Amsterdam, 12 December 2019
EMA/694594/2019
Committee for Medicinal Products for Human Use (CHMP)

Assessment report for paediatric studies submitted in accordance with article 46 of regulation (EC) No 1901/2006, as amended

Kalydeco/Symkevi

International non-proprietary name: TEZACAFTOR/IVACAFTOR

Procedure no.: EMEA/H/C/002494/P46/028 and EMEA/H/C/004682/P46/007

Note

Assessment report as adopted by the CHMP with all information of a commercially confidential nature deleted.

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8613
E-mail info@ema.europa.eu **Website** www.ema.europa.eu
An agency of the European Union

LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	anatomic class
BMI	body mass index
CDC	Center for Disease Control and Prevention
CF	cystic fibrosis
CFTR	CF transmembrane conductance regulator protein
CFTR	*CF transmembrane conductance regulator gene*
CFQ-R	Cystic Fibrosis Questionnaire-Revised
CI	confidence interval
CPK	creatine phosphokinase
CRF	case report form
CT	computed tomography
CYP	cytochrome P450
DBP	diastolic blood pressure
DMC	data monitoring committee
ECG	electrocardiogram
EDC	electronic data capture
FAS	full analysis set
FDA	Food and Drug Administration
$FEF_{25-75\%}$	forced expiratory flow at 25-75% of the pulmonary volume
FEV_1	forced expiratory volume in 1 second
FSH	follicle-stimulating hormone
FVC	forced vital capacity
GCP	Good Clinical Practice
GGT	gamma-glutamyl transpeptidase
GI	gastrointestinal
HR	heart rate
ICF	informed consent form
ICH	International Council on Harmonisation
IDMC	independent data monitoring committee
IEC	independent ethics committee
IPD	important protocol deviation
IRB	institutional review board
IV	intravenous
IVA	ivacaftor
KM	Kaplan-Meier
LDCT	low-dose computed tomography
LFT	liver function test
LLN	lower limit of normal
LS	least square

LSM	least squares mean
LUM	lumacaftor
MCID	minimum clinically important difference
MedRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
PE	physical examination
PEx	pulmonary exacerbation
$ppFEV_1$	percent predicted forced expiratory volume in 1 second
PR	PR interval, segment
PT	preferred term
q12h	every 12 hours
qd	once daily
QRS	the portion of an ECG comprising the Q, R, and S waves, together representing ventricular depolarization
QT	QT interval
QTc	QT interval corrected
QTcF	QT interval corrected by Friderica's formula
SAE	serious adverse event
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SE	standard error
SF-36	36-Item Short Form Survey
SI	SI units (International System of Units)
SOC	system organ class
SOP	standard operating procedure
TE	Treatment Emergent
TEAE	treatment-emergent adverse event
TEZ	tezacaftor
ULN	upper limit of normal
US	United States
WHO-DDE	World Health Organization-Drug Dictionary Enhanced

Table of contents

1. Introduction

On 03 July 2019, the MAH submitted a completed paediatric study, Study VX15-661-112 for Kalydeco (ivacaftor) 150 mg film-coated tablets, in accordance with Article 46 of Regulation (EC) No1901/2006, as amended.

The same Article 46 submission has been submitted in parallel for Symkevi (EMEA/H/C/00004682) which is a FDC of tezacaftor 100 mg /ivacaftor 150 mg film coated tablets in a combination regimen with ivacaftor 150 mg tablets.

These data are also submitted as part of the post-authorisation measure(s).

A short critical expert overview has also been provided.

2. Scientific discussion

2.1. Information on the development program

The MAH stated that Study VX15-661-112 (Study 112), a Phase 2, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Effect of VX-661 in Combination With Ivacaftor on Chest Imaging Endpoints in Subjects Aged 12 Years and Older With Cystic Fibrosis, Homozygous for the *F508del-CFTR* Mutation is a stand-alone study.

TEZ/IVA is approved as a combination regimen of Symkevi® with IVA 150-mg tablets (Kalydeco®) for the treatment of patients with CF 12 years of age and older who are homozygous for the F508del-CFTR mutation or who are heterozygous for the *F508del-CFTR* mutation and have 1 of the following mutations in the *CFTR* gene: *P67L*, *R117C*, *L206W*, *R352Q*, *A455E*, *D579G*, *711+3A→G*, *S945L*, *S977F*, *R1070W*, *D1152H*, *2789+5G→A*, *3272-26A→G*, and *3849+10kbC→T*.

