FDA Briefing Document
Arthritis Advisory Committee Meeting

July 25, 2019

NDA 205832/S-012
Nintedanib

Proposed indication: systemic sclerosis associated interstitial lung disease

Boehringer Ingelheim

DISCLAIMER STATEMENT

The attached package contains background information prepared by the Food and Drug Administration (FDA) for the panel members of the advisory committee. The FDA background package often contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division or Office. We have brought the supplemental New Drug Application for nintedanib for the treatment of patients with systemic sclerosis associated interstitial lung disease to this Advisory Committee in order to gain the Committee's insights and opinions, and the background package may not include all issues relevant to the final regulatory recommendation and instead is intended to focus on issues identified by the Agency for discussion by the advisory committee. The FDA will not issue a final determination on the issues at hand until input from the advisory committee process has been considered and all reviews have been finalized. The final determination may be affected by issues not discussed at the advisory committee meeting.

Table of Contents

1 Division Memo 8
1.1 Introduction 8
1.2 Background 8
1.3 Regulatory History 9
1.4 Clinical Program 10
1.5 Efficacy 11
1.5.1 Primary Efficacy Variable-FVC 11
1.5.2 Secondary Efficacy Measures 15
1.6 Safety 15
1.7 Benefit Risk Considerations 16
2 Draft Points to Consider 17
3 Clinical and Statistical Review 19
3.1 Introduction 19
3.2 Brief Clinical Background 19
3.3 Product Information and Regulatory Background 20
3.4 Development Program 23
3.4.1 Source of Clinical Data 23
3.4.1.1 Study 1199.214 24
3.4.1.1.1 Study Objectives 24
3.4.1.1.2 Study Design 24
3.4.1.1.3 Study Population 25
3.4.1.1.4 Study Treatments 27
3.4.1.1.5 Study Endpoints 28
3.4.1.1.6 Statistical Analysis Plan 28
3.5 Review of Efficacy 34
3.5.1 Efficacy Review Approach 34
3.5.2 Patient Disposition 34
3.5.3 Demographics and Baseline Characteristics 36
3.5.4 Efficacy Results – Primary Endpoint 39
3.5.4.1 The Primary Analysis 40
3.5.4.2 Sensitivity Analyses on Missing Data Assumptions 41
3.5.4.3 Annual Rate of Decline in FVC in Percent Predicted – A Secondary Endpoint 44
3.5.4.4 Categorical Endpoints Representing Proportion of Patients with a Decline in FVC at Week 52 – Further Endpoints 46
3.5.5 Efficacy Results – Key Secondary endpoints 49
3.5.6 Other Efficacy Results – Selected Endpoints 51
3.5.7 Additional Analyses 53
3.5.7.1 Gender, Race, Age, and Geographic Region 53
3.5.7.2 Other Special/Subgroup Populations 54
3.5.8 Efficacy Summary 57
3.6 Review of Safety 57
3.6.1 Overall Exposure 57
3.6.2 Deaths 58
3.6.3 Serious Adverse Events 59
3.6.4 Dropouts and/or Discontinuations Due to Adverse Effects 60

3.6.5 Severe Adverse Events 61
3.6.6 Treatment Emergent Adverse Events and Adverse Reactions 61
3.6.7 Safety Review Approach 62
3.6.8 Submission Specific Safety Concerns 63
3.6.9 Safety Summary 64
3.7 Benefit/Risk Considerations 65
4 Appendix 68
4.1 Schedule of Events 68
4.2 Recommendations for Treatment Interruption or Dose Reduction 70
4.3 Responder Analysis on FVC in Percent Predicted 71

List of Tables

Table 1: Annual Rate of Decline in FVC in mL over 52 Weeks Primary Analysis (Treated Set) 12
Table 2: Annual Rate of Decline in FVC in Percent Predicted Over 52 Weeks Primary Analysis (Treated Set)..... 13
Table 3: Proportion of Responders with Certain Thresholds Over 52 Weeks (Treated Set)..... 14
Table 4: Annual Rate of Decline in FVC in mL Over 52 Weeks by Region and Mycophenolate Mofetil Use at Baseline (Treated Set)..... 15
Table 5: Summary of Clinical Program..... 23
Table 6: Primary and Sensitivity Analyses for Missing Data Handling (Pattern Mixture Model Approaches)..... 31
Table 7: Analysis Datasets..... 35
Table 8: Patient Disposition..... 35
Table 9: Demographics..... 37
Table 10: Baseline Disease Characteristics..... 38
Table 11: Annual Rate of Decline in FVC in mL over 52 Weeks Primary Analysis (Treated Set) 40
Table 12: Annual Rate of Decline in FVC in mL over 52 Weeks, Tipping Point Analysis (Treated Set)..... 44
Table 13: Annual Rate of Decline in FVC in Percent Predicted over 52 Weeks Primary Analysis (Treated Set)..... 45
Table 14: Proportions of Responders with Certain Thresholds Over 52 Weeks (Treated Set) .. 47
Table 15: Absolute Change from Baseline in mRSS at Week 52 (Treated Set)..... 50
Table 16: Absolute Change from Baseline in SGRQ Total Score at Week 52 (Treated Set)..... 51
Table 17: Cox Proportional Hazards Analysis of Time to Death Over the Entire Study Duration (Treated Set)..... 52
Table 18: Exposure to Study Drug Over 52 Weeks, Treated Set..... 58
Table 19: AEs Leading to Death over Entire Study by Preferred Term, Treated Set..... 59
Table 20: Serious Adverse Events by Preferred Term for > 1% of Patients in Either Treatment Group Over 52 Weeks..... 60
Table 21: Study 1199.214: Significant AEs Over 52 Weeks..... 64
Table 22: Allowed Treatment Reduction or Interruption Periods of Nintedanib..... 70
Table 23: Recommendations for Managing Liver Enzyme Elevations..... 70

List of Figures

Figure 1: Mean (95% CI) Observed Change from Baseline in FVC in mL Over 52 Weeks by Treatment (Treated Set) 12
Figure 2: Study Design 25
Figure 3: Mean (95% CI) Observed Change from Baseline in FVC in mL over 52 Weeks by Treatment (Treated Set) 41
Figure 4: Forest-plot for Rate of Decline in FVC [mL/Year] over 52 Weeks Sensitivity Analyses (Treated Set) 42
Figure 5: Mean (95% CI) Observed Change from Baseline in FVC in Percent Predicted over 52 Weeks by Treatment (Treated Set) 45
Figure 6: Histogram of Percent Change from Baseline in FVC in mL at Week 52 (Treated Set) 48
Figure 7: Proportions of Responders at Various Cutoffs among Treated Patients, Based on Analysis of Percent Change from Baseline in FVC in mL at Week 52 (Treated Set) 49
Figure 8: Kaplan-Meier Estimate of Probabilities of Survival (Treated Set) 52
Figure 9: Forest Plot of the Rate of Decline in FVC (mL/year) Over 52 Weeks by Demographic Subgroup (Treated Set) 54
Figure 10: Forest Plot of the Rate of Decline in FVC (mL/year) Over 52 Weeks by Baseline Disease Characteristics Subgroup (Treated Set) 55
Figure 11: Forest Plot of the Rate of Decline in FVC (mL/year) over 52 weeks by Region by MMF Use Subgroup (Treated Set) 56
Figure 12. Histogram of Absolute Change from Baseline in FVC in Percent Predicted at Week 52 (Treated Set) 71
Figure 13: Proportions of Responders at Various Cutoffs among Treated Patients, Based on Analysis of Change from Baseline in FVC in Percent Predicted at Week 52 (Treated Set) 72

Division Memorandum

Department of Health and Human Services
Public Health Service
Food and Drug Administration
Center for Drug Evaluation and Research
Division of Pulmonary, Allergy, and Rheumatology Products (DPARP)
M E M O R A N D U M

Date: Friday, June 28, 2019

From: Sally Seymour, MD, Director
Division of Pulmonary, Allergy, and Rheumatology Products, CDER, FDA

Nikolay P. Nikolov, MD, Associate Director, Rheumatology
Division of Pulmonary, Allergy, and Rheumatology Products, CDER, FDA

To: Chair, Members and Invited Guests
Arthritis Advisory Committee

Subject: Overview of the FDA background materials for New Drug Application (NDA) 205832, supplement 12, for nintedanib (trade name OFEV), sponsored by Boehringer Ingelheim, for the proposed indication of systemic sclerosis associated interstitial lung disease.

1 Division Memo

1.1 Introduction

Thank you for your participation in the Arthritis Advisory Committee (AAC) meeting to be held on July 25, 2019. As members of the AAC, you provide important expert scientific advice and recommendations to the U.S. Food and Drug Administration (the Agency) on the regulatory decision-making process related to the approval of a drug or biologic product for marketing in the United States. The upcoming meeting is to discuss supplement 12 of the New Drug Application (NDA) 205832 from the Applicant, Boehringer Ingelheim (BI), for nintedanib for the proposed indication of systemic sclerosis associated interstitial lung disease (SSc-ILD).

Currently, there are no approved therapies for patients with systemic sclerosis or SSc-ILD, so there is no established regulatory precedent. BI has submitted the results from a single clinical trial (Study 1199.214) to support the approval of nintedanib for the treatment of patients with SSc-ILD. The focus of the AAC discussion will be data from Study 1199.214, also referred to as Safety and Efficacy of Nintedanib in Systemic Sclerosis Study (SENSCIS)[1]. We ask for your input on the efficacy results, including the clinical meaningfulness of the results and the benefit-risk assessment of nintedanib for the proposed indication. This Division Memorandum provides a brief overview of the application and an introduction to the main issues for discussion, which are addressed in more detail in the attached review.

1.2 Background

Nintedanib
Nintedanib is a small molecule, oral capsule, kinase inhibitor, indicated for the treatment of idiopathic pulmonary fibrosis (IPF). It inhibits multiple receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases (nRTKs). The proposed dosing regimen is 150 mg twice daily, the same as the dosing for the currently approved indication for the treatment of idiopathic pulmonary fibrosis (IPF).

SSc-ILD
Systemic sclerosis (SSc) is a rare, multisystem, connective tissue disease involving the skin, underlying tissues, blood vessels, and major organs that affects approximately 100,000 people in the United States.[2] It is characterized by microvascular damage and fibrosis of the skin and of various internal organs, including the lung, heart, kidneys and the gastrointestinal tract. SSc is a serious disease associated with increased morbidity and mortality with a 10-year survival rate less than 70% from the time of diagnosis.[3] The primary causes of SSc-related death are pulmonary fibrosis, pulmonary arterial hypertension, heart failure, or cardiac arrhythmia.

[1] Distler O, et al, Nintedanib for Systemic Sclerosis–Associated Interstitial Lung Disease, N Engl J Med. 2019; 380(26): 2518-2528
[2] https://www.scleroderma.org/
[3] Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972−2002. Ann Rheum Dis 2007;66:940−4

Interstitial lung disease (ILD), as detected by high resolution computed tomography (HRCT), is present in 55 to 65% of patients with SSc.[4] Severe ILD usually presents relatively early in the disease course within the first 3 years from time of diagnosis.[5] Median survival is 5 to 8 years in SSc-ILD.[6]

Available Therapies
Systemic sclerosis and SSc-ILD are conditions with high unmet medical need as there are no FDA-approved therapies. In clinical practice, patients with SSc are treated based on expert-derived recommendations for the management of organ-specific manifestations and empirically with off-label products used for other rheumatic diseases, such as cyclophosphamide. The 2017 update of European League Against Rheumatism (EULAR) recommendations for the treatment of SSc, and 2016 British Society for Rheumatology (BSR) guidelines for the treatment of SSc, recommend consideration of immunosuppressives such as cyclophosphamide and mycophenolate mofetil (MMF) for treatment of SSc-ILD.[7,8] Such therapies have inherent toxicities including cytopenias, infections, malignancies, among others.

1.3 Regulatory History

On October 15, 2014, nintedanib was approved for the treatment of idiopathic pulmonary fibrosis (IPF).

The first communication on the proposed clinical development in SSc-ILD, occurred in February 2015 when BI proposed to conduct a single confirmatory study in patients with SSc-ILD. At the 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 natural decline in forced vital capacity (FVC) 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. Respectively, the Applicant was advised to follow the patients to the conclusion of the study, to include all-cause mortality as an endpoint, to use observed FVC rather than FVC % predicted, and to include secondary endpoints that measure how patients feel and function.

On September 30, 2015, IND 124707 was opened with the proposed study. On July 6, 2016, nintedanib was granted orphan designation for the treatment of systemic sclerosis (including the associated interstitial lung disease).

---

[4] Launay D, Remy-Jardin M, Michon-Pasturel U, et al. High resolution computed tomography in fibrosing alveolitis associated with systemic sclerosis. J Rheumatol 2006;33(9):1789-801
[5] Steen VD, Medsger TA Jr. Severe Organ Involvement in Systemic Sclerosis with Diffuse Scleroderma. Arthritis Rheum 2000:43:2437−44
[6] Herzog EL, Mathur A, Tager AM, Feghali-Bostwick C, Schneider F, Varga J. Interstitial lung disease associated with systemic sclerosis and idiopathic pulmonary fibrosis: how similar and distinct? Arthritis and Rheumatology, Accepted Article, Accepted: May 08, 2014, doi:10.1002/art.38702; 2014. p. 1967-1978
[7] Update of EULAR recommendations for the treatment of systemic sclerosis, April 2017
https://ard.bmj.com/content/annrheumdis/early/2017/04/25/annrheumdis-2016-209909.full.pdf
[8] BSR and BHPR guideline for the treatment of systemic sclerosis, June 2016

The Applicant submitted supplement 12 for the treatment of SSc-ILD on March 7, 2019 and the application was granted priority review based on the criteria outlined in the 2014 Guidance for Industry: *Expedited Programs for Serious Conditions-Drugs and Biologics*.[9]

1.4 Clinical Program

The nintedanib clinical development program for SSc-ILD consists of a single study, Study 1199.214, which is a double blind, randomized, placebo-controlled, parallel-group design to evaluate the efficacy and safety of oral nintedanib in patients with SSc-ILD. In Study 1199.214, 576 patients were randomized 1:1 to nintedanib 150 mg by mouth twice daily or matching placebo. The primary endpoint was the annual rate of decline in FVC in mL over 52 weeks. Key secondary endpoints included absolute change in modified Rodnan Skin Score (mRSS) at Week 52 and absolute change in Saint George's Respiratory Questionnaire (SGRQ), a patient reported outcome (PRO) at Week 52. Additional secondary endpoints included time to death, Health Assessment Questionnaire Disability Index (HAQ-DI), and Functional Assessment of Chronic Illness Therapy (FACIT) dyspnea scale.

The main efficacy analysis was assessed at Week 52, but patients could remain on treatment up to a maximum of 100 weeks to collect follow-up safety and efficacy information. Patients were evaluated for safety assessments at Weeks 2, 4, 6, 12, 24, 36, 52, 68, 84, and 100. A follow-up visit was scheduled 28 days after the End of Treatment Visit. Patients who experienced clinically significant deterioration of SSc could receive rescue therapy. Permitted medications for management of deterioration include prednisone > 10 mg/day, azathioprine, cyclophosphamide, cyclosporine A, hydroxychloroquine, colchicine, D-penicillamine, sulfasalazine, rituximab, tocilizumab, abatacept, leflunomide, tacrolimus, tofacitinib, and potassium para-aminobenzoate. Patients who permanently discontinued study medication were asked to return for future visits as planned; patients who declined further follow-up visits were asked for vital status assessment at 52 weeks and 100 weeks after their randomization, or at the time the last full visit would have been scheduled, whichever occurred earlier.

The study was conducted as planned. A total of 576 patients, predominantly females (75%), were randomized and treated, 288 in each treatment arm. In addition to SSc-ILD, patients had a history of other SSc manifestations including pulmonary hypertension (9%), digital ulcers (39%), diarrhea/malabsorption/bacterial overgrowth (18%), esophageal dysphagia/reflux (74%), synovitis (24%), friction rubs (9%), and Raynaud phenomenon (97%), which were similar by treatment group. At baseline, 48% of the patients received treatment with mycophenolate and 7% received methotrexate. Use of mycophenolate and methotrexate was similar by treatment group. Approximately half of the patients were enrolled at sites in Europe, 25% were enrolled in Canada and the United States, and 23% in Asia. Overall, the patient demographic characteristics were balanced and representative of the intended patient population.

[9] https://www.fda.gov/media/86377/download

Of the 576 patients, 94% completed visits up to Week 52, the study primary endpoint; the nintedanib group had a numerically higher study withdrawal rate (8%) compared with the placebo group (5%). Treatment discontinuations occurred in 15% of patients: the nintedanib group had a numerically higher treatment discontinuation rate (19%) compared with the placebo group (11%). The most common reason for study withdrawal was adverse event.

1.5 Efficacy

1.5.1 Primary Efficacy Variable-FVC

The primary endpoint was the annual rate of decline in FVC in mL over 52 weeks. This endpoint was selected by the Applicant based on their experience with the IPF program which used the same primary endpoint. Further analysis of data from IPF clinical development programs has demonstrated that patients with less FVC decline also demonstrated an associated decrease in mortality. .[10,11] However, as noted in the Regulatory History section above, the Agency cautioned about the uncertainty of the proposed endpoint to alter natural decline in FVC in a one-year study in SSc-ILD, in the absence of preliminary information on the effects of nintedanib in this patient population. We note that FVC has been proposed as a validated outcome measure in patients with SSc according to the principles of Outcome Measures in Rheumatologic Clinical Trials (OMERACT).[12] While FVC is a surrogate endpoint that does not directly measure how a patient feels, functions, or survives, it has been demonstrated to reliably predict clinical benefit in IPF, a related condition. The clinical benefit from altering the rate of decline in lung function in patients with IPF, as measured by FVC over 52 weeks, has been shown to be consistent in two larger clinical programs, using two different products with different mechanisms of action, nintedanib and pirfenidone.[13,14]

Primary Endpoint Analysis

The annual rate of decline in FVC in mL over the 52-week treatment period (with measurements at Week 2, 6, 12, 24, 36 and 52) was compared between the two treatment groups. The adjusted annual rate of decline in FVC over 52 weeks was lower in the nintedanib group (-52 mL/year) than in the placebo group (-93 mL/year), with a treatment difference of 41 mL/year, Table 1:

[10] Karimi-Shah BA, Chowdhury BA, Forced vital capacity in idiopathic pulmonary fibrosis--FDA review of pirfenidone and nintedanib, N Engl J Med. 2015 Mar 26;372(13):1189-91

[11] Paterniti MO, et al, Acute Exacerbation and Decline in Forced Vital Capacity Are Associated with Increased Mortality in Idiopathic Pulmonary Fibrosis, Ann Am Thorac Soc. 2017 Sep;14(9):1395-1402

[12] Merkel P, Clements PJ, Reveille P, et al. Current status of outcome measure development for clinical trials in systemic sclerosis. J Rheumatol 2003;30:1630−47.

[13] FDA-approved nintedanib labeling

[14] FDA-approved pirfenidone labeling

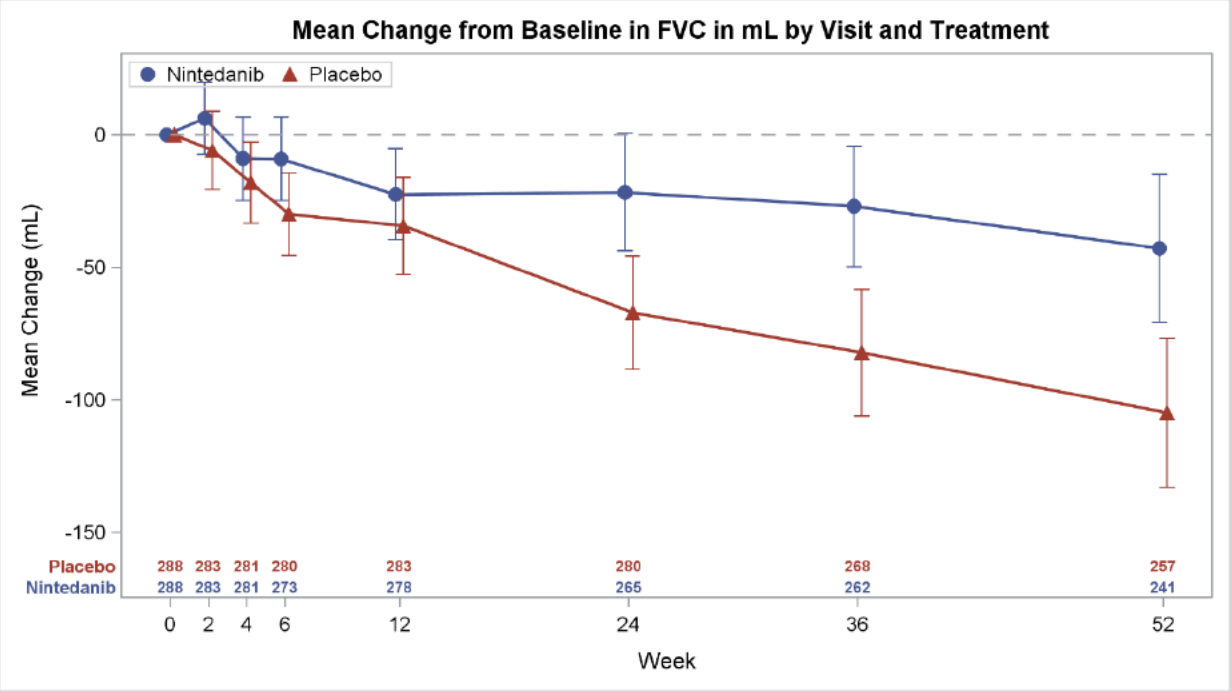
Table 1: Annual Rate of Decline in FVC in mL over 52 Weeks Primary Analysis (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Number Analyzed	288	287
Adjusted Annual Rate of Decline, mL/Year (SE)	-93.3 (13.5)	-52.4 (13.8)
Nintedanib vs. Placebo		
Difference (SE)		40.9 (19.4)
95% CI		2.9, 79.0
p-value		0.035

Abbreviations: N: sample size in Treated Set; SE: standard error; CI: confidence interval
Source: FDA Statistical Reviewer

The mean change from baseline in FVC in mL over 52 weeks by treatment group is shown in Figure 1. Data are observed values; vertical lines represent 95% confidence intervals. Numbers of patients with non-missing FVC data by week are displayed above the x-axis.

