Decline in forced vital capacity as a surrogate for mortality in patients with pulmonary fibrosis

Toby M. Maher\textsuperscript{1,2} | Susanne Stowasser\textsuperscript{3} | Florian Voss\textsuperscript{4} | Elisabeth Bendstrup\textsuperscript{5} | Michael Kreuter\textsuperscript{6} | Fernando J. Martinez\textsuperscript{7} | Patricia J. Sime\textsuperscript{8} | Christian Stock\textsuperscript{4}

\textsuperscript{1}Keck School of Medicine, University of Southern California, Los Angeles, California, USA
\textsuperscript{2}Imperial College London, London, UK
\textsuperscript{3}Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany
\textsuperscript{4}Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim am Rhein, Germany
\textsuperscript{5}Centre for Rare Lung Diseases, Department of Respiratory Diseases and Allergy and Department of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark
\textsuperscript{6}Center for Interstitial and Rare Lung Diseases, Pneumology and Respiratory Care Medicine, Thoraxklinik, University of Heidelberg, Member of the German Center for Lung Research, Heidelberg, Germany
\textsuperscript{7}Weill Cornell Medicine, New York, New York, USA
\textsuperscript{8}Department of Internal Medicine, Virginia Commonwealth University, Richmond, Virginia, USA

Correspondence
Toby M. Maher
Email: toby.maher@med.usc.edu

Funding information
Boehringer Ingelheim International GmbH

Associate Editor: Michael P Keane; Senior Editors: Phil Bardin and Paul Reynolds

Abstract

Background and Objective: Surrogate endpoints enable determination of meaningful treatment effects more efficiently than applying the endpoint of ultimate interest. We used data from trials of nintedanib in subjects with pulmonary fibrosis to assess decline in forced vital capacity (FVC) as a surrogate for mortality.

Methods: Data from the nintedanib and placebo groups of trials in subjects with idiopathic pulmonary fibrosis, other forms of progressive pulmonary fibrosis, and pulmonary fibrosis due to systemic sclerosis (NCT00514683, NCT01335464, NCT01335477, NCT01979952, NCT02999178, NCT02597933) were pooled. Using joint models for longitudinal and time-to-event data, we assessed the association between decline in FVC % predicted and time to death over 52 weeks. The rate of change in FVC % predicted and the current value of FVC % predicted were modelled longitudinally and estimates applied as predictors in time-to-event models.

Results: Among 2583 subjects with pulmonary fibrosis, both a greater rate of decline in FVC % predicted and a lower current value of FVC % predicted were associated with an increased risk of death over 52 weeks (HR 1.79 [95% CI: 1.57, 2.03] and HR 1.24 [1.17, 1.32] per 5-percentage point decrease, respectively). Associations between the rate of change in FVC % predicted and the risk of death were consistent between patients with IPF and other ILDs.

Conclusion: Data from clinical trials in subjects with pulmonary fibrosis of diverse aetiology demonstrate a strong association between decline in FVC % predicted and mortality over 52 weeks, supporting FVC decline as a surrogate for mortality in these patients.

KEYWORDS
clinical trial, interstitial lung disease, pulmonary function tests, vital capacity

INTRODUCTION

The use of surrogate endpoints enables the determination of meaningful treatment effects more efficiently than applying the endpoint of ultimate interest.\textsuperscript{1} The dismal prognosis of progressive pulmonary fibrosis\textsuperscript{2–5} means that mortality is an outcome of particular importance. A decline in forced vital capacity (FVC) has been associated with mortality in patients with idiopathic pulmonary fibrosis (IPF)\textsuperscript{6–8} and other forms of pulmonary fibrosis.\textsuperscript{3,8–11} Change in FVC has been accepted as a surrogate for mortality by authorities.\textsuperscript{12,13} However, further data are needed to establish FVC decline as a surrogate for mortality in patients with pulmonary fibrosis, particularly in populations other than IPF, and to quantify the relationship between FVC decline and mortality.

Joint modelling enables analysis of longitudinal and time-to-event data in a single model.\textsuperscript{14,15} These models are often used to determine the effect of a time-dependent
marker on the time to occurrence of clinically meaningful events. We applied a joint modelling approach to data from clinical trials of nintedanib to assess the patient-level association between longitudinal changes in FVC and mortality (and composite outcomes including mortality). Following a widely accepted three-step framework for surrogate endpoint validation, the present analysis primarily addresses level 2, which is established when a strong association is demonstrated between the surrogate and the final endpoint across cohorts or at patient level. These analyses add to the evidence supporting FVC decline as a measure of progression of pulmonary fibrosis in clinical practice, and have implications for the design and interpretation of clinical trials in patients with progressive pulmonary fibrosis.

