing document.

While the applicant considers this imputation scenario conservative, we consider all three pattern mixture model imputation approach plausible alternatives to missing at random. As shown in this forest plot, all three 95 percent

confidence intervals cross the zero reference line, which indicates no significant treatment effect in any of them.

As previously mentioned, a tipping-point analysis was preformed to evaluate how robust the primary results were across a more comprehensive range of scenarios than was assumed in the pattern-mixture-modeling analysis.

In the analysis, the departures from missing at random assumption were investigated using the delta adjustment method; that is subjects who discontinued early would have, on average, efficacy outcomes after discontinuation shifted by some amount of delta compared to otherwise similar subjects, with observed data in their treatment arm.

The results over a relatively comprehensive range by arm shift values are summarized in this table. The header rows show the shifts applied to the dropouts in the placebo arm, which negative 60 means an additional 60 milliliter per year decline was imposed on the assumed background, the missing-

at-random rate of decline in placebo.

Similarly, the first columns show the same range of shifts applied to dropouts in the nintedanib arm. The body of the table provides p-values for the comparisons for the nintedanib group to the placebo group for the corresponding shifts. The blue box cell in this table corresponding to shifts of zero in both arms is analagous to the primary analysis under the missing-at-random assumption.

The pink region shows shifts, which are sufficient to tip the rate of decline conclusion; that is, the results are no longer statistically significant at 0.05 level. The blue shaded region shows cases where significance was maintained.

The red boxes correspond to a relative shift of a 45 milliliter per year in favor of placebo. From the previous primary analysis, remember, we saw a treatment effect of about 41 milliliter per year. So if the dropouts in nintedanib were assumed to progress at the rate seen in placebo, then nintedanib will not have a significant effect

in the overall trial. We would ask you to weigh in on the clinical plausibility of this relative shift.

Analysis results with FVC in percent predicted is consistent with the primary analysis on FVC in mL. In the treated set, the adjusted rate of decline in FVC in percent predicted was 1.4 in the nintedanib group versus 2.6 in the placebo group. Compared with placebo, patients treated with nintedanib showed a statistically significant reduction in rate of decline with an estimated rate difference of 1.2 percent per year and a p-value of 0.033.

In the protocol and SAP, the applicant also looked at response rates using the following two response definitions, where patients were considered responders if they had either a relative change from baseline, in FVC in mL of greater than 5 percent, or an absolute change from baseline, FVC in percent predicted of greater than 10 percent at week 52.

We consider responders as patients in the

opposite direction; that is, patients with a relative change from baseline in FVC greater or equal than a threshold. For example, patients with a relative decline from baseline in FVC in mL at week 52 of less than or equal to 5 percent were defined as responders, and we also examined the different thresholds, 5, 10, 15, for both FVC in mL and in present predicted.

As appointed to earlier in the applicant's presentation, we took a different approach in handling missing data. In this analysis, patients with the missing data at week 52 were categorized as non-responders.

We used the Cochran-Mantel- Haenszel model, adjusting for ATA status for each responder variable. The adjusted odds ratio with associated 95 percent confidence interval and the nominal p-values are reported here. None of the odds ratios are significantly different from 1.

Next, I'm going to use a graphical approach to illustrate the comparative treatment effect through empirical distribution plots. In doing so,

we could get a better view of how treatment effects that are measured in continuous form are translated to categorical or binary responder proportions.

This histogram shows the distribution of percent change from baseline in FVC in mL at week 52 by treatment. In this plot, missing data were represented in the group on the left, reflecting the assumption that missing data are worst outcomes. There are no obvious differences between the two arms.

To visually aid in the understanding of the responder analysis, these figures displace the proportions of responders at various response thresholds; that is, proportion of patients who percent change from baseline were greater or equal than certain thresholds, where missing data were imputed as a decline worse than that threshold.

For example, with a threshold of 10 percent decline, 72 percent of patients in the nintedanib group and 74 percent of patients in the placebo group had no more than 10 percent decline from baseline, in FVC in mL at week 52, indicating that

placebo is numerically favorable over nintedanib. On the other hand, with the thresholds of 5 percent decline, 59 percent on nintedanib and 52 percent on placebo had no more than 5 percent decline from baseline, indicating that nintedanib is numerically favorable over placebo.

Next, I will present the results for selected secondary endpoints. For the first the key secondary endpoint of absolute change from baseline in mRSS at week 52, there was a negative 0.2 difference between nintedanib and the placebo. This was not statistically significant given the sequential testing plan and any subsequent secondary endpoint were not considered statistically significant.

For the key secondary endpoint of absolute change from baseline in SGRQ total score at week 52, there was a 1.7 difference between the nintedanib group and the placebo group. These comparisons favors placebo.

Of the total 576 treated set patients, survival status at the end of the study was

followed up for 570 patients, with 6 lost to follow-up, one patient in the placebo group and 5 in the nintedanib group. There were 19 deaths in total across the two treatment groups at the end of the study, with the rest of the patients being censored.

This table summarizes the analysis results for the mortality endpoint through two approaches, crude rate of death and Cox proportional hazard regression model for time to death. The crude probability of death was 3 percent in the placebo group and 3.5 percent in the nintedanib group. The hazard ratio of the nintedanib group versus placebo group was 1.2 favoring placebo.

The applicant pre-planned the subgroup analysis for the primary and both key secondary efficacy endpoints with subgroups based on ATA status: age, gender, race, geographical region, MMF use at baseline, and SSc subtype. No significant interaction was found between treatment in any these subgroups at the 0.05 level of statistical significance.

As clinical practice may differ across countries, we also performed a subgroup analysis with subgroups defined by a cross-classification of region and the baseline MMF use to evaluate the influence of stable background MMF use to study treatment by region. The displayed forest plots show subgroup analysis by MMF use at baseline by region and by the cross-classification of region and baseline MMF use. There were smaller point estimates for U.S. and the Canada patients for MMF users at baseline.

To give an overview of the collective evidence provided by nintedanib for SSc-ILD phase 3 program, this table summarizes the efficacy analysis results for primary endpoint and the selected secondary endpoints in terms of estimated treatment effects, associated confidence intervals, and the p-values.

There was a statistically significant improvement for the primary endpoint in FVC in percent predicted. The difference for the first key secondary efficacy endpoint was not

statistically significant. The point estimate for SGRQ favored placebo and the responder analysis odds ratios close to 1.

In summary, the primary analysis result was statistically significant. Pattern mixture modeling sensitivity analysis, assuming certain missing not at random assumptions, showed a lack of robustness in the primary analysis result. Tipping-point analysis result needs clinical interpretation.

From analysis on other measures of FVC, results of FVC in percent predicted is consistent with the primary analysis result. In categorical analysis, defined by selected thresholds, treatment effects were not statistically significant. In subgroup analyses, smaller point estimates of treatment effect were observed in U.S. and Canada patients in patients who were MMF use at baseline. Results from secondary endpoints were not supported.

Thank you for listening, and now back to Dr. Habal.

FDA Presentation - Nadia Habal

DR. HABAL: Thank you, Dr. Wang.

I will be delivering the last presentation for the FDA this morning. Here's the outline for my presentation. I will provide an overview of the safety in study 214, including deaths, serious adverse events, treatment-emergent adverse events, and labeled adverse events. I will then summarize the agency conclusions on nintedanib and SSc-ILD. I will conclude by providing a framework upon which a discussion of overall benefit versus risk can be initiated.

The analysis of adverse events was based on treatment-emergent adverse events defined as all adverse events with an onset after the first dose of study medication, up to the end of the residual effect period of 28 days. This slide presents a safety summary for the first 52 weeks of the study. In addition,
the slide provides deaths for the patients on treatment and following discontinuation of treatment.

There were 19 deaths overall in the treated set balanced between treatment groups. During the treatment period, there were 11 patients with treatment-emergent adverse events leading to death, including 6 patients in the nintedanib group and 5 patients in the placebo group. During the post-treatment period, which was 29 days and over after the last drug intake, there were 4 adverse events leading to death in each treatment arm. I will discuss causes of death on the next slide.

There were more serious adverse events in the nintedanib group. The most common severe adverse events that occurred more frequently in the nintedanib group were diarrhea and pneumonia. The other severe adverse events occurred in 1 to 3 patients each. The incidence of adverse events leading to drug discontinuation and dose decrease was higher in the nintedanib group than in the placebo group. The most common reason for drug discontinuation and dose decrease in the nintedanib group was diarrhea.

I will elaborate more on serious adverse

events and any adverse events on subsequent slides. First, I will discuss the causes of deaths observed in the study. Overall, the types and frequencies of adverse events leading to death were balanced by treatment group in this study.

The causes of death were mostly related to cardiac and respiratory events. The causes of death in both groups were adjudicated as cardiovascular deaths, respiratory deaths, undetermined deaths, and non-cardiovascular or non-respiratory deaths. The types of events were consistent with the expected causes of death in this patient population.

Next, I will talk about serious adverse events. The most frequently reported serious adverse events in both groups were lung related and included interstitial lung disease, pulmonary hypertension, dyspnea, and pulmonary fibrosis. There were 8 patients with SAEs of pneumonia in the nintedanib group compared to one patient in the placebo group.

The difference in SAEs of pneumonia was not

observed in the IPF program. Differences in serious infections in the SSc-ILD program were driven by differences in pneumonia.

(Pause.)

DR. HABAL: Overall, adverse events of infections were similar between treatment groups in study 214. Infections occurred more frequently in patients with SSc-ILD than observed in the pooled IPF studies. This may be explained by the concomitant immune suppressive therapy of the patients with SSc-ILD.

The most frequently reported types of infections were nasal pharyngitis, upper respiratory tract infection, urinary tract infection, bronchitis, and influenza. Other than pneumonia, the types and frequencies of serious adverse events are balanced by treatment group in the treatment-emergent period.

Next, I will discuss treatment-emergent adverse events. The proportions of patients who had treatment-emergent adverse events were similar between the two treatment arms. However, higher

proportions of patients in the nintedanib treatment group had gastrointestinal adverse events, including diarrhea, nausea, vomiting, and abdominal pain.

In addition, over 52 weeks, based on weight measurements, more patients in the nintedanib group lost greater than 10 percent of their body weight at some point during the first 52 weeks of treatment.

Next, I will discuss adverse events that are labeled warnings and precautions in the approved nintedanib prescribing information. The label adverse events that I will focus on today include the adverse events identified in the clinical studies in IPF and included in the nintedanib labeling. These include elevated liver enzymes and drug-induced liver injury; diarrhea, nausea, and vomiting; arterial thromboembolic events, bleeding events; and gastrointestinal perforation.

In this table, more patients in the nintedanib group had elevated liver enzymes, diarrhea, nausea and vomiting, and bleeding events.

The most common bleeding events in both groups were epistaxis and skin contusion. Arterial thromboembolic events were rare and balanced. The labeled adverse events were consistent with nintedanib in idiopathic pulmonary fibrosis.

I will finish the safety portion of my talk with the safety conclusions on the next slide. The safety in study 214 was generally consistent with the known safety profile of nintedanib. Deaths were balanced between the treatment groups. Other than from pneumonia, the types and frequencies of serious adverse events are balanced by treatment group. The most frequently reported treatment-emergent adverse events in the nintedanib group were consistent with those known for nintedanib.

I will now provide a framework upon which a discussion of the overall benefit versus risk can now be initiated. SSc-ILD is a rare and serious disease associated with high morbidity and high mortality. It is also a disease with high unmet need for new therapies. FVC was selected as an

endpoint based on the experience with nintedanib and other products in IPF. Slowing of FVC decline in IPF was supported by clinically relevant secondary endpoints, including IPF exacerbations, the St. George's Respiratory Questionnaire, and trends improved mortality.

The SSc-ILD program showed a decrease in adjusted annual FVC decline. As previously noted, the observed decrease in FVC decline was not supported by improvement in other measures of pulmonary function or differences in mortality. However, the relative slowing of the rate of FVC decline at approximately 45 percent, compared with placebo and the SSc-ILD program, was similar to that seen in the IPF program, where support by clinically relevant secondary endpoints have been established.

I will next discuss risks. In study 214, the safety profile was generally consistent with the known safety profile of nintedanib in idiopathic pulmonary fibrosis. In the SSc-ILD program, there were more adverse events and serious

adverse events of pneumonia in the nintedanib group over 52 weeks. This was not observed in the IPF program. There were more serious infections reported in the nintedanib group driven by the differences in pneumonia. However, overall infections were similar in both treatment groups.

In summary, the adjusted annual rate of decline in FVC over 52 weeks was lower in the nintedanib group than in the placebo group, with a treatment difference of 41 mL per year. A less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients on mycophenolate at baseline, of 27 mL per year, and patients from the U.S. and Canada of 10 mL per year.

The decrease in the adjusted annual rate of decline in FVC was not supported by improvement in key secondary endpoints. Over 52 weeks of treatment, there were no differences observed between treatment groups in assessments of pulmonary symptoms, including SGRQ, DLCO, and FACIT dyspnea score. There were no differences in

assessments of SSc disease activity, including mRSS, digital ulcer net burden, and ACR CRISS. In addition, there was no difference observed in change in function or activities of daily living as assessed by the HAQ-DI. Mortality was also similar between treatment groups.

The safety of nintedanib in study 214 was generally consistent with the known safety profile of nintedanib, which includes risks of liver toxicity and GI disorders. In addition, there were more SAEs of pneumonia in the nintedanib treatment group as compared to the placebo group in the patients with SSc-ILD.

The overall risk-benefit for the use of nintedanib in SSc-ILD are rare and serious disease with unmet medical need, for which there are no approved therapies, are the primary topics of discussion for this AC meeting.

Clarifying Questions

DR. SOLOMON: Thank you.

We have a little time now for some clarifying questions, and we'll start with

clarifying questions for the FDA inquiry. We'll take a list. There are hands going up. I'm going to start over here with Dr. Nason.

DR. NASON: Thank you. Martha Mason. I actually have two questions. One is just to make sure I understand how the tipping-point analysis is done. I suppose that's a question for the statistician. But I just wanted to make sure that I understood that this was being compared to last observation carried forward for those people who were missing or to not including them in the analysis, and that the tipping point, the adjustment you would make would be to change the slope of the decrease relative to the time period that that data were missing.

For instance, someone who was only missing one month of data, 11 out of 12 months would be used, and then just the slope at the end would be changed in the way you're describing versus that the whole 12 months would be imputed. I just wanted to make sure I was understanding that right.

I'll just ask the second question in case

you guys can answer both. In the very beginning of this presentation, there was a description of rescue therapy being allowed under certain circumstances. That's the first time I've seen that mentioned, and I wanted to know a little bit more about that as far as what that meant, what was used, and also if anyone had looked at how that might differ between the groups, if one of the groups -- if rescue therapy was used more in either the placebo or treatment group.

DR. YU WANG: Thank you, Dr. Nason. Your first question was whether worst-case scenario was carried forward. No, it was not. For the tipping-point analysis, the background rate of decline was assumed to be the same as missing at random. So it was assumed as the same, with everyone who completed the week 52 trial.

DR. NASON: But still. the slope Was only changed at the point where they went missing; correct?

DR. YU WANG: The second question, yes. Basically, in my presentation, I simplified as a

scenario. Basically, during the imputation progress, there were several steps. One step was, as you have seen in the tipping-point analysis, the header row and the columns, those are shifts in units of milliliter per year. So these were translated to deltas for each visit, where the delta was determined by both the slope and the time lapse between two visits.

DR. NASON: Thank you. That's more clear, but I guess my follow-up would be, is there any information on whether the discontinuations were sort of spread out through the year evenly for the two groups, or I guess for the treatment group, those discontinuations were -- sorry, lost to follow-up as far as data; not discontinuation of the drug.

Did they in both groups spread through the year or were they concentrated at the beginning for the treatment group?

DR. YU WANG: I have something in my backup slides. Could you go to backup slide number 3 and number 4? No. Sorry; number 5. Yes, number 5 and

number 6. I will use number 5 first.

This is a missing pattern summary table for the nintedanib group. As you can see, the columns corresponds to study follow-up at scheduled visits and the rows corresponding to different missing patterns. The missing pattern was determined by both data availability and timing.

So you can see, at week 52, basically for each cell, X means data was available and the dot denotes missing. On the frequency row, highlighted in blue are the counts of missing for that pattern. So if you'll sum up all the dots across the column that corresponds to week 52, you have 47 patients in the nintedanib arm who had their week 52 observation missing. From the dot and X combination, you can see the missing patterns.

The next slide will be the same summary table for placebo. I can show you --

DR. SOLOMON: Dr. Weisman, next question?

DR. NASON: Sorry. There was a second question in there, though, that I'd asked about the rescue therapy.

DR. SOLOMON: Rescue therapy; sure.

DR. HABAL: Hi. This is Nadia Habal. I will answer the second question regarding rescue therapy. The options for rescue therapy were prednisone, over 10 milligrams per day; colchicine; azathioprine; cyclophosphamide; cyclosporine A; hydroxychloroquine; hydroxychloroquine D; penicillamine and sulfasalazine; rituximab; tocilizumab; abatacept; leflunomide; tacrolimus; tofacitinib; and potassium para-aminobenzoate.

I think when Dr. Becker asked the applicant before about other treatments, they said 9 on placebo and 10 on nintedanib, and they said they had a slide with who got what. So I will defer to the applicant for that slide if they have it.

DR. SOLOMON: Sure. Let's see the slide.

DR. TETZLAFF: Yes. We will be happy to show the slide now.

DR. SOLOMON: Okay, great.

DR. TETZLAFF: Our apologies for not showing it earlier. I'd ask Dr. Clerisme-Beaty to come up to the podium because we have some more details

that are displayed here, and I think this is what Dr. Becker asked for.

DR. CLERISME-BEATY: Emmanuelle Clerisme-Beaty, Boehringer Ingelheim. This table shows the restricted medication that was initiated during the treatment period, as well as active patient and those who discontinued treatment. You can see that the first two set of columns show those that was initiated while the patient was on study drug. For mycophenolate, it was 4 and 4; methotrexate was 1 in 1. In addition to the percentages reported on baseline, 4 more patients were initiated on mycophenolate.

For the other restricted medication, those that were not allowed at baseline, those word 9 in placebo and 11 in nintedanib, with the breakdown shown at the bottom. Then for patients who discontinued the study drug, that's what's referred in the other set of columns.

DR. NASON: Sorry. A quick follow-up [inaudible - off mic].

DR. SOLOMON: Go ahead.

DR. NASON: Sorry. A quick follow-up. Is there also information on the reasons for the rescue therapy? Because I noticed the reasons given in the FDA slide, one of them is absolute decline in FVC, which obviously could be related to the primary endpoint, and another is deterioration in other organ systems or clinical parameters, which that's pretty broad, to me anyway. That maybe could be related to the nausea and diarrhea people experienced or something else like that.

So is there any breakdown of the reasons the rescue therapy was given?

DR. CLERISME-BEATY: Within the protocol, investigators were guided that they could initiate additional therapy with, of course, the disease worsening. While we provided in the protocol a criteria for definition for a significant deterioration, at the end, the physician made the decision whether or not the patient needed therapy.