Figure 1: Mean (95% CI) Observed Change from Baseline in FVC in mL Over 52 Weeks by Treatment (Treated Set)



Abbreviations: mL: milliliter.
Source: FDA Statistical Reviewer

The annual rate of decline in FVC in percent predicted over the 52-week treatment period was also compared between the two treatment groups. The adjusted annual rate of decline in FVC in percent predicted over 52 weeks was lower in the nintedanib group (-2.6%/year) than in the placebo group (-1.4%/year), with a treatment difference of 1.2%/year, Table 2.

Table 2: Annual Rate of Decline in FVC in Percent Predicted Over 52 Weeks Primary Analysis (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Number Analyzed	288	287
Adjusted Annual Rate of Decline, %/year (SE)	-2.6 (0.4)	-1.4 (0.4)
Nintedanib vs. Placebo		
Difference (SE)		1.2 (0.5)
95% CI		(0.1, 2.2)
p-value		0.033

Abbreviations: N: sample size in Treated Set; SE: standard error; CI: confidence interval
Source: FDA Statistical Reviewer

Sensitivity analyses for handling of missing data.
To assess the robustness of the primary analysis, the FDA review team conducted sensitivity analyses to address missing data resulting from treatment and study discontinuation. While the results for the primary endpoint were statistically significant based on the pre-specified analysis, the sensitivity analyses showed mixed results, as detailed in the attached review, primarily because the magnitude of the effect size was small.

Responder analyses
Given the uncertainty on what difference in FVC is considered clinically important in SSc-ILD, the FDA review team conducted responder analyses on relative decline in FVC using various thresholds for response such as "relative decline ≤10% (or 5%) from baseline at Week 52". In the FDA analyses, patients with missing data at Week 52 were categorized as non-responders. These endpoints are summarized in Table 3.

The proportion of responders with 5% threshold (relative decline ≤5%) was numerically higher in the nintedanib group (59%) than in the placebo group (52%), favoring nintedanib over placebo; the odds ratio was 1.37 (95% CI 0.98, 1.89; nominal p-value = 0.066). The proportion of responders with 10% threshold (relative decline ≤10%) was numerically lower in the nintedanib group (72%) than in the placebo group (74%), not favoring nintedanib over placebo; the odds ratio was 0.93 (95% CI 0.65, 1.35; nominal p-value = 0.704). Responder analyses using absolute decline in FVC percent predicted showed a similar pattern with the analyses using relative decline in FVC.

Table 3: Proportion of Responders with Certain Thresholds Over 52 Weeks (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)	Comparison vs. Placebo*			
	n (%)	n (%)	Odds ratio	95% CI		Nominal p-value
				Lower	Upper	
Responder definition using relative decline from baseline in FVC in mL at Week 52						
Relative decline ≤5%	149 (52%)	171 (59%)	1.36	0.98	1.89	0.066
Relative decline ≤10%	212 (74%)	208 (72%)	0.93	0.64	1.34	0.704
Responder definition using absolute decline from baseline in FVC in percent predicted at Week 52						
Absolute decline ≤5%	186 (65%)	196 (68%)	1.16	0.82	1.64	0.386
Absolute decline ≤10%	236 (82%)	227 (79%)	0.82	0.54	1.24	0.348

Abbreviations: N: sample size in Treated Set; n: number of patients within category; SE: standard error; CI: confidence interval
Note: Patients with missing data at Week 52 were considered as non-responders.
* Based on Cochran-Mantel-Haenszel tests stratified on baseline ATA status.
Source: FDA Statistical Reviewer

The Applicant conducted similar type of analyses but presented as "non-response" with similar thresholds such as "relative decline >10% (or 5%) from baseline at Week 52". We note that, the FDA and Applicant analyses convey identical information because non-response is just a complementary event of response. Of note, unlike FDA, the Applicant used the worst observation carried forward approach to imputing missing data at Week 52 and then applied the threshold to define non-response.

Subgroup analyses:
To evaluate the influence of stable background immunosuppressive therapy to study treatment, the protocol pre-specified subgroup analyses also included mycophenolate mofetil/sodium use at baseline. A less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients on mycophenolate mofetil at baseline (treatment difference 27 mL/year) and patients from the U.S. and Canada (treatment difference 10 mL/year) (Table 4).

Table 4: Annual Rate of Decline in FVC in mL Over 52 Weeks by Region and Mycophenolate Mofetil Use at Baseline (Treated Set)

	N (%)	Group Mean Rate		Mean Difference (95% CI)	p-value
		Nintedanib	Placebo		
Overall	575 (100)	-52	-93	41 (3, 79)	
MMF Use at Bassline					0.452
No MMF Use	297 (52)	-64	-119	56 (0, 111)	
MMF Use	278 (48)	-40	-66	27 (-26, 79)	
Region					0.350
US and Canada	142 (25)	-42	-52	10 (-69, 89)	
Other than US and Canada	433 (75)	-56	-108	51.6 (8, 95)	
Region by MMF Use at Baseline					0.600
US and Canada, No MMF Use	28 (5)	-143	-119	-24 (-242, 194)	
US and Canada, MMF Use	114 (20)	-15	-32	17 (-68, 101)	
Other than US and Canada, No MMF Use	269 (47)	-56	-120	64 (7, 122)	
Other than US and Canada, MMF Use	164 (29)	-57	-89	32 (-35, 99)	

Abbreviations: N: sample size in Treated Set; SE: standard error; CI: confidence interval, MMF: mycophenolate mofetil.
Source: FDA Statistical Reviewer

1.5.2 Secondary Efficacy Measures

The results from the key secondary endpoints, absolute change in modified Rodnan Skin Score (mRSS) and absolute change in Saint George's Respiratory Questionnaire (SGRQ) at Week 52, were not statistically significantly different between treatment and placebo-treated patients.

The results did not suggest differences in other secondary endpoints, including FACIT-dyspnea score and DL_{CO} at Week 52. Additionally, there were no differences in other disease-related secondary endpoints, including, number of digital ulcers, or HAQ-DI.

Overall mortality was also similar between treatment groups.

1.6 Safety

In Study 1199.214, adverse events in the nintedanib group were consistent with those known for nintedanib including diarrhea, nausea, vomiting, and liver abnormalities. Deaths and serious adverse events (SAEs) were balanced between the treatment groups. A greater proportion of patients in the nintedanib group developed pneumonia (8 vs. 1). There were more AEs leading to drug decrease and discontinuation in the nintedanib group which were mostly from gastrointestinal complaints. In addition, patients in the nintedanib group lost more weight than the placebo group.

1.7 Benefit Risk Considerations

The purpose of this Advisory Committee meeting is to discuss the results of Study 1199.214. The study showed a statistically significant lower rate of decline of FVC with nintedanib compared with placebo over 52 weeks. However, key secondary endpoints were not supportive of a direct treatment benefit for nintedanib over placebo at Week 52 in this study. Thus, the clinical significance of the treatment effect of lower rate of decline by 41 mL/year (or approximately 1.2% predicted) remains a question for discussion.

We ask you to discuss the strength of the available FVC data to support a treatment effect on FVC. Important considerations for this discussion include the size of the treatment effect and the impact of missing data on the robustness of the effect.

The next point of discussion is the clinical meaningfulness of the changes in FVC, given the lack of supportive efficacy from other secondary endpoints, including endpoints that directly assess how a patient feels, functions, or survives. In addressing this discussion point, it is important to consider the current understanding about the association of FVC decline, mortality, and other clinical outcomes, which are discussed in the attached review.

Finally, we ask you to discuss the benefit–risk assessment of nintedanib for the treatment of SSc-ILD.

We acknowledge that 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. However, we want to ensure that new products have a favorable benefit-risk assessment for patients. Given the modest treatment effect on FVC and the lack of support from key secondary endpoints (i.e. endpoints that directly measure clinical benefit), the benefit-risk of nintedanib for the treatment of SSc-ILD is important to discuss with this Advisory Committee.

We thank you for your participation in this Advisory Committee meeting and look forward to the discussion.

2 Draft Points to Consider

On July 25, 2019, the Committee will discuss the New Drug Application (NDA) 205832, supplement 12, for nintedanib (trade name OFEV), sponsored by Boehringer Ingelheim, for the proposed indication of systemic sclerosis associated interstitial lung disease. The Agency is seeking input from the Committee on whether the application provides substantial evidence of efficacy for the proposed indication, and overall benefit-risk considerations in SSc-ILD, as a rare and serious disease.

The following are draft points to consider for discussion at the upcoming AC.

- Discuss the efficacy data for nintedanib for the treatment of patients with SSc-ILD
    - Discuss the clinical meaningfulness of the changes in FVC with nintedanib treatment in the population studied
    - Discuss the FVC data from the following subgroups and the implications for use of nintedanib in patients in the US:
        - US and Canada subgroup compared to the overall study population
        - Patients on background MMF vs. no background MMF treatment
    - Discuss if the data provide substantial evidence of the efficacy of nintedanib in the population studied
- Discuss if the safety profile of nintedanib is adequate to support approval of nintedanib in patients with SSc-ILD
- Discuss if the benefit-risk is adequate to support approval of nintedanib for the proposed indication of treatment of patients with SSc-ILD

Clinical and Statistical Review

Department of Health and Human Services
Public Health Service
Food and Drug Administration
Center for Drug Evaluation and Research
Division of Pulmonary, Allergy, and Rheumatology Products (DPARP)

Date: June 28, 2019

From: Nadia Habal, MD, Primary Clinical Reviewer
Rachel Glaser, MD, Clinical Team Leader
Division of Pulmonary, Allergy, and Rheumatology Products, CDER, FDA

Yu Wang, PhD, Primary Statistical Reviewer
Yongman Kim, PhD, Statistical Team Leader
Division of Biometrics II, Office of Biostatistics, CDER, FDA

To: Chair, Members and Invited Guests
Arthritis Advisory Committee

Subject: FDA background materials for the Supplemental New Drug Application (sNDA)

3 Clinical and Statistical Review

3.1 Introduction

Boehringer Ingelheim submitted a supplemental new drug application (sNDA) 205832/supplement 12 on March 7, 2019 for nintedanib for the treatment of systemic sclerosis associated interstitial lung disease (SSc-ILD). The product is for oral administration. The proposed dose is 150 mg twice daily; the proposed dose in patients with mild hepatic impairment (Child Pugh A) or for temporary management of adverse reactions is 100 mg twice daily.

Nintedanib was approved for the treatment of idiopathic pulmonary fibrosis (IPF) on October 15, 2014 under NDA 205832. The proposed dosing regimen in SSc-ILD is the same as the approved dose for treatment of IPF. If approved, nintedanib would be the first approved treatment for SSc-ILD in the US.

3.2 Brief Clinical Background

Systemic sclerosis (SSc) is a rare, multisystem, connective tissue disease involving the skin, underlying tissues, blood vessels, and major organs that affects approximately100,000 people in the United States.[15] It is characterized by microvascular damage and fibrosis of the skin and of various internal organs, including the lung, heart, kidneys and the gastrointestinal tract. SSc is a serious disease associated with increased morbidity and mortality with a 10-year survival rate less than 70% from the time of diagnosis.[16] The primary causes of SSc-related deaths are pulmonary fibrosis, pulmonary arterial hypertension, heart failure, or cardiac arrhythmia. Interstitial lung disease (ILD), as detected by high resolution computed tomography (HRCT), is present in 55-65% of patients with SSc.[17] Severe ILD usually presents relatively early in the disease course within the first 3 years from time of diagnosis.[18] Median survival in SSc-ILD is 5-8 years.[19]

There are no FDA-approved therapies for treatment of systemic sclerosis or SSc-ILD. In clinical practice, patients with systemic sclerosis are treated based on expert-derived recommendations for the management of organ-specific manifestations and empirically with off-label products used for other rheumatic diseases. The Update of EULAR recommendations for the treatment of systemic sclerosis, recommends consideration of cyclophosphamide for

[15] https://www.scleroderma.org/
[16] Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972−2002. Ann Rheum Dis 2007;66:940−4
[17] Launay D, Remy-Jardin M, Michon-Pasturel U, et al. High resolution computed tomography in fibrosing alveolitis associated with systemic sclerosis. J Rheumatol 2006;33(9):1789-801.
[18] Steen VD, Medsger TA Jr. Severe Organ Involvement in Systemic Sclerosis with Diffuse Scleroderma. Arthritis Rheum 2000:43:2437−44
[19] Herzog EL, Mathur A, Tager AM, Feghali-Bostwick C, Schneider F, Varga J. Interstitial lung disease associated with systemic sclerosis and idiopathic pulmonary fibrosis: how similar and distinct? Arthritis and Rheumatology, Accepted Article, Accepted: May 08, 2014, doi:10.1002/art.38702; 2014. p. 1967-1978

treatment of SSc-ILD, particularly in patients with progressive ILD.[20] The BSR and BHPR guidelines for the treatment of systemic sclerosis recommends treatment of extensive or progressive ILD with immunosuppression, including intravenous cyclophosphamide. Mycophenolate mofetil (MMF) may also be used as an alternative or after cyclophosphamide.[21]

3.3 Product Information and Regulatory Background

Product Information

Nintedanib is a small molecule available as an oral capsule that inhibits multiple receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases (nRTKs). Nintedanib inhibits the following RTKs: platelet-derived growth factor receptor (PGFR) α and β, fibroblast growth factor receptor (FGFR) 1-3, vascular endothelial growth factor receptor (VEGFR) 1-3, and Fms-like tyrosine kinase-3 (FLT3). In addition, nintedanib inhibits the following nRTKs: Lck, Lyn, and Src kinases. Nintedanib binds competitively to the adenosine triphosphate (ATP) binding pocket of these receptors and blocks the intracellular signaling needed for the proliferation, migration, and transformation of fibroblasts involved in fibrotic tissue remodeling. Nintedanib oral capsules are available in 100 mg and 150 mg strengths.

Nintedanib is approved in the United States for the treatment of Idiopathic Pulmonary Fibrosis (IPF) since 2014 at a dose of 150 mg twice daily. In addition, it is also approved in the EU and other countries in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or locally recurrent non-small cell lung cancer of adenocarcinoma tumor histology after first-line chemotherapy. The proposed dosing and dosing regimen for treatment for SSc-ILD is the same as that approved for treatment of IPF.

SSc-ILD and IPF

To provide further context for the Committee's discussion, this subsection will discuss some of the regulatory considerations from the currently approved indication for nintedanib, IPF, as they may relate to the proposed indication of SSc-ILD.

IPF is a chronic progressive, diffuse parenchymal lung disease of unknown etiology that results in pulmonary fibrosis. It is characterized by scarring of the lungs, non-productive cough, and progressive dyspnea. Median survival time in patients with IPF is estimated to be from 3 to 5 years, with respiratory failure being the most frequent cause of death. Nintedanib, the subject of this review, and pirfenidone, were both approved in 2014 for the treatment of IPF. The approvals of these agents were primarily based on a demonstration of slowing of lung function decline as measured by forced vital capacity (FVC). In the nintedanib IPF development program, there were concerns about the use of FVC as the primary efficacy endpoint, as it had not been established as a validated clinical surrogate for clinically important outcomes. Additionally, there was a lack of information on what difference is considered clinically important. However, in a disease that is marked by a progressive decline in lung function, FVC was considered to be a logical primary endpoint, and was therefore considered acceptable for

[20] Update of EULAR recommendations for the treatment of systemic sclerosis, April 2017 https://ard.bmj.com/content/annrheumdis/early/2017/04/25/annrheumdis-2016-209909.full.pdf

[21] BSR and BHPR guideline for the treatment of systemic sclerosis, June 2016

the clinical development program. Due to the uncertainty around several aspects of the primary endpoint, clinically important secondary endpoints and evaluation of mortality were considered in the assessment of efficacy, in order to support the primary endpoint.

Secondary endpoints in the IPF studies included mortality, IPF exacerbations, and the St. George's Respiratory Questionnaire (SGRQ). SGRQ has been used routinely as a patient reported outcome (PRO) in COPD development programs. In the absence of a validated PRO for use in IPF, SGRQ was tested in the nintedanib IPF program. Other clinically meaningful and key secondary endpoints assessed in the clinical studies provided support for the efficacy of nintedanib in IPF.

While FVC and SGRQ have precedent for use in respiratory applications, there is no regulatory precedent for their use in SSc-ILD. The discussion of the relevance of these endpoints to other diseases requires consideration of the similarities and differences of the diseases. Both IPF and SSc-ILD are chronic, progressive diseases that ultimately result in pulmonary fibrosis. However, there are differences between the diseases with regard to demographics, diagnostic findings, and prognosis. Demographically, IPF is a disease of older males, whereas SSc-ILD patients enrolled in Study 1199.214 were middle aged females. Prognostically, the median survival for IPF patients is 2 to 5 years with considerable variability (almost 25% of patients living beyond 10 years[22]); in comparison, the median survival for SSc-ILD patients is 5 to 8 years[23]. The findings on high resolution computed tomography (HRCT) are different. For IPF, the classic signs on HRCT for usual interstitial pneumonitis (UIP, the histopathologic correlate for IPF) include traction bronchiectasis with peripheral basilar predominant opacities and honeycombing, specifically excluding extensive ground glass opacities; in contrast, for SSc-ILD, non-specific interstitial pneumonitis (NSIP, the most common histopathologic pattern seen with SSc-ILD) is associated with peripheral ground glass opacities.

FVC is the primary efficacy variable in the nintedanib SSc-ILD clinical development program. While FVC is a surrogate endpoint that does not directly measure how a patient feels, functions, or survives, it has been demonstrated to reliably predict clinical benefit in IPF, a related condition. The clinical benefit from altering the rate of decline in lung function in patients with IPF, as measured by FVC over 52 weeks, has been shown to be consistent in two larger clinical programs, using two different products with different mechanisms of action, nintedanib and pirfenidone.[24,25] Slowing of FVC decline has been associated with a decrease in mortality and has been supported by other clinically meaningful endpoints in IPF.[26,27]

[22] Nathan SD et al. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. Chest. 2011;140(1):221

[23] Yasuoka H. Recent Treatments of Interstitial Lung Disease with Systemic Sclerosis. Clin Med Insights Circ Respir Pulm Med. 2015; 9(Suppl 1): 97–110

[24] FDA-approved nintedanib labeling

[25] FDA-approved pirfenidone labeling

[26] Karimi-Shah BA, Chowdhury BA, Forced vital capacity in idiopathic pulmonary fibrosis--FDA review of pirfenidone and nintedanib, N Engl J Med. 2015 Mar 26;372(13):1189-91

[27] Paterniti MO, et al, Acute Exacerbation and Decline in Forced Vital Capacity Are Associated with Increased Mortality in Idiopathic Pulmonary Fibrosis, Ann Am Thorac Soc. 2017 Sep;14(9):1395-1402

In the nintedanib IPF program, the treatment difference (nintedanib vs. placebo) in rate of decline in FVC in the three clinical studies (Studies 1, 2, and 3 in the FDA-approved nintedanib labeling) ranged from 94 to 131 mL/year. The change in FVC was supported by statistically significant decreases in IPF exacerbations and improvement in SGRQ scores, in 2 of the 3 studies. Although not powered for survival, a numerical trend favoring nintedanib was seen for survival in both pre-specified and sensitivity analyses.

We acknowledge that SSc-ILD, as a disease process, may be sufficiently different from IPF such that a direct comparison between FVC changes in IPF patients may not be comparable or appropriate to FVC changes in SSc-ILD patients. However, the regulatory precedent with nintedanib in IPF provides some context as you consider the SSc-ILD program.

The observed FVC changes in other studies from published literature in SSc and SSc-ILD provide additional background information for this program. For example, the Scleroderma Lung Study (SLS) was a double-blind, randomized, placebo-controlled study of oral cyclophosphamide (CYC) treatment in 158 patients with active SSc-ILD. The primary endpoint was the adjusted percent predicted FVC change at 12 months.[28] The mean absolute difference in adjusted 12-month percent predicted FVC between the cyclophosphamide and placebo groups was 2.53%. The observed treatment effect of cyclophosphamide on changes in lung function were supported by improvement in the transitional dyspnea index and HAQ-DI, supporting clinical meaningfulness of slowing the rate of decline in FVC.

This information may be useful to keep in mind as you consider the results of Study 1199.214.

Regulatory Background

The following timeline highlights the pertinent regulatory interactions between BI and the Agency:

- October 15, 2014, Approval for the treatment of idiopathic pulmonary fibrosis.
- February 12, 2015, Pre-IND written responses: Given the slowly progressive nature of SSc-ILD, FDA recommended continued observation of patients until the conclusion of the study, not only through week 52, to assess response over a longer duration. The Agency also recommended the inclusion of all-cause mortality, as well as, secondary endpoints that directly 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 would depend on the persuasiveness of the treatment effect.
- September 11, 2015, IND opened: The Applicant submitted the protocol for Study 1199.214. The IND was deemed safe-to-proceed.
- June 21, 2018, Pre-sNDA written responses: General issues regarding format and content of the supplemental NDA (sNDA) were discussed. In addition, the Agency acknowledged the difficulties with enrollment in the terminated nintedanib/hormonal contraceptive drug-drug interaction (DDI) study in patients with non-small cell lung cancer. The Agency agreed that the proposed sNDA could be submitted in the absence

[28] Tashkin DP et al. Cyclophosphamide versus Placebo in Scleroderma Lung Disease. N Engl J Med 2006;354:2655-66.

of the DDI study and recommended the Applicant evaluate the DDI in a dedicated study in the SSc-ILD population as part of an ongoing or future study.

In addition, orphan drug designation was granted for the treatment of systemic sclerosis (including the associated interstitial lung disease) on July 6, 2016. Fast track designation for SSc-ILD was granted on March 7, 2018.

The Applicant submitted supplement 12 for the treatment of SSc-ILD on March 7, 2019. The application was granted a priority review based on the criteria outlined in the 2014 Guidance for Industry: *Expedited Programs for Serious Conditions-Drugs and Biologics*.[29]

3.4 Development Program

Reflective of the rarity of condition, the clinical development program in SSc-ILD consisted of a single phase 3, double-blind, placebo-controlled study was conducted in 576 patients with SSc-ILD randomized to receive nintedanib 150 mg twice daily or placebo for at least 52 weeks. Patients who completed Study 1199.214 on treatment and attended a follow-up visit after end of treatment were eligible to participate in an open-label long-term extension Study 1199.225; this study is ongoing.