METHODS

Subjects

Data were pooled from subjects who participated in the placebo-controlled, double-blind periods of the TOMORROW (NCT00514683), INPULSIS-1 and -2 (NCT01335464 and NCT01335477), or Phase III trial (NCT01979952) of nintedanib in subjects with IPF, the INBUILD trial (NCT02999178) in subjects with other forms of progressive pulmonary fibrosis, and the SENSCIS trial (NCT02597933) in subjects with pulmonary fibrosis due to systemic sclerosis (SSc). The designs of these trials have been published. In each trial, subjects were randomized to receive nintedanib or placebo. Acute exacerbations and hospitalizations were reported by investigators as adverse events. In the trials in patients with IPF, but not in the INBUILD trial, for an acute worsening to be classified as an acute exacerbation, infection was to be excluded as a potential cause as per routine clinical practice and microbiological studies. Data on acute exacerbations were not collected in the SENSCIS trial. These trials were carried out in compliance with the protocol and with the principles of the Declaration of Helsinki and the Harmonized Tripartite Guideline for Good Clinical Practice of the International Conference on Harmonization. The trials were approved by the local authorities. All subjects provided written informed consent prior to their participation.

Statistical analysis

We used joint models for longitudinal and time-to-event data. We assessed associations between the estimated rate of change in the percentage of the predicted value for FVC (FVC % predicted) over 52 weeks and time to death, acute exacerbation or death, and hospitalization or death, over 52 weeks. We also assessed the association between the estimated current value of FVC % predicted and time to death over 52 weeks. Thus, we looked at the effect of an estimated slope in FVC % predicted and the effect of an estimated current value of FVC % predicted at any given time on the risk of the event. The slope parameterization assumed linearity and reflected the change over 52 weeks. The linearity assumption was verified by descriptive analyses and by statistical models allowing for non-linear FVC trajectories. FVC % predicted was preferred over FVC in mL as this adjusts for some prognostic variables (i.e., age, sex, race/ethnicity, height).

FVC % predicted was modelled longitudinally using linear mixed-effects models. Individual estimates of the time-dependent slope, that is, the annual rate of change in FVC % predicted or the current value of FVC % predicted, were applied as predictors in survival models. FVC data collected up to 7 days after the end of treatment were included and no imputation was performed. The longitudinal sub-model was a random intercept and slope model that assumed separate slopes for subjects receiving nintedanib or placebo and was adjusted for baseline FVC % predicted and the effects of individual studies. The time-to-event sub-model assumed a parametric (piecewise constant) baseline hazard function, was adjusted for the effects of individual studies and, depending on the analysis, included either an effect of the estimated annual rate of change in FVC % predicted or

<table>
<thead>
<tr>
<th>Table 1 Baseline characteristics of the pooled dataset.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nintedanib</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td><strong>Male, n (%)</strong></td>
</tr>
<tr>
<td><strong>Age, years, mean (SD)</strong></td>
</tr>
<tr>
<td><strong>Body mass index, kg/m², mean (SD)</strong></td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Black or African-American</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Missing</td>
</tr>
<tr>
<td><strong>FVC, mL, mean (SD)</strong></td>
</tr>
<tr>
<td><strong>FVC, % predicted, mean (SD)</strong></td>
</tr>
<tr>
<td><strong>DLco, % predicted, mean (SD)</strong></td>
</tr>
<tr>
<td><strong>SpO₂, %, mean (SD)</strong></td>
</tr>
</tbody>
</table>

*aData from subjects who selected one race.

*bCorrected for haemoglobin; n = 1305 in nintedanib group, n = 1091 in placebo group.*
the estimated current value of FVC % predicted. In sensitivity analyses, we assessed associations between the rate of change in FVC % predicted and time to death in subgroups by FVC <75% and ≥75% predicted at baseline, by diagnoses of IPF and non-IPF and by treatment group.