2.2. Information on the pharmaceutical formulation used in the study

The test product was the same as the commercially approved product, Symkevi, for patients 12 years of age and older (TEZ 100-mg/IVA 150-mg fixed-dose combination tablets). The test product was administered to study subjects orally at a dose of TEZ 100-mg once daily (qd)/IVA 150-mg every 12 hours (q12h), which is also the approved dose of Symkevi. Subjects randomized to placebo received an inactive matching regimen.

2.3. Clinical aspects

2.3.1. Introduction

The MAH submitted a final report for:

Study VX15-661-112, a phase 2, randomized, placebo-controlled, double-blind study to evaluate the effect of VX-661 in combination with ivacaftor on chest imaging endpoints in subjects aged 12 years and older with cystic fibrosis, homozygous for the *F508del-CFTR* mutation.

2.3.2. Clinical study

Clinical study number and title

Study VX15-661-112, a phase 2, randomized, placebo-controlled, double-blind study to evaluate the effect of VX-661 in combination with ivacaftor on chest imaging endpoints in subjects aged 12 years and older with cystic fibrosis, homozygous for the *F508del-CFTR* mutation.

Description

Methods

Objective(s)

The primary objective of the study was to evaluate the treatment effect of TEZ/IVA on chest imaging endpoints as evaluated using low-dose computed tomography (LDCT) at Week 72 in subjects with CF 12 years of age and older who are homozygous for the *F508del-CFTR* mutation.

The secondary objective was to evaluate the safety of TEZ/IVA through Week 72.

Other objectives were as follows:

• To evaluate the effect of TEZ/IVA on radiologic subscore assessments on LDCT at Week 72

• To explore the effect of TEZ/IVA on chest imaging endpoints as evaluated on Chest magnetic resonance imaging (MRI) through Week 72

• To explore the correlation between the LDCT and Chest magnetic resonance imaging (MRI) Brody/Cystic Fibrosis-Computed Topography (CF-CT) scores (total and subscores)

• To evaluate the effect of TEZ/IVA on percent predicted forced expiratory volume in 1 second (ppFEV1)

Assessor's comments

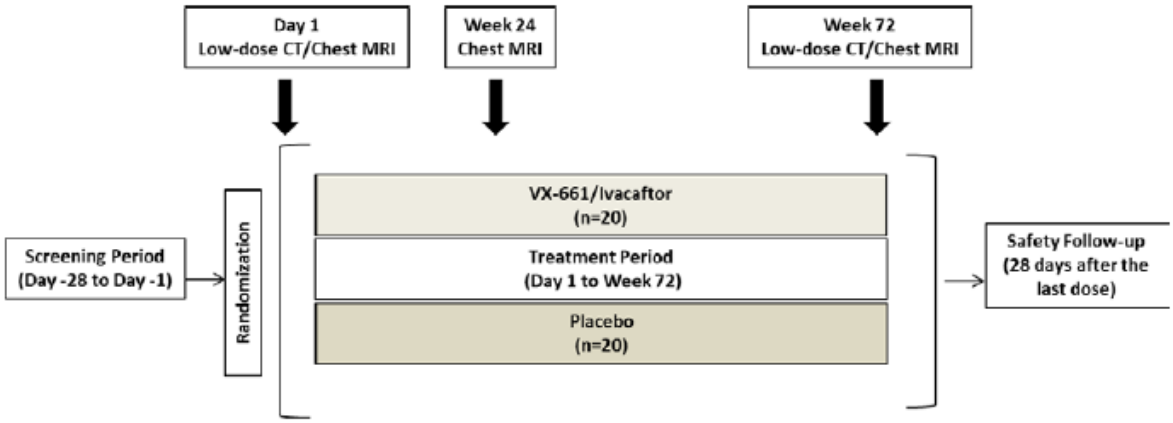
Study 112 was designed to explore the treatment effect of TEZ/IVA on chest imaging endpoints during 72 weeks of treatment. LDCT was used for chest imaging and the images were evaluated using the Brody/CF-CT scoring system. Magnetic resonance imaging (MRI) was planned, but the equipment needed to conduct ultra-short echo time chest MRI was unavailable at the start of enrolment, and consequently, MRI scans were not performed on any subject during the study.