3.4.1 Source of Clinical Data

The clinical data reviewed in this document are derived from a single study conducted in SSc-ILD, Study 1199.214. Efficacy and safety data were derived from the first 52 weeks of treatment. Individual patients remained on blinded study treatment up to 100 weeks, until the last patient completed 52 weeks of treatment, providing additional supportive data.

Table 5: Summary of Clinical Program

Study No.	Description	Subjects	Design	Treatment	Duration	Endpoints
1199.214	Phase 3 efficacy and safety	576 patients with SSc-ILD	R, DB, PC, PG	Nintedanib 150 mg BID, dose reduction to 100 mg BID PBO	52 weeks	- change in FVC (mL) - change in mRSS - change in SGRQ

Abbreviations: R: randomized; DB: double blind; PC: placebo controlled; PG: parallel group; BID: twice daily; FVC: forced vital capacity; SGRQ: St. George's Respiratory Questionnaire; mRSS: modified Rodnan skin score.

[29] https://www.fda.gov/media/86377/download

3.4.1.1 Study 1199.214

Study Title:
Study 1199.214: A double-blind, randomized, placebo-controlled study evaluating efficacy and safety of oral nintedanib treatment for at least 52 weeks in patients with 'Systemic Sclerosis associated Interstitial Lung Disease' (SSc-ILD)

Study Dates: November 30, 2015 to November 28, 2018
Study Sites: 194 sites (with screened patients) in 32 countries in Asia, Australia, Europe, North America, and South America

3.4.1.1.1 Study Objectives

Primary objective: The primary objective was to demonstrate a reduction in the annual rate of decline in FVC in mL over 52 weeks in the nintedanib treatment group compared with the placebo group.

Secondary objectives: The main secondary objectives were to demonstrate efficacy regarding skin fibrosis as assessed by the modified Rodnan Skin Score (mRSS) at Week 52 and to demonstrate an improvement of patient's symptoms as measured by the Saint George's Respiratory Questionnaire (SGRQ) total score at Week 52.

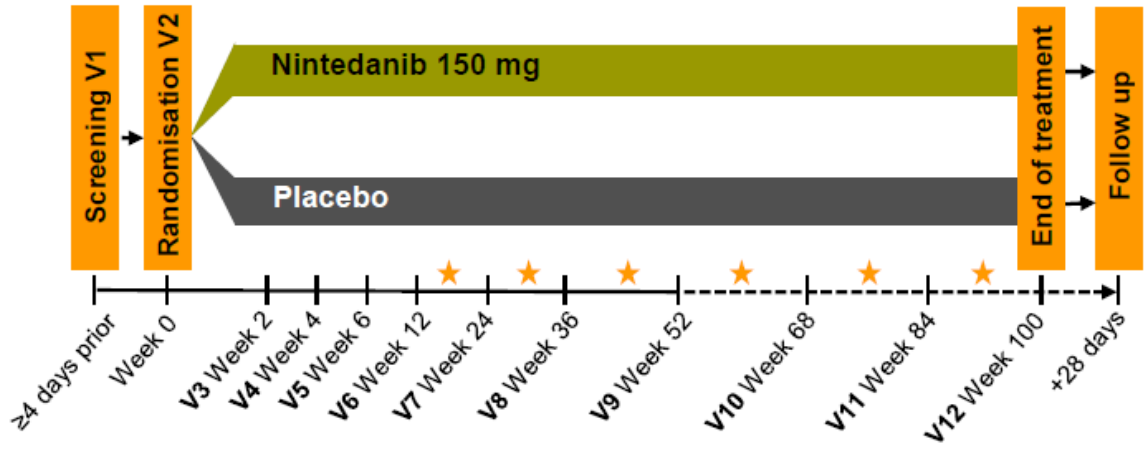
Other objectives were to assess safety and tolerability, mortality, the effects on different systemic organ manifestations of SSc, pharmacokinetics, and the effects of nintedanib on patient's perception of the disease.

3.4.1.1.2 Study Design

The study was a randomized, multicenter, double-blind, placebo controlled, parallel group design to investigate the safety and efficacy of nintedanib treatment in SSc-ILD. After signing informed consent before or at Visit 1 (Screening), a high-resolution computed tomography (HRCT), if not performed within the prior 12 months, was performed and sent for central eligibility review. Patients with a confirmed SSc-ILD diagnosis were randomized 1:1 to either nintedanib 150 mg BID or matching placebo; randomization was stratified by antitopoisomerase (ATA) antibody status (positive or negative).

The Study Design is presented in Figure 2. The main efficacy analysis was assessed at Week 52, but patients could remain on treatment up to a maximum of 100 weeks to collect follow-up safety and efficacy information. Patients were evaluated for safety assessments at Weeks 2, 4, 6, 12, 24, 36, 52, 68, 84, and 100. A follow-up visit was scheduled 28 days after the End of Treatment Visit. Patients who permanently discontinued study medication were asked to return for future visits as planned; patients who declined further follow-up visits were asked for vital status assessment at 52 weeks and 100 weeks after their randomization, or at the time the last full visit would have been scheduled, whichever occurred earlier. The schedule of assessments is summarized in the Appendix 4.1.

Figure 2: Study Design



An asterisk indicates an intermediate laboratory control visit. V = Visit
Source: CSR Figure 9.1:1

In the event of adverse events (AEs) or liver enzyme elevations, dose reduction from 150 mg BID to 100 mg BID was to be considered (Appendix 4.2). For AEs considered drug-related, treatment could be interrupted for up to 4 weeks; resumption of study drug at 100 mg BID was recommended with re-escalation within 4 weeks to 150 mg BID. If AEs were not considered drug-related, treatment could be interrupted for up to 8 weeks; resumption of the same dose of study drug was recommended. If AEs persisted at the lower dose, or if they were severe while on 150 mg BID, treatment discontinuation was to be considered. For elevated liver enzymes without signs of hepatic injury, dose reduction or interruption (AST or ALT increase to ≥3x to <5x ULN) and dose interruption (AST or ALT increase to ≥5x to <8x ULN) was recommended. If repeat liver enzymes were ≥3x ULN, permanent discontinuation was recommended.

3.4.1.1.3 Study Population

The study was conducted in adult patients 18 years of age or older with SSc of less than 7 years duration from first non-Raynaud symptom, with confirmed SSc-ILD based on HRCT.

Key Inclusion Criteria:

1. Patient ≥18 years at time of informed consent
2. Patients had to have fulfilled the 2013 ACR/EULAR classification criteria for SSc
3. SSc disease onset (defined by first non-Raynaud symptom) within 7 years of Visit 1
4. SSc-related ILD pattern confirmed by HRCT performed within 12 months of Visit 1. The extent of fibrotic disease in the lung had to be ≥10% on HRCT, assessed by central review
5. FVC ≥40% of predicted normal at Visit 2
6. DLco (corrected for hemoglobin [Visit 1]): 30% to 89% of predicted at Visit 2

Key Exclusion Criteria:

1. AST, ALT, Bilirubin >1.5 times upper limit of normal (ULN)
2. Creatinine clearance <30 mL/min
3. Airway obstruction (pre-bronchodilator FEV1/FVC <0.7) at Visit 2

4. In the opinion of the investigator, other clinically significant pulmonary abnormalities
5. Significant pulmonary hypertension (PH) defined by any of the following:
    - Previous clinical or echocardiographic evidence of significant right heart failure
    - History of right heart catheterization showing a cardiac index ≤2 L/min/m^2
    - PH requiring parenteral therapy with epoprostenol/treprostinil
6. Cardiovascular diseases, including:
    - Severe hypertension, uncontrolled under treatment (≥160/100 mmHg) within 6 months of Visit 1
    - Myocardial infarction within 6 months of Visit 1
    - Unstable cardiac angina within 6 months of Visit 1
7. More than 3 digital fingertip ulcers at Visit 2 or a history of severe digital necrosis requiring hospitalization or severe other ulcers at discretion of investigator
8. Bleeding risks, including:
    - Known genetic predisposition to bleeding
    - Patients who require
        - Fibrinolysis, full-dose therapeutic anticoagulation (e.g. vitamin K antagonists, direct thrombin inhibitors, heparin, hirudin)
        - High dose antiplatelet therapy. Prophylactic low dose heparin or prophylactic use of antiplatelet therapy (e.g. acetyl salicylic acid up to 325 mg/day, or clopidogrel at 75 mg/day, or equivalent doses of other antiplatelet therapy) are not prohibited.
    - History of hemorrhagic central nervous system (CNS) event within 12 months of Visit 1.
    - Any of the following within 3 months of Visit 1:
        - Hemoptysis or hematuria
        - Active gastro-intestinal (GI) bleeding or GI ulcers
        - Major injury or surgery (investigators judgment)
    - Coagulation parameters: International normalized ratio (INR) >2x ULN, prolongation of prothrombin time (PT) and partial thromboplastin time (PTT) by >1.5 x ULN at Visit 1
9. History of thrombotic event (including stroke and transient ischemic attack) within 12 months of Visit 1
10. Known hypersensitivity to study medication or its components
11. Other disease or condition that may interfere with testing procedures or in the judgment of the investigator could have interfered with study participation or put patient at risk from participation
12. Life expectancy of <2.5 years for disease other than SSc in investigator assessment
13. Patients with clinical signs of malabsorption or needing parenteral nutrition
14. Previous treatment with nintedanib or pirfenidone
15. Other investigational therapy received within 1 month or 6 half-lives (whichever was greater) before Screening Visit (Visit 1)
16. Treatment with:
    - Prednisone >10 mg/d or equivalent received within 2 weeks before Visit 2
    - Azathioprine, hydroxychloroquine, colchicine, D-penicillamine, sulfasalazine, received within 8 weeks before Visit 2

    - Cyclophosphamide, rituximab, tocilizumab, abatacept, leflunomide, tacrolimus, newer anti-arthritic treatments like tofacitinib and cyclosporine A, potassium para-aminobenzoate, received within 6 months before Visit 2
17. Unstable background therapy with either mycophenolate mofetil/sodium or methotrexate (combined therapy was not allowed). Patients had to be either:
    - Not on mycophenolate mofetil/sodium or methotrexate within at least 8 weeks before Visit 2, or
    - On stable therapy with either mycophenolate mofetil/sodium or methotrexate for 6 months before Visit 2 and were to remain stable on this background therapy for at least 6 months after randomization
18. Previous hematopoietic stem cell transplantation (HSCT), or HSCT planned within the next year
19. Major surgical procedures planned to occur during study period
20. Women who were pregnant, nursing, or who planned to become pregnant while in the study
21. Women of childbearing potential not willing or able to use highly effective birth control methods for 28 days before and 3 months after nintedanib administration
22. Active alcohol or drug abuse, in the opinion of the investigator
23. Patients with underlying chronic liver disease (Child Pugh A, B, C hepatic impairment)
24. Patients with a history of scleroderma renal crisis

3.4.1.1.4 Study Treatments

During the 52-week treatment period the treatment arms were as follows:

- Nintedanib 150 mg soft gelatin capsule by mouth twice per day (BID)
- Control product: matching placebo

Concomitant Medications

Continuation of stable doses of mycophenolate mofetil or methotrexate was permitted in the study. Corticosteroid doses ≤ 10 mg/day prednisone or equivalent were allowed if the dose was stable for at least 8 weeks prior to Visit 2. Other permitted immunosuppression was allowed in the event of clinical deterioration (see Rescue Therapy). Use of pirfenidone or nintedanib (outside of the study) was not permitted throughout the study. Treatment with full dose therapeutic anticoagulation and high-dose antiplatelet therapy during the treatment period was not permitted. Close monitoring was advised with concomitant use of P-gp and CYP3A4 inhibitors that may increase exposure to nintedanib.

Rescue Therapy

Patients who experienced clinically significant deterioration of SSc could receive additional treatment. A clinically significant deterioration was defined as:

- Absolute decline in FVC percent predicted > 10% compared to baseline
- A relative change from baseline in mRSS of >25% and absolute change from baseline >5 points, or
- Clinically significant deterioration in other organ systems or clinical parameters at the discretion of the investigator

Other causes for FVC decline were to be excluded. Permitted medications for management of deterioration include prednisone > 10 mg/day, azathioprine, cyclophosphamide, cyclosporine

A, hydroxychloroquine, colchicine, D-penicillamine, sulfasalazine, rituximab, tocilizumab, abatacept, leflunomide, tacrolimus, tofacitinib, and potassium para-aminobenzoate.

3.4.1.1.5 Study Endpoints

Primary Endpoint:

- Annual rate of decline in FVC in mL over 52 weeks.

Key Secondary Endpoints:

- Absolute change from baseline in the mRSS at Week 52
- Absolute change from baseline in SGRQ total score at Week 52

Other Secondary endpoints:

- Annual rate of decline in FVC in percent predicted over 52 weeks
- Absolute change from baseline in FVC in mL at Week 52
- Relative change from baseline (%) of mRSS at Week 52
- Time to all-cause mortality
- Absolute change from baseline at week 52 in CRISS index score
- Absolute change from baseline in DL_{CO} in percent predicted at Week 52
- Absolute change from baseline in digital ulcer net burden at Week 52
- Absolute change from baseline in HAQ-DI score at week 52.
- Absolute change from baseline in FACIT dyspnea score at Week 52

Further endpoints:

- Proportion of patients with a relative decline from baseline in FVC in mL at Week 52 of >5%
- Proportion of patients with a relative decline from baseline in FVC in mL at Week 52 of >10%
- Proportion of patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of >5%
- Proportion of patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of >10%
- Absolute change from baseline in SHAQ domain scores at Week 52

3.4.1.1.6 Statistical Analysis Plan

Analysis Sets

The following analysis sets were defined in the Statistical Analysis Plan (SAP):

- Randomized set (RS): This set included all randomized patients, whether treated or not.
- Treated set (TS): This set included all randomized patients who received at least one dose of study medication. This set was used for all analyses of efficacy and safety endpoints.

Analysis Definitions for Periods

The SAP defined the following periods while not necessarily all of them existed for every patient. The last day of each of the following periods was excluded from the respective period. It defined the first day of the subsequent period.

- Screening: from informed consent to randomization
- Post-randomization: from randomization to first study drug intake in treatment period
- Treatment period: from first study drug intake (or re-start of treatment if interruption) to last study drug intake (or the day before start date of interruption if interruption) plus one day
- Off-treatment: from start date of interruption to re-start of treatment
- Residual effect period: from the last study drug intake plus one day to last study drug intake plus 28 days plus one day or to date of first study drug intake in extension study, whichever occurred earlier
- Follow-up: from last study drug intake plus 29 days up to the beginning of post-study period. This period was only created if last study drug intake took place more than 28 days before study completion, or for patients having prematurely discontinued the treatment and still continuing the study
- Post-study: from the latest between last study drug intake plus 29 days, 'date of study completion' (from the study completion part of the eCRF) plus one day, and 'date of Informed Consent in extension study' (if applicable). This period was not created if date of first study drug intake in extension study was before last study drug intake plus 28 days

For Safety Data:

- For safety analyses, data from the treatment period, possible off-treatment periods and residual effect period were considered as on-treatment.

For Efficacy Data:

- For efficacy analyses, data from randomization date up to week 52 were considered.
- For efficacy descriptive analyses over the whole study period, all data collected after randomization date were considered.

Estimands

While no estimand was referenced or defined throughout the protocol or SAP, the overall approach of the primary analysis of the primary efficacy endpoint and the supporting sensitivity analyses implicitly targeted the de facto or treatment policy estimand as mentioned later in the study report.

Primary Efficacy Endpoint

Primary Analysis: The annual rate of decline in FVC in mL over the 52-week treatment period (with measurements at Week 2, 6, 12, 24, 36 and 52) was compared between the two treatment groups with a restricted maximum likelihood- based approach using a random coefficient regression model. This model included the fixed categorical effects of treatment group, ATA status, and gender, fixed continuous effects of time, baseline FVC (mL), age and height as well as the interaction terms treatment group-by-time and baseline-by-time.

An unstructured variance-covariance structure was used to model the random slope and intercept. The Kenward-Roger approximation was used to estimate denominator degrees of freedom and adjust standard errors. Least squares (LS) means of slope for each treatment group and mean treatment group difference, standard error (SE), 95% confidence intervals (CIs) and the p-value for the treatment group effect were to be presented. The primary treatment comparison of slopes was assessed through the treatment-by time interaction coefficient. The primary analysis was performed on the TS (according to randomized treatment), using all available data from baseline (excluded) up to Week 52 (after time-windowing), including visits done after premature treatment withdrawal, EOT visits and follow-up visits done before Week 52.

Multiple sensitivity analyses were performed for the primary endpoint, including:

- Sensitivity analyses to investigate the potential effect of missing data assumption on the results of the primary analysis:
    1. On-treatment Analysis
    2. Pattern Mixture Model (PMM) Approaches
    3. Tipping Point Analyses (Added during FDA Review of the sNDA)
- Sensitivity analyses to investigate the patient level linear decline in FVC model assumption on the results of the primary analysis.

Sensitivity Analysis for Missing Data Handling 1 (On-treatment Analysis):
This analysis was the same as the primary analysis for the primary efficacy endpoint, except that only on-treatment measurement of FVC (mL) were used. This approach implicitly assumes data were missing at random (MAR) and that patients who discontinued treatment would have behaved similarly to those who remained on treatment. Because this assumption for the missingness mechanism is rather strong, results for Sensitivity Analysis 1 are not presented in this document.

Sensitivity Analysis for Missing Data Handling 2 (PMM Approaches):
To investigate the potential impact of missing data on the treatment effect, patients were classified into four different patterns depending on the availability of data:

- Patients with a 52- week FVC value:
    1. those who received study drug until 52 weeks (defined as patients who did not prematurely discontinue the study medication before 52 weeks (pattern 1)
    2. those who prematurely discontinued study drug before 52 weeks but who were followed up until week 52 (pattern 2)
- Patients without a 52- week FVC value:
    3. those who were alive at 52 weeks (pattern 3)
    4. those who died before 52 weeks (pattern 4)

These four patterns were used in sensitivity analyses to estimate the treatment effect under differing assumptions regarding the persistence of efficacy post withdrawal of randomized treatment. As described in Table 6, missing data were imputed (resulting in 1000 multiply imputed datasets) and three resulting alternative analyses were defined. For each imputed

dataset, the same statistical model as defined for the primary analysis was used for the analysis. The results were pooled following the standard multiple imputation procedure.[30]

Table 6: Primary and Sensitivity Analyses for Missing Data Handling (Pattern Mixture Model Approaches)

	Pattern 3: Missing week 52 data in patients still alive at 52 weeks		Pattern 4: Missing week 52 data in patients who died before 52 weeks	
Analysis	**Handling of missing week 52 data**	**Underlying assumption regarding persistence of efficacy post-withdrawal**	**Handling of missing week 52 data**	**Underlying assumption regarding persistence of efficacy after death**
Primary	No imputation	Assumes MAR	No imputation	Assumes MAR
Pattern Mixture Model 1	Based upon the slope (SE) estimates in Drug and Placebo in patients of pattern 2, multiple imputation of missing week 52 data in the respective treatment group	Rate of decline in patients with missing week 52 data is similar to rate of decline in patients of pattern 2 in the respective treatment group (e.g. treatment effect persists in same manner as for pattern 2 patients after study drug discontinuation)	Multiple imputation of missing 52 week data due to death based on the same slope (SE) estimates in Placebo patients of pattern 2, but truncated to force the slope in patients who died to be more severe than in those who survived	Assuming that deaths observed in the study will likely be related to worsening of SSc, it seems reasonable to assume that the unobserved FVC values should on average be lower than those in patients who did not die prior to week 52. Rate of decline in patients who died before week 52 is similar to rate of decline in the Placebo patients of pattern 2 with most severe slopes.
Pattern Mixture Model 2	Based upon the slope (SE) estimates in Placebo patients of pattern 2: multiple imputation of missing week 52 data in all patients regardless of treatment group	Rate of decline in all patients with missing week 52 data is similar to rate of decline in Placebo patients of pattern 2 (e.g. treatment effect does not persist after study drug discontinuation)		
Pattern Mixture Model 3	Based upon the slope (SE) estimates in Placebo patients from the primary analysis model, i.e. in patients from pattern 1 or 2: multiple imputation of missing week 52 data in all patients regardless of treatment group	Rate of decline in all patients with missing week 52 data is similar to rate of decline in all Placebo patients (e.g. treatment effect does not persist after study drug discontinuation)	Multiple imputation of missing 52 week data due to death based on the same slope (SE) estimates in all Placebo patients (i.e. in patients from pattern 1 or 2), but truncated to force the slope in patients who died to be more severe than in those who survived	Assuming that deaths observed in the study will likely be related to worsening of SSc, it seems reasonable to assume that the unobserved FVC values should on average be lower than those in patients who did not die prior to week 52. Rate of decline in patients who died before week 52 is similar to rate of decline in the Placebo patients with most severe slopes.

Abbreviations: SSc: systemic sclerosis; FVC: full vital capacity; SE: standard error; MAR: missing at random
Source: SAP Table 7.4.2.1.2:1

[30] Rubin D. Multiple Imputation for Nonresponse in Surveys, John Wiley & Sons, 1987.

Sensitivity Analysis for Missing Data Handling 3 (Tipping Point Analysis)

While the above Pattern Mixture Model (PMM) sensitivity analyses represent reasonable assumptions alternative to the assumption of the primary analysis, they do not comprehensively explore the plausible space of missing data assumptions. Therefore, the FDA review team requested additional analyses that systematically and comprehensively explore the space of plausible missing data assumptions. In particular, we recommended the inclusion of tipping point (TP) analyses that vary assumptions about the missing outcomes on the two treatment arms. The analyses should be two-dimensional, i.e., should allow assumptions about the missing outcomes on the two arms to vary independently, and should include scenarios where dropouts on nintedanib have worse slopes than dropouts on placebo. The goal is to explore the plausibility of missing data assumptions under which the conclusions change, i.e., under which there is no longer evidence of treatment effect. These analyses should include all observed data, regardless of whether measurements were made on- or off-treatment.