So technically, these patients could have met the protocol criteria or their physician felt that they needed it. So it's not directly

corresponding to those criteria, but there was guidance provided in the protocol. At the end, the investigator decided based on their judgment.

DR. SOLOMON: Dr. Weisman?

DR. WEISMAN: I have a question for the statistician.

Now, remember you're talking to me. Simplify it a little bit. I'm just a country doctor from Beverly Hills trying to understand complicated numbers.

You applied the same penalty to missing data regardless of whether it was active or placebo. It's just that the active had more missing data, so the penalty was harder on the active, and it invalidated the p-value. It adjusted it.

Did I get that right?

DR. YU WANG: So you are referring to the tipping-point analysis.

DR. WEISMAN: Well, just the overall assessment of missing data, you applied the same penalty to all of it. It's just that there was more missing data in the active.

DR. YU WANG: I have to see; not necessarily so for penalty -- more penalty to the study treatment arm.

We conducted two types of analyses. One is pattern-mixture modeling approach. Another is tipping-point analysis. Let's use pattern-mixture modeling approach number 1, as example. For missing data, for patients under the nintedanib arm, imputed rate was considered to be similar to retrieve the dropout in the same treatment arm, which means we assume patients' outcomes study discontinuation will be similar to the trend observed in patients who discontinued the treatment but continued the study follow-up in the same treatment arm. We consider this is a reasonable assumption.

DR. WEISMAN: The sponsor has said, well, there wasn't really that much more missing data. They adjusted their missing data in their presentation, saying, well, we collected it outside of a window. So if we apply that back to the window, there was less missing data.

Is that an interpretation of what the sponsor was telling us early?

DR. YU WANG: Yes.

DR. WEISMAN: So what do you think of that adjustment, and did that adjustment that the sponsor is suggesting take away some of the bite from the difficulty with the missing data that they were penalized for in your analysis?

DR. YU WANG: I think I can answer your question with two steps. First, yes. I consider the post week-52 data we'll be supportive, however, there are caveats. First, data quality may be in question. Second, the data will be very scarce, like limited -- sparse I should say -- compared within the first 52-week period.

That's my first reason for not using -- it's a post hoc analysis, so our review didn't take that approach to utilize the post week 52 data.

The second reason is there are reasons we consider lung function declined profiles comparable between the study drug arm, and there's a placebo arm post-discontinuation. I can show you a time

profile we found in pattern number 2 patients. Those patients are who discontinued the treatment during the first 52 weeks but complete with their week 52 follow-up.

Would you please show my backup slide number 2, so second of the backup slide.

This is similar to my presentation slide for the overall population. This is observed visit-wise, mean change from baseline, in FVC in mL over 52 weeks in pattern 2. You can see aside from the small sample size because these retrieved dropout patients is very small, 12 in placebo and 24 in nintedanib.

Their curves are parallel or entwined together; you cannot separate them, or my point estimate actually favors placebo. I haven't confirmed these numbers with sponsor, so they can correct me if I'm wrong. My point estimate for rate of decline in pattern number 2 is nintedanib had a negative 154 milliliter per year decline, and the placebo had a negative 96 change from baseline; not decline, 96 change from baseline. Thank you.

DR. WEISMAN: Thank you.

DR. SOLOMON: Dr. Kim is standing. I don't know if you wanted to make some comments. If not, we can keep going.

DR. KIM: I'll try to add perhaps something to Dr. Wang's comment, but she answered I think succinctly.

DR. SOLOMON: So it's 12. We're going to keep going to keep going for a couple minutes. I Have Jeff Curtis; Dr. Katz; and Dr. May as last questions. Dr. Curtis?

DR. CURTIS: Thank you. I don't know if this is for the sponsor or for FDA, but is there an understanding or an estimate of the coefficient of variation or measurement error in FVC? It just seems the magnitude of the effect we're talking about is in the range of a couple percent. I think the coefficient of variation in the Scleroderma Lung Study was about 5 percent for the within subject coefficient of variation, and I wanted to understand more about the reliability of the primary outcome in this study.

DR. YU WANG: I will defer this to the applicant.

DR. TETZLAFF: Yes, we'll be happy to take it, and I ask our clinical expert, Dr. Maher, to step up and provide some insight on the variability of coefficient of variation.

DR. MAHER: Hi. Ted Maher, Imperial College, London. The short answer, the Coefficient of variation in centrally-read, standardized, spirometer FVC in phase 3 clinical trials is about 1 percent these days. So things that move forward from SLS I, where sites performed spirometry on their own spirometers, now we perform them on standardized parameters, and those are overread in real time by a remote physiologist who looks at the flow-volume loop and determines whether FVC has been appropriately performed. And as a consequence, we've got our measurement error pretty much as small as it can be.

DR. SOLOMON: I'm sorry. Is that the case in this trial?

DR. MAHER: Yes.

DR. SOLOMON: Okay. Thanks. Dr. Katz?

DR. KATZ: James Katz. I think this is for Dr. Wang. You stated that there's an assumption that the background rate of decline of FVC is the same between treatment and placebo.

DR. YU WANG: To clarify, not just a background FVC decline, but the decline assumed for the dropout patients may be the same.

DR. KATZ: Okay. Let me move to a second follow-up question. Are you happy with using the FVC as a surrogate outcome in a population that has a high rate of diarrhea and weight loss when FVC is sensitive; the measurement of FVC is sensitive to weight loss?

DR. YU WANG: I'll defer this to my clinical colleague.

DR. KATZ: Otherwise, can you control for the weight loss effect that may or may not attenuate your interpretation.

DR. SOLOMON: I'm just wondering if we want to just hold that off until after lunch. As a clarifying question, I think it's a little broad.

So maybe we'll just Dr. May with the final question between us and lunch.

DR. MAY: Hopefully, this is quick. Susanne May. I have a question with regard to the statistician, Dr. Wang. The applicant had presented additional analysis that I believe you didn't incorporate in your tipping-point analysis, where they include those 28 that have right-out-of-the window measurements. Do you have any other comments with regard to that approach or concerns?

Then for the applicant, or maybe you can answer this, how many of those 28 were in the placebo and how many were in the treatment arm?

DR. YU WANG: Can you repeat the second question?

DR. MAY: So the second one was just for the numbers, which is 28, where it's just out of the window that were included, but there's no breakdown between treatment arms. How many of those 28 were in the treatment group versus the placebo group?

DR. YU WANG: Can you go back to backup

slide number 3?

In the bottom corner, 16 out of 47 patients in the nintedanib arm had a post week-52 follow-up out of the 47 patients who didn't have week-52 follow-up.

Next please? Twelve patients in placebo out of 31 had post week-52 follow-up. So your first question, I think I answered when I answered earlier. There are two reasons we consider those data as are supportive but with caveat.

First is the data quality may be of concern. Second, the data as far as -- like you don't really know which time point those post week-52 are from. So the sponsor's proposal, they used the closest one. So you can be one year close, maybe if a trial allowed, or one day close, so we don't know. That's the reason we didn't take that into consideration. Thank you.

DR. SOLOMON: On this same point, do we have another clarifying -- Dr. Becker?

DR. BECKER: I thought I recalled in the presentation that there was a median of 9 days

after that 52nd visit. Would anybody be able to give us the range?

DR. SOLOMON: So perhaps the applicant can address these questions.

DR. TETZLAFF: Yes, and I'd ask our statistical expert, Dr. Carroll, to come to the podium and provide some insight.

DR. CARROLL: Thank you. Kevin Carroll, statistical consultant, Boehringer. I'm a paid consultant today, but I have no financial interest in the outcome of this meeting.

If I can just put this slide up just to address this specific question in relation to the study -- I think that's it -- real briefly, the average number of days was 8 and 9, as you can see. These were just outside of the window, and suggestions of endpoints being a year outside of the window are just untrue; that isn't the case.

It was close, and in fact, maybe after break -- because I know we want to go for lunch -- you can look at the profiles for these subjects, the 12 and the 16. We can look at them,

and you can see how many data points they have. And these patients continue in this study well beyond week 52.

So it's not just one extra value they have; there's a string of those values. I think it's not quite appropriate to say there's something wrong with the quality of those data. The study was designed for a 52-week primary endpoint assessment, but it was prospectively defined, on the patient's consent and the investigators who conducted the study, to collect data right the way through up to a maximum of a hundred weeks.

The first part of this study in terms of its rigor of its conduct, there's absolutely no different to the rigor of conduct after 52 weeks. So the data are reliable, and it's not just one data point in many of these subjects.

Maybe I'll just, very briefly, share an example of that. It's a little complex, but let's pop it up. So very briefly, each one of these is a single patient; this is what you're looking at. You have timer on the bottom, and the two vertical

lines that you can see represent the window.

You can see that patients who had values very close to that third vertical line, you can see how close they were. And these are some of the data points that we're talking about. Note that many patients have follow-up well beyond the 52-week time point, which is the window you're looking at. You can see how far the lines extend to the right.

So I think it's not unreasonable when you're in a situation of trying to minimize your missing data as per the NRC guidelines; in fact the FDA initiated, which is to try and include as much real data as you possibly can. It's not unreasonable to include some patients who had an actual value who did not drop out, and that value was very close to the end of the window.

When we do that, the tipping point that we talked about, which, very briefly, we should realize that tipping point number of 45, is assuming that the interstitial patients will have a detriment to the tune of 45, while the placebo

patients have no detriment coming off therapy.

Somehow the detriment only applies to nintedanib, so it's a little bit skewed, that approach. But if we include all the relevant data that we have, then the tipping-point analysis shows a delta of 120 mL will be required, which in the light of the date from the overall analysis -- but clinical colleagues can comment -- seems rather implausible from a statistical point of view. But just be clear about -- because I don't think the data lack robustness in the way that may have been suggested, if we consider all of the data.

DR. MAY: Quick question. Do you have the same slide for the placebo group?

DR. CARROLL: Yes, we do. It should come up in a second. There we go. So if I didn't tell you which was placebo and which was nintedanib, you wouldn't be able to tell because the pattern is just the same. We have patients just outside of the window, and many of them have extended follow-up. That's because the trial was designed to continue to follow patients the maximum of a

hundred weeks with equal rigor, the first part and the second part. There's no difference in the quality of the data from this trial, before or after 52 weeks.

DR. SOLOMON: Okay. Why don't we break for lunch now unless you --

DR. YU WANG: I can do it later.

DR. SOLOMON: Great. We're going to break for lunch. We're going to reconvene at 1:00. So we're going to have a slightly shortened lunch. Please take any personal belongings you may want. Committee members, please remember no discussion of the meeting during lunch amongst yourselves, with the press, or with any member of the audience. We'll see you soon. Thanks.

(Whereupon, at 12:13 p.m., a lunch recess was taken.)

A F T E R N O O N S E S S I O N

(1:00 p.m.)

Open Public Hearing

DR. SOLOMON: It's 1:00, and we're going to reconvene.

This is the open public hearing portion. Both the FDA 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, FDA believes 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 the sponsor, its product, and if known it's direct competitors. For example, this financial information may include the sponsor'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.

The FDA and the committee place great importance in the open public hearing process. The insights and comments provided can help the agency and the committee in their consideration of the issues before them.

That said, in many instances and for many topics, there will be a variety of opinions. One of our goals today is for the open public hearing to be conducted in a fair and open way, where every participant is listened to carefully and treated with dignity, courtesy, and respect. Therefore, please speak only when recognized by myself.

Thank you for your cooperation. Will speaker number 1 step up to the podium and introduce yourself? Please state your name and any

organization you are representing for the record.

DR. EVNIN: I'd like to thank the FDA advisory committee for allowing me to speak today. It's a very important topic of nintedanib and its possible approval for scleroderma patients afflicted with interstitial lung disease.

My name is Luke Evnin. I am the current chairman of the board of the Scleroderma Research Foundation, a post that I've held since 2002. I've been on the there since 1999. I'm also a patient. I was diagnosed with scleroderma in 1998. Professionally, I'm a co-founder and current managing director of MPM Capital, a tech biotech-focused venture capital firm. I received my technical training at UCSF.

I've listed my disclosures below. Personally, I have none. The Scleroderma Research Foundation, either in the past or currently, is supported by some corporate partners, including Boehringer Ingelheim.

The Scleroderma Research Foundation is a nonprofit organization. It's been dedicated, since

inception, to fund and facilitate the most promising, highest quality research aimed at improved therapies and ultimately a cure for scleroderma.

We've raised over $40 million since inception. We've made direct grants in excess of $30 million across our basic translational clinical programs. We continue to be guided by a world-class independent scientific advisory board. And among our many accomplishments, we're the underwriter and organizer of national consortia, including GRASP and CONQUER.

I'd like to start just by reframing what you've already heard, which is Sword of Damocles that hangs over the head of every scleroderma ILD patient. I've excerpted two charts from the recent Volkmann, et al. paper. This is additional follow-up on the original Scleroderma Lung Study I, SLS I. That's an academic study that originally enrolled 158 patients randomized to oral cyclophosphamide or placebo, treating for a year.

This paper updates the follow-up to a median

of 8 years, at which point in time 42 percent of the patients had passed away. The top chart tracks placebo versus cyclophosphamide, time to death; the bottom chart, the placebo versus cyclophosphamide on the composite endpoint of death for organ failure.

As you can see, the inescapable conclusion that once fibrosis has initiated in our patients, that's the in-bold patient population here, the prognosis is on the one hand very poor and current treatments are simply not effective.

I'd like to try to put a face on some of those numbers, and this is a story of Matt Dobie. Matt was the son of SRF board member Sharon Dobie. Matt was diagnosed at age 25 with scleroderma ILD and put on cyclophosphamide. As I believe the committee is aware, there are no drugs currently approved for treating scleroderma ILD, although cyclophosphamide and mycophenolate are commonly used, and have been shown to be roughly equivalent in terms of their efficacy.

Unfortunately, in 2015, with his treatment

failing, Matt went ahead with aggressive chemoablation followed by stem cell rescue that put him into temporary remission, but in January, 2017, at age 31, Matt contracted influenza pneumonia, which on the background of his scleroderma ILD killed Matt. And unfortunately as you can see from the prior chart, this is an all too common outcome for our scleroderma ILD patients

Scleroderma ILD affects a broad cut of our scleroderma patients. It afflicts our patients regardless of ethnicity, gender, and age. And stepping back to review the therapeutic landscape for our scleroderma patients, there are very limited options, and as you have heard, there are no new drugs.

Unfortunately, scleroderma is rare and heterogeneous, and the disease burden has been a challenge to quantify. At least in part due to these factors, industry interest in the disease has been relatively tepid despite the dire unmet medical need. The clinical heterogeneity substantially complicates running efficient

clinical trials, and historically, disease metrics have focused on the skin manifestations such as Rodnan skin score, but that has required in turn to focus mostly on newly diagnosed patients rather than the prevalent pool.

Turning our attention to nintedanib and its possible approval in scleroderma ILD, BI sponsored a large extended and unbiased phase 3 study and enrolled the largest number of patients ever in a scleroderma clinical trial, as you've heard, 576 patients from 32 different countries, and exposed those patients over an extended duration.

Moreover, it enrolled an all comers population, including roughly half of the patients on concomitant therapy of mycophenolate. Safety and efficacy were both demonstrated. The trial hit its primary endpoint with a p less than 0.04, and the benefit was clinically meaningful; in fact, equivalent to that seen for the approved indication of IPF in terms of percent protection of FVC decline, although, of course, IPF patients do lose more lung function over the course of a year than

scleroderma ILD patients.

The safety was excellent, as good in the scleroderma population as in the approved indication of idiopathic pulmonary fibrosis, which brings me to my appeal to this advisory committee, to recommend to the FDA for approval of nintedanib for patients with ILD.

Again, there are no drugs available for our scleroderma ILD patients, and in fact, there are very few clinical and very few novel agents that have the prospect of bending the survival curve for this unmet medical need. Nintedanib is safe, as safe in this population as in other approved indications, and nintedanib is effective. It hits primary endpoint in an all comers trial, including those on background therapy.

So please, enable doctors and their patients working together to make their own assessment of the suitability of nintedanib for their use. Thank you.

DR. SOLOMON: Will speaker number 2 step up to the podium and introduce yourself? Please state

your name and any organization you are representing for the record.

MS. MARKOFF: Good afternoon. My name is rosemary Markoff, and I am an active volunteer for the Scleroderma Foundation. I was diagnosed with scleroderma, otherwise known as systemic sclerosis, 23 years ago. I had never heard of this autoimmune disease and was concerned when I found out that it had no known cause or cure. But I was hopeful that living in the United States, with the best science in the world, that that would change soon. Unfortunately, that still is the case today.

Since diagnosed, I became very active in the Scleroderma Foundation, the largest patient advocacy organization for people with scleroderma. My volunteer work has been in various capacities, including running a support group for 20 years and advocating on Capitol Hill. I also was appointed to a four-year term for the NIH 18-member NIAM's advisory council as a patient representative.

My most rewarding experience as a volunteer has been working directly with other patients in

our support groups, and more recently, this past weekend at the 21st Annual Scleroderma Foundation Conference in Chicago with over 700 attendees, and I may note that many of these attendees needed oxygen support.

As such, I'm honored to speak to you today, on behalf of 100,000 Americans with systemic sclerosis, about the importance and promise that nintedanib has for our community. As stated, scleroderma still has no known cause or cure. However, research has provided drugs that can help with many of the most severe aspects of this disease. Kidney involvement and systemic sclerosis is primarily manifested by scleroderma renal crisis. Formally, it was the most severe complication in scleroderma and was the most frequent cause of death in these patients.

More than 30 years, with the development of angiotensin converting enzyme or ACE inhibitors, scleroderma renal crisis became a very treatable complication of scleroderma. Although there are still many patients who do not survive and have

poor outcomes, early diagnosis of renal crisis and prompt therapeutic intervention can achieve excellent outcomes.

So kidney involvement should not be the cause of death today as it was 30 years ago, but pulmonary involvement has taken its place. Pulmonary disease and systemic sclerosis, mainly comprises interstitial lung disease and pulmonary arterial hypertension. Over the past 40 years, the mortality rate for people living with systemic sclerosis has not changed significantly.

Today, lung disease and systemic sclerosis causes approximately 50 percent of deaths, and of that number, interstitial lung disease accounts for 33 percent of systemic sclerosis related deaths, according to a study published in the European Respiratory Review.

Interstitial lung disease related to scleroderma is progressive and debilitating. It robs people from leading normal lives, including such simple tasks of going to the grocery store or playing with one's children or grandchildren. It

often leads to disability, losing one's income and career, and in many cases leads to death, as previously stated.

It can be identified, and early detection is key to improving outcomes for people with the disease. However, there are no targeted therapies available specifically to address this critical need in the systemic sclerosis population. That is until now.

The results in the phase 3 SENSCIS trial, in which nintedanib was studied in patients with interstitial lung disease related to scleroderma, are very compelling. Not only do they provide hope for people who suffer the effects of this condition, they provide a path forward for a targeted therapy that was shown to slow the annual rate of decline in lung function. And while we still do not have a cure for scleroderma, perhaps nintedanib could have the same impact for interstitial lung disease that ACE inhibitors has had for kidney involvement.