Subjects with ≥1 post-baseline FVC value and data on time to the respective event were included in the analysis. Analyses were implemented using the SAS® macro %JM.23

RESULTS

Subjects

The pooled dataset comprised 2583 subjects, of whom 1399 received nintedanib and 1184 received placebo; 1344 subjects (52.0%) had IPF, 663 (25.7%) had other forms of progressive pulmonary fibrosis, 576 (22.3%) had SSc-ILD. The baseline characteristics of subjects in the individual trials have been published.18–22 In the pooled dataset, most subjects were male (60.0%) and white (66.5%); mean (SD) age was 63.6 (10.9) years and FVC was 75.3 (17.7) % predicted (Table 1).

Associations between FVC % predicted and death

The difference between the nintedanib and placebo groups in the rate of change in FVC % predicted over 52 weeks was 2.84 (95% CI: 2.20, 3.49; p < 0.0001). Over 52 weeks, 4.6% of subjects in the nintedanib group and 5.5% of subjects in the placebo group died. Deaths by trial and treatment are presented in the Supporting Information Table S1. A greater rate of decline in FVC % predicted was associated with an increased risk of death over 52 weeks (HR 3.20 [95% CI: 2.48, 4.14] per 10-percentage point decrease [e.g., a change from 80% predicted to 70% predicted]; Table 2; Figure 1). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by diagnosis of IPF versus other fibrosing ILDs were consistent with findings in the overall population (HR 3.56 [95% CI: 2.54, 5.02]; Table 3; Supporting Information Table S2). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by FVC <75% and ≥75% predicted at baseline were consistent with findings in the overall population (HR 3.27 [95% CI: 2.32, 4.61] and HR 3.44 [95% CI: 2.25, 5.25] per 10-percentage point decrease in subjects with FVC <75% and ≥75% predicted, respectively; Supporting Information Table S2). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by diagnosis of IPF versus other fibrosing ILDs were consistent with findings in the overall population (HR 3.56 [95% CI: 2.54, 5.02]; Table 3; Supporting Information Table S2). The difference between the nintedanib and placebo groups in the rate of change in FVC % predicted over 52 weeks was 2.84 (95% CI: 2.20, 3.49; p < 0.0001). Over 52 weeks, 4.6% of subjects in the nintedanib group and 5.5% of subjects in the placebo group died. Deaths by trial and treatment are presented in the Supporting Information Table S1. A greater rate of decline in FVC % predicted was associated with an increased risk of death over 52 weeks (HR 3.20 [95% CI: 2.48, 4.14] per 10-percentage point decrease [e.g., a change from 80% predicted to 70% predicted]; Table 2; Figure 1). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by diagnosis of IPF versus other fibrosing ILDs were consistent with findings in the overall population (HR 3.56 [95% CI: 2.54, 5.02]; Table 3; Supporting Information Table S2). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by FVC <75% and ≥75% predicted at baseline were consistent with findings in the overall population (HR 3.27 [95% CI: 2.32, 4.61] and HR 3.44 [95% CI: 2.25, 5.25] per 10-percentage point decrease in subjects with FVC <75% and ≥75% predicted, respectively; Supporting Information Table S2). Associations between the rate of change in FVC % predicted and the risk of death in subgroups by diagnosis of IPF versus other fibrosing ILDs were consistent with findings in the overall population (HR 3.56 [95% CI: 2.54, 5.02]; Table 3; Supporting Information Table S2).
4.98] and HR 2.80 [95% CI: 1.83, 4.30] per 10-percentage point decrease in subjects with IPF and with other fibrosing ILDs, respectively; Supporting Information Table S3). A lower current value of FVC % predicted was associated with an increased risk of death over 52 weeks (HR 1.55 [95% CI: 1.37, 1.75] per 10-percentage point decrease; Table 3).

### Associations between FVC % predicted and time to acute exacerbation or death or hospitalization or death

Over 52 weeks, 8.0% of subjects in the nintedanib group and 10.9% of subjects in the placebo group had an acute exacerbation or died. Descriptive analyses of patients with acute exacerbation or death by trial and treatment are provided in the Supporting Information Table S1. A greater rate of decline in FVC % predicted was associated with an increased risk of acute exacerbation or death (HR 2.99 [95% CI: 2.32, 3.85] per 10-percentage point decrease; Table 4; Supporting Information Figure S1). Over 52 weeks, 23.4% of subjects in the nintedanib group and 22.6% of subjects in the placebo group were hospitalized or died. Descriptive analyses of patients with hospitalization or death by trial and treatment are provided in the Supporting Information Table S1. A greater rate of decline in FVC % predicted was associated with an increased risk of hospitalization or death (HR 2.28 [95% CI: 1.89, 2.75] per 10-percentage point decrease; Table 5; Supporting Information Figure S2).