Sensitivity to the Analysis Model

Sensitivity to linearity assumption and sensitivity to covariates analyses results were consistent with the primary analysis model and will not be presented in this document.

Multiplicity Control Procedure

A hierarchical testing procedure was used, in that if results from the primary analysis for an endpoint were found to be statistically significant at the two-sided significance level of 0.05, the following endpoint in the hierarchy was to be tested at the same significance level in its primary analysis. If results for any of these endpoints were found to be not statistically significant, formal hypothesis testing was not performed for any remaining endpoints in the hierarchy. The procedure began with the primary efficacy endpoint, and the hierarchy was as shown below:

- Absolute change from baseline in the mRSS at Week 52
- Absolute change from baseline in SGRQ total score at Week 52

Primary Analyses for the Key/Hierarchical Secondary Efficacy Endpoints

Absolute change from baseline in the mRSS at Week 52: A restricted maximum likelihood (REML) based mixed effect model for repeated measures (MMRM) model was used for the analysis of continuous longitudinal secondary endpoints. The model included the fixed, categorical effects of treatment, ATA status, visit, and treatment-by-visit interaction, and continuous, fixed covariates of baseline and baseline-by-visit interaction. An unstructured variance-covariance structure was used to model the within patient measurements. Missing data were not imputed and assumed as missing-at-random.

Absolute change from baseline in the total SGRQ score at Week 52: This endpoint was analyzed in the same manner as in the primary analysis of the absolute change from baseline in mRSS at Week 52, using the same missing data handling methods.

Primary Analysis for Selected Other Secondary Efficacy Endpoints

Annual rate of decline in FVC in percent predicted over 52 weeks: This endpoint was analyzed in the same manner as in the primary analysis of the primary efficacy endpoint, using the same missing data handling methods.

Primary Analysis for Selected Further Efficacy Endpoints

In the analysis of binary endpoints, the categorical endpoints representing proportion of patients were summarized descriptively. In the SAP, it was pre-planned that Wilson 95% confidence interval and a nominal p-value were to be calculated for each proportion of patients. Patients with missing data were considered as non-responders. In the applicant's CSR, analyses of these endpoints were performed using a Cohchran-Mantel-Haenszel (CMH) model adjusting for ATA status. Adjusted Mantel-Haenszel odds ratios (OR) with 95% confidence intervals (CI) were used to quantify the treatment effect of nintedanib relative to placebo.

Safety Analyses

In general, safety analyses were descriptive in nature. No inferential statistical testing was planned on the safety data.

Protocol Amendments

There were 3 global amendments of study 1199.214.

Global amendment 1 (March 02, 2016) included the following key protocol changes:

- Women of childbearing potential had to perform a pregnancy test every 4-6 weeks. Urine dipstick tests were done at every visit and then were provided for at home pregnancy testing as soon as visit intervals were > 6 weeks
- The inclusion criterion was revised to change the reference time point for historical HRCT within 12 months to Visit 1 instead of Visit 2
- Exclusion criteria were updated to clarify exclusion for:
    - Severe ulcers other than digital ulcers, at discretion of the investigator
    - Severe GI symptoms due to SSc
    - Underlying chronic liver disease (Child Pugh A, B, C hepatic impairment).

Global amendment 2 (January 26, 2017) included the following key protocol changes:

- The inclusion criteria revised for SSc disease onset from within 5 years to within 7 years of Visit 1
- Exclusion criteria revised as follows:
    - To exclude patients not on mycophenolate mofetil / sodium or methotrexate within at least 8 weeks prior to Visit 2
    - Reference time point for airway obstruction assessment was changed from Visit 1 to Visit 2
    - History of Scleroderma Renal Crisis added as exclusion
- Clarification that all fatal cases would be reviewed by independent adjudication committee
- Absolute change from baseline at Week 52 in CRISS index score was added as a secondary endpoint
- Added analysis of rate of decline in FVC in percent predicted in the same way as the primary endpoint, including ATA status and baseline FVC% predicted as covariates

- The study sites were increased from 170 to 230 and in 33 instead of 20 countries worldwide.

Global amendment 3 (February 15, 2018) included the following changes:

- Clarification of end of study management for patients who prematurely discontinued study medication
- Definition of clinically significant deterioration extended to other clinical parameters than mRSS and FVC
- Based on the half-life of the study drug, adverse events that occur between the start of treatment and up to 7 days after the date of the last dose of study medication will also be analyzed.

The changes contained in the protocol amendments are not expected to impact study outcomes in a biased manner.

3.5 Review of Efficacy

3.5.1 Efficacy Review Approach

In the nintedanib for the treatment of SSc-ILD clinical development group, Study 1199.214 serves as the single study supporting efficacy. The efficacy review focuses on examining the robustness of the primary analysis result of the primary efficacy endpoint and the efficacy results on the key secondary efficacy endpoints.

3.5.2 Patient Disposition

A total of 819 patients were screened for eligibility. Of these, 580 patients passed the initial screening test and were randomized (290 nintedanib, 290 placebo). Among them, 4 patients (2 nintedanib, 2 placebo) were not treated due to not fulfilling the inclusion/exclusion criteria and determined post randomization. Thus, the Treated Set (TS) consisted of 576 patients (288 nintedanib, 288 placebo). These data are summarized in Table 7.

Table 7: Analysis Datasets

	Placebo	Nintedanib	Total
Screened Population			819 (100%)
Not Randomized			239 (29.2%)
Adverse Event			2 (<1%)
Consent withdrawn (not due to adverse event)			23 (2.8%)
Inclusion/Exclusion criteria not met			196 (23.9%)
Other			12 (1.5%)
Randomized Population	290	290	580 (70.8%)
Patients Not Treated	2 (<1%*)	2 (<1%*)	4 (<1%*)
Treated Set	288 (99.3%*)	288 (99.3%*)	576 (99.3%*)

Source: FDA Statistical Reviewer
*: Proportion with respect to the Randomized Population.

Of the 576 treated patients, 94% completed visits at Week 52, the time at which the primary endpoint was measured. The nintedanib group had a numerically higher study withdrawal rate (8%) compared with the placebo group (5%). The nintedanib group also had a numerically higher treatment discontinuation rate (19%) compared with the placebo group (1%). This contributed to differences in missing data between the two treatment groups. The impact of these differences was assessed via sensitivity analyses, detailed in this review.

The most common reason for study withdrawal was adverse event. Other reasons included *Patient refusal to continue taking study medication*, *Non-compliant with protocol*, and *Other*. With regard to treatment discontinuations, the reasons were similar to those for study withdrawal, with the most common reason also being *Adverse event*. These data are summarized in Table 8.

Table 8: Patient Disposition

	Placebo N=288	Nintedanib N=288	Total N=576
	n (%)	n (%)	n (%)
Prematurely Discontinued from Study Medication before 52 Weeks and Reason			
No	257 (89.2%)	232 (80.6%)	489 (84.9%)
Yes	31 (10.8%)	56 (19.4%)	87 (15.1%)
Reasons for Discontinuing Study Medication			
Adverse Event	21 (7.3%)	40 (13.9%)	61 (10.6%)
Patient refusal to continue taking study medication	7 (2.4%)	9 (3.1%)	16 (2.8%)
Non-compliant with protocol	1 (<1%)	1 (<1%)	2 (<1%)

	Placebo N=288	Nintedanib N=288	Total N=576
	n (%)	n (%)	n (%)
Other	2 (<1%)	6 (2.1%)	8 (1.4%)
Completed Visits up to 52 Weeks and Reason for Study Discontinuation			
Completed	275 (95.5%)	264 (91.7%)	539 (93.6%)
Did not Complete	**13 (4.5%)**	**24 (8.3%)**	**37 (6.4%)**
Reasons for not Completing Visits up to 52 Weeks			
Adverse Event	8 (2.8%)	14 (4.9%)	22 (3.8%)
Consent withdrawn, not due to adverse event	2 (<1%)	2 (<1%)	4 (<1%)
Non-compliant with protocol	1 (<1%)	0	1 (<1%)
Other	2 (<1%)	8 (2.8%)	10 (1.7%)

Abbreviations: N: sample size in Treated Set; n: sample size in corresponding category.
Source: FDA Statistical Reviewer

Protocol Violations/Deviations

In study 1199.214, important protocol deviations (iPDs) were defined as those protocol deviations that could potentially impact the efficacy assessments or the patients' rights or safety. Important protocol deviations were pre-defined in the SAP and assessed before the locking and unblinding of data; note that the term "deviation" is a synonym for "violation", which is used in the SAP. As no per protocol set was defined in this study, none of the iPDs led to exclusion of patients from any analyses. Over 52 weeks, 18.9% of patients were reported with iPDs (20% in the nintedanib group and 18% in the placebo group). Overall compliance was the largest protocol deviation in the nintedanib group (6% vs 3% in the placebo group). Overall compliance was defined as between 80% and 120% and the calculation was based on capsule count. In cases where capsules were not returned or returned incomplete to the study site, actual compliance could not be calculated. The decision, whether or not this constitutes an iPD, was based on investigator's assessment. The other protocol deviations were proportionally similar in both treatment groups.

3.5.3 Demographics and Baseline Characteristics

In study 1199.214, the demographic characteristics of the two treatment arms were generally balanced as summarized in Table 9. As expected for the target population of SSc-ILD, this was a predominantly female (75.2%) population. The patients were predominantly White (67.2%) and Asian (24.8%), of non-Hispanic ethnicity (93.1%), with a median age of 54 years, and similar by treatment group. The proportion of patients by age group were also similar by treatment group. The geographic contributes from the study sites are shown below. Approximately 46% of the patients were enrolled at sites in Europe, while 25% were enrolled in Canada and the United States, and 23% in Asia. Overall, the patient demographic characteristics were balanced and representative of the intended patient population.

Table 9: Demographics

		Placebo N=288	Nintedanib N=288	Total N=576
		n (%)	n (%)	n (%)
Sex	Female	212 (73.6%)	221 (76.7%)	433 (75.2%)
	Male	76 (26.4%)	67 (23.3%)	143 (24.8%)
Race	American Indian or Alaska Native	3 (1.0%)	2 (<1%)	5 (<1%)
	Asian	81 (28.1%)	62 (21.5%)	143 (24.8%)
	Black or African American	16 (5.6%)	20 (6.9%)	36 (6.3%)
	Multiple	2 (<1%)	2 (<1%)	4 (<1%)
	Native Hawaiian or other Pacific Islander	0	1 (<1%)	1 (<1%)
	White	186 (64.6%)	201 (69.8%)	387 (67.2%)
Ethnicity	Hispanic/Latino	18 (6.3%)	22 (7.6%)	40 (6.9%)
	Not Hispanic/Latino	270 (93.8%)	266 (92.4%)	536 (93.1%)
Region	Asia	71 (24.7%)	59 (20.5%)	130 (22.6%)
	Canada and United States	73 (25.3%)	69 (24.0%)	142 (24.7%)
	Europe	126 (43.8%)	140 (48.6%)	266 (46.2%)
	Rest of World	18 (6.3%)	20 (6.9%)	38 (6.6%)
Age (Years)	Median (Minimum, Maximum)	54.0 (21, 78)	57.0 (20, 79)	55.0 (20, 79)
Weight (Unit: kg)	Median (Minimum, Maximum)	68.0 (36, 124)	68.0 (37, 127)	68.0 (36, 127)
Height (Unit: cm)	Median (Minimum, Maximum)	163.0 (144, 192)	162.0 (145, 193)	162.0 (144, 193)

Abbreviations: N: sample size in Treated Set; n: sample size in corresponding category; kg: kilogram; cm: centimeter.
Source: FDA Statistical Reviewer

Patients had met criteria for SSc as defined in the inclusion criteria. In the overall population, mean FVC in L, FVC in percent predicted and DL_{CO} in percent predicted were 2.5 L, 72.5% and 53.0%, respectively and were generally balanced by treatment group. The median time since first onset of non-Raynaud symptoms was 3.41 years in the overall population, and similar by treatment group. Approximately half of the patients had diffuse cutaneous SSc and approximately half of the patients had limited cutaneous disease, and 60% of the patients had anti-topoisomerase antibodies. Anti-RNA polymerase III antibodies and anti-centromere antibodies were present in a minority of patients (8.5% and 7.1%, respectively). Interstitial lung disease was assessed by HRCT within 12 months of screening. The mean extent of fibrosis on HRCT, as determined by centralized reading, was 36%; the majority of patients had evidence of ground glass opacities (84.5%), and the minority had honeycombing (15.5%). Baseline disease characteristics were generally balanced by treatment group. Disease characteristics are presented in Table 10.

In addition to ILD, patients had a history of other SSc manifestations including pulmonary hypertension (9.0%), digital ulcers (38.7%), diarrhea/malabsorption/bacterial overgrowth (17.9%), esophageal dysphagia/reflux (74.3%), synovitis (24.1%), friction rubs (9.4%), and Raynaud phenomenon (96.7%). Prior digital ulcers were reported by 42.4% of nintedanib-treated patients as compared to 35.1% of placebo treated patients. Other disease manifestations were similar by treatment group.

At baseline, 48.4% of the patients received treatment with mycophenolate and 6.6% received methotrexate. Use of mycophenolate and methotrexate was similar by treatment group. In the analysis of all concomitant therapies (baseline, on-treatment, and post-study drug discontinuation), a greater proportion of patients in the nintedanib group than the placebo group used prednisone (26.4% vs. 20.1%), while a greater number in the placebo group than the nintedanib group were treated with prednisolone (19.1% vs. 17.4%). Use of proton pump inhibitors and nifedipine/amlodipine was similar by treatment group. Few patients received additional immunosuppression with cyclophosphamide (4.2% nintedanib, 1.4% placebo) or rituximab (1.4% nintedanib, 0.7% placebo) during or post-treatment.

Table 10: Baseline Disease Characteristics

		Placebo N=288	Nintedanib N=288	Total N=576
		n (%)	n (%)	n (%)
Lung Function				
FVC (L)	Mean (SD)	2.54 (0.82)	2.46 (0.74)	2.50 (0.78)
FVC (% predicted)	Mean (SD)	72.7 (16.6)	72.4 (16.8)	72.5 (16.7)
DL_{CO} (% predicted)	Mean (SD)	53.2 (15.1)	52.8 (15.1)	53.0 (15.1)
SSc-ILD History				
Time since first onset of non-Raynaud symptom (Years)	Mean (SD)	3.50 (1.78)	3.48 (1.62)	3.49 (1.70)
	Median (Min, Max)	3.47 (0.36, 7.16)	3.40 (0.26, 7.07)	3.41 (0.26, 7.16)
SSc subtype	Diffuse cutaneous SSc	146 (50.7%)	153 (53.1%)	299 (51.9%)
	Limited cutaneous SSc	142 (49.3%)	135 (46.9%)	277 (48.1%)
Autoantibody Status				
Anti-topoisomerase antibodies	Negative	111 (38.5%)	115 (39.9%)	226 (39.2%)
	Positive	177 (61.5%)	173 (60.1%)	350 (60.8%)
Immunosuppressive Agent Use at Baseline				
Mycophenolate mofetil	No	148 (51.4%)	149 (51.7%)	297 (51.6%)
	Yes	140 (48.6%)	139 (48.3%)	279 (48.4%)
Methotrexate	No	273 (94.8%)	265 (92.0%)	538 (93.4%)

		Placebo N=288	Nintedanib N=288	Total N=576
		n (%)	n (%)	n (%)
	Yes	15 (5.2%)	23 (8.0%)	38 (6.6%)
Mycophenolate Mofetil Use at Baseline by Region* (US and Canada vs. Rest)				
Mycophenolate mofetil Use at baseline*	US and Canada	73 (25.3%)	69 (24.0%)	142 (24.7%)
	No MMF Use	16 (5.6%)	12 (4.2%)	28 (4.9%)
	MMF Use	57 (19.8%)	57 (19.8%)	114 (19.8%)
	Other than US and Canada	215 (74.7%)	219 (76.0%)	434 (75.3%)
	No MMF Use	132 (45.8%)	137 (47.6%)	269 (46.7%)
	MMF Use	83 (28.8%)	82 (28.5%)	165 (28.6%)
HRCT Assessment Results				
Extent of fibrotic disease in the lung (%)	Median (Min, Max)	30 (5, 90)	30 (10, 95)	30 (5, 95)
Ground Glass Opacities	Missing	6 (2.1%)	5 (1.7%)	11 (1.9%)
	No	36 (12.5%)	42 (14.6%)	78 (13.5%)
	Yes	246 (85.4%)	241 (83.7%)	487 (84.5%)

Abbreviations: N: sample size in Treated Set; n: sample size in corresponding category; SD: standard deviation, Min: minimum; Max: maximum.
*: The common denominator in the *Mycophenolate mofetil use at baseline by Region* categories are the counts of the overall Treated Set population, not by Region.
Source: FDA Statistical Reviewer

3.5.4 Efficacy Results – Primary Endpoint

The primary endpoint was the annual rate of decline in forced vital capacity (FVC) in mL over 52 weeks. FVC has been used previously to support regulatory decision making for drugs approved for the treatment of IPF. While there is no regulatory precedent for its use in SSc-ILD, FVC has been proposed as a validated outcome measure in patients with SSc according to the principles of Outcome Measures in Rheumatologic Clinical Trials (OMERACT).[31]

In this section, we present the four steps the FDA review team took in assessing the treatment effect of nintedanib on FVC in Study 1129.214: evaluating the effect size through the primary analysis, assessing the robustness of the effect through sensitivity analyses on missing data assumptions, evaluating the effect size in terms of the FVC percent predicted, and interpreting the data using responder analyses of FVC change from baseline. All the four steps are based on pre-planned analyses in the study protocol or SAP with addition of tipping point analyses and graphical approaches to facilitate understanding.

[31] Merkel P, Clements PJ, Reveille P, et al. Current status of outcome measure development for clinical trials in systemic sclerosis. J Rheumatol 2003;30:1630–47.

3.5.4.1 The Primary Analysis

In Study 1199.214, there was a statistically significant difference in annual rate of decline in FVC in mL over 52 weeks, when comparing nintedanib to placebo (p=0.035). The adjusted mean difference between nintedanib and placebo was 40.9 mL/year (95% CI: 2.9 to 79.0) (Table 11).

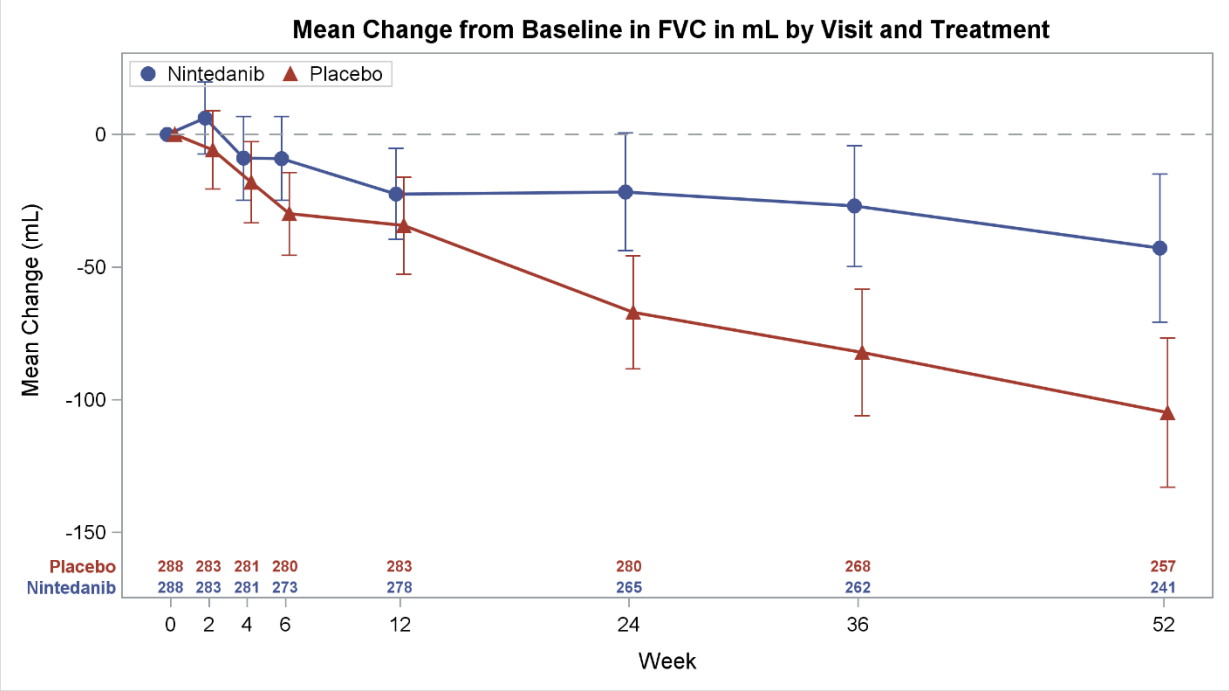
Table 11: Annual Rate of Decline in FVC in mL over 52 Weeks Primary Analysis (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Number Analyzed	288	287
Adjusted Annual Rate of Decline, mL/Year (SE)	-93.3 (13.5)	-52.4 (13.8)
Nintedanib vs. Placebo		
Difference (SE)		40.9 (19.4)
95% CI		2.9, 79.0
p-value		0.035

Abbreviations: N: sample size in Treated Set; SE: standard error; CI: confidence interval
Source: FDA Statistical Reviewer

The mean change from baseline in FVC in mL over 52 weeks by treatment group is shown in Figure 3. Data are observed values; vertical lines represent 95% confidence intervals. Numbers of patients with non-missing FVC data by week are displayed above the x-axis.

Figure 3: Mean (95% CI) Observed Change from Baseline in FVC in mL over 52 Weeks by Treatment (Treated Set)



Abbreviations: mL: milliliter.
Source: FDA Statistical Reviewer

3.5.4.2 Sensitivity Analyses on Missing Data Assumptions

Multiple sensitivity analyses to missing data assumptions were performed for the primary endpoint, including a series of analyses that utilized the Pattern Mixture Model (PMM) approach with multiple imputation and the tipping point analysis.