Having facilitated a scleroderma support

group for many years, I know firsthand how scleroderma patients suffer from lung involvement. I also know my doctor's concerned for me, as he has recently sent me for more tests, checking for any sign of an increase in lung involvement. With the prospect of nintedanib being a potential targeted therapy, the medical community will be able to react more quickly with an interstitial lung disease diagnosis, which is so very important for the success of treatment.

The systemic sclerosis community has few tools with which to fight this insidious disease and it's multiple comorbidities. I welcome the opportunity to provide the voice of a patient at this advisory committee and to speak why our community needs better therapies to combat interstitial lung disease. I thank you for this opportunity to speak to you.

DR. SOLOMON: Thank you. Will speaker number 3 step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

DR. FOX-RAWLINGS: Thank you for the opportunity to speak today on behalf of the National Center for Health Research. I am Dr. Stephanie Fox-Rawlings. Our center analyzes scientific and medical data to provide objective health information to patients, health professionals, and policymakers. We do not accept funding from drug or medical device companies, so I have no conflicts of interest.

There is a critical need for new treatments for SSc-ILD. We all hope that nintedanib will help slow the rate of decline, but the data is not yet sufficient. There is a statistically significant reduction in decline of FVC after one year for patients randomized to the drug, however, there are questions about whether it's clinically meaningful for patients.

The questions about the impact on patient's health is reinforced by the lack of improvements in secondary endpoints that measured patient-centered outcomes such as quality of life. If those modest changes in FVC were meaningful, we expect that the

quality-of-life measures and other patient-centered outcomes would also have improved, but they don't.

In addition, the evidence for efficacy regarding FVC comes from a single clinical trial. While the trial seems well designed to answer this question, including randomizing over 570 patients to drug and placebo and following them for 52 to 100 weeks, it is still a single trial. Replication is a key to scientific evidence. Independent trials could have a smaller or larger effect due to differences in demographic, or treatment profiles of patients, or other factors.

It is important to have clear evidence that this drug slows decline before approval. That will take additional studies. It may be that studies need to be longer or that this drug is only beneficial for certain patients. It is essential to have that information before approval because once the drug goes on the market, it is often impossible to compare it to placebo.

Clear evidence of efficacy is especially needed because this drug has risks, and patients

and the doctors should have enough information to weigh the benefits and risks in order to decide whether or not to try it.

We understand the desire to approve treatments with more uncertainty for conditions without good treatment options. However, approving drugs with questionable efficacy raises the cost of healthcare, potentially exposes patients to risks without the possibility of benefit, discourage the development and scientific evaluation of new and more effective treatments, and can prevent patients from seeking treatments that do work.

To continue to be the gold standard for approval, FDA needs to maintain high standards of evidence for approval. Thank you for considering our analysis of the data for this product.

DR. SOLOMON: Thank you. Will speaker number 4 step up to the podium and introduce yourself? Please state your name and any organization you are representing for the record.

DR. COSGROVE: Certainly. My name is Gregory Cosgrove. I'm a physician at National

Jewish Health, and I'm also chief medical officer for the Pulmonary Fibrosis Foundation. The Pulmonary Fibrosis Foundation does receive grant support from the sponsor that's currently being considered, as well as other members of the scientific community.

I myself have served as a consultant for the sponsor under consideration today, as well as other members in the community such as Genentech and Apple and PhRMA [ph]. I do not believe that my prior consultancy has any relationship or conflict with the work I will discuss today.

I believe, certainly, it's been demonstrated through several presentations that there's a definitive and demonstrable unmet need in systemic sclerosis-associated interstitial lung disease, so I won't dwell on that issue because it's, I think, quite well accepted. What is clear is that moving forward, we need to identify resources to help these individuals.

As part of my role in the Pulmonary Fibrosis Foundation, I believe I can speak not only for the

group of individuals who have pulmonary fibrosis and interstitial lung disease, but as we look at the registry, which encompasses 2000 individuals across the United States, the second largest individual group of patients in the registry are those with scleroderma.

So therefore, we represent a large proportion of individuals in addition to other diseases such as idiopathic pulmonary fibrosis and chronic hypersensitivity. What again is clear is it is a rare disease, but it's an incredibly important one, and interstitial lung disease, as has been mentioned, is a harbinger of significant morbidity, as well as mortality.

Understanding ways in which to care for patients has been limited, mostly due to the rarity of the disease, the complexity in the way in which it manifests in a variable penetrance throughout patients, whether it be systemic disease of the skin, the lungs, or the kidney. What is quite clear, though, in the advent of lung fibrosis, mortality is dramatically enhanced. And

unfortunately, the impact upon that mortality has not changed as previously alluded to.

Treatment options remain quite limited, and in contrast to other patients with interstitial lung disease, an option for transplant, which was alluded to earlier today, is even further limited availability for patients with systemic sclerosis due to the systemic nature of the disease and unfortunately complications associated as a result of that systemic disease.

As such, there are limited options for individuals. While there are expert recommended treatments that can be utilized, their efficacy remains limited in the vast majority of patients. Therefore, the importance of today's discussion and the appropriate and rigorous evaluation of therapies that can impact the lives of individuals with systemic sclerosis and interstitial lung disease cannot be underscored.

As suggested by the prior speakers, I'll also give you insight into the patient perspective, which I think is incredibly important as you make

your decisions and recommendations. In the survey conducted of over 1068 individuals with pulmonary fibrosis, a simple question was posed.

What is the goal of any therapy that you would accept for the treatment of your disease? Thirty-five percent of those responding suggested, "I would like to stop the progression of my disease." The second most frequent answer, which again, 35 percent, was to slow the rate of progression.

So understanding the clinical impact of any therapy that slows the progression, while it may be debatable from an epidemiologic and statistical perspective, from the patient perspective, should you slow the disease that is impacting their lives and more likely than not taking their lives, that is a clinically meaningful endpoint to them.

So I leave you with that thought, and thank you for the attention and your evaluation for this important study.

Clarifying Questions (continued)

DR. SOLOMON: Thank you.

This is the conclusion of open public hearing. We have a little time before we go to the charge to the committee, and there was a request to go back to some clarifying questions.

Nikolai, do you have some questions that you wanted to ask?

DR. NIKOLOV: Thanks, Dr. Solomon. It's not that much of a question, but just to continue the discussion on the interpretation of the data, particularly handling of the missing data assumptions, because I think there was some probably lack of understanding of what analyses were done and what data were used.

I don't know if our statistical colleague, Dr. Wang, might have just a response to a previous question, and then I'll continue.

DR. YU WANG: So just to clarify, one imputation rule, basically, if you --

DR. SOLOMON: Could you help orient us? What was the question that you're now answering? Lunch has wiped out our memory.

(Laughter.)

DR. YU WANG: Let's go back to the claim of the detrimental effect, that more detrimental effect was imposed on the nintedanib arm as compared to the ones imposed on the placebo arm. This I cannot agree.

Basically, if you still remember the tipping-point analysis, in the blue box, we used the zero-zero cell as the reference point. So for that cell, for unretrieved dropouts, data were imputed using MAR assumption, which means there was a comparative 41 milliliter per year for the nintedanib arm versus those ones for the placebo.

So all the added deltas are imposed on top of this difference.

DR. SOLOMON: Would it be useful to put up the slide?

DR. NIKOLOV: It's slide 18.

DR. SOLOMON: Thank you.

DR. YU WANG: I used this cell as the reference point. This cell is analogous to the missing at random assumption. Assuming this, the primary analysis gives us a 41 milliliter per year

difference in favor of nintedanib. So if you look at those red box cells, the detrimental effect of negative 45 for nintedanib, if you add those up on to the 41, it's about zero.

So basically, for those red cells, for unretrieved dropouts, we imputed them, similarly. I ended my answer.

DR. NIKOLOV: And maybe I can continue from here. I think these tipping-point analyses are based on the primary analysis per specified analysis. We certainly want to have the least amount of missing data so we don't have to use different methods to account for the missingness and for the assumptions.

What we heard from the sponsor's presentation is that there are additional patients that had missing data at week 52 for the primary analysis, the prespecified primary analysis, but they had additional measurements within a window of 28 days. And what we would like the committee to discuss, whether this is reasonable or how reasonable, or unreasonable it is, to consider

those data for the comparisons.

I don't know if the sponsor can have the data for these analyses to present, and how the tipping-point analysis, including these patients, that didn't really have missing data within that window, would actually change the interpretation.

DR. KIM: This is Yongman. Before you respond, I found from your previous --

DR. SOLOMON: Would you mind identifying who you are?

DR. KIM: Oh, I'm sorry. This is Yongman Kim from FDA statistics. I found from your previous response through tipping-point analysis, you sent us -- including the data taken after the 52 weeks. According to your analysis, 44.4 milliliter difference in the analysis.

Is that correct? I don't know.

DR. TETZLAFF: I'd like Dr. Carroll to come up to the podium to speak to this.

DR. CARROLL: Hi. Kevin Carroll, statistical consultant. Yes, that's correct. But I just would quickly add that what the FDA asked us

for over lunch, and what we're currently working on, was to re-do the analysis if we could, but only including patients who had additional values between just after the window plus 28 days.

But there's only 4 patients who fall out of that window anyway, so when we do get that analysis -- hopefully we will; it's still being worked on -- it will be virtually identical to what you heard just said.

DR. KIM: Okay. Thanks.

DR. SOLOMON: You still had to ask the sponsor for some more clarifying points, or was that the point that you wanted to make?

DR. NIKOLOV: It was more of if the sponsor has the data, to provide for the comparisons when they include these additional patients that had their FVC measured right outside of the prespecified window of 8 days. And that would be fair if they don't have it because we haven't really required this before or ask for those before.

DR. TETZLAFF: It seems that we don't have

these data right now, but we'll be happy to provide this to the FDA.

DR. SOLOMON: We're going to have some time for a few more clarifying questions. Dr. Geller?

DR. GELLER: Make a comment --

DR. SOLOMON: Please.

DR. GELLER: -- about what was just said. Nancy Geller. I worry about bias being introduced in who did give data that was missing at week 52, who did give data later. I just worry about whether those patients are the same as those who did give data up to week 52. So that's my concern about an unplanned, as well as the fact that this is a post hoc analysis.

I also have a question, and it concerns the change from baseline in FVC presented by BI. That's slide CE-19. And I wanted to compare that to a similar graph presented by the FDA, which has much larger confidence intervals, and I'd like to understand why that is.

DR. TETZLAFF: Should we respond to this? I'd like to ask our project statistician, which has

to do with standard error analysis, the confidence interval.

DR. GELLER: Yes.

DR. VOSS: Florian Voss from Boehringer Ingelheim.

DR. GELLER: Slide 14 of the FDA.

DR. VOSS: Yes. In our plot, you can see the standard error, whereas in the FDA plot, the confidence interval is displayed. In addition, our plot is based on the primary analysis model, but if you would display the confidence intervals, it would look very similar.

DR. YU WANG: I agree. That's true.

DR. GELLER: Thank you. Of course, the standard errors make the data look much further apart.

DR. SOLOMON: Dr. Stoller?

DR. STOLLER: I have a clarifying question for the sponsor. It really regards the ascertainment of the outcome. The context is, of course, that spirometry enforced vital capacity, as we all understand, is a very technique-dependent

measurement, highly dependent upon the adequacy of achievement of end-of-test criteria; that is to say for the non-lung docs, obviously the duration of time on exhales is an important determinant of the total accumulated exhale volume.

So if I truncated my expiration early, my forced vital capacity would be much less than if I satisfied the end-of-test criteria, which for the American Thoracic Society would be 6 seconds with a 2-second expiratory plateau, as well known.

Now, I was reassured to understand that there was central monitoring, but one of the comments about the nature of the central monitoring confused me greatly, which is that there was an oversight of the flow-volume loops. But of course the ascertainment of achievement of end-of-test criteria has nothing to do with the flow-volume loop; it's all about the volume time tracing.

So what leads up to my question, which is tell me about the satisfaction of end-of-test criteria by central review of the spirometry measurements, and tell me whether this varied by

venue, particularly because this could systematically bias the data in either direction, honestly, depending on the methodologic adequacy of the test ascertainment, could this account for variation between the United States and Canada and other centers?

Is my question clear?

DR. TETZLAFF: That is clear, and these are all valid points. What I can tell you is that we applied criteria to lung function testing as robust as it can be, matching a standard that is common in robust clinical trials. We had different levels of quality control.

The first level was that we centralized not only the spirometry in terms of the readout, and that is what you refer to you, we also provided the spirometers. That is the first level. So the equipment was provided to the sites. The spirometers as equipment had quality control to observe. For example, the 6-second breath out that you're referring to, if this was not sufficient, the equipment itself would indicate that the

maneuver was insufficient. And the third level, finally, was the central overread that you alluded to.

DR. STOLLER: But just a follow-up question, I understand the feedback from the spirometer, but of course, the spirometer doesn't perform the test. So the question really directly is, do you have data upon the percent satisfaction of end-of-test criteria stratified by country, centers, et cetera? Is there some variation in the achievement of the end of test? You had many measures to look at it, but my question regards the outcome. It's a very precise question, if that makes sense.

DR. TETZLAFF: Thank you. I'd ask Dr. Stowasser to come up to the podium to respond to that directly.

DR. STOWASSER: Susanne Stowasser, Boehringer Ingelheim. Dr. Stoller, I cannot give you precisely the data you request, but what I can tell you is we had in total 6.7 percent of lung spirometry data from more than 6,000 measurements that did not qualify the ATS-ERS Miller 2005

criteria for acceptability and reproducibility.

What I cannot provide is a split by region, but what we have done -- and all these data were included in the primary model, in the primary analysis. But we have done a sensitivity analysis excluding this 6.7 percent of data, and the sensitivity analysis shows basically the same result as the primary analysis.

(Dr. Stoller nods in affirmative.)

DR. SOLOMON: I just want to follow up with this general theme of the by country variation that we're seeing in the data. And this is a little bit more than a clarifying question, but I'm just going to ask it because I'm the chair.

(Laughter.)

DR. SOLOMON: Does the sponsor have some good explanation for why we see this post hoc -- I understand it's post hoc, but this variation by country, we are representing the U.S. FDA, and I think we're all kind of interested in the fact that the effects were so different by country.

DR. TETZLAFF: Can I have the forest plot

from the main presentation? So what we presented in our presentation today was not necessarily by country but by region in order to make sure whether there was a difference between regions. And you are referring, of course, as we talked about already, to the Canadian and U.S. region.

However, you do see -- and this was the primary reason of including region here, to see whether there was any heterogeneity in the treatment effect of nintedanib caused by any of these subgroups, because, again, the study itself was not powered for any of these subgroups. And the p-values here clearly indicate that there was a lack of heterogeneity.

I'd like Dr. Carroll to come up and provide us with some thoughts on the difficulties of the interpretation when it comes to particulars subgroups.

DR. CARROLL: Kevin Carroll, statistical consultant. I'm trying to keep this brief. I think everybody on the panel knows, well the difficulties associated with subgroup analyses.

They're ubiquitous. They're in every phase 3 trial, but they still cause difficulty in the interpretation.

So when we look at this particular forest plot -- I'm sorry. I should say also the interpretation is really rather difficult when powered or sized or designed to look at subgroups. So when you look at the subgroups here -- and of course all sponsors evaluate their data in subgroups. It's natural to do that, but what you see is some variability from subgroup to subgroup, which is actually what you would expect.

But what it also shows is that we have broad and overlapping confidence intervals, and it also shows that the interaction p-values on the right-hand side, they measure the statistical evidence for true difference. There's not really a lot of compelling evidence that there are real differences here.

So we just have to be a little mindful in these kinds of analyses where we're kind of looking for consistency, where we haven't predefined and

inferential subgroup, and when we don't have any interaction, really the best estimate of the treatment effect in any one of these subgroups is, in fact, the result from the trial for which the study was powered and designed.

So I make those comments -- just a final comment -- not to dismiss subgroup findings. They're right here. The data are what they are. The sponsor is extremely transparent about the data, but more to offer a context within which the data can be considered and interpreted from a clinical perspective.

I guess the last thing I'd say is if we covered the left-hand side of the graph here, if we didn't have labels, if we covered that up -- imagine they don't exist -- it's debatable as to whether anybody would say, hey, there's one particular subgroup that is clearly different from the rest. It's natural that we look to the U.S. subgroup because, obviously, it's where we are. But I think it's important to remember that you will expect variability.

I think the last little comment might be helpful, is you only need 6 subgroup analyses in any trial, 6 independent subgroup analyses, for there to be about approximately a 50 percent probability that one of them will appear to have a negative point estimate, even when, in truth, there is total consistency. That was published by Professor Sen [ph] a number of years ago, just to highlight the difficulties with interpreting subgroups.

But as I say, not to dismiss, they are what they are, totally transparent, but I do think some caution is needed. And to rely on the point estimate alone I think is not correct. When you're dealing with these analyses, you have to take into account the confidence bound and to what extent that overlaps with the overall; otherwise, I think we'll have some misinterpretation. Thank you.

DR. SOLOMON: Thank you. Dr. Richards?

DR. RICHARDS: Thank you. John Richards. In the long-term extension of the INPULSIS trial, you followed up the patients for 68 months, what

percentage of the patients remained on the drug at 68 months? What was the persistence of therapy?

DR. TETZLAFF: Dr. Stowasser?

DR. STOWASSER: What I can show you here is, in total, 700 -- let me put it this way; 703 patients in total rolled over to the INPULSIS-1 open-label extension study. Of those, we had lung function data up to 192 weeks.

I realize this is not the question you asked. You wanted to know the number of patients who stayed on treatment, right? This is a difficult question because what happened during INPULSIS-1, that once the drug was approved in countries, in some countries, it was mandatory, for example, in Japan, to switch the patients that were in the INPULSIS-1 study to a commercial drug.

So that's why I cannot answer exactly your question because your underlying question probably is how many terminated due to adverse events due to tolerability reasons.

DR. RICHARDS: Correct, yes.

DR. SOLOMON: Dr. May?

DR. MAY: Sorry. Going back to one thing that we already discussed, I just want to make clear my interpretation of slide 18 for the tipping-point analysis. You nicely described that the zero-zero cell is representative of missing at random, in that in each of the groups for the people that have missing data, they are essentially imputed as if they would have been part of that group and continued in the same way as before.

So if I want to interpret this table, then, one reasonable way to look at it might be to say I'm focusing on placebo, the zero column, because placebo on the zero column would mean that they would just continue as the rest of the trial, as they look like. And there's maybe a variability of them to decrease or increase, but in general, that would probably be very realistic to say in placebo -- well, I don't expect much of a change, so that would get me to the column of zero.

Then if I interpret, for example, the cell that has zero for placebo shift and minus 45 for the treatment group, could that be interpreted as

similar to, well, those people who had missing data -- and this is what we're worried about with respect to the missing data and that there is more in the treatment arm -- they might actually end up being similar to the placebo rather than being similar to the other treated patients, which is the assumption of missing at random.