Study design

This was a Phase 2, randomized, placebo-controlled, double-blind, parallel-group, multicentre study in subjects with CF who were homozygous for the *F508del-CFTR* mutation.

As shown in figure below, this study included a Screening Period (28-days), a 72-week Treatment Period, and a Safety Follow-up Visit (approximately 28 days after the last dose). Low-dose CT scans were performed at Day 1 and week 72. The Week 72 CT scan may have been delayed for up to 60 days if the subject was recovering from a pulmonary exacerbation. The scan was done after the pulmonary exacerbation was resolved and at least 28 days after the antibiotic regimen for the treatment of pulmonary infection had been completed. If this antibiotic regimen was not completed by the end of the 60-day extension, the subject may have completed the CT scan within the first 30 days of enrolling in Study VX14-661-110; or within 1 week of the end of the 60-day extension if they did not enroll in Study VX14-661-110. No extension was permitted for any other assessment.

Figure 9-1 VX15-661-112: Schematic of Study Design



CT: computed tomography; MRI: magnetic resolution imaging
Notes: The Safety Follow-up Visit was not required for subjects who completed the Week 72 Visit and enrolled in the TEZ/IVA open-label extension study, VX14-661-110, within 28 days after the last dose of study drug.

The study was monitored by an independent data monitoring committee (IDMC), which conducted periodic reviews of the safety data. The IDMC charter was finalized before the first subject was screened. The IDMC evaluated accumulating safety data from the study and made recommendations to Vertex on the conduct of the study.

Low-dose CT scans and Chest MRIs were to be performed on the same day, but no chest MRI was collected. Equipment to conduct ultra-short echo time chest MRI was unavailable at the start of enrolment and was not collected for any subjects during the study; thus, it will not be discussed further in this report.

Assessor's comments

The MAH states that the 72-week duration was selected to allow adequate exposure of TEZ/IVA to assess safety of treatment and chest imaging endpoints based on review of the natural history progression of radiographic changes demonstrated on longitudinal studies seen on CT scans. In a 2-year natural history study in 48 children with CF, with a mean age of 11 years at the start of the study, the mean change in the Total Brody score was +2.2 points per year (de Jong PA. et al, 2004). Based on this and other longitudinal studies (which are not discussed by the MAH), it was estimated that the progression of disease observed in the proposed study population would be a 3.3 point increase in the Total Brody score at Week 72. There has been 1 study that has examined the impact of CFTR modulation on structural lung disease. In 10 subjects (aged 10 to 44 years) who carried the CFTR-G551D mutation, Sheikh et al (2015) reported a within-treatment annual improvement of 13.6 points in the Total Brody/CF-CT score after 1 year of IVA treatment.

Chest imaging is recognized as essential in the assessment of respiratory disease of children and adults with CF. The ECFS 2018 revision of the best practice guidelines (Castellani C., Duff AJA., Bell SC. et al, 2018) recommends that chest X-rays are routinely performed on an annual basis as well as at times of clinical deterioration. Other imaging modalities, such as high resolution CT scanning, should be available (in the cystic fibrosis centre), and are used routinely in some CF centres. No clear recommendations are given, however, at what age CT scanning of the lung to monitor lung disease should be started, and how often this should be done.

Chest CT offers advantages over other pulmonary endpoints because it has greater sensitivity to detect early structural lung disease even while the commonly used endpoint, ppFEV1, is within normal range

and respiratory symptoms are absent. Additionally, CT can demonstrate the structural abnormalities that are specific to CF lung disease and can be performed similarly in all age groups. However, concerns about radiation exposure and need for sedation/anaesthesia (particularly in young children) apparently limited a general recommendation for routine use of CT in patients with cystic fibrosis in the EU. Magnetic resonance imaging (MRI) of the lung as a radiation-free imaging technique which enables not only morphological imaging, but also visualisation and regional measurement of functional qualities of the lung has emerged as an alternative to CT scanning but it is not generally recommended likely because structural limitations (reduced spatial resolution), need for a specific training, availability and cost.

Study population /Sample size

This study was conducted at 9 sites in Australia.

A total of 41 subjects were randomized: 20 subjects in the TEZ/IVA group and 21 subjects in the placebo group.

Assessor's comments

The study was exploratory in nature and the sample size was not based on statistical power.