PMM Approaches with Multiple Imputation

All the three PMM analyses assumed unfavorable scenarios for the nintedanib group, to different degrees. By dividing the observed FVC data up to Week 52 into 4 complete/missing patterns, as discussed in section 3.4.1.1.6, these PMM approaches were

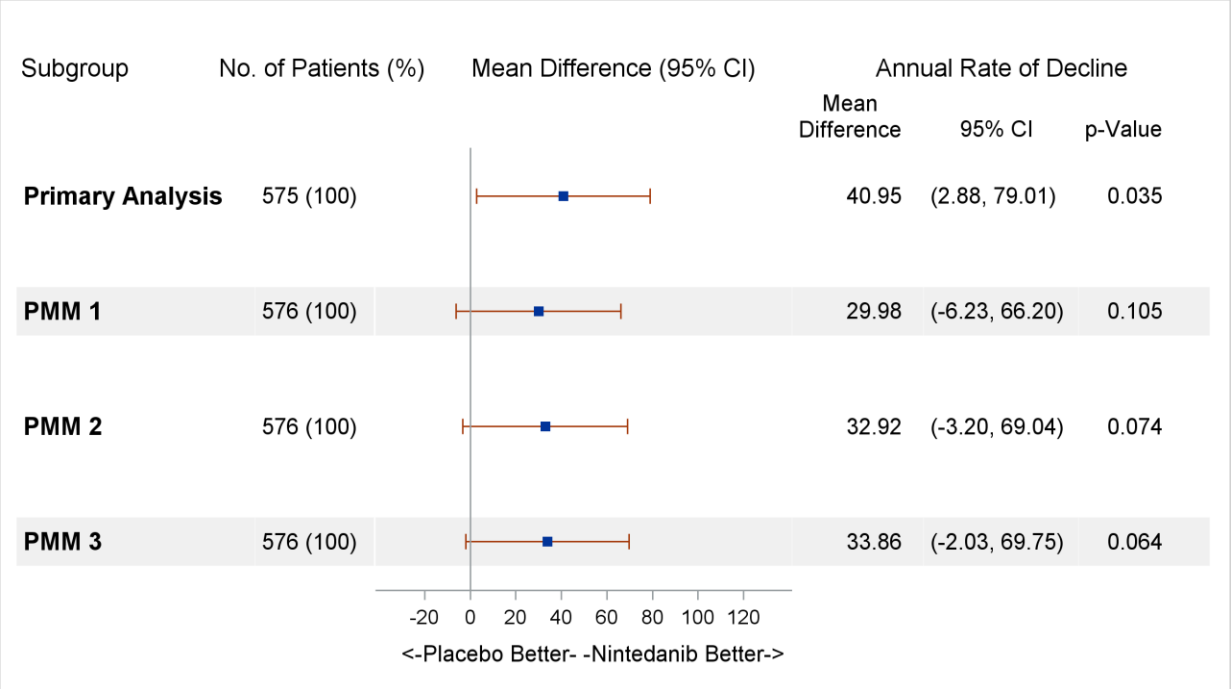
- PMM1:
  - Patients who were alive at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 3) was assumed to be similar to the rate of decline in patients who discontinued treatment but continued study visits till Week 52 (Pattern 2) in the ***respective treatment*** group;
  - Patients who were dead at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 4) was assumed to be similar to the rate of decline in the worst half of the patients who discontinued treatment but continued study visits till Week 52 (Pattern 2) in the ***placebo*** group.
- PMM2:

- Patients who were alive at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 3) was assumed to be similar to the rate of decline in patients who discontinued treatment but continued study visits till Week 52 (pattern 2) in the ***placebo*** group;
- Patients who were dead at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 4) was assumed to be similar to the rate of decline in the worst half of the patients who discontinued treatment but continued study visits till Week 52 (Pattern 2) in the ***placebo*** group.

- PMM3:
  - Patients who were alive at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 3) was assumed to be similar to the rate of decline in patients who continued study visits till Week 52 (Patterns 1 and 2) in ***placebo*** group;
  - Patients who were dead at Week 52: the rate of decline in patients who discontinued treatment and study visits (Pattern 4) was assumed to be similar to the rate of decline in the worst half of the patients who continued study visits till Week 52 (Pattern 2) in the ***placebo*** group.

All the three analyses resulted in reduced treatment effects, confidence intervals that include 0, and p-values greater than 0.05. The results of the three PMM analyses are summarized in Figure 4.

Figure 4: Forest-plot for Rate of Decline in FVC [mL/Year] over 52 Weeks Sensitivity Analyses (Treated Set)



Abbreviations: No.: Number; PMM: pattern mixture model; CI: confidence interval
Source: FDA Statistical Reviewer following the applicant's proposed analyses

In all three alternative PMM approach analyses, nintedanib is no longer statistically significantly separated from placebo.

Tipping Point Analysis

A tipping point analysis was then performed to evaluate how robust the primary analysis results were across varying missing data assumptions, a more comprehensive range over scenarios assumed in the PMM analyses. The objective of this analysis was to more precisely identify the point at which the conclusion changes. In this analysis, missing data with monotone missingness patterns were first multiply imputed assuming that missingness was at random among those in the same treatment group, with the same sex, age, height at baseline, ATA status, and with comparable FVC values from baseline through discontinuation. These imputed values were then shifted by modifying the slope of the decline after the last recorded FVC (within the 52-Week period) by the shift parameter (S) corresponding to the patient's treatment arm. The results over a relatively comprehensive range of by-arm shift (S) values are summarized in Table 12. The boxed cell in the table (Shift in Placebo = 0, Shift in Nintedanib = 0) can be read as a reference point: the analysis in this cell assumes missing-at-random mechanism as employed in the primary analysis and no shift (shift = 0) was applied to either arm; the result in this cell is nearly identical to the primary analysis result except for minor differences due to multiple imputation simulations. If FVC values after study discontinuation in the placebo group followed the same trend as those of comparable placebo patients who remained in the study through Week 52, then in order to tip to a lack of statistical significance, FVC values after study discontinuation in the nintedanib group would have to be roughly greater than 30 mL/year worse than those of comparable nintedanib patients who remained in the study through Week 52. Across the range of by-arm shifts explored by the applicant, it required shifts of 15 to 45 mL/year in the nintedanib group, relative to those applied to the placebo group, to tip the conclusion. If these shifts and relative shifts are clinically implausible, then missing data would be considered to have only minimal impact on the study conclusions.

Table 12: Annual Rate of Decline in FVC in mL over 52 Weeks, Tipping Point Analysis (Treated Set)

		Shift in Placebo (Unit: mL/Year)						
		-60	-45	-30	-15	0	15	30
Shift in Nintedanib (Unit: mL/Year)	-60	**38.3** **0.3, 76.3** **0.048**	37.7 -0.4, 75.7 0.052	37.0 -1.0, 75.0 0.056	36.4 -1.6, 74.4 0.060	35.8 -2.2, 73.8 0.065	35.2 -2.8, 73.2 0.070	34.6 -3.4, 72.6 0.075
	-45	39.4 1.4, 77.3 0.042	**38.7** **0.8, 76.7** **0.046**	38.1 0.1, 76.1 0.049	37.5 -0.5, 75.5 0.053	36.9 -1.1, 74.9 0.057	36.9 -1.7, 74.3 0.061	35.7 -2.3, 73.6 0.066
	-30	40.4 2.5, 78.4 0.037	39.8 1.9, 77.8 0.040	**39.2** **1.3, 77.2** **0.043**	38.6 0.6, 76.5 0.046	38.0 0.0, 75.9 0.0498	37.4 -0.6, 75.3 0.054	36.7 -1.2, 74.7 0.058
	-15	41.5 3.6, 79.5 0.032	40.9 3.0, 78.8 0.035	40.3 2.4, 78.2 0.037	**39.7** **1.8, 77.6** **0.040**	39.1 1.1, 77.0 0.044	38.4 0.5, 76.4 0.047	37.8 -0.1, 75.8 0.051
	0	42.6 4.7, 80.5 0.028	42.0 4.1, 79.9 0.030	41.4 3.5, 79.3 0.032	40.8 2.9, 78.7 0.035	**40.2** **2.2, 78.1** **0.038**	39.5 1.6, 77.4 0.041	38.9 1.0, 76.8 0.044
	15	43.7 5.8, 81.6 0.024	43.1 5.2, 81.0 0.026	42.5 4.6, 80.4 0.028	41.9 4.0, 79.7 0.030	41.2 3.3, 79.1 0.033	**40.6** **2.7, 78.5** **0.036**	40.0 2.1, 77.9 0.039
	30	44.8 6.9, 82.7 0.020	44.2 6.3, 82.1 0.022	43.6 5.7, 81.4 0.024	42.9 5.1, 80.8 0.026	42.3 4.4, 80.2 0.029	41.7 3.8, 79.6 0.031	**41.1** **3.2, 79.0** **0.034**

Abbreviation: mL: milliliter.
Each cell contains treatment difference, 95% confidence interval, and p-value.
Source: Analysis results by FDA Statistical Reviewer per the Applicant's Response to FDA Information Request

While the results for the primary endpoint were statistically significant based on the pre-specified analysis, the sensitivity analyses showed mixed results, mainly because the magnitude of the treatment effect size was small. Whether this effect is clinically meaningful is an issue we would like the AC panel to discuss.

3.5.4.3 Annual Rate of Decline in FVC in Percent Predicted – A Secondary Endpoint

Annual rate of decline in FVC in percent predicted was analyzed using a similar model as in the analysis of the primary endpoint. There was a statistically significant difference in annual rate of decline in FVC in percent predicted over 52 weeks, when comparing nintedanib to placebo (p=0.033). The adjusted mean difference between nintedanib and placebo was 1.2 %/year (95% CI: 0.1 to 2.2) (Table 13).

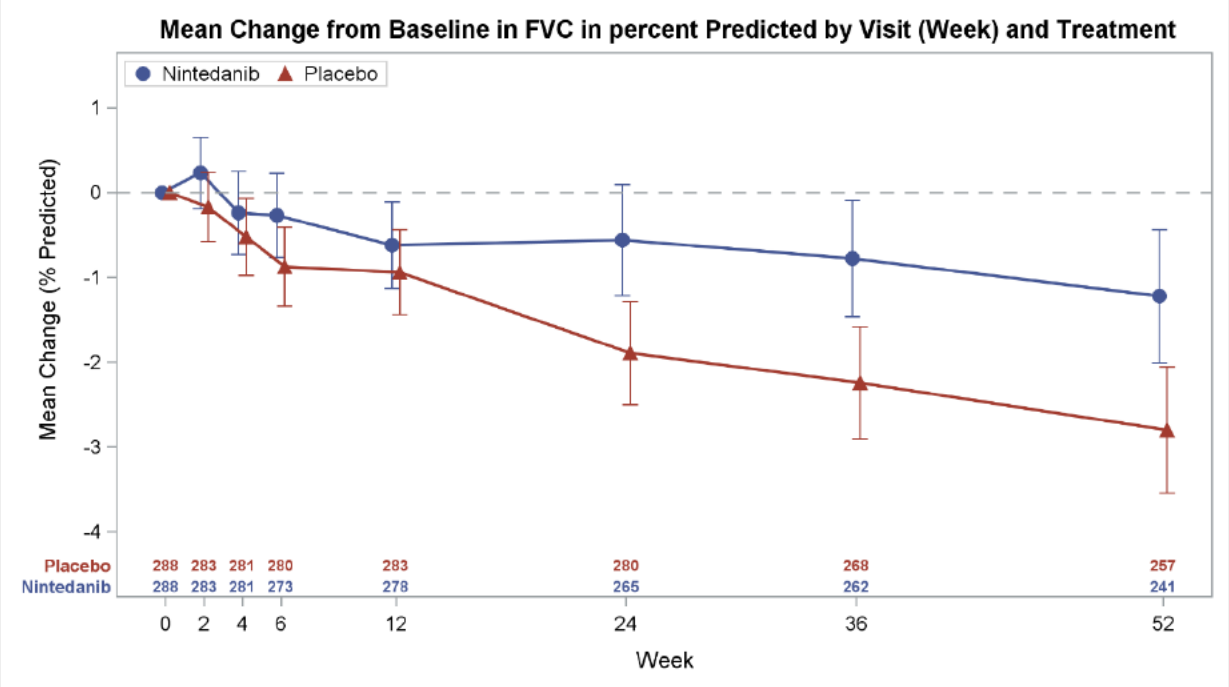
Table 13: Annual Rate of Decline in FVC in Percent Predicted over 52 Weeks Primary Analysis (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Number Analyzed	288	287
Adjusted Annual Rate of Decline, %/year (SE)	-2.6 (0.38)	-1.4 (0.4)
Nintedanib vs. Placebo		
Difference (SE)		1.2 (0.5)
95% CI		(0.1, 2.2)
p-value		0.033

Abbreviations: N: sample size in Treated Set; SE: standard error; CI: confidence interval
Source: FDA Statistical Reviewer

The mean change from baseline in FVC in percent predicted over 52 weeks by treatment group is shown in Figure 5. Data are observed values; vertical lines represent 95% confidence intervals. Numbers of patients with non-missing FVC data by week are displayed above the x-axis.

Figure 5: Mean (95% CI) Observed Change from Baseline in FVC in Percent Predicted over 52 Weeks by Treatment (Treated Set)



Source: FDA Statistical Reviewer

3.5.4.4 Categorical Endpoints Representing Proportion of Patients with a Decline in FVC at Week 52 – Further Endpoints

In Further Endpoints, the protocol/SAP also defined categorical "non-responder" variables based on either relative change from baseline in FVC in mL or absolute change from baseline in FVC in percent predicted, as follows:

- patients with a relative decline from baseline in FVC in mL at Week 52 of >5%
- patients with a relative decline from baseline in FVC in mL at Week 52 of >10%
- patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of >5%
- patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of >10%

However, to facilitate the interpretation of benefit (response), as opposed to no benefit (non-response), the FDA review team converted the protocol-specified endpoints to "responder" variables corresponding to favorable outcomes:

- patients with a relative decline from baseline in FVC in mL at Week 52 of ≤5%
- patients with a relative decline from baseline in FVC in mL at Week 52 of ≤10%
- patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of ≤5%
- patients with an absolute decline from baseline in FVC in percent predicted at Week 52 of ≤10%

These endpoints are summarized in Table 14. In these analyses, patients with missing data at Week 52 were categorized as non-responders. To explore the treatment effect of nintedanib group relative to placebo, a Cochran-Mantel-Haenszel (CMH) model adjusting for ATA status was performed for each responder variable. The adjusted Mantel-Haenszel odds ratios (OR) with associated 95% confidence intervals (CI), and nominal p-values are also reported in Table 14.

The proportion of responders with 5% threshold (relative decline ≤5%) was numerically higher in the nintedanib group (59.4%) than in the placebo group (51.7%), favoring nintedanib over placebo; the odds ratio was 1.37 (95% CI 0.98, 1.89; nominal p-value = 0.066). The proportion of responders with 10% threshold (relative decline ≤10%) was numerically lower in the nintedanib group (72.2%) than in the placebo group (73.6%), not favoring nintedanib over placebo; the odds ratio was 0.93 (95% CI 0.65, 1.35; nominal p-value = 0.704), indicating an inconsistent direction of the treatment effect, likely due to the small effect size and the disproportionately higher missing data in the nintedanib group and the assumption that missing data represent worse outcome. Responder analyses using absolute decline in FVC in percent predicted showed a similar pattern with the analyses using relative decline in FVC in mL (Table 14).

Table 14: Proportions of Responders with Certain Thresholds Over 52 Weeks (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)	Comparison vs. Placebo*			
	n (%)	n (%)	Odds ratio	95% CI		Nominal p-value
				Lower	Upper	
Responder definition using relative decline from baseline in FVC in mL at Week 52						
Relative decline ≤5%	149 (51.7%)	171 (59.4%)	1.36	0.98	1.89	0.066
Relative decline ≤10%	212 (73.6%)	208 (72.2%)	0.93	0.64	1.34	0.704
Responder definition using absolute decline from baseline in FVC in % predicted at Week 52						
Absolute decline ≤5%	186 (64.6%)	196 (68.1%)	1.16	0.82	1.64	0.386
Absolute decline ≤10%	236 (81.9%)	227 (78.8%)	0.82	0.54	1.24	0.348

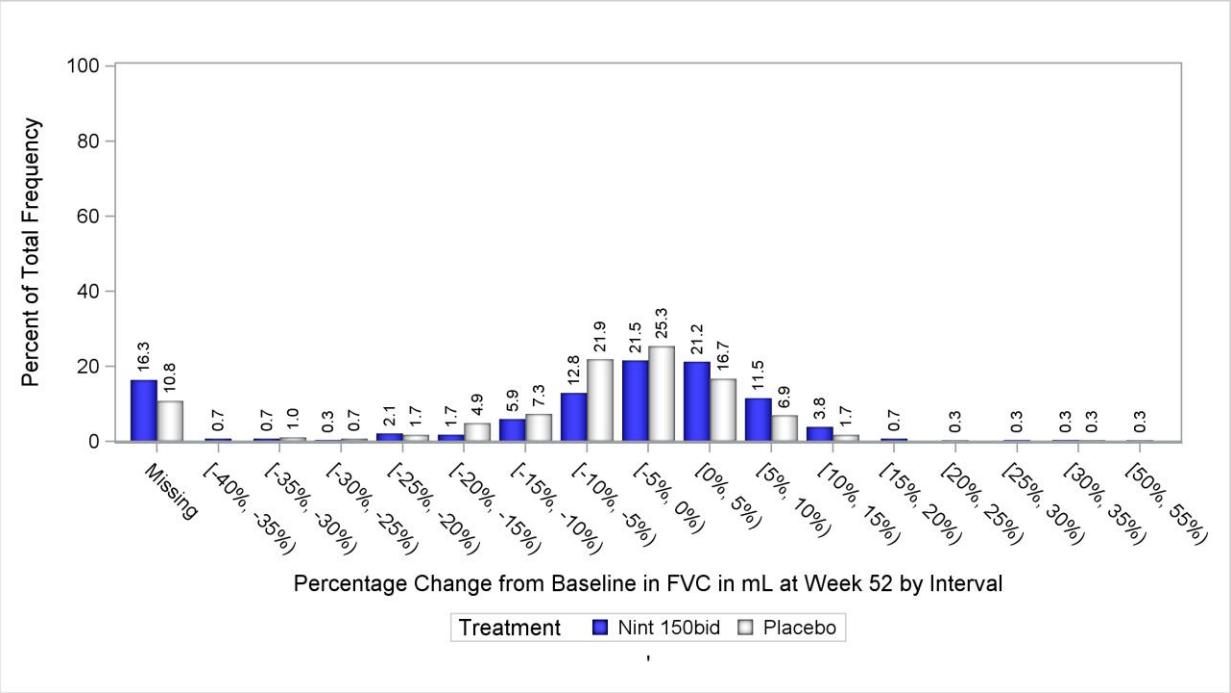
Abbreviations: N: sample size in Treated Set; n: number of patients within category; SE: standard error; CI: confidence interval
Note: Patients with missing data at Week 52 were considered as non-responders.
* Based on Cochran-Mantel-Haenszel tests stratified on baseline ATA status.
Source: FDA Statistical Reviewer

The Applicant conducted similar analyses on the "non-responder" variables. We note that, the FDA and Applicant analyses convey identical information because non-response is just a complementary event of response. Of note, unlike FDA, the Applicant used the worst observation carried forward approach to imputing missing data at Week 52 and then applied the threshold to define non-response.

We also used graphical approaches to investigate the distribution of change from baseline in FVC by treatment group.

Figure 6 displays this distribution as a histogram of percent change from baseline in FVC in mL at Week 52. In this plot, missing data were represented in a group on the left, reflecting the assumption that missing data have worse outcome.

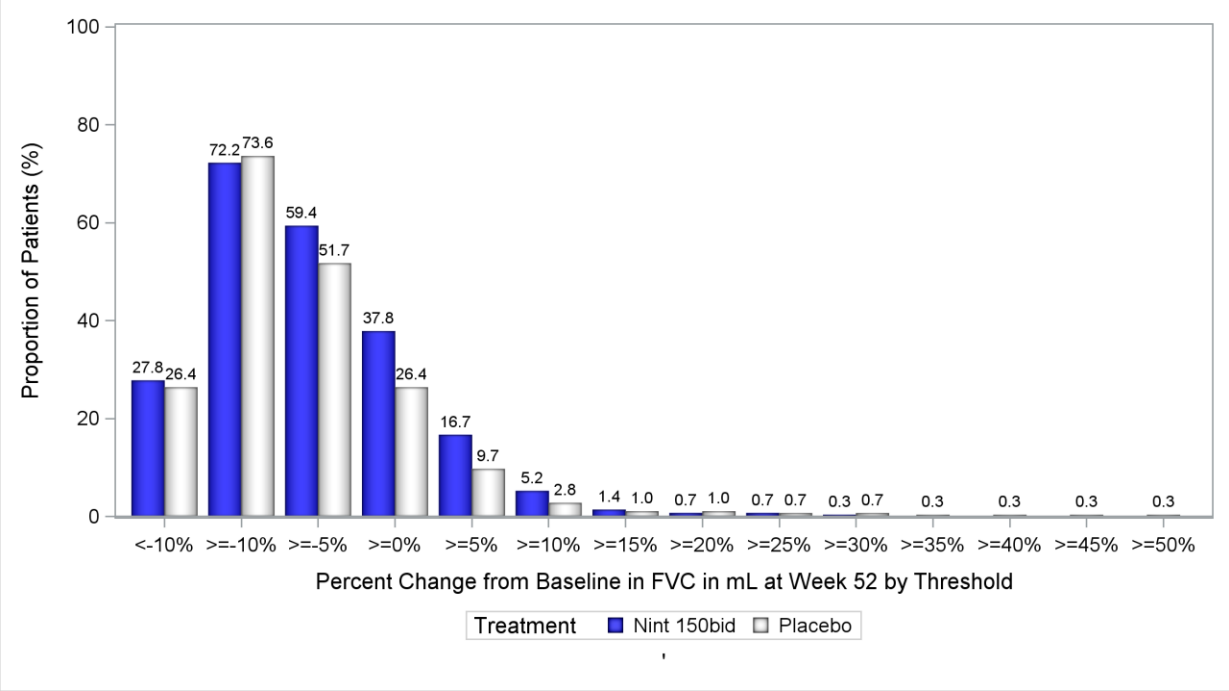
Figure 6: Histogram of Percent Change from Baseline in FVC in mL at Week 52 (Treated Set)



Abbreviations: mL: milliliter
Source: FDA Statistical Reviewer

To visually aid the understanding of the responder analysis, in which responders were defined as patients with no worse outcome than certain thresholds, Figure 7 displays the proportions of responders at various response definitions; that is, proportions of patients whose percent change from baseline were greater than certain cutoffs, where missing data was imputed as a decline worse than 10%. For example, with cutoff of -10%, 72.2% and 73.6% of patients in the nintedanib and placebo arms respectively had 10% or less decline from baseline in FVC (mL) at Week 52, indicating that placebo is numerically favorable over nintedanib. On the other hand, with cutoff of -5%, 59.4% on nintedanib and 51.7% of patients on placebo had 5% or less decline from baseline in FVC (mL) at Week 52, indicating that nintedanib is numerically favorable over placebo. Finally, with cutoff of 0%, 37.8% and 26.4% of patients on nintedanib and placebo respectively had an improvement at Week 52 compared with baseline, indicating that nintedanib is numerically favorable over placebo. In this plot, missing data were represented in a group on the left combined with the group of >10% decline, again reflecting the assumption that missing data have worse outcome.