So if they were similar to the placebo for the ones that are missing -- and they are known to have lot of adverse events, so they might have withdrawn, et cetera -- that this would be representative of the placebo staying the same, the missing data being the same, and then the missing data in the treatment arm being similar to what we're seeing in placebo once their missing.

Is that correct?

DR. YU WANG: Correct.

DR. SOLOMON: Okay. Dr. Calhoun is up next.

DR. CALHOUN: I've got two questions, one for the agency and one for the sponsor, just to follow up on Dr. May's question.

Is it the position of the agency that it's

appropriate to essentially eliminate the therapeutic benefit when you're imputing data? That was my understanding, too, Dr. May. With the minus 45 diagonal red boxes, you have imputed data that eliminates any therapeutic benefit of nintedanib. Right?

DR. YU WANG: I will just state my --

DR. CALHOUN: Because it's 41 mLs difference, so you've applied negative 45 to that difference. So essentially, by doing that, you've imputed missing data as having absolutely no effect across that.

So the question is, for the agency, is that reasonable, if you impute absolutely no effect when it's just a fraction of the data altogether? You don't impute on the basis of the data you actually have; you impute no effect at all. Is that a reasonable approach?

DR. KIM: This is Yongman Kim, FDA. I try to -- Dr. Wang's comment. The main purpose of the tipping analysis is to assess the penalty for the active [indiscernible], if there's plausible

assumption or not. The table, the column or row, penalizing the shift, slope, declining rate is assumption,, not the -- they [indiscernible] dependent value.

So the main purpose was 45 milliliter apparently imposed on the active arm is really plausible critically to offset the statistical significance with the [indiscernible] assumption.

DR. CALHOUN: Okay. I'm not a biostatistician, so I'll take your word for that.

The question for the sponsor is, on several occasions you've talked about a strategy of dose reduction from a 150 twice a day to 100 twice a day in order to mitigate some of the adverse effects. The question I've got is what's the relative effect on efficacy of that dose reduction? Do you have any data on that?

DR. TETZLAFF: Yes, we do, and we are happy to share these data with you. I ask Dr. Stowasser to come up and present this.

DR. STOWASSER: Susanne Stowasser, Boehringer Ingelheim. We have looked at this in

two ways. The first way is the SENSCIS trial investigated a dosing regimen that allowed dose reduction and treatment interruption. As you have seen, a significant proportion of patients, more than one-third, has had a dose reduction or a treatment interruption, and the primary endpoint was met. The SENSCIS trial investigating a dosing regimen was positive. This is one way to look at it.

The other way to look at this is we looked at the annual rate of declines in patients treated with nintedanib by dose reduction or treatment interruption resulting in a lower or higher dose intensity. Of course, these are not randomized treatments anymore, but this is an exploratory analysis that suggests or is supportive that patients who have dose reductions or treatment interruptions still benefit from the drug.

As you can see, the declines in the nintedanib treatment patients across those subgroups is similar to the overall population.

DR. SOLOMON: Jennifer Horonjeff?

MS. HORONJEFF: Thank you. Jen Horonjeff. I want to circle back to the PROs that I know that we didn't see any difference in, and just get some clarity from the sponsor. In something like the HAQ, I would understand why we wouldn't necessarily see that for more ADLs, but I'm a little bit curious about the SGRQ, which I wasn't familiar with, but I looked up the questions on the questionnaire over the break.

While I also know it's not validated in this population, I was curious if the sponsor had looked at perhaps why we didn't see any difference here. I understand it's not powered for that, but was there something about time since diagnosis, the duration of disease, that might explain why we didn't see any difference between those populations?

DR. TETZLAFF: We'll be happy to comment on this. We do think, as said before, that the SGRQ is not the most appropriate instrument. Let me just use the opportunity because it was said previously that the SGRQ was positive or showed a

signal in the IPF trials. At best, we can say it had inconsistent results. In fact, one trial was positive, one trial was negative for the SGRQ. The pooled data were negative, just for correction of this issue.

Since this is a clinical question on the PRO and it's utilities, I'd like to ask our clinical expert, Dr. Maher, to provide some insights into the difficulties of this instrument.

DR. MAHER: Ted Maher, Imperial College, London. Can I have the OMERACT slide, actually, please?

I was involved in the trial steering committee, and obviously we fully recognize the importance that the FDA and obviously patients and patient groups put on function and feeling. And you'll recognize in the OMERACT document for systemic sclerosis-associated ILD, they put in there the importance of PROs and health related quality-of-life tools without actually naming any. So one of the challenges we've had in designing the study was that there are no validated measures.

I think we can extrapolate also some of our understanding about the St. George's Respiratory Questionnaire and some of the other respiratory questionnaires from idiopathic pulmonary fibrosis, and I think what we realize is that they behave inconsistently across the course of disease.

For somebody with well-preserved lung function to lose even 500 mLs of FVC has relatively little impact on lung function. If you take someone of the extreme end of the spectrum with perhaps an FVC of 50 percent predicted, if that individual loses 500 mLs, they have a massive decrement in quality of life.

So what we see in the IPF studies is if we take the segments of the population with the most severe level of impairment, then actually we can start to show changes in quality of life over a 12-month window.

I think if we then just show the results of the St. George's Respiratory Questionnaire from the SENSCIS study, in essence what you see is that the change in the placebo group is actually within the

measurement error of the instrument. So this is a naught to a hundred scale, and a 0.88 change is next to nothing.

In COPD, where the instrument was validated, the MCID is considered to be somewhere in the region of 4. And given that we were looking to stabilize disease, without seeing a change in the instrument in the placebo group, it was always going to be impossible to show a benefit in the treatment group.

So I think it's really a weakness of the tools in patients with chronic, slowly progressive disease, where your goal is to arrest disease decline. And I think, unfortunately, we still don't have a good tool that we could put into this population and reliably expect to show benefits over a 12-month window.

DR. SOLOMON: Todd Gilligan?

MR. GILLIGAN: Todd Gilligan. My question I have is twofold, is first on the dose introduction. So I'm assuming everybody came in at 150 milligrams --

DR. TETZLAFF: That is correct.

MR. GILLIGAN: Then do you have data of those you backed off to 100 milligrams? Did they come back up to 150 milligrams, and how they tolerated if you increased the level after that?

DR. TETZLAFF: I'd ask Dr. Kohlbrenner to share some insights that we have on these populations.

DR. KOHLBRENNER: Veronika Kohlbrenner, Boehringer Ingelheim. Yes, we looked at patients who treatment interrupted and what they did. In order to manage side effects, treatment interruption was often employed, particularly for diarrhea, and dose reduction was employed likewise, then. After treatment interruption, the dose was resumed at the reduced dose.

What you can see here is that for 20 percent -- so 25 for the 117 that had reduced dose, for 20 percent of patients, a dose increase was attempted, however, half of those patients then reduced the dose again. However, with these mitigation strategies, as I presented

earlier, the majority of patients were able to continue in the trial through 52 weeks and beyond.

MR. GILLIGAN: Can I follow up to that, just on the end of this? How many of the patients who dropped out were on mycophenolate first, and then also of those who caught pneumonia were on mycophenolate as well? Do you have those numbers?

DR. KOHLBRENNER: Among the -- well, first of all, let me say, we also recognize the numerical imbalance in serious pneumonia reports among the nintedanib treated patients. Among the 8 patients with serious pneumonia, 5 of the 8 were on immunosuppressants, mycophenolate; 2 of those 8 also were on cyclophosphamide.

DR. SOLOMON: Dr. Curtis?

DR. CURTIS: I had a follow-up clarifying question to the one about the patient-reported outcomes that was just addressed. It relates to slide CP-16. That's the categorical analysis of various magnitudes of shift in response between the two treatment arms.

I understand that the PROs overall weren't

different between the groups, but for the people that improved or worsened beyond certain thresholds, be that 5 or 10 percent, did those patients feel better? I understand that's a subgroup analysis, but the Scleroderma Lung Study did something like that, where they said if you worsened more than 3 percent or improved more than 3 percent, they could show a difference in the St. George's Respiratory Questionnaire, the HAQ-DI, et cetera.

So the idea being that if you change more than this certain threshold, that there is actually a difference in people's PROs, and I'm wondering if that was done for this analysis in the SENSCIS trial.

DR. TETZLAFF: I'd ask Dr. Stowasser to respond to this directly.

DR. STOWASSER: Susanne Stowasser, Boehringer Ingelheim. We have not done specifically this analysis from your point of view, how you mentioned it, but what we have done is, as part of our a priori PRO validation analysis is to

look into the change of SGRQ or change of FACIT dyspnea score by a different categorical threshold of change in FVC.

The data is here, and what you can see is that, basically, there is not much difference in a one-year clinical trial in this patient population with relatively preserved lung function at baseline, across these different thresholds of change in FVC.

If the data suggests anything, they would suggests that you need a decline of at least 10 percent predicted to show some ability to detect change in the SGRQ on the FACIT dyspnea score, which is, by the way, very similar to an analysis which we did in the IPF population.

DR. SOLOMON: Dr. Kerr gets the last question before we have the charge.

DR. KERR: I was interested in the 100-milligram dose, and you showed efficacy on that. But most of the reduction resulted from GI side effects and diarrhea. And I wondered if you were able to associate that with the scleroderma

patients who had prior GI involvement.

Also, given that the secondary outcomes, specifically the Rodnan score didn't change, were you able to stratify higher doses of Rodnan scores with a muted response in the FVC?

DR. TETZLAFF: I think these are two questions, and for the first question, I ask Dr. Kohlbrenner to step up to the podium, and we will subsequently respond to your second.

DR. KOHLBRENNER: Yes, recognizing that scleroderma patients bear the burden of gastrointestinal disease, esophageal dysmotility, gastroesophageal reflux, diarrhea already, we did an analysis that grouped patients according to whether they came in with this predisposition, which is shown on the right, which is obviously the majority of the patients, versus those who did not report SSc related gastrointestinal symptoms at baseline.

What you can see here, interestingly, is that for diarrhea, among patients with and without predisposition, the numbers look very, very

similar, that there is not an additional potentiation for those patients with diarrhea.

DR. TETZLAFF: For the second question, I have to state that we will look into this and need to provide this once we have it analyzed.

DR. SOLOMON: Okay. Well, thank you.

I neglected to read the final statement around the open public hearing, and I'm going to read that into the record now.

The open public hearing portion of this meeting has been concluded a while ago, and we'll no longer take comments from the audience. The committee will now turn its attention to address the task at hand, the careful consideration of the data before the committee, as well as the public comments. I also just want to make a comment that the first row of the audience is for the press only. Please move to another seat if you are not with the press. Thank you.

Dr. Glaser?

Charge to the Committee - Rachel Glaser

DR. GLASER: Good afternoon. Thank you all

for an engaging discussion, both this morning and this afternoon. As we prepare for the committee discussion and voting, I want to provide a brief reminder and overview of the scientific issues, the regulatory framework upon which our decision-making is based, and the questions to be discussed and voted upon.

Now that you have heard all the presentations and had an opportunity to ask clarifying questions, we ask you to carefully consider whether the efficacy results are robust. As you have heard today, study 214 showed a statistically significant lower annual rate of decline of FVC with nintedanib treatment compared with placebo over 52 weeks.

The observed decrease in FVC decline was not supported by improvement in other measures of pulmonary function, disease activity, or physical function, including endpoints that directly assess how a patient feels, functions, or survived. In addition, the treatment effect was less robust in subgroups, including patients from the U.S. and

Canada, as well as the subgroup on mycophenolate at baseline.

We ask you to consider the clinical significance of the treatment effect of a decrease in FVC decline of 41 milliliters per year in the absence of supportive efficacy from other secondary endpoints.

With regard to safety considerations, the safety profile was generally consistent with the known safety profile of nintedanib in IPF. Deaths and serious adverse events were balanced between the treatment groups. Adverse events, adverse events leading to dose decrease, and adverse events leading to drug discontinuation were more frequently reported in the nintedanib treatment group and were most frequently related to gastrointestinal events. These are described in the nintedanib labeling.

As Dr. Habal presented, there was a numerical imbalance in serious adverse events of pneumonia in the nintedanib group, however, overall adverse events of infections were similar between

the treatment groups. Other than the increase in pneumonia, there were no new safety signals.

Systemic sclerosis ILD is a rare and serious disease associated with high morbidity and mortality. It is also a disease with high unmet need for new therapies. Study 214 demonstrated a statistically significant decrease in the annual rate of decline of FVC with nintedanib treatment compared with placebo.

As previously noted, the observed decrease in FVC decline was not supported by improvement in other measures of pulmonary function, such as SGRQ or FACIT dyspnea scale, in other measures of disease activities such as mRSS or differences in mortality.

FVC is an endpoint that does not directly measure how a patient feels, functions or survives. In IPF, a decrease in decline in FVC was demonstrated to result in clinical response, while the treatment difference in the SSc-ILD study was less than that in the nintedanib IPF program, which ranged from 94 to 131 milliliters per year, as

compared to 41 milliliters per year in study 214. The relative difference in FVC decline, comparing nintedanib to placebo, was similar between the two diseases.

To what extent the treatment effect in IPF can be relied upon to support the modest effect observed in systemic sclerosis. ILD population is for the committee's consideration today. In considering the risks of nintedanib, the warnings and precautions for nintedanib are listed on the right side of the slide, along with the noted numerical increase in SAEs of pneumonia observed in the clinical study. Overall, the safety of nintedanib in SSc-ILD is generally consistent with the established safety profile of nintedanib in IPF and with the safety described in the prescribing information.

In summary, while the efficacy data are consistent with the treatment effect of nintedanib versus placebo, the committee has asked to discuss the clinical meaningfulness of the efficacy observed in the study. As we have discussed, there

are situations where a single study of a new treatment may be sufficient to support a marketing application; in particular, when there's independent substantiation from related supportive study data, and/or when evidence from the single study is both clinically and statistically very persuasive.

The considerations of the single-study approach for nintedanib for SSc-ILD include that SSc-ILD is a rare disease. IPF and SSc-ILD are both chronic, progressing lung diseases, and the studies in IPF were similar in design but with a larger sample size than the study in SSc-ILD.

Based on the studies in IPF, which demonstrated decrease in decline in FVC, decrease in exacerbations, and trends to improvement in mortality, FVC is an accepted endpoint in IPF development programs. The relevance of the findings in IPF to provide context for the findings in SSc-ILD are for your consideration today.

The next few slides are included for reference of the regulatory framework used by the

agency in the review and regulatory decision-making for drugs. FDA's decision to approve an application depends on the determination that the drug meets the statutory standards for safety and effectiveness, manufacturing controls, and labeling.

The focus of today's meeting is the safety and effectiveness piece of the application. In the questions that follow, you will see that you will have the opportunity to vote on the adequacy of the efficacy and safety data separately. For the benefit-risk assessment and approval, your vote should reflect your assessment of both efficacy and safety together for the proposed indication.

The efficacy standard describes the need for substantial evidence from adequate and well-controlled investigations supporting the language and labeling. With respect to safety, an application can be refused to be approved in one of several circumstances as listed on this slide. These include information that the drug is unsafe or that there is insufficient information about the

drug to determine whether the product is safe for use under the conditions prescribed, recommended, or suggested in its proposed labeling.

With this background, the first question for the committee to discuss is the efficacy data for nintedanib for the treatment of systemic sclerosis interstitial lung disease. We ask that you include a discussion of the clinical meaningfulness of the changes in FVC observed with nintedanib treatment in the population studied.

The next question for the committee to discuss is the FVC data for nintedanib for the following subgroups: the U.S. and Canada subgroup as compared to the overall study population, as well as the patients on background mycophenolate treatment at baseline versus the patients who did not receive background mycophenolate at baseline. Discuss the implications, if any, of the results of these subgroups for use of nintedanib in patients in the U.S.

The remaining questions are voting questions. The committee will be asked to vote

whether the data provides substantial evidence of the efficacy of nintedanib for the treatment of systemic sclerosis ILD. If you voted no, we ask that you discuss what additional data, if any, will be needed. If you voted yes, please provide comments.

Then the committee will be asked to vote on whether the safety data are adequate to support approval of nintedanib for the treatment of systemic sclerosis ILD. If you voted no, we ask that you discuss what additional data, if any, will be needed, and if you voted yes, please also provide any comments.

The last voting question is whether the benefit-risk profile is adequate to support approval of nintedanib 150 milligrams twice daily for the proposed indication of the treatment of systemic sclerosis ILD. If you voted no, we ask that you discuss what additional data, if any, will be needed. If you voted yes, please also provide your comments.

Thank you, and I will now turn the meeting

back to you, Dr. Solomon.

Questions to the Committee and Discussion

DR. SOLOMON: Great. Thank you.

We will now proceed with the questions to the committee and panel discussions. I'd like to remind public observers that while this meeting is open for public observation, public attendees may not participate except at the specific request of the panel.

So again, I'll read the first question. This is a discussion question, not a voting question. We're to discuss the efficacy of nintedanib for the treatment of patients with systemic Sclerosis ILD and to discuss the clinical meaningfulness of the changes in FVC with nintedanib treatment in the population studied.

We've been discussing this for quite a while now, but we're going to discuss it further.

Dr. Weisman, I definitely would like to hear the pulmonologists, who kind of live with these measurements and these symptoms, weigh in heavily here. Dr. Weisman, did you want to kick off?

DR. WEISMAN: I agree. Let's hear from the pulmonologists, and then I'd like to make some comments after that and questions.

DR. SOLOMON: Any pulmonologists? Dr. Garibaldi?

DR. GARIBALDI: Hi. Brian Garibaldi. I guess it's hard for me to answer that question without considering the population in which we're going to be using this drug, which gets to the second part of that question, which is the subgroup analysis looking at mycophenolate.

I'd like to hear a little bit more about how both side effects and also treatment effects are distributed geographically, and how that relates to the mycophenolate issue, because in reality, most of us in practice are using mycophenolate as opposed to cyclophosphamide at this point to treat our patients with scleroderma ILD.

I worry that the blunt of the blunted effect that we're seeing in the U.S. subpopulation, as well as the less effect that we see in the patients who are already on treatment with mycophenolate, I

would say that a 27-milliliter difference over the course of years is probably not meaningful clinically.

Now, is that different from a 41-millimeter difference? I'm not sure. But in reality, when we're going to be using this drug, it's most likely going to be used in concert with mycophenolate at this point in time. So I have concerns about whether or not that truly is a meaningful difference.

DR. SOLOMON: Dr. Weisman?

DR. WEISMAN: Well, this brings up the difficult question of what are we actually treating with this drug in scleroderma ILD? We have a drug that was approved by the FDA, based upon some substantial biology of this drug attacking these profibrotic pathways in IPF, and on that basis, this drug was approved for that condition, and it's recognized that that condition is a bit different from scleroderma ILD, and that condition is a different set of demographics. It has a different change in trajectory. Patients get worse rather

quickly. And the measurements probably were a bit easier to deal with because of the rapid change and the ability to show a difference.

So now we're talking about another disease where the fibrotic pathway is there, but is it in the beginning of the disease or is it at the end of the disease? The patients have a slower trajectory, and what's its relationship to inflammation?