Key inclusion criteria

- Subjects (males and females), aged 12 years or older on the date of informed consent or, where appropriate, date of assent.
- Homozygous for the *F508del-CFTR* mutation, genotype was confirmed with testing performed at the Screening Visit. If the *CFTR* screening genotype result was not received before randomization, a previous *CFTR* genotype laboratory report may have been used to establish eligibility. *CFTR* genotyping may have waived if the subject had a documented result from a previous Vertex study.
- Confirmed diagnosis of CF defined as a sweat chloride value ≥60 mmol/L by quantitative pilocarpine iontophoresis. A sweat chloride test was performed at the Screening Visit if an eligible sweat chloride value was not available in the subject's medical records and the Screening Visit value was needed to establish eligibility. For subjects using sweat chloride values documented in their medical records to establish eligibility, the sweat chloride test at the Screening Visit was optional.
- ppFEV1 ≥70% of predicted normal for age, sex, and height (equations of Wang et al. or Hankinson et al.) during screening. Spirometry measurements met American Thoracic Society/European Respiratory Society criteria for acceptability and repeatability.
- Stable CF disease as judged by the investigator.

Key exclusion criteria

- History of any comorbidity that, in the opinion of the investigator, might confound the results of the study or posed an additional risk in administering study drug to the subject such as history of cirrhosis with portal hypertension, and/or history of risk factors for Torsade de Pointes, obesity, acute neurologic events, and autonomic neuropathy.
- Abnormal liver function at Screening defined as

    - any 2 or more of the following: ≥3 x upper limit of normal (ULN) aspartate aminotransferase (AST), ≥3 x ULN alanine aminotransferase (ALT), ≥3 x ULN gamma-glutamyl transpeptidase (GGT), ≥3 x ULN alkaline phosphatase (ALP), or ≥2 x ULN total bilirubin.
    - Abnormal liver function defined as any increase of ≥5 x ULN in AST or ALT.
- Abnormal renal function defined as glomerular filtration rate ≤50 mL/min/1.73 m2 (calculated by the Modification of Diet in Renal Disease Study Equation) for subjects ≥18 years of age and ≤45 mL/min/1.73 m2 (calculated by the Counahan-Barratt equation) for subjects aged 12 to 17 years (inclusive).
- For an acute upper or lower respiratory infection, pulmonary exacerbation (PEx), or changes in therapy (including antibiotics) for pulmonary disease before Day 1 (first dose of study drug), antibiotic regimen for the treatment of a pulmonary infection must have been completed at least 28 days before Day 1 (first dose of study drug).
- History of solid organ or haematological transplantation.
- History or evidence of cataract, lens opacity, Y-suture, or lamellar rings determined to be clinically significant by the ophthalmologist during the ophthalmologic examination during the Screening Period.
- Colonization with organisms associated with a more rapid decline in pulmonary status (e.g., *Burkholderia cenocepacia*, *Burkholderia dolosa*, and *Mycobacterium abscessus*). For subjects with a history of a positive culture in the past, suggested criteria were provided to the investigator to consider the subject free of colonization.
- Any contraindication to undergoing LDCT or Chest MRI, as per the site's institutional guidelines.

Assessor's comments

Inclusion and exclusion criteria are similar to those used in the pivotal study 106 in subjects homozygous for F508del. No specific criteria are related to the primary endpoint of the study other than trying to ensure that subjects have stable disease at the time of screening/inclusion into the study and the exclusion criterion related to the presence of contraindications to undergoing LCDT or Chest MRI.

Treatments

The test product was administered to study subjects orally at a dose of TEZ 100-mg once daily (qd)/IVA 150-mg every 12 hours (q12h), the approved dose of Symkevi. Subjects randomized to placebo received an inactive matching regimen.

Study drugs were administered within 30 minutes after starting a meal with fat-containing food such as a standard CF high-fat, high-calorie meal or snack by the subject, according to the guidelines outlined in the protocol.

Outcomes/endpoints

Primary endpoint

Absolute change in Total Brody/CF-CT score from baseline at Week 72 using LDCT.

This score provides both localization and quantification of 5 abnormalities characteristic of CF lung disease: bronchiectasis, air trapping, mucus plugging, bronchial wall thickening, and parenchymal changes (includes parenchymal opacities, ground glass opacities, and cysts/bullae).