### Further analyses

The joint models yielded consistent associations for all endpoints when fitted separately by treatment group (data not shown). Treatment effect estimates are not of interest for the specific objective of this study, but have general relevance for surrogacy validation. Results for death are provided in the Supporting Information Table S4.

### DISCUSSION

These analyses of data from clinical trials of nintedanib in subjects with pulmonary fibrosis demonstrate a strong association between decline in FVC and death over 52 weeks. According to the three levels of surrogacy outlined by Taylor and Elston16 and guidance provided by the International Conference on Harmonization,24 these findings, plus the biological plausibility of a link between FVC decline and mortality,25 support FVC decline as a surrogate for mortality at level 2. Demonstration of the highest level of surrogacy would require evidence from more randomized controlled trials that effects of treatment on the surrogate have commensurate effects on the final patient-related outcome.

In our analyses, an increase in the estimated rate of decline in FVC % predicted of 1 percentage point over 52 weeks increases the risk of mortality by 12% over the same period. The risk of mortality associated with this decline in FVC was of a similar magnitude in subjects with...
IPF and with other forms of pulmonary fibrosis, supporting previous studies showing that progressive fibrosing ILDs have commonalities in clinical course.\textsuperscript{2,3,11,26–28}

Data from trials of nintedanib in subjects with IPF suggest that 60%–70% of deaths had a respiratory cause.\textsuperscript{18–20} We investigated all-cause mortality, as opposed to respiratory-related mortality, as this was assessed in a consistent fashion across the trials and is the endpoint of ultimate interest. Although estimates of the risk of mortality based on the rate of decline in FVC and the current value of FVC both support FVC as a surrogate, they differ in their interpretation. For example, for a subject who lost 5 percentage points in FVC % predicted at week 52, the model based on the rate of decline in FVC, which assumes a linear decline in FVC, would estimate a 79% increase in the risk of mortality at each timepoint compared with a subject with no decline in FVC. For the same patient, the model using the current value of FVC would estimate a 24% increase in the risk of mortality at week 52. This is because this model only considers the absolute magnitude of the change at a specific time, irrespective of how rapidly the change occurred. The estimate based on the rate of decline in FVC may be seen as a better reflection of the situation in an individual subject, as it is based on subject-level trajectories, and supports the importance of regular monitoring of FVC.

The strong association between decline in FVC and death supports the need for prompt initiation of therapies that slow decline in FVC in patients with progressive pulmonary fibrosis. A meta-analysis of data from placebo-controlled clinical trials demonstrated a consistent effect of nintedanib on reducing the rate of FVC decline across progressive fibrosing ILDs.\textsuperscript{29} The use of nintedanib was associated with a reduction in mortality in analyses of pooled data from patients with IPF.\textsuperscript{30} In addition, real-world observational studies have suggested improved survival in patients with IPF who receive antibiotic therapy.\textsuperscript{31–34} A meta-analysis of data from 12,956 patients with IPF showed a relative risk of mortality of 0.55 (95% CI: 0.45, 0.66) in patients who received versus did not receive antibiotic therapy.\textsuperscript{33}

Low FVC and decline in FVC have been associated with an increased risk of hospitalization and of acute exacerbation in subjects with IPF and other fibrosing ILDs.\textsuperscript{13,35–37} In our analyses, an increase in the rate of decline in FVC % predicted of 1 percentage point over 52 weeks increased the risk of acute exacerbation or death by 12%, and the risk of hospitalization or death by 9%. An analysis of the SENSCIS trial using similar methodology found a similar association between decline in FVC % predicted and risk of hospitalization or death over 52 weeks.\textsuperscript{38}

Strengths of our analyses include the large and heterogeneous population of subjects, the standardized procedure for measuring FVC and the use of a joint modelling approach that allows for estimation of time-dependent effects that considers temporality (i.e., FVC changes precede events at the individual level). Our analyses were limited by the time frame over which FVC decline and mortality were assessed (52 weeks), the use of different definitions for an acute exacerbation across the trials, and the lack of collection of data on acute exacerbations in the SENSCIS trial.