So I'm struggling with an understanding of how to place this data in the context of a scleroderma patient that has a different trajectory, and we don't know exactly where and when the fibrotic pathway actually takes place. So even if this drug was approved, when would we actually use it; at what point in the disease?

So maybe I'm asking a question to this panel to think about, and perhaps we could be enlightened a bit by maybe asking a couple of our scleroderma experts that are here in the room to tell us what they think, even if this drug were approved, where would it be in the projected trajectory of a

scleroderma patient, and how does it relate to the pathophysiology of inflammation and fibrosis; and are these different and independent phenomena?

Help us understand this a bit.

DR. SOLOMON: Are there folks on the panel -- before we go to other experts in the room, are there folks on the panel that want to discuss that specific question? Todd, please.

MR. GILLIGAN: If I can speak to that as someone with scleroderma ILD, and walking you through what I believe many of you medical professionals know and I've learned since November of 2017, is that 5 or 6 years ago, if I would have been diagnosed with this disease, I would have been on cyclophosphamide, or Cytoxan, at that time, and I wouldn't have been given salts after [indiscernible] mycophenolate, which I'm currently taking.

That drug existed years ago for transplant patients, completely a non-fibrotic reason that we have mycophenolate. And we crossed over and allowed that for use now, which is where I started

taking that drug in March and introduced that drug replacing the Cytoxan, which you alluded to you're doing with your patients, and I've heard you talked about using it maybe in conjunction with or somewhere in.

So as I go into my PFTs and my FVC is falling, as my diagnosis has come, and I'm losing 17 percent of my forced vital capacity a year, and I don't know how long that rate will decrease, I would, from my end of looking at this -- and we can talk about minimal effectiveness, as you look at it, used in conjunction with or at the same time that you introduced mycophenolate, from a patient perspective, I'm willing -- I'm not a medical professional -- you folks, and that I heard people saying things, use it along with your doctor to make that decision, that's where I would see it being used.

Right now, I have a visit at the Mayo Clinic coming up here August 12th, and my next lung function test is coming, and that's what I see, because if this decrease continues, I need other

options; otherwise I'd just continue down the path that I'm on.

So as I throw it out there from a patient perspective, I'm living it; that's what I see going on.

DR. SOLOMON: That's very helpful.

I just wanted to stay on this topic of inflammation versus fibrosis. Is that what you want?

DR. REDLICH: Yes. I would say, as a pulmonologist who has a cold, I see a range of patients with interstitial pulmonary fibrosis. The idea that you UIP is this distinct entity that we understand completely the pathogenesis of and exactly when to start treatment on, there's just a lot of overlap between all of these ILDs, and they're really all a combination of inflammation of fibrosis.

I've been at this for too many years. For all of the money spent on all of the mechanistic research, it's still a mush of inflammation of fibrosis with very similar mediators across all of

these processes. So to me, the fact that the scleroderma is progressing slower means that it's harder to show the impact in the year. So the fact that you have shown in impact of a change says something. But you may well be on that drug for more years with a slower progression.

So you may say, well, 50 cc's in a year, but that's 102 years potentially; we don't know. Realistically, with a rare disease to do a study that's three or four years long would be really challenging. So I do think that I look at that as, yes, at least from the data, it doesn't seem that that effect -- we don't have a reason to think it would wear off after a year.

DR. WEISMAN: Can I respond to her question, to her answer?

DR. SOLOMON: Yes.

DR. WEISMAN: But the question is what is the meaning of FVC here? It can be affected by inflammation. It can be affected by fibrosis. It's a surrogate marker for something that has multiple pathways. So that's what's being

addressed in this question here.

So how do we know when -- is this helping us understand when to be able to initiate treatment, if in fact what we've heard is this one-year treatment, the data was not quite as robust as everyone wanted, and not as robust as the data in IPF -- is this something that we needed to wait two years to see?

This is the question that I'm raising to the committee here, to understand the meaning, or the meaningfulness, of this change in FVC.

DR. REDLICH: Well, I don't think one necessarily has to understand the pathogenesis of everything to decide that something may have efficacy in terms of -- there are lots of things that impact your FVC, it's true. It's not so easy to actually show a change in lung function. So the fact that you are able to, even if it's a small change, I think does say something.

DR. SOLOMON: Dr. Stoller, and then Dr. Calhoun.

DR. STOLLER: Well, I'd like to respond to

Dr. Weisman's and invoke his country doctor.

At some level, in response to your question, I think the utilization of this drug, were it approved, would default to our usual clinical reflexes, which is to say, to echo Dr. Redlich's remarks, we will never, in the context of a rare disease that requires recruitment of large numbers of patients to do subset analysis -- I admire the question, but I would regard it as relatively unanswerable in terms of the temporal sequencing of utilization of drugs.

So what we all do in our practices, whether rheumatologic or medical, is to contextualize the agent. We have mycophenolate. Most of us currently use that as a first-line agent for reasons that have been nicely articulated. Again, it falls outside the bounds, but I think it's not lost on any of us who are clinicians, that this drug is quite expensive, for example.

One would likely, were it available in my practice, to answer your question, probably offer a patient, with a conversation, of course -- probably

offer a patient mycophenolate, watch their slope for a while. Recognize that if they stabilized on mycophenolate in a way that was consistent with age-expected loss of FVC, would probably not be inclined to offer another agent that had, although well-defined toxicity, nonetheless clear toxicity in terms of GI toxicity. And in the face of progressive loss of lung function, would be inclined to then add a second agent to demonstrate what is now a 27 mL as opposed to a 41 mL decline.

I think that's a simple minded, kind of country doctor approach to how this will actually be used in clinical practice were it to be approved. But it doesn't reflect --

DR. REDLICH: It doesn't --

DR. STOLLER: -- the ability to answer your question, which is beautifully articulated, but in my view unanswerable.

DR. REDLICH: No, I agree. I think it's a really good question, but we aren't able to answer that for UIP either. At our weekly ILD conference,

it's sort of this discussion, well, should we start another agent? Should we wait and recheck the PFTs in 3 months, and 6 months see how they're doing?

It would be nice to be able to make these decisions with more data, but I think as Dr. Stoller described, is what happens in practice.

DR. SOLOMON: Dr. Calhoun?

DR. CALHOUN: Dr. Stoller has nicely articulated an approach, and I think it's a reasonable one. I think the other factor to put in here to Dr. Weisman's question is that as I recall the data from the sponsor, when they split response rates by percent predicted FEV1, breaking at 70, response rates were pretty similar.

So if that were to be the case, then you don't need to wait until someone is at 55 percent predicted in order to initiate an additional treatment. Oftentimes in lung diseases, in particular, it's the people who have the more severe disease in whom it's easiest to show a response, and that's not true with this drug, which was kind of interesting, that those who had minimal

effect, minimal decrement in vital capacity, had as robust an effect on reducing the decline in lung function as did those who had more severe disease.

DR. SOLOMON: I think going along with the comments of our pulmonary colleagues about where to sequence the drug and also Todd Gilligan's comments about MMF may not be working fully well, and let's add it to MMF, it does draw my mind to thinking about the subgroup analysis of people on MMF. While we've heard why subgroups are hazardous, it seems like that's actually the subgroup of most interest.

So it's just something that I think we have to kind of think about as we think about the clinical meaningfulness of the results of the trial before us.

Other comments?

DR. GELLER: I have a question. Nancy Geller. I have a question, which is perhaps clinical. This is a young population, and if we approve this drug based on one-year data, these people are going to be taking it probably for the

rest of their lives.

What do you think about that, clinically? Is the effect going to be maintained or attenuated over time, or should we say that we don't know, and we have no data now?

DR. SOLOMON: Dr. Calhoun?

DR. CALHOUN: That's something later in my list of things to talk about. But I believe the sponsor showed data at 52 weeks, and then some -- it's a smaller end, but at 100 weeks. If I read the data correctly, there was the 41 mLs in the first year, and then it was more like 21 mLs in the second year.

Is that correct? And the question is whether that actually is an estimate of the real effect or whether that's being driven by the small sample size.

DR. TETZLAFF: We did a variety of exploratory analysis on this, and the effect size -- and I'd ask Dr. Carroll to maybe speak to this -- was in a range between 40 and 60 mL, in these exploratory analyses.

DR. CARROLL: Kevin Carroll, statistical consultant. Trying to be brief, yes, we did look at the 100-week data. The study wasn't designed for that, but still you have some data to look at, in an exploratory sense. And the treatment effect of 2 years was like 65 mL. It's right here; let's put it up very briefly.

So that effect, the cumulative effect, was 50 percent more than it was in the first year, and that's the best analysis we think we can do. It's on an ITT basis.

DR. CALHOUN: That 65 mLs is not cumulative from beginning of trial? That's the added between 52 weeks and 100 weeks?

DR. CARROLL: No, that's cumulative.

DR. CALHOUN: That's cumulative --

DR. CARROLL: As I just said, that's cumulative. The cumulative effect at the 2-year time point, approximately.

DR. CALHOUN: So if you've got 41 mLs in the first year, you got another 23 in the second.

DR. CARROLL: Yes, you could kind of look at

it like that. What we did was one year, you have a difference, and then the curves continue to separate to the tune of 65 mL at 2 years. So there is some added benefit in it. It's about half. It's about 50 percent of what you had at one year.

DR. CALHOUN: They just don't continue to diverge at the same rate.

DR. CARROLL: Well, the difficulty is --

DR. CALHOUN: You have a small N.

DR. CARROLL: Yes. The study is designed in a way that makes it really difficult to know. It's the best that we can do with the data that we have.

DR. SOLOMON: Dr. Redlich?

DR. TETZLAFF: I'm sorry. Just to add that we do have some efficacy data from the IPF experience and the [indiscernible] trial we've talked about. We have evidence that the treatment effect is sustained over 68 weeks.

DR. REDLICH: Just to go back to the point about what if people are on this for many years, I think unlike hypertension or elevated lipid levels, that people could be on those medications for 20 or

more years, if someone ended up -- systemic sclerosis may progress slower than UIP, but it's still a progressive disease. As we've heard, a median survival of, whatever, 5 to 9 years.

So if someone like that were on the medicine for 10 or 15 years, that would be great that they had extended longevity. Although I think that is a legitimate concern, I think it's overridden by the mortality.

DR. SOLOMON: Dr. Garibaldi, did you want to make comment?

DR. GARIBALDI: Yes. I think I was going to make a similar comment that we don't know what the long-term side effects of nintedanib beyond 10 years. I mean, we only have experience in the U.S. for 5 years with IPF patients. But I would approach it in the same way. If we can get you to live long enough to have your secondary side effect that we don't know about yet, then that would not be such a bad thing.

I think one of the questions about rate of progression, scleroderma lung disease is different

in IPF. We don't typically think of flares as being as common, and we know that in IPF, flares might be reduced by being on antifibrotic drugs. And that may be something that might lead to that same benefit in reducing the rate of decline.

One question that I have in my mind is in the adverse event groups, the classification of pneumonia, were those truly infectious events or were they flares or the underlying scleroderma lung disease? One of the things I worry about a lot in our scleroderma patients is aspiration, particularly in people who have significant GI side effects with a much higher rate of vomiting.

That signal there, with an increase in pneumonia, I don't know what it means, but I'd be interested in wondering what happens long term in people who stay on the drug, who continue to have side effects, who already have gastric motility issues. Is that something that's going to affect lung function decline over time, and I don't know the answer to that.

DR. SOLOMON: Dr. May?

DR. MAY: Following up on that, I was wondering -- and I can understand that this is a study of a rare disease, and it's difficult to get enough numbers. And it's probably even more difficult to judge safety with regard to this.

Excuse me. This is a non-clinician question. There was a presentation, the first presentation from the public, where somebody who was depicted who had pneumonia and then died from that pneumonia. I was wondering with the increased rate of pneumonia and us not following up for a longer time, could this be a safety issue with respect to increasing the potential for pneumonia for a relatively modest treatment effect; so speaking to the risk-benefit ratio of this.

DR. SOLOMON: You know what? We want to focus on the FVC, and we want to focus on the data that we have, not the data that we wish we had. There's a lot of data that we wish we had, but we only have what we have here.

DR. MAY: But we are asked about the risk-benefit ratio, right?

DR. SOLOMON: We'll get there.

DR. MAY: Okay.

DR. SOLOMON: We'll get there. I just want to keep us on task.

DR. NIKOLOV: Dr. Solomon?

DR. SOLOMON: Yes?

DR. NIKOLOV: Maybe a point of clarification of why we specifically asked this question, it has to do with the fact that we see a small treatment effect that was not as robust or as big as with the IPF program. We didn't really see a whole lot of supportive efficacy from endpoints that measure how patients feel, function, or survive, and we think this is important.

So that's one of the reasons we brought this question, but we also wanted to get the impression from the committee, what do clinicians follow to make decisions how to treat these patients? Is it pulmonary function tests or is it how patients function or symptoms? The impression I get so far is that it sounds like decisions are made based on pulmonary function tests over time. But I want to

make this a point of discussion.

DR. SOLOMON: Dr. Becker?

DR. BECKER: So I would totally agree. We do a lot of pulmonary function tests, even in kids, as challenging as they are. I think that in that regard, I take these data to be important. Even though the secondary patient-reported outcomes did not show a difference, I still find that's what I follow primarily, the diffusion capacity or the FVC as I start to follow these patients.

I think particularly what was meaningful to me is we know that ILD is chronic and progressive, and at times lethal, and every little bit we get to slow down to me is important.

DR. SOLOMON: Dr. Katz?

DR. KATZ: James Katz. I don't just rely on FVC. I find it a worrisome surrogate marker. It's affected chest-wall restriction that happens in scleroderma. It's affected by chest-wall muscle weakness. As I mentioned before, I've learned that it's affected by weight loss.

So I can't just use the FVC. I need to use

the high-resolution CT scan and the DLCO to factor in my thinking. So I would ask the discussants do they feel that the FVC is a misleading surrogate in this case?

DR. KERR: Well, in practice, we all do different things, but as per OMERACT, this is the tool they've given us to use for these bit of data. But even when we see a patient and we check the FVCs, DLCO, the 6-minute walk test, et cetera, even when we see that, we still have to go through the differential, whether it's pneumonia, whether it's venous thromboembolism, et cetera in these patients.

That I think is what I'm struggling with here, in that the patients we're looking at, we're looking at an antifibrotic drug, that we know the sequelae or the pre-event to this is inflammatory disease. That's where the MMF, et cetera, comes in. And the question is, in this cohort, whether there was still a component of inflammation; even though they're on 6 months of MMF, was it telling us that the ground glass opacities indicated more

inflammatory disease and fibrotic disease? And maybe that's why there wasn't such a robust response compared to the IPF group of patients.

It goes against what we're accustomed to in rheumatology for early diagnosis and prevention of progression, where fibrosis tends to be more the end stage in this disease. And that's where we're trying to time, where do we do this? Do we look at the slope over time of these patients and then apply this drug? And I don't think we have that answer here today.

DR. SOLOMON: Todd Gilligan?

MR. GILLIGAN: From the patient perspective, again, I've had this conversation with my rheumatologist and pulmonologist, and we decided to wait 6 months before we started the mycophenolate the first time around just to see if by chance the disease would slow itself, what the progression was, and to take another pulmonary function test.

We looked at DLCO, both. And I'm guessing you folks are bright enough in here to know there's probably a correlation between your FVC and DLCO

somewhere in there for anybody with this disease. Whether they're on the same path and track, we get it; they're going to be different. But that was a decision that we made.

So I echo, again, that the FVC. looking at it, and again,

conversation with doctors, this gives you another viable option. And you folks, it's called practicing medicine for a reason, and we would engage in that conversation with the patient at that time to discuss which one you use, or both, or one or the other first.

I think the FVC is a nice indicator. I know for my own -- I'll say I'm a sample of one; I get that. But we are a small group, in a small population, that is living out here with this disease.

DR. SOLOMON: Dr. Calhoun?

DR. CALHOUN: I just wanted to comment a little on the effect of weight on vital capacity. So yes, it can be affected by weight. It's affected by obesity, so if the BMI is high. It's

also affected by low weight, but only insofar as it affects muscle weakness. So if someone loses weight but isn't weak, that won't necessarily affect their forced vital capacity.

I take your point that there are a lot of factors that influence that, and it may not be the sole thing, but the fact of the matter is that was the primary outcome that, as I understand, the agency agreed to. Right? You agreed to the vital capacity is the outcome for this trial. Right?

[Dr. Habel nods yes.]

DR. NIKOLOV: Correct. That was consistent with the endpoint used.

DR. CALHOUN: Correct. So you agreed to that. That's the outcome we've got. I'm a little less concerned about weight loss adversely impacting vital capacity because there was a relatively small fraction of people that lost as much as 10 percent, and 10 percent probably wouldn't, in and of itself, again -- unless there was muscle weakness associated with that. And to the extent that systemic sclerosis induces muscle

weakness, et cetera, et cetera, you're absolutely right.

DR. SOLOMON: Any new comments or should we summarize?

DR. REDLICH: Well, I just had a quick question for the sponsor. I believe that to be enrolled in the study, you had a CT scan to document that you had lung disease. Were any follow-up CT scans done? I assume not, but I didn't know if that was on a subgroup or on India, because in clinical practice, as was commented, we usually use other information in addition to an [indiscernible] EC.

DR. TETZLAFF: We do have a substudy running that is not analyzed by this time point, but I would ask Dr. Seibold to come to the podium and speak on how we use these measures.

DR. SEIBOLD: I wonder if I might have HRCT slide pulled up. I think this has been a very interesting conversation, but I'd just like to reflect on a few things.

First, for the rheumatologists, the

pulmonary community went through a decade or more of thinking that anti-inflammatory immunosuppressives would work for IPF, suspecting that there was inflammatory component. And they either did harm in their studies or some no effect.

One end of the spectrum may have a non-inflammatory and more purely fibrotic disease. In scleroderma, Mike's question about when is it inflammatory and when is it fibrotic, I think at face value, it sounds like a great question, except that all of these patients have fibrosis. So the question then becomes, is it a fibrotic disease that also has some inflammation, or is it an inflammatory disease that initiates fibrosis? There are mechanotransductive effects, that once fibrosis is established, it's self-perpetuating, because fibroblast biology is changed by the environment in which it's living.

So if you wanted to be a purist here, you could go to your HRCT. And if you had a patient that had mainly ground glass and no reticular change, you might opt for an immunosuppressive drug

because that patient was apparently dominantly inflammatory; or if you went to your HRCT and it was dominantly reticular change, and there was minimal ground glass, you would be hard-pressed to argue that that patient would benefit from an anti-inflammatory therapy.

One of the things we learned in Scleroderma Lung Study I when we did bronchoalveolar lavage is that the level of cellularity in the lavage offered no predictive value about whether or not there would be response to cyclophosphamide or a change in pulmonary function.

Then I think the last thing is kind of clear, that people need to be a little bit clear about it, so thinking through this data set. This is not a head-to-head comparison with mycophenolate. These are patients who were mandated to be on mycophenolate for at least 6 months before they got into the study, so they're mycophenolate survivors. Those that couldn't tolerate mycophenolate or objectively failed mycophenolate, they're not in the study.