CT images were scored by 2 blinded, independent, qualified readers. Readers were randomly assigned images as a primary reviewer or secondary reviewer. The primary reviewer's score was selected if the scores were concordant. Assessments that had discordant scores by more than 2 points in any subdomain were reviewed by a third independent, expert reader for an adjudication read. The third reader selected 1 of the original reads to record as the final assessment (the adjudicator did not have the option of rescoring the case).

The primary analysis was based on an analysis of covariance (ANCOVA) model with the change from baseline of Total Brody/CF-CT score at Week 72 as the dependent variable, and treatment, sex (male vs. female), and age (12-17 years vs. 18 years and older) as covariates. LDCT images were scored by 2 central reader(s) blinded to treatment group and visit time point. The difference between the TEZ/IVA group and the placebo group in mean change from baseline in Total Brody/CF-CT score at Week 72 was estimated using the FAS.

The primary result obtained from the model was the estimated treatment effect in each treatment group (with a 95% CI), the estimated between-group difference in treatment effects, and a 95% CI for the difference. For missing data, no imputation was performed.

Secondary and other endpoints and other assessments

- Change from baseline in Brody/CF-CT subscores (bronchiectasis, bronchial wall thickening, mucus plugging, parenchyma, and hyperinflation) as measured by Brody/CF-CT score at Week 72 using LDCT

  Analysis of change from baseline scores for each Brody/CF-CT subscore was based on the FAS using a similar model to the primary endpoint, i.e., an ANCOVA model was fit to each subscore, with treatment, sex, and age (12-17 years vs. 18 years and older) as covariates with no imputation of missing data.

- Absolute change from baseline in percent predicted forced expiratory volume in 1 second (ppFEV1) at Week 72.

- Absolute change in additional spirometry variables (FEV1, Forced Vital Capacity [FVC], Percent predicted FVC, FEV1/FVC (ratio), and Forced expiratory flow [FEF25-75%]).

- Response in CFQ-R Respiratory Domain

  In study 112 3 different versions of CFQ-R forms were used as follows:

  - CFQ-R for Children Ages 12 and 13 had a total of 35 questions to form 8 domains. All questions were scored 1, 2, 3, or 4.

  - CFQ-R for Adolescents and Adult (subjects 14 years and older) had a total of 50 questions to form 12 domains. Question 43, which was scored 1, 2, 3, 4, or 5, was not used in calculating any domain; all the other 49 questions were scored 1, 2, 3, or 4.

  - CFQ-R for Parents/Caregivers (subjects 13 years and younger) had a total of 44 questions to form 11 domains. Question 37, which was scored 1, 2, 3, 4, or 5, was not used in calculating any domains; all the other 43 questions were scored 1, 2, 3, or 4.

For all 3 CFQ-R versions, to calculate the score for each domain, the response scores on the negatively phrased questions were reversed (reversed scores = 5 − response scores) so that 1 always represented the worst condition and 4 always represented the best condition.

The scaled score for each domain ranged from 0 (worst condition) to 100 (best condition). It was calculated as follows:

Scaled score for a domain = 100 × (mean [scores of all questions in that domain] − 1) / 3

The scaled score for a specific domain was not calculated if more than half of the questions in the domain had missing scores.

- Response in Other CFQ-R Domains
- Absolute change from baseline in BMI and BMI z-score through Week 72 (the latter only for subjects < 20 years old).
- Absolute change from baseline in weight and weight z-score through Week 72 (the latter only for subjects < 20 years old).
- Absolute change from baseline in height and height z-score through Week 72 (the latter only for subjects < 20 years old).
- Pulmonary exacerbations and hospitalizations analysed as number of events and as time-to-event.

  The analysis period for all variables related to pulmonary exacerbations and hospitalisation started from the first dose of study drug and ended on the date of the last assessment up to Week 72. For each treatment group, the annualized number of events was calculated as the Total number of events x 336 / Total number of days on study; (48 weeks = 336 days) where the Total number of events = Total number of events through the study end date and the Total number of days on study = Analysis period end date − the date of first dose + 1. The Total number of years (48 weeks) on study is defined as the Total number of days on study/336.
- 36-Item Short Form Survey (Version 2)
- Safety: analysis based on adverse events, clinical laboratory values (hematology, serum chemistry, coagulation studies, vitamin levels, lipid panel, and urinalysis), standard 12-lead ECGs, vital signs, pulse oximetry.

  Safety endpoints were analyzed based on the Safety Set (for each applicable treatment period). Only descriptive analyses of safety were performed and no statistical testing was performed.