In conclusion, data from clinical trials in over 2500 subjects with pulmonary fibrosis demonstrate a strong association between decline in FVC % predicted and mortality over 52 weeks, supporting the use of FVC decline as a surrogate for mortality in clinical trials and a measure of the progression of pulmonary fibrosis in clinical practice. A visual abstract summarizing the data in this manuscript is available at: https://www.globalmedcomms.com/respiratory/FVCDeclineAsSurrogateForMortality

**AUTHOR CONTRIBUTIONS**

Toby M. Maher: Conceptualization (supporting); writing – review and editing (supporting). Susanne Stowasser: Conceptualization (supporting); writing – review and editing (lead). Florian Voss: Conceptualization (supporting); formal analysis (equal); methodology (supporting); writing – original draft (supporting). Elisabeth Bendstrup: Conceptualization (supporting); writing – review and editing (supporting). Michael Kreuter: Conceptualization (supporting); writing – review and editing (supporting). Fernando J. Martinez: writing – review and editing (supporting). Patricia J. Sime: writing – review and editing (supporting). Christian Stock: Conceptualization (lead); formal analysis (equal); methodology (lead); visualization (lead); writing – original draft (lead).

**ACKNOWLEDGEMENTS**

We thank the subjects who participated in these trials. The trials were funded by Boehringer Ingelheim International GmbH (BI). Statistical programming support was provided by Show Reddy Maramreddy of BI and Gert van Beek of Elderbrook Solutions.

**CONFLICT OF INTEREST STATEMENT**

The trials mentioned in these analyses were funded by Boehringer Ingelheim International GmbH (BI). Julie Fleming and Wendy Morris of FleishmanHillard, London, UK, provided writing assistance, which was contracted and funded by BI. BI was given the opportunity to review the manuscript for medical and scientific accuracy and intellectual property considerations.

Toby M. Maher is a member of the Advisory Board at Respiratory and a co-author of this article. He was excluded from all editorial decision-making related to the acceptance of this article for publication. He also reports consultancy fees from AstraZeneca, Bayer, Blade Therapeutics, Boehringer Ingelheim (BI), Bristol Myers Squibb, Galapagos, Galecto, GlaxoSmithKline, IQVIA, Plant, Respirant, Roche/Genentech, Theravance, Veracyte and fees for presentations from BI and Roche/Genentech. Susanne Stowasser, Florian Voss and Christian Stock are employees of BI. Elisabeth Bendstrup reports a grant from BI; fees for presentations from BI and Chiesi; support for travel from BI and Roche; has participated on Data Safety Monitoring Boards
(DSMs) or advisory boards for AbbVie and BI. Michael Kreuter reports grants, consultancy fees and fees for presentations from BI and Roche; and holds leadership roles with Deutsche gesellschaft für Pneumologie, European Respiratory Society, German Respiratory Society. Fernando J. Martínez reports consultancy fees from AbbVie, BI, Bristol Myers Squibb, Bridge Biotherapeutics, CSL Behring, Dev-Pro, IQVIA, Genentech, Sanofi, Shionogi, twoXAR, Veracyte; fees for presentations from United Therapeutics; has served on DSMs, advisory boards or steering committees for Affirent/Merck, Bayer, Biogen, BI, Netto, Resivipant, Roche, Veracyte. Patricia J. Sime reports consultancy fees from BI, FibroGen, UCB and Three Lakes Foundation; Galecto stock; and has served on a DSM for Roche.

DATA AVAILABILITY STATEMENT
To ensure independent interpretation of clinical study results and enable authors to fulfill their role and obligations under the ICMJE criteria, Boehringer Ingelheim grants all external authors access to clinical study data pertinent to the development of the publication. In adherence with the Boehringer Ingelheim Policy on Transparency and Publication of Clinical Study Data, scientific and medical researchers can request access to clinical study data after publication of the primary manuscript in a peer-reviewed journal, regulatory activities are complete and other criteria are met. Researchers should use https://vivli.org/ to request access to study data and visit https://www.mystudywindow.com/msw/datasharing for further information.

HUMAN ETHICS APPROVAL DECLARATION
The clinical trials which providing data for this study were approved by the local authorities. All subjects provided written informed consent prior to their participation. Clinical trial registration details are available in the Section ‘Methods’ of this article.

ORCID
Toby M. Maher https://orcid.org/0000-0001-7192-9149
Elisabeth Bendstrup https://orcid.org/0000-0002-4238-6963

REFERENCES


SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.