Figure 7: Proportions of Responders at Various Cutoffs among Treated Patients, Based on Analysis of Percent Change from Baseline in FVC in mL at Week 52 (Treated Set)



Abbreviations: mL: milliliter
Note: Missing data were considered as having a relative decline in FVC in mL of >10%.
Source: FDA Statistical Reviewer

Similar graphs were also generated for absolute change from baseline in FVC in percent predicted at Week 52 presented in Appendix 4.3. Interpretation of the graphs should be similar to those with relative decline from baseline in FVC in mL.

3.5.5 Efficacy Results – Key Secondary endpoints

The key secondary efficacy endpoints were as follows:

- Absolute change from baseline in the mRSS at Week 52
- Absolute change from baseline in SGRQ total score at Week 52

Key secondary endpoints were analyzed in a hierarchical manner such that if the previous endpoint failed to reach statistical significance, the subsequent endpoints were not considered statistically significant.

Absolute Change from Baseline in mRSS at Week 52

Modified Rodnan Skin Score (mRSS)[32] is a widely accepted measure of skin involvement in SSc. It consists of an evaluation of patient's skin thickness rated by clinical palpation using a 0–3 scale for each of 17 surface anatomic areas of the body. The minimum clinically important

[32] Steen VD, Medsger TA Jr. Improvement in skin thickening in systemic sclerosis associated with improved survival. Arthritis Rheum 2001:44:2828−35.

difference (MCID) of the mRSS has been estimated in the range of 3.2 - 5.3 mRSS units anchored on physician global assessment.[33] A negative change indicates improvement in skin thickening.

For the first key secondary endpoint of absolute change from baseline in mRSS at Week 52, there was a -0.21 (95% CI: -0.94 to 0.53) difference between the nintedanib group compared to placebo, based on the FDA statistician's analysis. This was not statistically significant (p=0.579). Given the hierarchical analysis structure, the subsequent secondary endpoint was not considered statistically significant.

Table 15: Absolute Change from Baseline in mRSS at Week 52 (Treated Set)

	Placebo (N=288)	**Nintedanib (N=288)**
Number Analyzed	286	288
Baseline Mean mRSS Score (SD)	10.91 (8.81)	11.33 (9.18)
Adjusted Mean Change from Baseline (SE)	-1.96 (0.26)	-2.17 (0.27)
Nintedanib vs. Placebo		
Difference (SE)		-0.21 (0.37)
95% CI		-0.94, 0.53
p-value		0.579

Abbreviations: N: sample size in Treated Set; SD: standard deviation; SE: standard error; CI: confidence interval; mRSS: Modified Rodnan Skin Score.
Source: FDA Statistical Reviewer

The absolute change from baseline in mRSS was similar in both the nintedanib and placebo treatment groups and did not exceed the MCID.

Absolute change from baseline in SGRQ total score at Week 52
For the second key secondary endpoint of absolute change from baseline in SGRQ total score at Week 52, there was a 1.69 (95% CI: -0.73 to 4.12) difference between the nintedanib group compared to placebo, based on the FDA statistician's analysis. This comparison shows a numerically worse outcome in the nintedanib group as compared with the placebo group.

[33] Khanna D, Furst DE, Hays RD, et al. Minimally important difference in diffuse systemic sclerosis: results from the D-penicillamine study. Ann Rheum Dis 2006;65:1325–9.

Table 16: Absolute Change from Baseline in SGRQ Total Score at Week 52 (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Number Analyzed	283	282
Baseline Mean SGRQ total Score (SD)	39.4 (20.9)	40.7 (20.2)
Adjusted Mean Change from Baseline (SE)	-0.88 (0.87)	0.81 (0.88)
Nintedanib vs. Placebo		
Difference (SE)		1.69 (1.24)
95% CI		-0.73, 4.12

Abbreviations: N: sample size in Treated Set; SD: standard deviation; SE: standard error; CI: confidence interval; SGRQ: St. George's Respiratory Questionnaire.
Source: FDA Statistical Reviewer

In summary, the key secondary endpoints were not supportive of a direct treatment benefit for nintedanib over placebo at Week 52 in this study.

3.5.6 Other Efficacy Results – Selected Endpoints

Time to Death over the Whole Trial

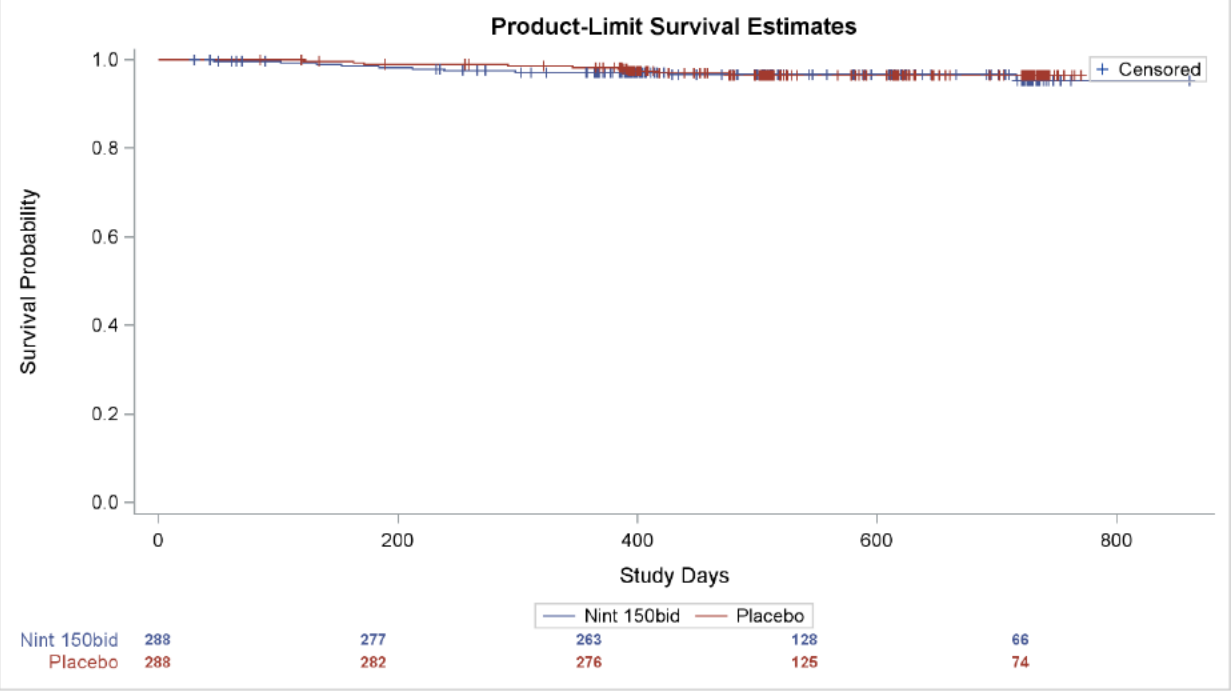
Of the total 576 TS patients, survival status at the end of study was followed-up for 570 patients, with 6 lost to follow-up (1 patient in placebo group, and 5 in nintedanib group). There were 19 deaths in total across the two treatment groups at the end of study, with the rest of patients censored. Table 17 summarizes the analysis results for the mortality endpoint through 2 approaches: Crude rate of death, and Cox proportional hazard regression model for time to death. The crude probability of death was 3.1% in the placebo group, and 3.5% in the nintedanib group. The hazard ratio of nintedanib group versus placebo group was 1.16 (95% CI: from 0.47 to 2.84), numerically favoring placebo. Figure 8 displays the probability of survival using Kaplan-Meier estimates.

Table 17: Cox Proportional Hazards Analysis of Time to Death Over the Entire Study Duration (Treated Set)

	Placebo (N=288)	Nintedanib (N=288)
Survival status at the end of study, n (%)		
Dead	9 (3.1%)	10 (3.5%)
Lost to follow-up (Vital status at the End of Study Unknown)	1 (0.3%)	5 (1.7%)
Alive (Censored at the End of Study)	278 (96.5%)	273 (94.8%)
Cox Proportional Hazard Model Analysis		
Nintedanib vs. Placebo		
Hazard Ratio		1.16
95% CI		0.47, 2.85
p-value		0.751

Abbreviations: N: sample size in Treated Set; CI: confidence interval.
Source: FDA Statistical Reviewer

Figure 8: Kaplan-Meier Estimate of Probabilities of Survival (Treated Set)



Abbreviations: Nint 150bid: nintedanib 150 milligram twice daily.
Source: FDA Statistical Reviewer

Of note, in the IPF clinical program, there were trends toward improvement in mortality.[34] This was not observed in Study 1199.214. This may be due to the relatively short duration of the

[34] FDA-approved nintedanib labeling

study to assess long term changes in a chronic disease, differences in the progression of the underlying diseases between IPF and SSc-ILD, and/or the smaller study sample size.

There were no differences in other secondary endpoints, including FACIT-dyspnea score and DL_{CO} at Week 52. Additionally, there were no differences in other disease-related secondary endpoints, including ACR CRISS responder, number of digital ulcers, or HAQ-DI.

3.5.7 Additional Analyses

The Applicant pre-planned subgroup analyses for the primary and both key secondary efficacy endpoints with subgroups based on antitopoisomerase (ATA) status, age, gender, race, geographical region, mycophenolate mofetil/sodium use at baseline, and SSc subtype. No significant interaction was found between treatment and these subgroups at the 5% level of statistical significance.

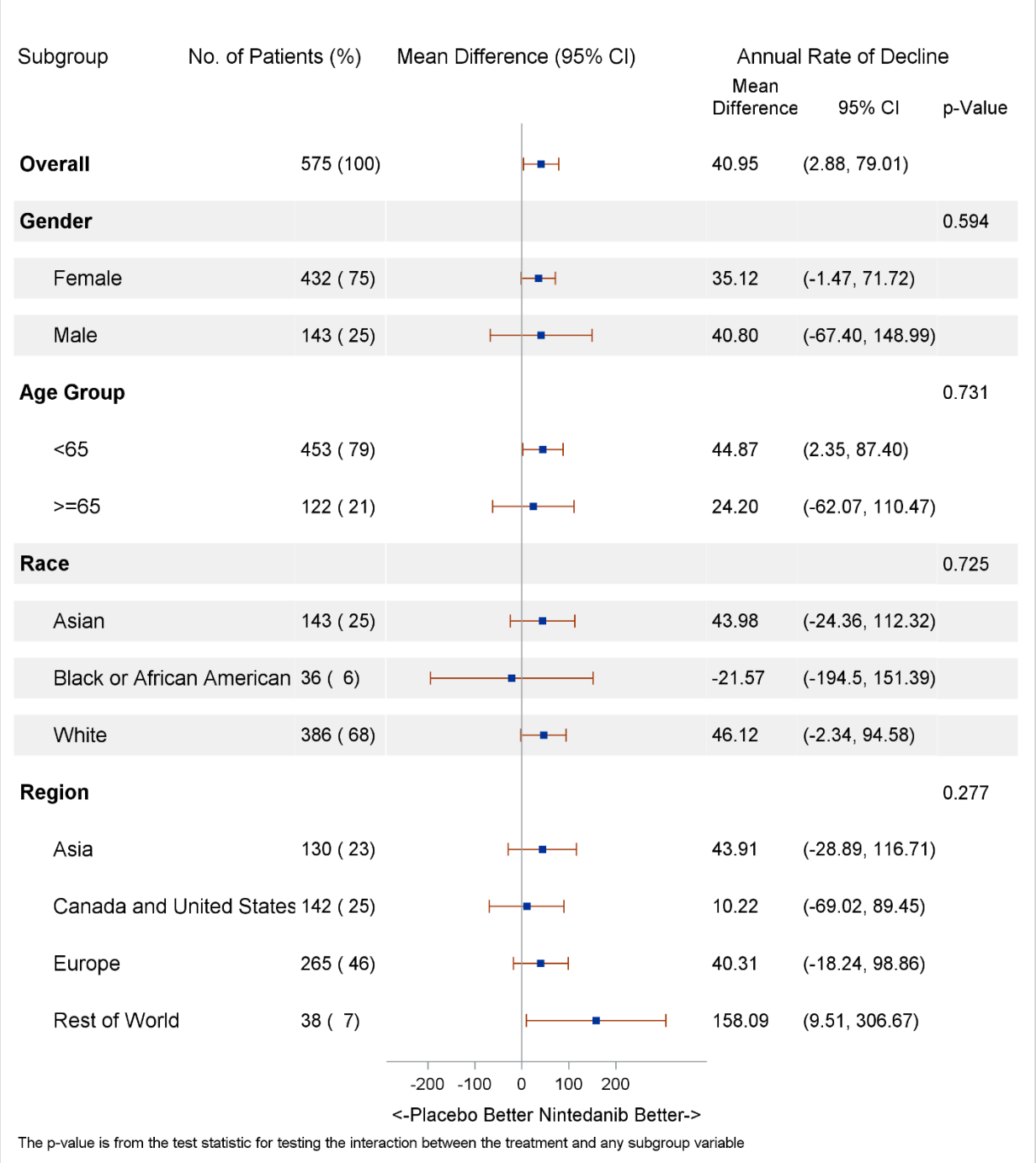
In this review, subgroup analyses were only performed for the primary efficacy endpoint. This section provides the reviewer's subgroup analyses by gender, race, age, and geographical region. Subgroup analyses were also performed for selected baseline SSc disease and treatment factors to inform the potential effect of these factors on efficacy.

3.5.7.1 Gender, Race, Age, and Geographic Region

For each subgroup factor, the model was adapted from the pre-specified primary efficacy analysis model. For the annual rate of decline in FVC endpoint, an interaction analysis was performed with the primary analysis random coefficient model by including the subgroup variable, the subgroup variable-by-time interaction, and the subgroup variable-by-time-by-treatment interaction as covariates. When a covariate in the model is the subgroup variable, it is replaced with the categorical version of itself when needed. By-subgroup mean annual rates of decline in FVC were estimated to illustrate the treatment effects under each subgroup. Under each subgroup, the mean difference estimate between the nintedanib group and the placebo group together with associated CIs was presented using a forest plot.

The rate of decline in FVC (mL/year) over 52 weeks by demographic subgroup is seen in Figure 9. Overall, the difference between the nintedanib and placebo groups was 40.95. The numbers are generally similar between females and males. There was less of a treatment difference in the over 65 years of age group category as compared to under 65. In addition, the treatment difference was not the same by race. White and Asian patients had a similar treatment difference that was consistent with the overall analysis, but the Black or African American subgroup of patients appeared to have less of a response to study drug. Definitive conclusions regarding this observation are limited due to the small subgroup sample size. Finally, there was also difference in subgroup by region. Notably, the patient population in Canada and the United States only had a mean treatment difference of 10.22 mL/year vs the overall of 40.95 mL/year.

Figure 9: Forest Plot of the Rate of Decline in FVC (mL/year) Over 52 Weeks by Demographic Subgroup (Treated Set)



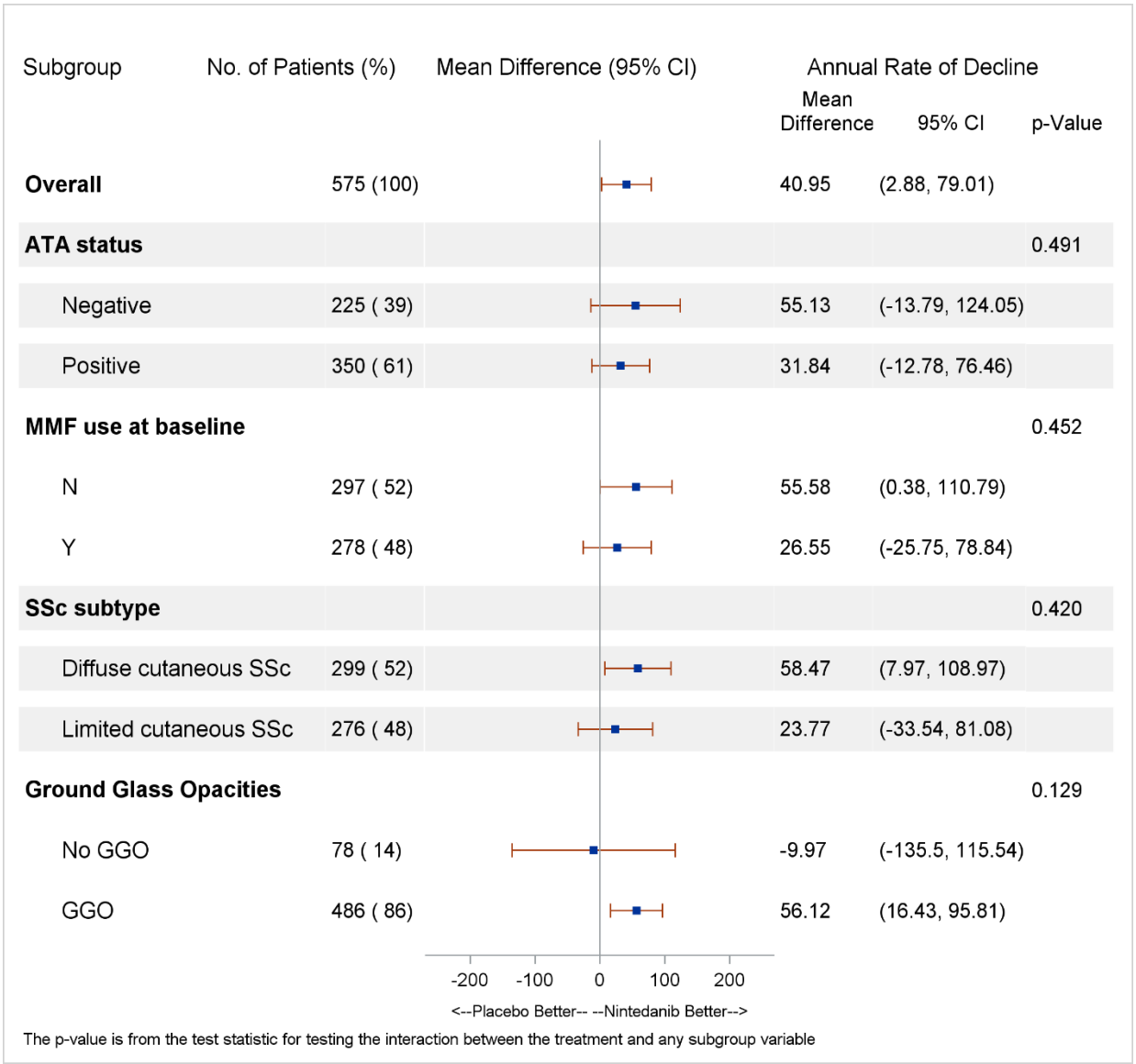
Abbreviations: No.: number; mL: milliliter; CI: confidence interval
Source: FDA Statistical Reviewer

3.5.7.2 Other Special/Subgroup Populations

The protocol pre-planned subgroup analyses on SSc-ILD baseline disease characteristic subgroup factors included ATA status and SSc subtype (Figure 10).

To evaluate the influence of stable background immunosuppressive therapy to study treatment, the protocol pre-planned subgroup analyses also included mycophenolate mofetil/sodium use at baseline. A less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients on MMF at baseline (treatment difference 26.6 mL/year). We also investigated the influence of baseline HRCT assessment patterns including ground glass opacities to nintedanib (Figure 10). The modeling approach was similar as that for the demographic subgroup analyses.