  Study drug administration was interrupted immediately (before confirmatory testing), and the medical monitor was notified, if any of the following criteria was met:

  - ALT or AST >8 × ULN
  - ALT or AST >5 × ULN for more than 2 weeks
  - ALT or AST >3 × ULN, in association with total bilirubin >2 × ULN and/or clinical jaundice

Assessor's comments

In Study 112 a low-dose CT (LDCT) algorithm was used to minimize exposure to ionizing radiation. The CT scans were optimized for low radiation dose and high image quality. Inspiratory and expiratory (the latter needed to quantify air trapping) LDCT scans were obtained at Day 1, Week 24, and week 72. These 2 CT scans each used about 2 to 2.5 mSv which is slightly above the estimated average of 1.5

mSv each year from natural sources to which Australians are exposed to and below the estimated 10 mSv up to which no direct evidence of human health effects has been seen. CT scans were obtained using a spirometry-controlled volumetric technique. Images were reconstructed at 5 mm thickness and at <1 mm thickness (exact thickness was specified for each manufacturer) using a high frequency reconstruction algorithm for 5 mm sections and standard reconstruction for <1 mm sections.

Study 112 was performed at 9 centres in Australia. The MAH states that additional technique information which are specific to each manufacturer and model of CT scanner were available but are not discussed (i.e., in relation to how these may impact on the total score and on the subscores). Comparisons with the longitudinal studies mentioned by the MAH (de Jong PA et al, 2004; Sheikh SI et al, 2015) or with other studies in which chest CT was assessed is hampered due to potential differences in the CT technique but also in relation to the population studied (e.g. young children versus adult subjects).

Quantitative evaluation of chest CT findings requires application of a scoring system to derive numerical values. Several CT scoring systems for CF have been developed for adults and children over the age of 6. The Brody's scoring system is a lobar scoring system, which semi-quantitatively scores the degree of structural lung disease by assigning a score to each lobe separately (including the lingula). It requires a complex method of calculation of the extent at different lung zones/segments and inclusion of weighting factors. To improve standardization and training, the CF-CT scoring system, based on the Brody II system, was developed which consists of a large training module and 7 training sets that were scored by Brody and de Jong (the most experienced observers at that time) to define the 'gold standard' ratings (Szczesniak R, 2017). The score provides both localization and quantification of 5 abnormalities characteristic of CF lung disease: bronchiectasis, air trapping, mucus plugging, bronchial wall thickening, and parenchymal changes (includes parenchymal opacities, ground glass opacities, and cysts/bullae). It has been used in multiple studies to validate chest CT as an outcome and therefore it is considered an acceptable choice within the available scoring systems although the clinical value of the numbers generated is difficult to understand. The Brody scoring system has been reported as a total score with a maximal possible value of 207 and as a score representing the average severity of each of the six lobes, including the lingula as a separate lobe, with a maximum of 40.5 (Sanders DB. et al, 2015).

Upon request, the MAH has clarified that the subscore ranges used in Study 112 were bronchiectasis 0 to 12; mucus plugging 0 to 6; peribronchial thickening 0 to 9; parenchymal opacities 0 to 5; and air trapping 0 to 4.5. The maximum possible score for a single lobe is 36.5 (12 + 6 + 9 + 5 + 4.5). Given that each lobe is scored separately, with the lingula considered a lobe (6 total), this results in a range for the total score of 0 to 219. For the purposes of Study 112, in which changes from baseline were compared, the absolute score was chosen as the primary outcome while in some publications the score is expressed on a percentage scale of 0 to 100, which allows for comparisons to scoring systems with different ranges. Two readers scored each lobe, and any subdomain score differing by more than 2 points was adjudicated by a third independent, expert reader who chose 1 of the scores to report. The expert reader noted that there was a greater than usual disagreement between the readers in Study 112, which added variability to the results.

Statistical Methods

As the study was exploratory and sample size was not based on statistical power, the analysis for all efficacy endpoints was an estimate of the treatment effect within-group and the difference between the treatment arms. No statistical hypothesis testing was performed and there was no multiplicity control

for any endpoint. No minimum clinically important difference (MCID) has been established for the Total Brody score and its subdomains.

Study results were analyzed using descriptive statistics which include n, mean, SD, SE, median, minimum, and maximum. Ninety-five percent confidence intervals (95% CI) were provided for the analysis of the primary variable (total score and subscores).