This is a different subset. So even in the setting of being on that therapy, we still saw added benefit. And although this is not a mycophenolate comparison, again, if you want to look at subsets, those patients not on mycophenolate lost about 116 milliliters. Those on mycophenolate lost about 66. So the data suggests that mycophenolate has some partial benefit, but nowhere near the benefit that's been reported in the primary mycophenolate trials.

So I think there's room for polypharmacy in some of these patients, but my real perspective on this is that this is a fibrotic disease. You have a drug here under consideration that has purely antifibrotic mechanisms. Then the best measure we have of interstitial lung disease, the FVC, moves in the right direction. Hence, it argues that we're having a bona fide antifibrotic effect. There may be other treatable aspects of the disease, but that what you should be focusing in on.

DR. SOLOMON: Thank you.

Other discussion before I summarize?

(No response.)

DR. SOLOMON: So I think it's been a robust conversation about FVC. The points that I heard that I think were perhaps most salient are the fact that there are a lot of questions about sequencing drugs. So it gets into this question of with or without MMF. There are a lot of subgroups that we don't have enough information on from this trial.

The role of FVC, clearly, the pulmonary function tests are important. FVC is a key component of the pulmonary function tests that many do focus on, even though the FVC has important confounds: weight, muscle strength, et cetera.

The role of HRCT in understanding these patients is also of some question. I think we had some important comments from Todd Gilligan about where patients see this fitting in, in the discussion with their provider and with slowing down the progression of the disease.

With that, we're going to go on to the next discussion question. Question 2, again, a

discussion question, not a voting question, to discuss the FVC data from the following subgroups and the implications for use in nintedanib in patients in the U.S. and Canada subgroup compared to the overall study population; A and B, patients on background MMF versus no background MMF.

We've had a lot of discussion about this already, but just to kind of zero in, some of the points I just want to bring up, I think the sponsor made it pretty clear that these are underpowered subgroups. These were not prespecified subgroups as far as I can tell, that there were some interesting trends, but we should just understand that they are subgroups.

I think we've already heard some discussion about the fact that background mycophenolate has some effect, so seeing the incremental effect on top of that might be slightly difficult. But I'll stop there with my editorial comments and let people ask or discuss.

Dr. Geller?

DR. GELLER: I thought the gentleman who

just spoke from BI said that these were MMF survivors.

DR. SOLOMON: No, I don't think that's --

DR. GELLER: I don't think that's true.

DR. SOLOMON: That's not true. Can you just clarify that not everybody had to have a --

DR. GELLER: Half. It looks like half.

DR. SOLOMON: Fifty percent were on it at baseline.

DR. SEIBOLD: [Inaudible - off mic]. But those on MMF --

DR. SOLOMON: But not everybody in the trial had been on MMF.

DR. SEIBOLD: But had shown an ability [inaudible - off mic].

DR. NIKOLOV: If you can speak to the microphone, please.

DR. SOLOMON: Did someone from BI want to just repeat that?

DR. NIKOLOV: For the record.

DR. TETZLAFF: Half of population of the patients included in the trial were at baseline on

MMF worldwide, and that meant that they were on stable MMF for 6 months preceding the entry visit.

DR. SOLOMON: Thank you.

Dr. Nason?

DR. NASON: Just a quick question. Was that also true in U.S. and Canada? Was it about 50/50 or were those two correlated where there was a lot more, I suppose, mycophenolate in U.S. and Canada than other parts?

DR. TETZLAFF: It was 80 percent?

DR. SOLOMON: But there was a slide in the presentation on the cross tabulation between MMF and region. I don't know if you guys want to pull it up again.

DR. GELLER: It's slide 28 of the FDA.

DR. SOLOMON: Okay. Slide 28 of the FDA would be helpful to kind of clarify this point.

DR. NIKOLOV: It would be Dr. Wang's presentation.

DR. SOLOMON: Dr. Wang, do you just want to describe it, or does someone from the FDA?

DR. YU WANG: For the third forest plot, we

conducted first a cross-classification of region, which is the U.S. and Canada versus the rest of the world and MMF use at baseline. So this cross-classification resulted in four groups. You can see the point estimate together with 95 percent confidence intervals for each group.

DR. SOLOMON: Thank you.

DR. YU WANG: That's post hoc.

DR. SOLOMON: Yes.

Dr. Geller, did you have another question?

[Dr. Geller gestures no.]

DR. SOLOMON: Dr. Stoller?

DR. STOLLER: I'll address the question about the U.S. versus Canada, recognizing, again, the limitations of subsets, and recognizing, to Dr. Redlich's point, perhaps any decrement in decline of FVC is important for those of us who follow it, but also recognizing that a difference of 40 mL is clinically small. I think we'd all have to recognize, at least over the context of one year, whether that amplifies over time is unanswerable.

The one question I think might enhance our understanding with the existing data set would be a more detailed analysis of the methodologic satisfaction of end-of-test criteria stratified by geography.

Having been in pulmonary function labs, albeit not in the context of clinical trials around the world, I am aware that there is high degree of variability in the attention to methodological rigor in the ascertainment of pulmonary function test data, and it would be important to know that the data was stratified by methodologic acceptability by geography to better understand the subsets.

DR. SOLOMON: With the caveat that we know this is a secondary subgroup analysis, it is still -- and I think that this is a plausible explanation for this difference. Even though the interaction terms aren't significant, the MMF by region of the world is kind of an interesting theory. But again, it is kind of curious that things do look different. Even though it's the

limitations of subgroup analysis, we know that the confidence intervals are very wide and overlapping, but from a biologic plausibility standpoint, it is a little confusing.

Dr. Redlich, did you have --

DR. REDLICH: I had the same question.

DR. SOLOMON: I didn't know if there's any insight.

Dr. Weisman?

DR. WEISMAN: Dan, you're raising the question because this study was all comers without much restriction all over the world, and it allowed prior use, concomitant use, of mycophenolate so that it was prespecified to do this subgroup analysis, U.S. and in Canada versus the rest. Is that correct? It was prespecified to do this analysis.

Am I wrong?

(No response.)

DR. WEISMAN: When the study was designed to be wide ranging --

DR. SOLOMON: Can the applicant tell us, was

this prespecified, this region of world?

DR. WEISMAN: U.S. and Canada versus the rest, Was that prespecified?

DR. TETZLAFF: Dr. Carroll?

DR. CARROLL: Hi. Kevin Carroll, statistical consultant. The analysis by region was prespecified in the context of looking for consistency, as was the analysis by prior MMF use for at least 6 months or not, but the cross-tabulation -- that we saw on an FDA slide -- is post hoc.

I would just point out -- sort of being said about subgroups -- we have to be careful with the MMF U.S. group. That has 5 percent the randomized patients in that subset, just 5 percent, so you have to be real carefully in interpreting that particular analysis. That's why it has a confidence interval from minus 200 to plus 200.

DR. WEISMAN: So it was recognized that it was important to address this issue because it was designed for all comers. Therefore, this question was important to be answered. So I think this

discussion here at the committee meeting is important. What's the meaning of it?

DR. SOLOMON: Right, I agree. And it's good that it's a discussion point. It is somewhat frustrating as the panel to have so little data to really have a more intelligent conversation about it. We're really just conjecturing at this point, and with wide confidence intervals, it's hard to make much of it.

I think for me this is one of the general themes of much of today's conversation, is that we are sitting trying to make decisions on relatively sparse data. That that is frustrating, and we see these kind of interesting secondary analyses and subgroups that we want to create a story around. But I think this consultant is appropriate in saying, look, these are wide confidence intervals, and they're not the primary analysis. However, the primary analysis is, it's so thin in some respects, the data. It is a positive study, clearly, but once you start kicking the tires of it with sensitivity analysis, et cetera, it seems less

robust than we would like.

DR. GELLER: This is Nancy Geller. You may not like this interpretation of that same slide, but it seems that the positive result is driven by other than U.S. and Canada and by no MMF at baseline.

DR. SOLOMON: Do you want to put up the slide again? Would people like to see that again? I think that was the slide from Dr. Wang's presentation.

DR. YU WANG: Slide number 28.

DR. SOLOMON: Dr. Geller, do you just want to walk us through how you came to those conclusions?

DR. GELLER: Sure. P-values for interactions are not significant, but it's very hard to get a significant p-value in an interaction because that's a much less powerful test. I'm looking at the mean difference in the confidence intervals, and I see, overall, just a little to the right of zero. That's what we have been talking about.

For MMF use at baseline, it's just hits zero; no MMF use. It just hits zero. For other than the U.S. and Canada, it's just to the right of zero. So I'm saying that those results drive the overall result.

DR. SOLOMON: Other comments regarding this issue of these subgroups? Dr. Weisman? Sorry.

DR. WEISMAN: That makes a very difficult decision. Is it going to be decided that the FDA can approve a drug for outside the United States and not taking mycophenolate? I mean, we have to look at this regulatory and statistical issue in the context of whether this drug is going to be prescribed in the United States. So this has to be a larger discussion.

DR. GELLER: And it's confounded by the fact that only 25 percent of the patients are from the U.S. and Canada.

DR. SOLOMON: Todd Gilligan?

MR. GILLIGAN: My only question to the medical community on this is, I've heard this conversation back and forth of whether it's

fibrotic or it's inflammation first. So we attack it with mycophenolate first from the inflammation immuno side. This gives you an option to attack it from the fibrotic side. And there's not a person I don't think in the room who's answered whether or not it's a fibrotic issue first, and we don't know where that is. But from a patient perspective, it gives you an option -- and I understand that mycophenolate seems to work as well, and it gives you an option above and beyond.

I've heard other doctor -- I'm just a simple guy from Iowa here, sitting here today. Other people in the room have used in conjunction with another option, and 40 milliliters doesn't sound like a great volume in your lungs, and I understand that. But when you're looking at mortality, I look at that and I go, okay, if I can get some more out of my lungs, and I don't know if that's another year, two years, down road as another option on top, that that seems like a good option to me.

DR. NIKOLOV: This is Nikolay Nikolov. Again, just to clarify the reasoning behind us

asking this question, it's pretty much we struggle with the same issues or questions that you do. And again, from a statistical perspective, all of these are post hoc analysis. Particularly the subgroup analysis is merely a curiosity, and we just wanted to get the opinion of the committee of what did you think about this.

Again, to Dr. Weisman's comment, we take these data at face value, but we don't make decisions based particularly on these. Ultimately, the primary endpoint was met for this study, and the rest of the data are, again, open for discussion. Subgroup analysis can be very challenging, tricky, and not necessarily easy to interpret.

DR. SOLOMON: Dr. Redlich?

DR. REDLICH: Well, I'm not a statistician, but my simplistic view is not that it's been shown not to be effective in the U.S., it's just that the sample size in the U.S. was too small to show benefit. And that's why we have the larger population.

Is that a simplistic way of looking at that?

DR. SOLOMON: I think the fact of the matter is the point estimate was moving towards the null, but you're absolutely right that it's in a small sample size.

DR. REDLICH: Then it's also further complicated by the grade or use of the other medication in the U.S.

DR. SOLOMON: Clearly. I think the point that Dr. Nikolov makes about it's the overall result is what is being presented here, that is the primary outcome that was prespecified. We are not -- I think that the efficacy conversation is around the totality of the data and not on the subgroups of the data. The subgroups, we like to look at them, but we have to come back to the totality of the data.

DR. GELLER: I have a statistical question. Was region and MMF use in the model, in the primary analysis?

DR. YU WANG: Yes, region and -- I need to check about MMF use. I don't think MMF use was in

the model. It's ATA status.

DR. GELLER: Right. That's a stratification variable, so of course it's in the model.

DR. YU WANG: I don't think -- maybe the applicant can correct me.

DR. TETZLAFF: Yes, we can. I ask Dr. Voss, the project statistician, to allude to the factors that were included in the model.

DR. VOSS: Florian Voss from Boehringer Ingelheim. We have not included region and MMF in the primary analysis model, but we did is we included MMF in a sensitivity analysis in the model, and it showed consistent results with the primary analysis.

DR. SOLOMON: Are there other comments that people want to make regarding this question? If not, I'm going to summarize.

(No response.)

DR. SOLOMON: It's a brief summary. There were a number of interesting secondary analysis, some of them prespecified, some of them post hoc, that do give some differences in point estimates,

confidence intervals widely overlapping. We don't have great explanations for why these are the case. There was some statistical conversation about what is and isn't in the model. I think that's really the summary of what we've heard.

We get to take a break. We will now take a 15-minute break. Panel members, please remember, no discussion of the meeting topic, and we will resume at 3:20. Thank you.

(Whereupon, at 3:03 p.m., a recess was taken.)

DR. SOLOMON: It's about 3:20, so why don't we gather?

Before we move to the voting questions, there is some further data that the sponsor has prepared with respect to this issue of observations of the FVC that might have occurred outside the 52-week window that they re-analyzed. So we'd love to have them have an opportunity to present those.

DR. TETZLAFF: Can we have the slide, please? And Dr. Carroll will speak to this.

DR. CARROLL: Hi. I'm Kevin Carroll,

statistical consultant. I'm just making sure this is the correct slide. Yes, this is the correct slide; just making sure.

Unfortunately, we did this in a rush. It's something the FDA asked us to produce, and just ignore the N's, but the bars and the treatment effects are all correct.

What is this? This was the primary analysis. We were asked to include those patients who had a value just outside of the window. You remember there was 28 of those. The FDA asked us to restrict the time period outside of the window to 28 days. So when we do that, we repeat the primary analysis. It doesn't include 28 patients anymore; it includes 24 because 4 of them, 2 in each arm, had a value more than 28 days.

So when we include those data, or as FDA asked us to do, then you can see the treatment effects, about 43 mL, 45 percent difference. I think that answers the question. I have two other quick things to add.

If we repeated the tipping-point analysis on

this kind of approach with these patients added in, the detrimental delta is no longer 45; it's 120, which I think changes one's view of the robustness of the primary endpoint results.

The last little thing to add, there's a lot of discussion about regional effects on subsets and so on. It may be of interest that in the previous two IPF trials, there was absolutely no evidence of any regional interactions. If anything, the treatment effect was slightly higher in U.S. patients. It's interesting to always look at some independent evidence, is there some regional issue generally, and it certainly wasn't seen before. Thank you.

DR. NIKOLOV: Dr. Solomon, this is Nikolay Nikolov. Just for clarification and for transparency, and for the record, we requested these additional analyses based on the discussion that happened earlier today and the questions that came from the impact of missing data on the primary analysis.

Based on this additional wider window of

capturing the primary endpoint, the amount of missing data decreases, and the impact on the tipping-point analysis changes dramatically. Notwithstanding Dr. Geller's comment that these might be different patients, but we just want to bring this up for the discussion and see if the committee considers this a reasonable look at the data.

DR. SOLOMON: That was very helpful; thank you. I'm not sure if there's further discussion that people want to have on this point. I don't know, Dr. Geller, if you have some thoughts.

DR. GELLER: I guess I'd like to maybe repeat the interpretation. So when you do this, you have less missing data. You don't know exactly who you're adding. But the p-value for the overall result goes down, get smaller, and that has a dramatic effect on the tipping-point analysis. The question is, this was not prespecified but it came up in discussion here, so that's the question.

DR, NIKOLOV: We're just trying to get as much discussion of any available data from this

program, and we appreciate the committee's input on this.

DR. SOLOMON: Okay. If there's no further comments -- Dr. Curtis, did you want to raise -- you had one question.

DR. CURTIS: Hi. It's Sean Curtis. I had a question to the sponsor, clarifying, again, on our topic of the observed smaller point estimate in the face of mycophenolate use. Could you just remind us of how the drug nintedanib is cleared; what the metabolism is; if there's been any drug-drug interactions with mycophenolate, for example; and whether you have any drug levels from the trial in the face of mycophenolate, just to help provide a little clin-pharm around this issue.

DR. TETZLAFF: Absolutely. We have looked at this, and our PK expert, Dr. Wind, will provide some of our insights with you.

DR. WIND: Sven Wind, clinical pharmacologist at Boehringer Ingelheim. You asked specifically about the drug-drug action potential, I think, for both. What we can say is that

nintedanib and mycophenolate, based on all the data that we have, do not have the potential to interact because they do not inhibit or induce the metabolism or transport of each other drug.

To underline this, we actually have also data from the SENSCIS trial. This is actually a box plot comparing the exposure data from patients on mycophenolate and without mycophenolate. You see that the exposure nicely overlaps for the trough levels.

DR. SOLOMON: Okay. Unless there are more comments, I think we can close the discussion portion and move to the voting.

I just want to read that we have a new electronic voting system for this meeting. Once we begin the vote, which we will do in just a little bit, the buttons will start flashing, and they'll continue to flash even after you have entered your vote. Please press the button firmly that corresponds to your vote. If you are unsure of your vote or you wish to change your vote, you may press the corresponding button until the vote is

closed.

After everyone has completed their vote, the vote will be locked in. The vote will then be displayed on the screen, and the DFO will read the vote from the screen into the record. Next, we'll go around the room, and each individual who voted will state their name and vote into the record. You can also state the reason why you voted as you did if you want to. We will continue in the same manner until all questions have been answered or discussed.

Any questions about the procedures?

DR. NASON: We had spoken over the break about possibly having some discussion.

DR. SOLOMON: Yes.

DR. NASON: Okay.

DR. SOLOMON: Before the vote is done, I will ask if there is any further discussion; but just to be straight on the voting procedures.

DR. STOLLER: Keep pressing?

DR. SOLOMON: I don't think you have to keep pressing; press once; just press. If you want to

change your vote, you can press it again, Michael.

So I'm going to read the first question, and then I'm going to open it up if there's discussion, and then we'll get to the voting. Question 3, a voting question is, do the data provide substantial evidence of the efficacy of nintedanib for the treatment of systemic sclerosis interstitial lung disease? If no, what further data are needed?

Does anyone have further discussion points around this voting question?

(No response.)

DR. SOLOMON: Seeing none, we can move to the voting.

(Vote.)

DR. WANG: For the record, for question number 3, we have 10 yeses, seven no's, zero abstain.

DR. SOLOMON: Okay. Now that the vote is complete, we'll go around the table and have everyone who voted state their name, their vote, and if you want to, you can state the reason why you voted as you did into the record.

Dr. Curtis is nonvoting, so Dr. Geller?

DR. GELLER: Nancy Geller. I voted no because I think the results are quite tenuous.

DR. STOLLER: Jamie Stoller. I voted yes. I'm aware that yes/no is a dichotomous outcome, but I'm always given to condition or qualify a level of energy behind my vote. And I would say that my vote was yes with a very weak level of affirmation. And I would condition that comment by saying this is the classic problem of trying to make dichotomous decisions with inadequate data, which simulates what we do in clinical medicine all the time,

That said, as is done in guideline documents, I want to be clear on what anchors my decision. And my decision is anchored by a deep appreciation of the difficulty of doing clinical trials in rare diseases. I take care of many patients with unusual lung conditions, where this issue abounds about inadequate data. I'm sure it's true with my rheumatologic colleagues as well.