Figure 10: Forest Plot of the Rate of Decline in FVC (mL/year) Over 52 Weeks by Baseline Disease Characteristics Subgroup (Treated Set)

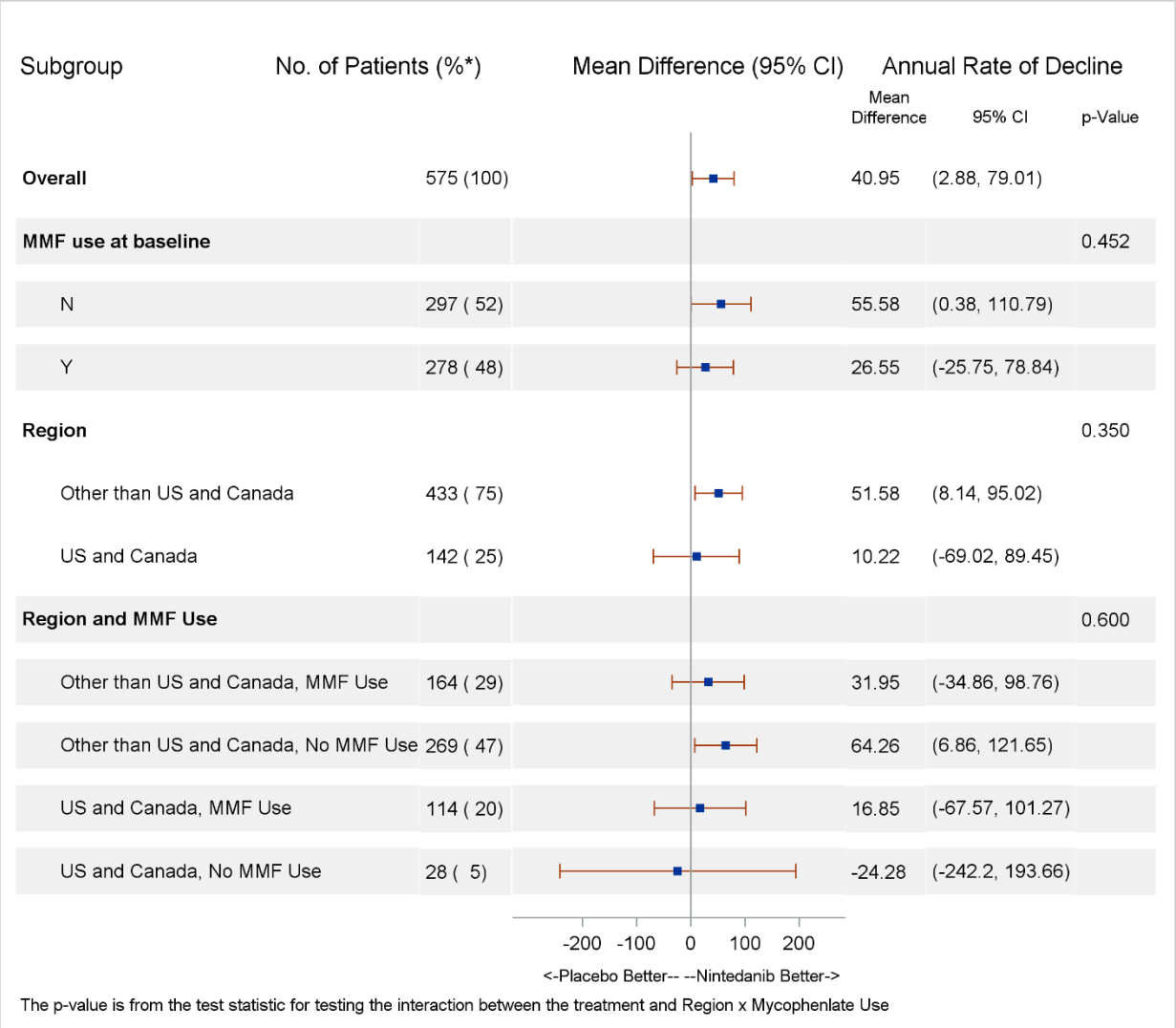


Abbreviations: No.: number; mL: milliliter; CI: confidence interval; MMF: Mycophenolate.
Source: FDA Statistical Reviewer

As clinical practice may differ across countries, the FDA review team also performed a subgroup analysis defined by both region (US and Canada, vs. Other than US and Canada) and

baseline MMF use to evaluate the influence of stable background MMF use to study treatment by region (Figure 11). A less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients from the US and Canada (treatment difference 10.2 mL/year). In the US and Canada subgroup, approximately 80% of the patients received MMF at baseline.

Figure 11: Forest Plot of the Rate of Decline in FVC (mL/year) over 52 weeks by Region by MMF Use Subgroup (Treated Set)



Abbreviations: No.: number; mL: milliliter; CI: confidence interval; MMF: Mycophenolate.
*: The common denominator N=576 was used in calculating the proportions of patients in each combination level defined by Region x MMF Use.
Abbreviation: MMF=Mycophenolate Mofetil.
Source: FDA Statistical Reviewer

Across the subgroup factors, there was no significant interaction between subgroups and treatment. However, lack of a significant treatment-by-subgroup interaction could be due to small subgroup sample size and should not be interpreted as evidence that no interaction exists.

3.5.8 Efficacy Summary

The adjusted annual rate of decline in FVC over 52 weeks was lower in the nintedanib group (-52.4 mL/year) than in the placebo group (-93.3 mL/year), with a statistically significant treatment difference of 40.9 mL/year (95% CI: 2.9 to 79.0; p=0.035). The adjusted annual rate of decline in percent predicted FVC over 52 weeks was -1.4% predicted/year in the nintedanib group and -2.6% predicted/year in the placebo group; the adjusted difference between groups was 1.2% predicted/year (95% CI: 0.1 to 2.2; p=0.033). While the results for the primary endpoint were statistically significant based on the pre-specified analysis, the sensitivity analyses on missing data assumptions and responder analyses with various thresholds showed mixed results, mainly because the magnitude of the effect size was small. Differences of this magnitude did not result in improvement in measures of direct clinical benefit related to pulmonary involvement, such as SGRQ or FACIT-dyspnea score at Week 52. Also, there were no differences in other disease-related secondary endpoints, including ACR CRISS responder, number of digital ulcers, or HAQ-DI. There was also no improvement in mortality. Additionally, a less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients on mycophenolate mofetil at baseline (treatment difference 26.6 mL/year) and patients from the U.S. and Canada (treatment difference 10.2 mL/year).

3.6 Review of Safety

3.6.1 Overall Exposure

In Study 1199.214, 288 patients were exposed to nintedanib 150 mg BID and 288 patients received placebo. Over 52 weeks, the mean (SD) exposure was 10.52 (3.43) months in the nintedanib group and 11.35 (2.39) months in the placebo group. Similarly, over the whole study, the mean exposure was lower in the nintedanib group (14.51 months) as compared to the placebo group (15.70 months).

Exposure data during the 52-week treatment period are summarized in Table 18. A greater proportion of patients in the nintedanib group had duration of exposure ≤3 months and 3-6 months, while a greater proportion of placebo-treated patients had exposures > 12-14 months, reflecting the greater number of patients who discontinued from the nintedanib treatment group. See Section 3.5.2 for discussion of patient disposition. A greater proportion of patients in the nintedanib treatment group than in the placebo group had at least 1 dose reduction (40.6% and 4.5%, respectively) or at least 1 treatment interruption (37.8% and 11.5%, respectively). This likely reflects tolerability issues with nintedanib treatment. See discussion below of AEs leading to treatment interruption.

Table 18: Exposure to Study Drug Over 52 Weeks, Treated Set

	Placebo N=288	Nintedanib N=288
Duration of exposure [1] (months)		
Mean (SD)	11.35 (2.39)	10.52 (3.43)
Median (min, max)	12.21 (0.4, 12.2)	12.21 (0, 12.2)
Duration of exposure, [1] categories (months), (N, %)		
≤3	10 (3.5)	25 (8.7)
>3–6	9 (3.1)	16 (5.6)
>6–12	75 (26.0)	73 (25.3)
>12–14	194 (67.4)	174 (60.4)
Total exposure [patient years]	273.0	253.0
Patients with at least 1 dose reduction (N, %)	13 (4.5)	117 (40.6)
Patients with at least 1 treatment interruption, (N, %)	33 (11.5)	109 (37.8)

[1]Duration of exposure over 52 weeks was defined from first study drug intake up to:

a. Last drug intake for patients who prematurely discontinued before 52 weeks (included) or
b. Week 52 (ie 372 days after first drug intake) for patients who did not prematurely discontinue the study drug before 52 weeks (included)

Adapted from Study 1199.214 CSR Table 10.5:1

3.6.2 Deaths

In Study 1199.214, there were 19 deaths overall in the treated set, balanced by treatment group (Table 19). During the treatment period, there were 11 patients with treatment-emergent AEs (TEAEs) leading to death, including 6 patients (2.1%) in the nintedanib group and 5 patients (1.7%) in the placebo group. TEAEs that lead to death in the nintedanib group included 1 patient each with: arrhythmia, scleroderma renal crisis with thrombotic microangiopathy, acute lung injury, pneumonia, lung adenocarcinoma, and malignant mesothelioma. These were adjudicated as 2 CV deaths due to sudden cardiac death and ischemic stroke, 2 respiratory deaths due to pneumonia, and 2 non-CV/non-respiratory deaths. In the placebo group, the TEAEs that lead to death included 1 patient each with: cardiac arrest, dyspnea, acute myocardial infarction, pneumonia, and interstitial lung disease. The causes of death in the placebo group were adjudicated as 3 CV deaths due to sudden cardiac death and acute MI, and 2 respiratory deaths due to underlying ILD.

During the post treatment period, there were 4 AEs leading to death in each treatment arm. Deaths in patients randomized to the nintedanib group included 1 patient each with: chest pain, circulatory collapse, respiratory failure, and small cell lung cancer. These were determined by adjudication to be undetermined death, respiratory death due to other respiratory causes, respiratory death due to underlying ILD, and non-CV/non-respiratory death, respectively. The 4 post-treatment deaths in the placebo group were due to 1 patient each with: cardiac arrest, lung neoplasm malignant, septic shock and sudden death. These were determined by adjudication to be respiratory death/ underlying ILD, undetermined death, non-CV/non-respiratory death, and respiratory death.

Table 19: AEs Leading to Death over Entire Study by Preferred Term, Treated Set

	Treatment Period	Post-Treatment Period
Nintedanib	Acute lung injury	Chest pain
	Arrythmia	Circulatory collapse
	Lung adenocarcinoma	Respiratory failure
	Mesothelioma malignant	Small cell lung cancer
	Pneumonia	
	Scleroderma renal crisis/thrombotic microangiopathy	
Placebo	Cardiac arrest	Cardiac arrest
	Acute myocardial infarction	Lung neoplasm malignant
	Interstitial lung disease	Septic shock
	Pneumonia	Sudden death
	Dyspnea	

Treatment Period is between the first dose of study drug until the last dose plus 28 days.
Post-Treatment Period is 29 days and over, after last drug intake.
Adapted from Study 1199.214 CSR, Table 12.1.1:1 and Table 12.2.1:2

Overall, the types and frequencies of AEs leading to death appear to be balanced by treatment group in the treatment-emergent and post-treatment periods. AEs leading to death by adjudicated cause were similar between treatment groups.

3.6.3 Serious Adverse Events

Serious adverse events (SAEs) were reported by 69 patients (24%) in the nintedanib and 62 patients (21.5%) in the placebo group. SAEs during the 52-week treatment period are summarized in Table 20. SAEs by PT (>1% of patients in either treatment group) that were more frequent in the nintedanib group than in the placebo group were pneumonia (2.8% vs. 0.3%), interstitial lung disease (2.4% vs. 1.7%), pulmonary arterial hypertension (1.0% vs. 0), and acute kidney injury (1.0% vs. 0.3%). Except for pneumonia, the differences between groups were generally due to small numbers of patients. SAEs within the Gastrointestinal Disorders system organ class (SOC) occurred in 11 nintedanib-treated patients (3.8%) on as compared to 5 placebo-treated patients (1.7%). Two patients in each treatment group had SAEs of diarrhea, while 2 patients in the nintedanib group had intestinal pseudo-obstruction, and 2 patients in the placebo group had vomiting. Other SAEs were singular by PT.

There were 6 patients with SAEs within the hepatobiliary disorders SOC, including 1 patient in each group with drug-induced liver injury, and one nintedanib patient each with hepatocellular injury and liver disorder. All patients recovered after treatment discontinuation or dose reduction. Other SAEs in the hepatobiliary disorders SOC included bile duct stone and cholecystitis, each in 1 nintedanib treated patient.

Serious cardiovascular events were similar between the treatment groups. In the nintedanib group there was 1 patient (0.3%) with MACE of arrhythmia, and in the placebo group there were 3 patients (1%) with MACE events of acute myocardial infarction, cardiac arrest, and cerebral infarction.

In the first 52 weeks of the study, in the nintedanib group there were 3 patients (1%) with acute kidney injury and 1 patient in the placebo group (0.3%). In the nintedanib group, one patient had diarrhea, hypovolemia, and hypotension which led to syncope. On hospitalization, a blood creatinine value of 2.4 mg/dL was noted, and acute kidney injury was diagnosed. The patient improved with fluid replacement. In another patient in the nintedanib group, the patient was on treatment with nintedanib for about 2 months. Nintedanib was discontinued due to diarrhea. During the subsequent weeks, the patient was reported with ANCA positive vasculitis and acute kidney injury. The third report of acute kidney injury in the nintedanib group experienced a subarachnoid hemorrhage as the result of an injury. The patient had diarrhea due to nintedanib, developed hypovolemia and hypotension, followed by syncope and a fall that caused the subarachnoid hemorrhage. The patient was also noted to have elevated blood creatinine and diagnosed with acute kidney injury. The patient recovered from subarachnoid hemorrhage and acute kidney injury.

Table 20: Serious Adverse Events by Preferred Term for > 1% of Patients in Either Treatment Group Over 52 Weeks

	Placebo N=288 (n, %)	Nintedanib N=288 (n, %)
Patients with ≥1 SAE	62 (21.5)	69 (24)
Interstitial lung disease	5 (1.7)	7 (2.4)
Pneumonia	1 (0.3)	8 (2.8)
Pulmonary hypertension	4 (1.4)	4 (1.4)
Dyspnea	5 (1.7)	3 (1.0)
Pulmonary fibrosis	4 (1.4)	3 (1.0)
Systemic sclerosis pulmonary	3 (1.0)	2 (0.7)
Acute kidney injury	1 (0.3)	3 (1.0)
Pulmonary arterial hypertension	0	3 (1.0)

Adapted from Study 1199.214 CSR, Table 15.3.1.1.1: 17

Aside from pneumonia, the types and frequencies of SAEs appear to be balanced by treatment group in the treatment-emergent period.

3.6.4 Dropouts and/or Discontinuations Due to Adverse Effects

Over 52 weeks, there were adverse events leading to dose decrease in 98 (34%) of patients in the nintedanib group and 10 (3%) of patients in the placebo group.

Over 52 weeks, there were adverse events leading to drug discontinuation in 46 (16%) of patients in the nintedanib group and 25 (9%) of patients in the placebo group.

In the nintedanib group, the most common reasons for decrease of dose was diarrhea (22%), nausea (2%), vomiting (2%), and elevation in alanine aminotransferase (1.4%).
The most common reasons for drug discontinuation in the nintedanib group was diarrhea (7%), nausea (2%), and vomiting (1.4%).

In the placebo group, the most common reasons for drug discontinuation was interstitial lung disease in 3 (1%) of patients. All of the other reasons for drug discontinuation in the placebo group were only seen in 1 patient each. The most common reason for decrease of dose in the placebo group was diarrhea in 3 patients (1%).

The main reason for treatment interruptions were AEs (79% of dose interruptions in the nintedanib group and 69% in the placebo group), in particular diarrhea (41% nintedanib, 19% placebo) and upper abdominal pain (10% nintedanib, 10% placebo).

The incidence of AEs leading to dose decrease and drug discontinuation was higher in the nintedanib group than in the placebo group. The most common reason for drug dose decrease, discontinuation, and interruption was diarrhea in the nintedanib group, consistent with the known safety profile of nintedanib.

3.6.5 Severe Adverse Events

During the 52 weeks treatment period, there were 52 (18%) patients with severe AEs in the nintedanib group and 36 (13%) patients with severe AEs in the placebo group. The most common severe AEs in the nintedanib group were: diarrhea, pneumonia, upper abdominal pain, and vomiting. The most common severe AEs in the placebo group were: diarrhea, dyspnea, and interstitial lung disease. The severe adverse events are consistent with the known safety profile of nintedanib.

3.6.6 Treatment Emergent Adverse Events and Adverse Reactions

The analysis of AEs was based on the concept of treatment-emergent AEs: all AEs with an onset after the first dose of study medication up to the end of the residual effect period (28 days) were considered on-treatment. There were 283 (98%) of patients in the nintedanib group and 276 (96%) of patients in the placebo group who had TEAEs. The most common TEAE by PT for either group was diarrhea 76% in the nintedanib group and 32% in the placebo group. Next was nausea in 32% of the nintedanib group and 14% of the placebo group. Vomiting was seen in 25% of the nintedanib group and 10% of the placebo group. Abdominal pain, decreased appetite, and weight decrease were also more common in the nintedanib group. There were more investigator defined drug related AEs in the nintedanib group (83%) than the placebo group (43%), with diarrhea being the most common. The TEAEs were consistent with the known safety profile of nintedanib.

3.6.7 Safety Review Approach

The assessment of safety is based on data from the single pivotal study, Study 1199.214, conducted in 576 patients with SSc-ILD, randomized 1:1 to treatment with nintedanib or placebo. Overall, the size of the safety database is adequate to assess the safety of nintedanib in SSc-ILD, in the context of the known safety profile in IPF.

All safety analyses were performed on the Treated Set (TS). The TS consisted of patients who were randomized and received at least one dose of study medication. The primary safety analysis was based on events occurring within the first 52 weeks of treatment (through Day 373). Additional supportive safety analyses were conducted for the whole study period, including the residual effect period, defined as 28 days after the last dose of study drug.

Categorization of Adverse Events

The definitions used for adverse events (AEs) and serious AEs (SAEs) were per 21CFR 312.32. All AEs that occurred between the first dose of study drug until the last dose plus 28 days were considered 'treatment-emergent.' AEs that started before the first dose of study drug and worsened during the treatment period were also considered as 'treatment-emergent.' Adverse events occurring between the start of an interruption of study drug and the end of the interruption of study drug were considered 'off-treatment.'

Adverse events that occurred after the last dose of study drug plus 28 days were assigned to 'follow-up' or 'post-study.' The post treatment period was defined by the Applicant as follow-up period (occurring between last drug intake +29 days and beginning of post-study period) and post-study period [occurring on or after last drug intake +29 days, date of completion +1 day, or date of informed consent in extension (whichever was latest)].

Adverse events of special interest were AEs relating to gastrointestinal perforation and hepatic injury. In addition to the specified adverse events of special interest, the study statistical analysis plan (TSAP) specified Gastrointestinal AEs (diarrhea, nausea, vomiting, dehydration, weight decrease, and decreased appetite) as AEs of particular note. The intensity of AEs was categorized as mild (awareness of signs or symptoms which are easily tolerated), moderate (enough discomfort to cause interference with usual activity), or severe (incapacitating or causing inability to work or to perform usual activities). In addition, AEs of diarrhea were classified according to the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.

An independent adjudication committee reviewed all deaths to adjudicate cause to cardiovascular death, respiratory related death, non-cardiovascular/non-respiratory death, or undetermined cause of death. The adjudication committee also reviewed all adverse events categorized as major adverse cardiovascular events (MACE).

In Study 1199.214, a 4-point MACE definition was used, including any fatal event in system organ class (SOC) of cardiac disorders, any fatal event in SOC vascular disorders, any fatal or nonfatal event in SMQ 'myocardial infarction (broad)', and 'any fatal or nonfatal stroke'.

3.6.8 Submission Specific Safety Concerns

The current prescribing information for nintedanib contains warnings and precautions regarding risks of the following: elevated liver enzymes and drug-induced liver injury, diarrhea, nausea, vomiting, arterial thromboembolic events, bleeding events, and GI perforation.[35]
Significant AEs over 52 weeks are shown in Table 21.

Elevated liver enzymes and drug-induced liver injury

- See discussion below

Diarrhea, nausea, vomiting

- The nintedanib group had 76% of patients with diarrhea compared to 32% of placebo patients.
- The nintedanib group had 32% of patients with nausea compared to 14% for placebo.
- The nintedanib group had 25% of patients with vomiting compared to 10% for placebo.

Arterial thromboembolic events

- There were 2 patients (0.7%) each in both the nintedanib and placebo groups with arterial thromboembolic events.

Bleeding events

- The most common bleeding events in both groups were epistaxis and skin contusion. In addition, the nintedanib group had 5 patients (2%) with rectal hemorrhage and two (0.7%) with hematochezia. There was also 1 patient in the nintedanib group who had cerebral amyloid angiopathy with cerebral microhemorrhage and 1 patient who had diarrhea due to nintedanib, developed hypovolemia and hypotension, followed by syncope and a fall that caused subarachnoid hemorrhage.

GI perforation

- While there were no GI perforations in the nintedanib group over the first 52 weeks, one patient in the nintedanib group had a large intestine perforation (sigmoid, rectosigmoid perforation) in the post-treatment period, 42 days following discontinuation of treatment. The AE was complicated by a diagnosis of anti-neutrophil cytoplasmic antibody positive vasculitis, hospitalization in the intensive care unit, enterococcus faecium infection, and clostridium difficile infection. The AE of GI perforation was assessed by the investigator as not related to the study drug.
- There was 1 patient in the placebo group with an SMQ of GI perforation, but the event was an anal abscess.

The nintedanib group had more: diarrhea, elevated liver enzymes, nausea, vomiting, and bleeding events over placebo.

In addition, the adverse events of special interest (AESI) in the 1199.214 study were AEs relating to gastrointestinal perforation and hepatic injury.

[35] FDA-approved nintedanib labeling

Over 52 weeks, there were 38 (13%) patients in the nintedanib group and 9 (3%) patients in the placebo group who had elevated liver enzymes. Ten patients (4%) in the nintedanib group had ALT ≥3 times upper limit of normal (ULN).
There was 1 patient (0.3%) in both the nintedanib and placebo groups with drug induced liver injury (DILI).

Over 52 weeks, 20.5% of patients in nintedanib group lost > 10% of their body weight at some point during the first 52 weeks of treatment vs 4.5% of the placebo group. In the nintedanib group, 47% of patients had weight loss of at least 5% of their body weight vs 19% in the placebo group.

Table 21: Study 1199.214: Significant AEs Over 52 Weeks

	Placebo N=288 (n, %)	**Nintedanib N=288 (n, %)**
Elevated liver enzymes	9 (3)	38 (13)
DILI	1 (0.3)	1 (0.3)
Diarrhea	91 (32)	218 (76)
Nausea	39 (14)	91 (32)
Vomiting	30 (10)	71 (25)
Arterial thromboembolic events	2 (0.7)	2 (0.7)
Bleeding	24 (8)	32 (11)
GI Perforation	1 (0.3)*	0

Adapted from Study 1199.214 CSR, Table: 15.3.1.1.1: 5 and Table: 15.3.1.1.5: 1
DILI: drug induced liver injury
GI: Gastrointestinal
*GI perforation based on SMQ analysis, PT anal abscess

3.6.9 Safety Summary

The most common AEs in the nintedanib group were consistent with those known for nintedanib including: diarrhea, nausea, vomiting, and liver abnormalities. Overall, the deaths and SAEs were balanced between the treatment groups. During the first 52 weeks, the most frequent SAEs were ILD (12 patients overall), pneumonia (9 patients overall), pulmonary hypertension (8 patients overall), and dyspnea (8 patients overall). There were more AEs leading to drug decrease and discontinuation in the nintedanib group which were mostly from gastrointestinal complaints. In addition, patients in the nintedanib group lost more weight than the placebo group. In this study, patients with clinical signs of malabsorption were excluded. In the patient population of systemic sclerosis, malabsorption can be part of the disease process. Gastrointestinal complaints and weight loss are important considerations in this patient population. In addition, in the nintedanib IPF program, patients with low body weight (<65 kg), and female patients had a higher risk of elevations in liver enzymes. Overall, the TEAEs are

consistent with the known safety profile of nintedanib and no new adverse drug reactions were identified for nintedanib in SSc-associated ILD.