Baseline value, unless otherwise specified, was defined as the most recent non-missing measurement (scheduled or unscheduled) collected prior to the first dose of study drug. For ECGs, the baseline was defined as the average of the 3 pre-treatment measurements (triplicate) on Day 1.

Three analysis sets were defined:

- The All Subjects Set was defined as all subjects who had been randomized or had received at least 1 dose of study drug. This analysis set was used in subject listings and the disposition summary table, unless otherwise specified.
- The Full Analysis Set (FAS) was defined as all randomized subjects who had received at least 1 dose of study drug. The FAS was used in efficacy analyses in which subjects were analyzed according to their randomized treatment group.
- The Safety Set was defined as all subjects who received at least 1 dose of study drug. The Safety Set was used for all safety analyses in which subjects were analyzed according to the treatment they received.

Assessor's comments

The analysis of the primary efficacy variable (total Brody/CF-CT score) and of the sub-scores was based on an analysis of the covariance. For these variables the estimated between-group difference in treatment effects was provided. For the remaining efficacy variables, only the absolute within-group changes from baseline at week 72 are shown. Although the table showing the results of the absolute change from baseline in the respiratory domain of the CFQ-R indicates that the results of an ANCOVA analysis are presented, the MAH has clarified that the title for Table 11-5 in the clinical study report (and therefore in this report, see further below) is mislabelled, i.e., absolute within-group changes are presented.

Results

Recruitment/ Number analysed

A total of 41 subjects were randomized: 20 subjects in the TEZ/IVA group and 21 subjects in the placebo group.

Table below shows the disposition of subjects.

Table 10-1 Subject Disposition, All Subjects Set

Disposition/Reason	TEZ/IVA n (%)	Placebo n (%)	Total n (%)
All Subjects Set (Randomized or Dosed)[a]	20	21	41
Randomized	20	21	41
Randomized but not dosed	0	0	0
Safety Set[b]	20	21	41
Full Analysis Set[c]	20	21	41
Completed study drug treatment	20 (100.0)	19 (90.5)	39 (95.1)
Prematurely discontinued study drug treatment	0	2 (9.5)	2 (4.9)
Reason for discontinuation of study drug treatment			
Adverse event	0	1 (4.8)	1 (2.4)
Subject refused further dosing (not due to AE)	0	1 (4.8)	1 (2.4)
Completed study	20 (100.0)	20 (95.2)	40 (97.6)
Prematurely discontinued from study	0	1 (4.8)	1 (2.4)
Reason for discontinuation from study			
Withdrawal of consent (not due to AE)	0	1 (4.8)	1 (2.4)

Source: Table 14.1.1

AE: adverse event; IVA: ivacaftor; n: size of subsample; TEZ: tezacaftor

Notes: Percentages were calculated relative to the number of subjects in the Full Analysis Set. If a subject discontinued TEZ/IVA (or matching placebo) and IVA (or matching placebo) tablets for different reasons, the subject was counted for both reasons, but only counted once in the total number of subjects who prematurely discontinued treatment.

[a] The All Subjects Set was defined as all subjects who were randomized or received at least 1 dose of study drug. Subjects in the all subjects set were summarized under planned treatment, but dosed and not randomized subjects (if any) are summarized under actual treatment.

[b] The Safety Set was defined as all subjects who received at least 1 dose of study drug. Subjects in the Safety Set were summarized under actual treatment.

[c] The Full Analysis Set was defined as all randomized subjects who received at least 1 dose of the study drug. Subjects in the Full Analysis Set, as well as randomized and randomized but not dosed rows, are summarized under the planned treatment assignment.

Of the 41 subjects who received at least 1 dose of study drug, 40 (97.6%) subjects completed the study and 39 (95.1%) completed study drug dosing. No subject in the TEZ/IVA treatment group and 1 (4.8%) subject in the placebo group discontinued treatment due to an AE.

Therefore, the FAS included 20 subjects in the TEZ/Iva group and 21 in the placebo group. The Safety Set included the same number of subjects.

Baseline data

Tables below show baseline demographic and disease characteristics data of subjects in the Full Analysis Set.