An appreciation of the profound -- as I

think has been amply articulated by many colleagues here, by the profound unmet needs, so while I'm not at all confident in the magnitude of impact or its robustness, based on the very slicing and dicing we've done, I would say they did in fact meet the primary outcome measure.

From my prior experience on this committee, we have anchored decisions on failure to meet the primary outcome measure with very marginal misses, and I think it's, therefore, difficult to discount having achieved a primary negotiated prespecified outcome measure, not withstanding all the qualifications on the data.

DR. WEISMAN: Hi. Michael Weisman. I voted yes because I've been so unhappy over 40 years of watching the failure of generalized immunosuppression and anti-inflammatory drug therapy treating anything in scleroderma. And it was only when we understood the biology of hypertension that we were able to eliminate scleroderma renal crisis. And now that we think we understand a lot of the biology of fibrosis, we're

able to use that to treat this organ involvement in scleroderma.

So I think, on balance, I think we headed in the right direction, and that's why I voted yes.

DR. KERR: Gail Kerr. I voted no because I thought that although statistically significant, I didn't think those millimeter changes were clinically significant to that population of patients. I am disappointed in that outcome because I think we're all looking towards some kind of targeted therapy.

The fact that it didn't do any benefit to the secondary outcomes, I think also negated that small efficacy when you consider the side effects of the drug, and some patients lost significant weight because of it. I think it's a challenge going forward to try and decide the exact patients in whom to give this drug to. And if we're thinking that fibrotic initiates inflammation, maybe the study to do is to compare it to mycophenolate in another trial.

DR. GARIBALDI: Hi. I'm Brian Garibaldi. I

voted no, and I really struggled with this. I think while, yes, the primary endpoint was met, I think given the small magnitude of the benefit and the uncertainty of its clinical significance, I'd like to see more data before I'm convinced that this is a drug that we should be using in patients with scleroderma.

DR. MAY: Susanne May. I voted no because even though it was statistically significant, it was on a surrogate marker and was not supported by some of the other secondary outcomes. The question asked about substantial evidence, I would consider this as some evidence. But a number of different factors that I would have wanted to see for this rare population, like the backup on secondary, more patient oriented outcomes would have convinced me otherwise. But without that, it was a no.

MR. GILLIGAN: Tom Gilligan, and I voted yes for some of the reasons that were mentioned before of the fibrotic side of the disease rather than the immunosuppressant side of the disease as another medical option for you all to decide when and where

to use with patients on that side.

Again, I understand there are some statistical questions still out there, but I look back, again, at the history of where this disease has come in a short period of time, and I know the population is small and the timeline is short for, I guess, mortality, looking at it from my perspective. Therefore, that was my reason for my, yes.

MS. HORONJEFF: Jen Horonjeff. I voted yes as well. They met the primary endpoint as was described by the FDA, as well as through OMERACT. While I am here as a consumer representative, pointing out that OMERACT does work with patients in coming up with those, so although they didn't meet some of the PROs, I'm confident that the outcome measures that were decided as a primary endpoint were still something that was significant for patients. And since that was met, and given the unmet need of the population, I voted yes.

DR. CALHOUN: Bill Calhoun. I voted yes. This study met its primary endpoint. I don't think

that it's fair to disparage forced vital capacity as the endpoint when it's been suggested as a principle endpoint by august groups who think about this. And the agency agreed that it was the right endpoint to look at. So I don't think that it's right to disparage vital capacity as just a surrogate. I think it's an clinically important endpoint that has implications.

I'm compelled by the orphan status of the drug. I'm compelled by the lack of alternatives. I'm compelled by the case statement that Dr. Seibold and Dr. Brown made. I think, frankly, this drug's already on the market, so however the approval process goes, my guess is that people will probably use this, perhaps off label -- if reimbursement can be had for it, they'll use it off label because there are no other alternatives.

I think the issue of the lack of secondary outcomes may be actually a strategic error in the design of the trial. The secondary outcomes of PROs that were evaluated haven't been validated in this disease, so why would you expect them to

change? The SGRQ, in particular, has a short recall time, so if you're looking at that over the course of a year, there's going to be baseline drift.

Again, the fact that none of the secondary outcomes hit bothers me not at all. The physiology is pretty hard to argue with. And although the magnitude of the effect is small, again, as Dr. Seibold and Dr. Brown pointed out, this is a heterogeneous disease, and some people will perhaps have really remarkable benefit from it. And for those who don't, they'll probably stop.

DR. KATZ: James Katz. I voted no. I'll echo the sentiments of Dr. May, who I think put it very nicely. It would help to see benefit in DLCO and high-resolution CT.

DR. SOLOMON: This is Dr. Solomon. I voted yes. The considerations that I had were related. This is a disease with tremendous unmet need with a hard endpoint that was prespecified and met. While the data are clearly not as robust as anyone would like to see, we're dealing with a disease that is

rare, a trial that is difficult to pull off, and I think we all would like to see further data, but I think for now, I believe it is efficacious.

DR. REDLICH: Carrie Redlich. I voted Yes for I think the reasons stated with some of the ambivalence that was also stated, but recognizing the challenges of doing studies in a rare disease, that the substantial may be a little less substantial than we would all like.

DR. CURTIS: Jeff Curtis. I voted no for I think reasons that have been stated. The magnitude of the effect size was small at best, and I think probably not very clinically relevant. Certainly within a year, symptoms, or mortality, or any other secondary endpoint that didn't really even have much supportive evidence or trends I think was honestly disappointing to me, so that's why I voted no.

DR. NASON: Martha Nason. I voted no. My reason, again, have largely been stated, especially by Dr. May in the sense that I had the same reaction to the word "substantial." If it had said

"moderate efficacy," for instance, I might have voted yes because I think there's some evidence here, but I don't think it's substantial evidence.

That was based on all sorts of things, including the sensitivity analysis with the missing data; including the secondary endpoints not lining up; including the effect size We'll get to the cost benefit -- sorry, I keep saying that -- risk-benefit question, but to me it's very hard to separate those, even just the 75 percent who were having vomiting and all of that when I think about efficacy.

DR. OLIVER: Alyce Oliver. I voted yes. I thought it was a difficult decision, as stated, because of the word "substantial," but it did meet its primary outcome. There are a lack of alternatives. This is a devastating disease, so I think we need any help that we can get to treat it. I was buoyed a little bit by the limited data at a hundred weeks, which did show a continued or sustained effect, which is better than what we're seeing with CellCept and Cytoxan.

DR. RICHARDS: John Richards. I voted yes. I was also concerned with the word "substantial," but in the end, vacillated towards yes. It is a rare disease with devastating consequences. I was not terribly concerned about it not meeting the secondary end endpoints. I don't think we have good endpoints in terms of patient-reported outcomes for this disease, so that did not disturb me.

I think looking at skin scores would be a different study. You may design it differently with earlier patient enrollment. We know that the skin scores tend to improve later on in the disease, and these patients had disease duration of several years at the start of it.

DR. BECKER: Hi. This is Mara Becker. I voted yes. At the risk of repeating, I agree with everyone who has already commented, but my rationale is that they met their primary outcome. I was further encouraged by the addition of the data within that first 28 days after the 52nd-week time point, and, frankly, we don't have any

options, and this to me is, in some ways, a safer option than some of the drugs that we use routinely in these patients.

So a little diarrhea, as I thought about, might be better than some of the other significant side effects we get with cyclophosphamide and mycophenolate, so I voted yes.

DR. SOLOMON: Now we're on to question 4, another voting question, and we'll proceed in a similar way, but we're going to start on your side, Dr. Becker, so just be prepared.

Is the safety profile of nintedanib adequate to support approval of nintedanib for the treatment of systemic sclerosis interstitial lung disease? If no, what further data are needed?

DR. GELLER: Did we do the if no, what further data are needed for the first question?

DR. SOLOMON: We didn't do it formally. I thought people had stated it in their discussion points, but maybe not. I guess not.

DR. GELLER: Well, I think a second trial, which might have a slightly different design to

just knock the ball home, would be a great way to get approval. I think there are several different ways to design such a trial. You can do an MMF plus or minus the nintedanib, or you can do MMF versus, and you can do this for different durations, longer than 52 weeks perhaps. I know it's hard to do, but I would like the data to be more convincing.

DR. SOLOMON: Is there anybody else who voted no that wanted to be more detailed, like Dr. Geller, about what other data would be useful?

DR. NASON: I guess I would just echo that I think it's the long term, slightly longer anyway; Maybe not 10 years, but maybe 2 years, a little bit more than the one year. The more we have to follow-up, especially since this is a slower disease than IPF. Their decrease is slower. I think it's the longer data that's needed.

DR. SOLOMON: Did you vote no?

DR. WEISMAN: I voted yes [off mic].

(Laughter.)

DR. SOLOMON: It said if no. Go ahead.

DR. WEISMAN: I was wondering whether or not the sponsor had any data that they've collected or interested in a new technique of quantitative HRCT analysis, which seems kind of exciting to look at the rate of change of ILD in scleroderma ILD patients. There has already been some interesting preliminary data that's been published already on that technique, and it seemed, at best, it would be a very sensitive way at looking for rate of change in a very specific marker of ILD.

I was wondering if the sponsor had any intention of doing those kinds of studies or maybe they've already looked at that. I don't know.

DR. SOLOMON: I don't think we're going back to the sponsor right now. I think we're just making suggestions for what other data may be collected in future studies.

Dr. Stoller?

DR. STOLLER: I had a highly qualified yes, so perhaps allows a comment. The recommendation for additional data, as I've said before, is to pay additional attention to the methodologic adequacy

of the spirometry measurements and centers, and to be able to respond to the question, recognizing that only 6.7 percent failed. The question is where are those 6.7 percent, and can we answer the question about geographic variation and methodologic caliber?

DR. SOLOMON: We're going to go on to the next voting question, and if you are no and you want to discuss, why don't you try to do it when you're giving your rationale, if that's okay. So let me just read it again.

The safety profile of nintedanib, is it adequate to support approval of nintedanib for the treatment of systemic sclerosis ILD? If no, what further data are needed?

Is there any discussion on this point? Okay.

DR. GELLER: I've been wondering about these digestive effects. Specifically, as soon as I opened all of the reading material, I saw the 75 percent diarrhea with this drug, and there's also nausea and vomiting. I just wonder how -- well,

one question is, how many patients in each arm had combinations of these? How often were these things repeated? Because it seems to me that they affect quality of life in a really pretty great way.

That's a very high percentage of diarrhea, and even though much of it is not less than moderate, I still am very concerned about what that means to the patient.

DR. SOLOMON: The sponsor I think might have some response.

DR. TETZLAFF: Thank you, Mr. Chairman. I think that's a very important question, and the management of diarrhea is something where we have management schedules that we offered for patients. But I think to answer really this question, the best context is really the clinical context, and I'd like to ask Dr. Maher to speak to the issue of how to handle patients with diarrhea, how it looks like, and what extent of an issue and non-issue this is.

DR. MAHER: Thank you. Ted Maher, Imperial College. Yes, as already stated, I'm a

pulmonologist. I look after both scleroderma ILD patients, and I look after patients with idiopathic pulmonary fibrosis. I've been using nintedanib, in fact, since the first INPULSIS trials in 2010, and I've got patients who've been on drug for 8 years. I've got over a thousand patients taking drug.

Just to give you some context, I think when you look at the 75 percent diarrhea figure, that's quite startling. But to put it in context, I think one has to understand what is meant by the term "diarrhea," because that can go all the way from torrential passing of stool, as we see with some oncology drugs, through to what we actually see with nintedanib, which for most patients is a change in bowel habits.

So often they'll be going to the toilet twice a day instead of once a day, or alternatively, the stools that they're passing will be softer than usual. For a very small proportion of patients, they do get some fecal urgency. So the 1 in 20 patients who really struggle, it's because when they have to go to the toilet, they

have to go pretty quickly. But for the vast majority, it's something that's easily manageable, either with lifestyle changes, so reducing high fat content in diet, using antidiarrheal drugs such as loperamide, pr sort of just learning to manage the symptoms day to day.

In our practice, we've gone from 5 or 6 years ago, having about 30 percent of patients discontinue drug over 12 months, to that dropping down to 10 to 15 percent within 12 months discontinuing because of side effects.

So I think the practical reality when you use the drug is that, for most patients, it's very tolerable, and the pure numeric figures that you see when we report the side effects make it look a lot worse than it is. And you've got to remember that diarrhea was very specifically asked about at every study visit, which is why 35 percent of the placebo group also reported diarrhea.

DR. SOLOMON: Thanks.

There were a couple other points. Dr. May?

DR. MAY: Yes. My apologies for the

previous mistiming of the other question regarding pneumonia, but I think it falls into the safety profile question, and for me to understand it as a non-clinician. Even though we might not have the data or see it, but just theoretically, could there be an increased risk of deaths because of the increased incidence of pneumonia in this subgroup?

Number two, a question, even though the secondary analysis were not statistically significant, on the quality of life, it was in the wrong direction. Could it be that it is in the wrong direction -- I don't understand the questionnaire well enough to answer this. Could it be in the wrong direction because of the increase in adverse events that they have, so that they have an impact on their quality of life because of the anticipated adverse events?

What's really interesting for me to hear is that a lot of the other drugs that are used in this patient population have much worse outcomes, but getting the opposite direction on the quality of life was surprising to me.

DR. TETZLAFF: Thank you again for giving us the chance to speak a little bit about the pneumonia since this is understandably a concern to the audience. I ask our safety specialist, Dr. Kohlbrenner, to speak to the pneumonia data.

DR. KOHLBRENNER: Veronika Kohlbrenner, Boehringer Ingelheim. So very specifically to your question about whether pneumonia could have an effect on deaths, I can speak very specifically to the one case of pneumonia that occurred, which led to death was a very complex, prolonged hospitalization in the patient I had previously mentioned with scleroderma renal crisis. While the scleroderma renal crisis had resolved, the patient was prolonged in the hospital and developed pneumonia and sepsis. So whether there was no nosocomial effect involved is a possibility.

There was a second patient. In terms of among the few patients who died within the SENSCIS study, the second patient was also a very unusual patient in that that patient died within 3 weeks of study initiation, also during an ICU stay with

ventilatory deterioration.

So there were definitely complex associations, so the correlation of the pneumonia events to the death outcome events, we do not see, albeit, these are fortunately a few cases.

DR. TETZLAFF: And when we try to increase the number of events by grouping for respiratory system organ classes, we don't see any imbalance any more. That was also something that was introduced in the presentation.

If I have to chance, there was the comment on the SGRQ going in the wrong direction. I guess Dr. Maher would also want to speak about this because we don't actually -- I'll let Dr. Maher speak.

DR. MAHER: Ted Maher, Imperial College. I think it's a very quick answer around the St. George's Respiratory Questionnaire. Insomuch as I alluded to earlier, it's a hundred point scale. The difference that we were seeing in the placebo group was minus 0.8, so less than a 1 percent change over a year. In the nintedanib group it's

plus 0.8 in the opposite direction. The noise of the instrument is about 2 and a half percent, so this is well within the range noise. I wouldn't read anything into that change, truth be told.

In the IPF studies, for instance, where we have a more rapid progression of disease, we see a 3 to 5-point change in the placebo group over 12 months. So I think the level of change we're seeing is just uninterpretable because it's within the noise of the instrument, and I think such a tiny change in either direction is insignificant clinically.

DR. SOLOMON: Dr. Katz?

DR. KATZ: James Katz. I wanted to go back to the question about diarrhea in this population and quality of life. I think it's really important to keep in mind that in this particular patient population, a rheumatologist confronted with a patient who has a change in bowel habits has to really think carefully because these patients get small bowel overgrowth, malabsorption, watermelon stomach, and wide-mouth diverticuli. These are the

patients that end up getting colonoscopy, endoscopy, and it's a big deal even if it's a tolerable diarrhea.

DR. SOLOMON: Dr. Redlich?

DR. REDLICH: This also relates to the rate of adverse effects, which, in my understanding, particularly the GI symptoms are also related to the percentage, the 40 percent or so that had a dose reduction, and also the pretty high rate of treatment interruption, 38 percent.

I was sort of curious and went back to look at the INPULSIS study and UIP, and there was a lower -- I think, as was mentioned, both need for dose reduction or dose interruption were lower. I guess my question is that my understanding is it's been attributed in at least large part to the greater likelihood of GI symptoms in this group.

I was also wondering, the other big difference between the group is the UIP group was over 70 percent male, and bigger people, their mean kilogram weight was higher. This study was largely, I guess around 70 percent, women with also

smaller mean body weight. We haven't really gotten into a dose discussion, and that wasn't one of the questions we were asked, the dosing of different size people, and is it possible that this is not just a GI component, but potentially in smaller size people, you might be dosing this at a lower dose.

The other thing, it ended up that a greater percentage of people were on the lower dose, I think the data showed, in this study compared to the INPULSIS, which made sense because people, their dose was reduced for a period of time.

So I don't know if there were just any thoughts about that or how that's managed.

DR. TETZLAFF: I'm not sure we fully understand the question here.

DR. REDLICH: I guess the question is, are you at all concerned about using the same dose for all-size people?

DR. TETZLAFF: I'd like Dr. Stowasser to respond to this directly.

DR. STOWASSER: Susanne Stowasser,

Boehringer Ingelheim. What we do know is that there is no need for dose reduction in any subgroup of patients that are characterized by factors that might impact exposure, which is older patients, lower body weight patients, or race, Asian patients. These are the groups that tend to have higher exposure.

This takes into account the high variability in exposure that we see in this drug that is around 50 to 80 percent the coefficient of variation. There's one exemption. These are patients with, as per labeling, as has already been mentioned today, with hepatic impairment, with hepatic impairment mild, Child-Pugh [ph]. As per label, it's recommended to start with a lower dose with a 100-milligram bid.

The reason why we do not recommend an a priori dose reduction in lower body weight in elderly patients is that we risk -- that we would risk an exposure that is non-efficacious because 150 milligram results in a plasma level that is close to the exposure of maximum efficacy. I

explained the variability of exposure, so that's why we would not recommend a starting dose of 100 milligram.

DR. REDLICH: Thank you.

DR. SOLOMON: Dr. Becker?

DR. BECKER: I'm sorry, Dr. Kohlbrenner that you're sitting down now, but I know you had mentioned earlier to a question of mine, and I wanted to reiterate to the folks who are not clinicians on the panel, from a pneumonia standpoint, a substantial number of those patients were on immune suppression therapy, correct?

I feel like that's kind of important when we're trying to decipher how much of that pneumonia is due to this drug versus all the other drugs and the comorbidity of having chronic lung disease on top of it.

So I was hoping you could just remind us of how many folks that had severe pneumonia were also on concomitant immune suppressant therapy.

DR. KOHLBRENNER: Again, Veronika Kohlbrenner. Among the 8 serious pneumonia cases,

there were 5 patients who were on concomitant mycophenolate, and two of those were also on concomitant -- or two of those also had recently added cyclophosphamide.