3.7 Benefit/Risk Considerations

Systemic sclerosis interstitial lung disease (SSc-ILD) is a serious condition associated with increased morbidity and mortality. There are no FDA-approved therapies for SSc or SSc-ILD. Further, there are no established endpoints for clinical studies in SSc or SSc-ILD. Currently, patients with SSc-ILD are treated with off-label products used for other rheumatic diseases, including cyclophosphamide and mycophenolate mofetil. These drugs are associated with significant potential toxicities including infections, malignancies, and cytopenias. There is a high unmet need for additional therapeutic options for SSc-ILD.

The nintedanib clinical development program for SSc-ILD consists of a single, double blind, randomized, placebo-controlled study to evaluate the efficacy and safety of oral nintedanib in 576 patients with SSc-ILD. The primary endpoint was the annual rate of decline in FVC in mL over 52 weeks. Key secondary endpoints were: absolute change in modified Rodnan Skin Score at Week 52 and absolute change in Saint George's Respiratory Questionnaire at Week 52. Additional secondary endpoints included time to death, HAQ-DI, CRISS index score, and FACIT dyspnea scale.

The adjusted annual rate of decline in FVC over 52 weeks was lower in the nintedanib group (-52.4 mL/year) than in the placebo group (-93.3 mL/year), with a treatment difference of 40.9 mL/year. A less robust treatment effect was observed in adjusted annual rate of decline in FVC in the subgroups of patients on mycophenolate mofetil at baseline (treatment difference 26.6 mL/year) and patients from the US and Canada (treatment difference 10.2 mL/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, DL_{CO}, 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/activities of daily living as assessed by the HAQ-DI. Mortality was also similar between treatment groups.

The efficacy results from Study 1199.214 indicate a modest observed treatment difference in the primary endpoint, the adjusted annual rate of decline in FVC over 52 weeks, between nintedanib and placebo, without observed benefit in the key secondary and other secondary endpoints during the 52-week comparisons in study.

For the interpretation of the results from Study 1199.214, it is important to consider additional contextual information summarized below.

While FVC is a surrogate endpoint that does not directly measure how a patient feels, functions, or survives, it has been demonstrated to reliably predict clinical benefit in IPF, a related condition. The clinical benefit from altering the rate of decline in lung function in patients with

IPF, as measured by FVC over 52 weeks, has been shown to be consistent in two larger clinical programs, using two different products with different mechanisms of action, nintedanib and pirfenidone.[36,37]

For example, in the nintedanib IPF program, the treatment difference (nintedanib vs. placebo) in the three clinical studies (Studies 1, 2, and 3 in the FDA-approved nintedanib labeling) ranged from 94 to 131 mL/year. The change in FVC was supported by statistically significant decreases in IPF exacerbations and improvement in SGRQ scores, in 2 of the 3 studies. Although not powered for survival, a numerical trend favoring nintedanib was seen for survival in both pre-specified and sensitivity analyses. SSc-ILD, as a disease process, may be sufficiently different from IPF such that a direct comparison between FVC changes in IPF patients may not be comparable or appropriate to FVC changes in SSc-ILD patients. However, the nintedanib effect on the relative slowing of the rate of decline in FVC was consistent between the nintedanib SSc-ILD and IPF programs. To what extent the treatment effect in IPF can be relied upon to support the modest effect observed in SSc-ILD in Study 119.214 is an important question for the Committee's consideration.

The observed FVC changes in other studies in SSc and SSc-ILD provide additional context for the discussion of the results from Study 1199.214. For example, the Scleroderma Lung Study (SLS) was a double-blind, randomized, placebo-controlled study of oral cyclophosphamide (CYC) treatment in 158 patients with active SSc-ILD. The primary endpoint was the adjusted percent predicted FVC change at 12 months.[38] The mean absolute difference in adjusted 12-month percent predicted FVC between the cyclophosphamide and placebo groups was 2.53%. The observed treatment effect of cyclophosphamide on changes in lung function were supported by improvement in the transitional dyspnea index and HAQ-DI, further supporting clinical meaningfulness of slowing the rate of decline in FVC.

The safety of nintedanib in Study 1199.214 was generally consistent with the known safety profile of nintedanib, which includes risks of liver toxicity and GI disorders. Deaths were balanced by number and adjudicated cause (CV death, respiratory death, non-CV/non-respiratory death) in both treatment groups. Aside from pneumonia, the types and frequencies of SAEs appear to be balanced by treatment group in the treatment-emergent period. Nintedanib patients more frequently experienced AEs leading to drug discontinuation, and AEs leading to dose decreases, largely driven by gastrointestinal AEs, consistent with the labeled warnings and precautions for nintedanib. In addition, AEs of weight loss were reported more frequently in the nintedanib treatment group and reported more frequently in the SSc-ILD patients as compared to the IPF patients.

Overall, the adjusted difference in the annual rate of decline in FVC in Study 1199.214 was modest and not supported by other assessments of pulmonary symptoms or function, or assessments of other disease manifestations. The subgroup of patients on MMF at baseline in both the nintedanib and placebo groups had a smaller treatment benefit over 52 weeks. The clinical meaningfulness of the observed reduction in the rate of decline in FVC with nintedanib

[36] FDA-approved nintedanib labeling
[37] FDA-approved pirfenidone labeling
[38] Tashkin DP et al. Cyclophosphamide versus Placebo in Scleroderma Lung Disease. N Engl J Med 2006;354:2655-66.

treatment in patients with SSc-ILD, in the absence of improvement in other efficacy endpoints, and the overall risk/benefit for its use in SSc-ILD, a 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.

4 Appendix

4.1 Schedule of Events

Visit		1	2	3	4	5	6	6a	7	7a	8	8a	9	Xa[18] (9a, 10a, 11a)	X[18] (10, 11)	12/ EOT	FU[19]
	Screening		Randomised Treatment Period*														FU
Weeks of treatment			0	2	4	6	12	18	24	30	36	44	52	60 + every 16wk	68 + every 16wk	100	EOT + 4wk
Day Time window	Before or at the latest at Visit 1	≥4d prior Visit 2	1	15 ±3	29 ±3	43 ±3	85 ±3	127 ±7	169 ±7	211 ±7	253 ±7	309 ±7	365 ±7	±7	±7	701 ±7	+28 +7
Informed Consent[1]	X																
Send HRCT to central review[2]	X																
Demographics		X															
Medical history		X	X														
Adverse events, conc. therapy		X	X	X	X	X	X		X		X		X		X	X	X
In-/exclusion criteria		X	X														
Questionnaires (SGRQ, FACIT-dyspnoea, SHAQ, EQ-5D-5L, patient global VAS)[3]			X						X				X			X	
Review questionnaires for completeness			X						X				X			X	
Physical examination, vital signs		X	X	X	X	X	X		X		X		X		X	X	X
Collect/review/dispense menstruation calendar	X		X	X	X	X	X		X		X		X		X	X	X

Visit		1	2	3	4	5	6	6a	7	7a	8	8a	9	Xa[18] (9a, 10a, 11a)	X[18] (10, 11)	12/ EOT	FU[19]
	Screening		Randomised Treatment Period*														FU
Weeks of treatment			0	2	4	6	12	18	24	30	36	44	52	60 + every 16wk	68 + every 16wk	100	EOT + 4wk
Day Time window	Before or at the latest at Visit 1	≥4d prior Visit 2	1	15 ±3	29 ±3	43 ±3	85 ±3	127 ±7	169 ±7	211 ±7	253 ±7	309 ±7	365 ±7	±7	±7	701 ±7	+28 +7
HCRU			X	X	X	X	X		X		X		X		X	X	
mRSS assessment			X				X		X		X		X		X	X	
Digital ulcer assessment			X				X		X		X		X		X	X	
Safety Laboratory (blood and urine)		X[4]	X	X	X	X	X	X[5]	X	X[5]	X	X[5]	X	X[5]	X	X	X
Pregnancy test[6]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PK sample[7]					X				X								
Autoantibody assessment[8]		(X)	X						X				X				
Biomarker samples[9]			X		X				X				X				
DNA banking sample[10]			X														
SpO_2 (earlobe or forehead, resting)			X						X				X			X	
Spirometry (FVC)[11]		X	X	X	X	X	X		X		X		X		X	X	X
DLCO[11]		X	X						X				X			X	
12-lead ECG		X	X[12]						X				X			X	
Echocardiography[13]		X											X				

Visit		1	2	3	4	5	6	6a	7	7a	8	8a	9	Xa[18] (9a, 10a, 11a)	X[18] (10, 11)	12/ EOT	FU[19]
	Screening		Randomised Treatment Period*														FU
Weeks of treatment			0	2	4	6	12	18	24	30	36	44	52	60 + every 16wk	68 + every 16wk	100	EOT + 4wk
Day Time window	Before or at the latest at Visit 1	≥4d prior Visit 2	1	15 ±3	29 ±3	43 ±3	85 ±3	127 ±7	169 ±7	211 ±7	253 ±7	309 ±7	365 ±7	±7	±7	701 ±7	+28 +7
Randomization			X														
IRT call/notification	X[14]		X		X		X		X		X		X		X	(X)	
Dispense trial medication			X		X		X		X		X		X		X		
Collect trial medication					X		X		X		X		X		X	X	
Compliance/drug accountability				X[20]	X	X[20]	X		X		X		X		X	X	
Physician global VAS			X						X				X			X	
Termination of trial medication[15]																X	
Vital status assessment[16]													X			X	
Conclude subject participation[17]																	X

Footnotes:

* In case of dose change (reduction or re-escalation) additional visits have to be included (refer to Section 6.2.4).
- In case of prematurely trial medication discontinuation, the patient completes end of treatment Visit (EOT and follow-up Visit (FU) 4 weeks later), the patient should be asked to come to future visits as planned (refer to Section 6.2.3).
- EOT assessments are the same as described for Visit 12. If EOT is performed at week 52, the Flowchart for Visit 9 is valid (i.e. Biomarker samples to be drawn).

[1] Before or at the latest at Visit 1. Informed consent (IC) needs to be signed before any procedure related to the trial is performed. All adverse events (AEs) and concomitant therapies (CTs) from the day of signing informed consent have to be recorded. The screening period (informed consent to Visit 2) must not be longer than 12 weeks.

[2] Review of high resolution computer tomography (HRCT) for extent of fibrotic disease in the lung (10% or more). Central review: a historical HRCT not older than 12 months should be sent; only if the patient does not have a HRCT within 12 months at Visit 1 but meets all other inclusion and no exclusion criteria, the HRCT can be performed for the purposes of participation in the trial (except for patients in Germany).

[3] Self-reported outcome questionnaires must always be done by the patients in a quiet place prior to any other visit procedure. Order of questionnaires: 1. SGRQ, 2. FACIT-dyspnoea, 3. SHAQ, 4. EQ-5D-5L, 5. Patient's global VAS.

[4] The safety lab of Visit 1 must be repeated if screening is longer than 6 weeks.

[5] Intermediate lab tests (a-Visit) do not necessarily need to be a site visit. Cautionary note: dependent on concomitant treatment additional safety monitoring should be considered at discretion of the investigator.

[6] β-HCG will be performed at Visit 2 only, at central lab. Urine dipstick pregnancy tests will be provided centrally and should be performed in all women of childbearing potential every 4-6 weeks: at least at every visit and if necessary, additionally at home or at a local doctor / laboratory. If urine test is not acceptable to local authorities, a blood test can be done at a local laboratory. Women of childbearing potential will be instructed accordingly.

[7] PK samples will be taken at Visits 4 and 7 just before drug administration. Date and exact clock time of drug administration and blood sampling must be recorded on the eCRF. patients will be provided (Visits 3 and 6) with a PK card to support the record of the exact clock time of medication intake three days preceding PK sampling.

[8] anti-Topoisomerase antibodies (ATA) will be assessed at Visit 2, 7, and 9 (and at V1 if historically not available); anti-RNA polymerase III antibodies ((anti-)RNA Pol III) and anti-centromere antibodies (ACA) will be assessed at Visit 2 only.

[9] Biomarker samples will be taken just before drug administration. Date and exact clock time of drug administration and blood sampling must be recorded on the eCRF.
- Samples for Protein Biomarkers will be taken at Visits 2, 4, 7, 9, just before drug administration.
- One sample for prespecified DNA analyses will be taken at Visit 2.
- Samples for RNA expression analyses will be taken at Visits 2, 7, and 9, just before drug administration.

[10] DNA (Desoxyribo Nucleid Acid) banking sample: one blood sample will be taken from those eligible patients who signed a separate informed consent at Visit 2 (or on a subsequent visit); Participation is voluntary and is no prerequisite for participation in the trial.

[11] Order of lung function measurements: same time each visit ± 90 min, reference time at Visit 2: 1. FVC followed by patients rest; 2. DLCO.

[12] ECG will be performed (if possible prior to blood draw) at Visit 2 prior randomisation (only if abnormal at Visit 1).

[13] Echocardiography will at least be performed in patients with a history of pulmonary hypertension at screening (time window Visit 1 to Visit 2) and after 1 year (time window Visit 9 to Visit 9a).

[14] IRT needs to be notified at time point of informed consent (at the latest at Visit 1) to trigger trial medication shipments; ATA status (historical) will be entered at randomisation (Visit 2) at the latest.

[15] Termination of trial medication data needs to be collected any time trial medication is permanently discontinued.

[16] Vital status at 52 weeks and at 100 weeks or at the timepoint when patient's last full visit (i.e. EOT or V9, V10, V11, V12) would have been scheduled, whatever occurs earlier should be available for all patients. Permission to contact withdrawn patients for vital status assessment should be requested by site.

[17] Trial completion:
- At the end of the follow-up Visit for patients who have completed the trial on treatment as planned.
- After early discontinuation (end of treatment [EOT]) and follow-up Visit), if a patient refuses to attend future visits as originally planned.
- At the end of Visit 12 or at the global end of the trial for patients who discontinued trial medication early but came to future visits as planned.

[18] Same scheme should be repeated as often as needed: Visit 'X' stands for Visit 10 and Visit 11; Visit 'Xa' stands for Visits 9a, 10a, and 11a.

[19] The follow-up (FU) visit should be planned for 28 days (+7 days window) after last drug intake (end of treatment [EOT]).

[20] Compliance / drug accountability only in case of dose reduction/increase

Source: Study 1199.214 protocol

4.2 Recommendations for Treatment Interruption or Dose Reduction

Table 22: Allowed Treatment Reduction or Interruption Periods of Nintedanib

	AEs considered drug-related	AEs not considered drug-related
Maximum interruption period	4 weeks	8 weeks
Recommended re-start	with reduced dose (100 mg bid)	with the same dose (100 mg bid or 150 mg bid)
Re-escalation	within 4 weeks to 150 mg bid	not applicable

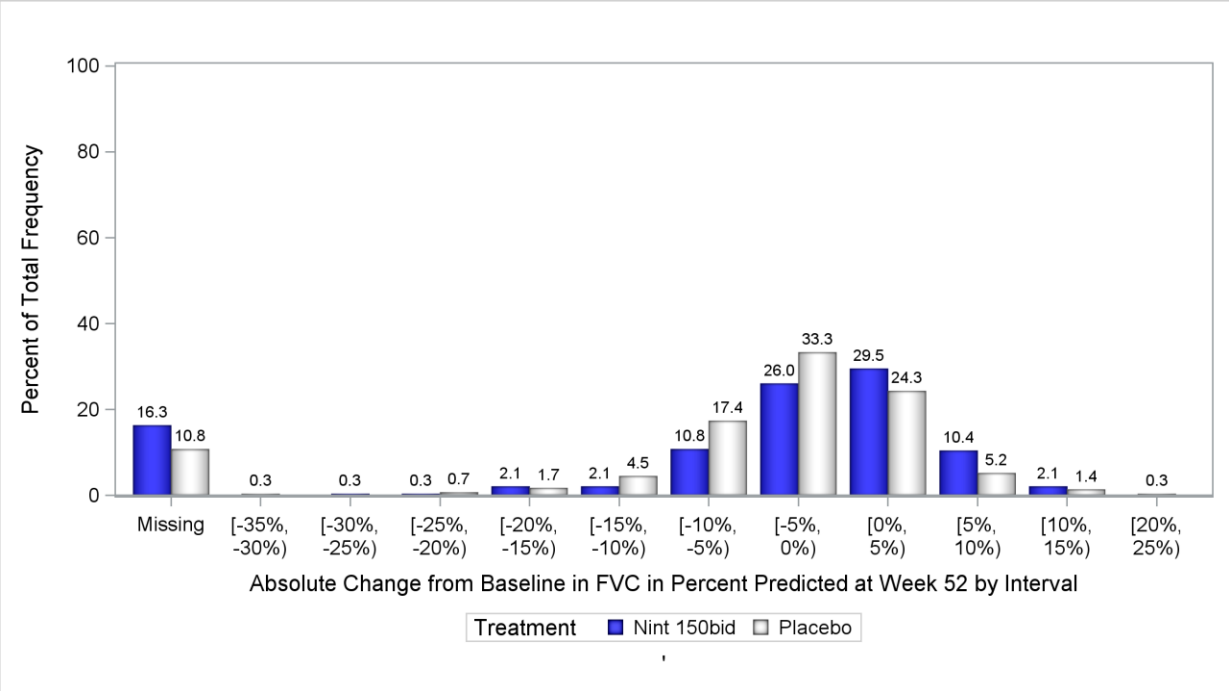
Source: CSR 199.214 Table 9.4.2.1:1

Table 23: Recommendations for Managing Liver Enzyme Elevations

Time point	AST or ALT increase to			Signs of hepatic injury[1]
	>1.5x to <3x ULN	≥3x to <5x ULN and no signs of hepatic injury[1]	≥5x to <8x ULN and no signs of hepatic injury[1]	
Visit 2 (randomisation)	Withdrawal of trial medication or continuation to be justified[2]	Withdrawal of trial medication	Withdrawal of trial medication	Withdrawal of trial medication
Any other Visit	Continuation as planned[3]	Dose reduction or interruption[4]	Interruption of trial medication	Withdrawal of trial medication
		Close observation[5] After 2 weeks or any time later	Close observation[5] After 2 weeks or any time later	Clinical evaluation of hepatic injuriy[1]
	<3x ULN	≥3x ULN	<3x ULN	≥3x ULN
	Dose reduced: return to initial dose Interrupted: restart at reduced dose. Bi-weekly monitoring for at least 8 weeks	Permanent discontinuation Close observation[5]	Restart at reduced dose Weekly monitoring for 4 weeks, then bi-weekly for at least 8 weeks	Permanent discontinuation Close observation[5]

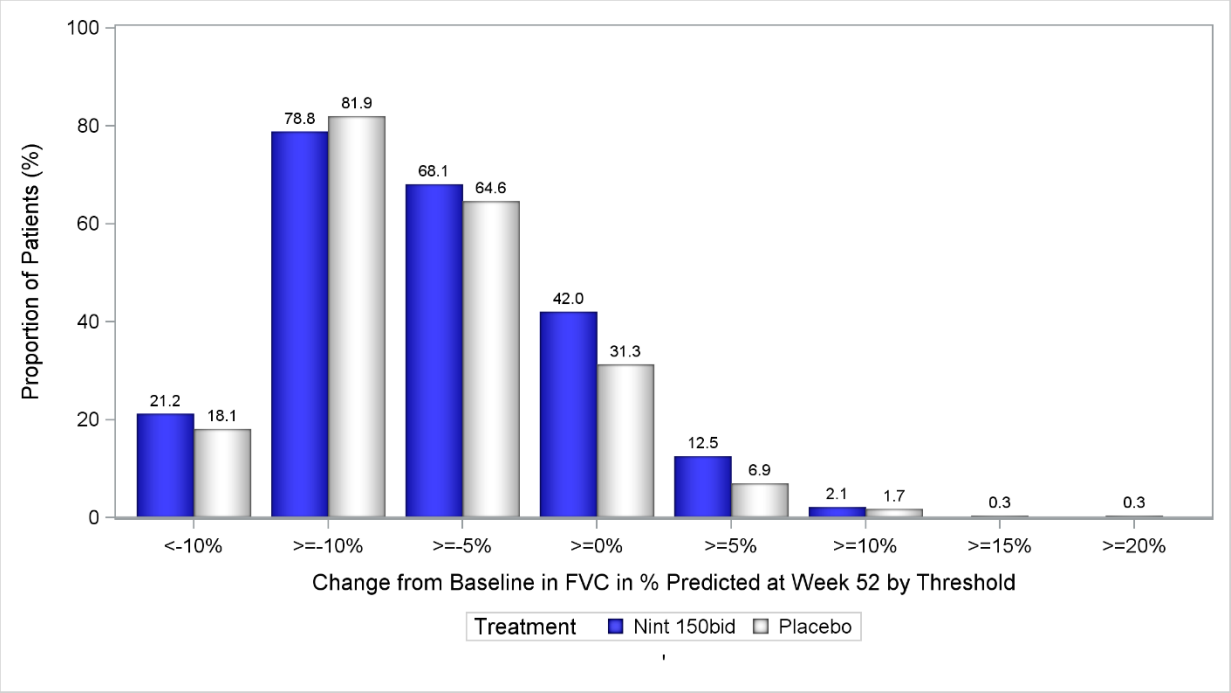
4.3 Responder Analysis on FVC in Percent Predicted

Figure 12. Histogram of Absolute Change from Baseline in FVC in Percent Predicted at Week 52 (Treated Set)



Abbreviations: FVC: forced vital capacity.
Source: FDA Statistical Reviewer

Figure 13: Proportions of Responders at Various Cutoffs among Treated Patients, Based on Analysis of Change from Baseline in FVC in Percent Predicted at Week 52 (Treated Set)



Abbreviations: FVC: forced vital capacity.
Note: Missing data were considered as having an absolute decline in FVC in percent predicted of >10%.
Source: FDA Statistical Reviewer