DR. SOLOMON: Todd, did you have a --

MR. GILLIGAN: I asked that question earlier, and then I had that down, and I didn't know if I could interject that. And my only other comment to anyone who's not a medical professional on that is I started my mycophenolate -- I started at 1000 milligrams for the diarrhea, the vomiting, and all of the symptoms listed for nintedanib the same way, to see if my stomach could tolerate that drug as well.

I don't know if I'm allowed to interject this at this point, but I don't see a lot of difference between the two on that side.

DR. SOLOMON: Good. This has been a good clarifying discussion.

Another point?

DR. NASON: It's actually more of a question to the clinicians. It was helpful to me, too, when

Dr. Becker made the comment about other drugs and their profiles because I don't really have that background. I guess as I sit here and think about how to vote on the next question, it's clear there's a safety signal in terms of diarrhea, pneumonia, hepatic changes and events, bleeding events, hypertension, weight loss.

There are a lot of safety signals, and I don't know how to think of those in terms of this population over several years, let's say, because they may, again, don't have quite as fast a disease course as the people this drug's already used her for and as compared to other things they might take.

I don't really know what to do. I guess I'm struggling with could we be causing more risk to them -- and maybe this is the next question -- with these, increased hypertension, changes in liver, changes in bleeding -- I don't know. I don't know how to put that into context, I suppose, and I don't know if any of the clinicians could give me any more insight or if I'm just destined to

struggle with it.

DR. SOLOMON: Dr. Becker?

DR. BECKER: From my perspective, I first looked at this safety profile as it related to the RLD label safety, like what we know about the safety, which reassuringly, there was nothing new in this population of patients. I think you've heard throughout the complications of these are people that can have significant GI disruption from their disease. They're at risk for immune suppression from other drugs that we put them on. They may have reflux and dysmotility just because of their underlying disease.

So it is sometimes hard to piece out, and I think that when I personally looked at the safety analysis of the data that were presented, I thought, well, that's pretty much in alignment with what has already been known. When you think about the grand magnitude of what these people are faced with from their disease burden, I still think that that's on the lighter side compared to what they have to deal with just by having scleroderma,

systemic sclerosis, which affects multiple organs in a major way.

DR. SOLOMON: Dr. Garibaldi?

DR. GARIBALDI: I just wanted to comment on that as well. I think this gets into what we've heard from both patients and other advocates in the room, is that this is a unique conversation between a patient and their own physician, particularly for some of the GI side effects, what they're willing to tolerate and how that can be managed by either dose reduction or other adjunctive therapies, particularly for the diarrhea issues.

So I don't see this as being any different from what we manage in IPF, recognizing that the likelihood of increased GI side effects is probably because there's an increased incidence of GI issues in patients with scleroderma to begin with.

So I don't see this as being -- I think these side effects, there'll be something that physicians and patients are going to have to deal with, and discuss, and make decisions about what people are willing to tolerate, but also remember

that patients with IPF -- obviously scleroderma patients are very sick as well, but IPF patients are a much older population with other comorbidities as well, and we tend to be able to manage these side effects with careful monitoring as long as you're checking liver function testing and checking with your patients. It doesn't seem to be something that's out of proportion of what we've seen in the IPF population.

DR. SOLOMON: Todd Gilligan, do you want to make a comment?

MR. GILLIGAN: My one last comment on that one is, again, the 3-week blood work becomes pretty common for those of us in the community with it on that end. To your point, between patient and doctor, those are the reasons I chose to wait 6 months before the mycophenolate. We eased into it, then upped my dosage from 1000 milligrams to 1500 because I could tolerate.

I didn't have the diarrhea symptoms, but again, the blood work becomes common on that side. If pneumonia symptoms, colds come on, we can

drop -- it's got to be that dialogue and conversation between patient and doctor should this become available; my 2 cents.

DR. SOLOMON: Okay.

MS. HORONJEFF: I'll just add to that. Jen Horonjeff. I think this is something that we wrestle with in rheumatology and oncology already with different medications, so these trade-offs between the side effects versus what's happening to the patients and the disease is of course something that needs to be weighed out, just echoing that conversation and shared decision-making.

But especially what we're hearing in this particular disease is that this has a lethal outcome if we aren't treating it, so sometimes when we're seeing this in rheumatology, it might not be as dire, but giving the patients those opportunities to figure what's best for them and their families, and what that means for them in their treatment.

DR. SOLOMON: Okay. Why don't I re-read the question, and then we're going to go to voting. Is

the safety profile of nintedanib adequate to support approval of nintedanib for the treatment of systemic sclerosis interstitial lung disease? And if no, what further data are needed? If you could put that in your discussions. So we'll go to vote now.

Do people want instructions again?

(No response.)

(Vote.)

DR. WANG: For the record, question number 4, we have 14 yeses, and 2 nos, and 1 abstain.

DR. SOLOMON: As promised, I'm going to start with Dr. Becker.

DR. BECKER: This is Mara Becker, and I voted yes. As mentioned, I think the adverse events that were reported in this trial were in line with the known safety profile already that has been already reported with IPF and on the current label. I also think that the pneumonia signal that we see is hard to interpret in light of the fact that many of these patients were already on immune suppression, which could complicate that finding,

so I voted yes,

DR. RICHARDS: John Richards. I voted yes. I think the safety profile is in keeping with what's already known about the drug. There are additional concerns in patients with scleroderma in that they do have a lot of GI symptoms as well. But again, I think that comes to a discussion between the patient and the doctor, and I think this patient group, as well as physicians, are kind of used to monitoring liver and other potential toxicities of this drug, and symptoms seem to abate with dose reductions and stopping.

DR. OLIVER: Alyce Oliver. I voted yes. I did not note any new safety signals compared to what is already known with use in IPF.

DR. NASON: Martha Mason. I hesitantly voted yes, based largely on this discussion we had just two minutes ago. I do agree, it seems like the safety here is in line with the safety profile from IPF. What I've struggled with is how that translates for people with a different disease and with a different time course of disease, so

therefore may be spreading those safety issues out over more years.

So the only thing I really want is longer term data. Short-term data, I don't think there's too much missing. I think it probably is similar to IPF, but I really would like that longer term data to see how this all plays out.

DR. CURTIS: Jeff Curtis. I voted yes. Although the safety profile of this drug isn't benign, there was nothing here that concerned me excessively that would be beyond the ability or even the comfort of rheumatologists, pulmonologists, and other specialists that manage this disease. There didn't seem to be anything here from a safety perspective that was vastly worse than mycophenolate and certainly not cyclophosphamide.

So if rheumatologists and pulmonologists are comfortable with that, and I think most are, this felt on par, compared to cyclophosphamide even, less toxic than that agent.

DR. REDLICH: Carrie Redlich. I voted yes.

I agree with the previous comments.

DR. SOLOMON: This is Dan Solomon. I voted yes. Again, similar to what others have said, the safety profile is in line with the known safety profile of the use of the drug in IPF. The fact that it's been on the market for IPF now for 5 or 6 years, and there hasn't been anything new in postmarketing surveillance, is also comforting. In the data that were presented, they are in line with many drugs that are used for scleroderma.

DR. KATZ: James Katz. I voted no. The danger is the assumption that the adverse effect profile is actually manageable, and that's fine if that's true. But if the drug precipitates renal crisis, if it causes me to miss malabsorption, if it results in weight loss that increases mortality, then only with more time are we going to know that this adverse profile is actually manageable.

DR. CALHOUN: It's Bill Calhoun. I voted yes, and I did so because the safety profile appears to be consonant with what's in the label, number one. And number two, the pneumonia signal I

think is probably expected, based on the degree of immunosuppression these people have, and the degree of esophageal dysmotility, and perhaps microaspiration that they've got. It's a concern, requires follow-up, but I don't believe that it rises to the level to warrant a no.

MS. HORONJEFF: Jen Horonjeff. I voted yes for a lot of the reasons that have already been stated, but I think, again, I'm just trying to empower patients and physicians to make these decisions together and decide what's best for their own treatment plan.

MR. GILLIGAN: Todd Gilligan, and I voted yes for the reasons that were previously stated. Again, I don't see any other risks that are here that aren't already available for the treatments for the disease today.

DR. MAY: Susanne May. I voted yes on the background that the median life expectancy is 5 to 8 years in a population that's relatively young, and other treatments don't seem to have -- sometimes you have worst side effects than

this one.

DR. GARIBALDI: Brian Garibaldi. I voted yes. Again, I have some concerns about the potential interactions of GI side effects with things like pneumonia or aspiration, but I don't think there is a clear signal in this data that that's happening. That's certainly something that needs follow-up, but based on the data presented, I voted yes.

DR. KERR: Gail Kerr. I voted no simply because the profile is similar to IPF, but it's the magnitude of the diarrhea that's not offset by the benefit of the drug that concerned me. I would therefore offer consideration that given most of the patients who had the side effect, you're able to reduce the dose in those patients and demonstrate efficacy despite your concern for lack of adequate plasma levels for efficacy. You might want to consider, in this population or a subset, actually going with a lower dose, 100 milligrams.

DR. WEISMAN: Michael Weisman, and I voted yes for the above-mentioned reasons.

DR. STOLLER: Jamie Stoller. I voted yes largely for reasons stated. I would agree with the recommendation for longer follow up and would simply comment on the pneumonia data to suggest this was unassociated with an increased mortality risk.

As I understood the data, it's difficult to ascribe that to immunosuppression alone because that was balanced, as I recall, between the control group, placebo group, and the treatment group. And in that regard, that question will only be answered by longer term follow-up. That's essential in my view.

DR. GELLER: Nancy Geller. I abstained because you guys seem to think the safety profile is manageable, and I think the adverse event tables indicate this is a drug that I couldn't imagine taking. So I decided I would sit this one out.

DR. SOLOMON: Now we're up to question 5, which is the final question, a voting question. Here we have the benefit and risk profile, and is the benefit-risk profile adequate to support

approval of nintedanib at the proposed dose of 150 milligrams twice daily for the treatment of systemic sclerosis interstitial lung disease? And if no, what further data are needed?

So before we go to the voting, do people want to discuss this balancing of risks and benefits? Would that be useful? I don't know if Dr. Nason, or Dr. Oliver?

DR. OLIVER: Alyce Oliver. I just have a question for the sponsor. Is the indication for nintedanib monotherapy or in combination with CellCept?

DR. TETZLAFF: The indication for nintedanib that is suggested is nintedanib for the treatment of systemic sclerosis-associated interstitial lung disease.

DR. SOLOMON: That satisfies?

[Dr. Oliver nods yes.]

DR. SOLOMON: Other points, other questions, further conversation about balancing risks and benefits?

(No response.)

DR. SOLOMON: I don't want to belabor it.

With the question in mind, I don't need to read it again, we should go to vote.

(Vote.)

DR. WANG: For the record, question number 5, we have 10 yeses; 7 nos.

DR. SOLOMON: Dr. Geller, I'm going to come back to you.

DR. SEYMOUR: Dr. Solomon, if we can make sure, when you go around, if folks who voted no can answer subpart A, which is what additional data is needed, that would be very helpful. Thank you.

DR. SOLOMON: Great.

DR. GELLER: Nancy Geller. I voted no because I think the benefit is not great, although it met its primary endpoint. I think the safety profile is not very impressive or impressive in a negative way. I think that we need another trial.

DR. STOLLER: This is Jamie Stoller. I voted yes. This is a numerator and denominator question, and it doesn't surprise me that the votes segregated on the prior assessments, the 10-7

split. Having said that, I will again qualify my yes by the level of confidence in that yes, which is quite low. And I think that in that context, longer term follow-up is needed and better attention to the methodologic, as I said before. Ascertainment of the primary outcome measure would help the interpretation of the data.

DR. WEISMAN: I voted yes because of my prior statement. This is now an advance in the management of a very difficult problem, based upon what I consider reasonable science on understanding fibrosis in interstitial lung disease.

DR. KERR: Gail Kerr. I am being consistent, and I made my suggestions regarding the lower dose, 100 milligrams twice a day, possibly.

DR. GARIBALDI: Brian Garibaldi. I voted no, and again, I struggled with this, recognizing the need for therapies for this disease, but I don't think the treatment effect rose to the level of pushing this through with just a single trial. We need more data to really understand what the benefit of this drug is going to be in scleroderma,

DR. MAY: Susanne May. I voted no because I really think that the risk-benefit ratio is not overwhelming or overwhelming enough to say a yes, particularly given relative moderate effect, not substantial evidence. The primary outcome that was a biomarker that is now supported by other secondary outcomes in the way of patient-centered outcomes suggest on the biomarker that has a questionable level of clinical significance in relationship to the side effects and the lack of patient-specific meaningful difference. That was the reason for the no.

MR. GILLIGAN: Todd Gilligan. I voted yes because, again, it's a drug that's already being used, albeit for another diagnosis in the market. The data met the endpoint, and in my opinion, it gives an option beside an immunosuppressant, chemotherapy type drug of an antifibrotic that doesn't exist for the treatment of this disease right now.

MS. HORONJEFF: Jen Horonjeff. I voted yes for reasons I've already stated, that I think that

this is something that is going to be a valuable tool for physicians and patients to be able to use to treat a disease that has a lot of unmet need. I do think it's reasonable to look at postmarket surveillance on this, as well as real-world evidence to see how is this actually coming into play, as we can see this because, of course, we're talking about having a longer term follow-up. But in the immediate need, I'd like to get this into the hands of patients that would help them.

DR. CALHOUN: Bill Calhoun. I voted yes because it was the logically consistent thing for me to do based on my other two votes. In addition, the outcome of patients who have this disorder, who've got an interstitial lung disease related to systemic sclerosis, doesn't look good. It's a fatal outcome, and we have nothing that really is effective in mitigating that. And even though the effect size of this particular agent is not huge, there is some evidence of benefit.

I think docs always need additional tools, and whether this tool is going to be effective for

every patient or not is a question that will be answered with additional data. My guess is that given the heterogeneity of the disease, it won't be right for everybody, but it may be right for some people, and docs need the flexibility to prescribe that.

It also empowers patients to have this on the market, to have the discussion with their physicians in a shared decision-making, and make the determination as to whether the side effect profile and the risks that are accompanying that side effect profile line up with the patients understanding what their disease is and what their life goals are.

In terms of where the company might go, additional studies are always nice. One of the things that they could do that would be very substantively helpful would be to develop and validate a patient-reported outcome that would be responsive to the kinds of changes in physiology that we see with fibrotic interstitial lung diseases.

DR. KATZ: James Katz, and I voted no.

DR. SOLOMON: Dan Solomon. I voted yes, but similar to Dr. Calhoun, I have very strong feelings that this -- I did so with a fair amount of apprehension. I fully support the needs of patients and providers regarding this morbid and mortal condition. However, the false hopes is not what we want to do, and having the data to figure out which patients are really going to benefit, and at what stage in their disease is what we really need to be able to say with certainty.

Clearly, this single study doesn't give us that confidence. It gives us enough confidence to say the drug, in my mind, has efficacy and a safety profile that are adequate and therefore should be approved. However, as far as how to use the drug, I really want to understand the subgroups of patients that will benefit. There was so much evidence in the subgroup analysis that this is not a universally positive drug. I think we really need to understand those subgroups.

So I'd put it upon the FDA and the

manufacturer to really have a very clear program for postmarketing surveillance, whether it's further phase 3 studies, or whether it's postmarketing surveillance studies that help us to define these patient subgroups; that the risk-benefit are inadequate and the drug should not be used as well as the patient groups that it really has the most benefit in. Again, that might take the form of further trials, phase 4 postmarketing trials or observational studies.

DR. REDLICH: Carrie Redlich. I voted yes also with some ambivalence for the same reasons that have been mentioned, more on the concern about efficacy.

In terms of additional studies or things that could be done potentially with data that already exists, I'd just follow up. I think it was Dr. Kerr who suggested further analysis of CT scans. It sounded like there was a subgroup that you might have some honor that potentially in the future would be another outcome to look at.

DR. CURTIS: Jeff Curtis. I voted no,

primarily, again, due to the effect size. I think the efficacy wasn't compelling, and the lack of any secondary endpoints, likewise, wasn't compelling. The fact that it met its primary endpoint at a year didn't really sway me because you can get a significant p-value for any tiny, tiny effect size if your trial is big enough, so I think that doesn't really speak to clinical relevance.

So the bit that we're asked to vote on here as our third vote, is this worth it given the side effects. I think it was hard, honestly, Todd, with some of your comments, for me to think I'm going to look somebody in the face and say, in a year I'm going to give you this drug. Your respiratory function is going to be better in what I can measure, but your symptoms won't be any better on expectation. Your physical function won't be better.

I haven't done anything that I have much evidence for to improve your mortality. If you have any meaningful difference in your symptoms, it's going to be that you have more

gastrointestinal symptoms and you have more serious infections. Is that a drug you want at the end of a year? Are you happy with that result?

Again, in deference to the risk-benefit discussion doctors have with individuals on balance from the SENSCIS trial, that was hard for me to say that that should be something that I am comfortable voting to approve.

DR. NASON: Martha Nason. I voted no for many of the reasons that have been stated either around this table so far after this question or after question number 3, the one about efficacy, because in many ways, it comes down to the same thing. I think the magnitude and level of evidence for efficacy were both marginal, given everything, given the statistics and sensitivity analyses, and the secondary endpoints and everything.

I think the side effect profile was relatively clear, which is why I voted that it was okay on the last question, but the fact that it's relatively clear to me doesn't mean that the trade-off is necessarily worth it. We were seeing

data that showed that maybe there were 40 or 50 percent, more percent of people were responders, or good things were happening on FVC. At the same, we were also seeing 40 and 50 percent increases in severe adverse events and things like that.

So that's a hard trade-off, and given that the efficacy I think is still marginal, that made this particular vote more clear, but it doesn't make the question more clear. There's still a lot more to figure out about the ratio or the net gain between the risk and the benefit.

DR. OLIVER: Alyce Oliver. I voted yes. I do think there needs to be longer studies to see if there is a sustained effect at two years and beyond and also studies of the subgroups to determine who will respond best to this medication given the side effect profile.

DR. RICHARDS: John Richards. I voted yes. I was concerned about the effect size but came down on voting yes, partly because of the severity of the disease and the difficulty with finding patients to do studies of this size and magnitude.

They went across the world to recruit the patients for this study, so undertaking another study is not going to be an easy task.

DR. BECKER: Hi. It's Mara Becker, and I voted yes. I don't think I have anything additional to add for what already has been said for all my yes groups. That's it.

DR. SOLOMON: Did we get enough explanation of nos?

DR. NIKOLOV: I think we were very happy with the robust discussion that we had. We did get a lot of useful, helpful feedback. and we have a lot of homework to do now. We certainly appreciate everyone's input. I don't think we have any additional questions other than thank you to everyone for taking the time to attend this important meeting and providing very helpful feedback.

Adjournment

DR. SOLOMON: Please take all your personal belongings with you, as the room is cleaned at the end of the day. All materials left on the table

will be disposed of. Please also remember to drop off your name badge at the registration table, and we will now adjourn. Thank you very much.

(Whereupon, at 4:30 p.m., the meeting was adjourned.)