Table 10-2 Subject Demographics, Full Analysis Set

Demographics	TEZ/IVA N = 20	Placebo N = 21	Total N = 41
Sex, n (%)			
Male	9 (45.0)	11 (52.4)	20 (48.8)
Female	11 (55.0)	10 (47.6)	21 (51.2)
Childbearing potential[a], n (%)			
Yes	11 (100.0)	8 (80.0)	19 (90.5)
No	0	2 (20.0)	2 (9.5)
Age at screening (years)			
Mean (SD)	20.4 (7.5)	20.1 (9.3)	20.2 (8.4)
Median	19.5	15.0	18.0
Min, Max	12, 41	12, 43	12, 43
Age category at screening (years)			
12 to <18	8 (40.0)	11 (52.4)	19 (46.3)
≥ 18	12 (60.0)	10 (47.6)	22 (53.7)
Ethnicity, n (%)			
Not Hispanic or Latino	20 (100.0)	21 (100.0)	41 (100.0)
Race, n (%)			
White	20 (100.0)	21 (100.0)	41 (100.0)

Source: Table 14.1.3

IVA: ivacaftor; n: size of subsample; N: total sample size; TEZ: tezacaftor

[a] % for the Yes/No categories under 'Childbearing potential' was calculated using number of female subjects as denominator.

Table 10-3 Baseline Characteristics, Full Analysis Set

Characteristic	TEZ/IVA N = 20	Placebo N = 21	Total N = 41
Weight (kg)			
n	20	21	41
Mean (SD)	59.7 (11.9)	59.6 (15.2)	59.7 (13.5)
Median	56.5	58.0	57.0
Min, Max	44.0, 83.0	30.0, 93.0	30.0, 93.0
Height (cm)			
n	20	21	41
Mean (SD)	164.7 (9.6)	166.8 (12.5)	165.8 (11.1)
Median	163.5	167.0	164.0
Min, Max	149.0, 181.0	135.0, 195.0	135.0, 195.0
BMI (kg/m^2)[a]			
n	20	21	41
Mean (SD)	21.93 (3.51)	21.10 (3.40)	21.51 (3.44)
Median	20.95	20.40	20.68
Min, Max	17.24, 28.96	16.03, 30.11	16.03, 30.11
BMI z-score (for subjects <20 years old)			
n	10	13	23
Mean (SD)	−0.03 (0.77)	−0.15 (0.64)	−0.10 (0.68)
Median	−0.14	−0.17	−0.16
Min, Max	−1.34, 1.64	−1.64, 0.83	−1.64, 1.64
Percent Predicted FEV_1			
n	20	21	41
Mean (SD)	91.4 (16.0)	86.6 (12.7)	89.0 (14.4)
Median	91.1	87.9	89.4
Min, Max	64.7, 128.0	66.9, 107.5	64.7, 128.0
Percent Predicted FEV_1 category, n (%)			
<70	2 (10.0)	2 (9.5)	4 (9.8)
≥70 to ≤90	8 (40.0)	9 (42.9)	17 (41.5)
>90	10 (50.0)	10 (47.6)	20 (48.8)
Sweat Chloride at Screening, mmol/L			
n	17	12	29
Mean (SD)	101.2 (11.4)	102.3 (8.9)	101.6 (10.3)
Median	97.5	103.8	103.0
Min, Max	77.0, 118.0	88.0, 115.0	77.0, 118.0
Use of dornase alfa[b], n (%)	13 (65.0)	17 (81.0)	30 (73.2)
Use of inhaled antibiotic[b], n (%)	3 (15.0)	0	3 (7.3)
Use of bronchodilator[b], n (%)	15 (75.0)	15 (71.4)	30 (73.2)
Use of inhaled bronchodilator[b], n (%)	15 (75.0)	15 (71.4)	30 (73.2)
Use of inhaled hypertonic saline[b], n (%)	12 (60.0)	10 (47.6)	22 (53.7)
Use of inhaled corticosteroids[b], n (%)	11 (55.0)	10 (47.6)	21 (51.2)

Characteristic	TEZ/IVA N = 20	Placebo N = 21	Total N = 41
Colonization of Pseudomonas aeruginosa, n (%)			
Positive	11 (55.0)	13 (61.9)	24 (58.5)
Negative	9 (45.0)	8 (38.1)	17 (41.5)
CFQ-R Respiratory Domain Score			
n	20	21	41
Mean (SD)	76.8 (13.3)	82.4 (11.7)	79.7 (12.7)
Median	80.6	83.3	83.3
Min, Max	50.0, 94.4	61.1, 100.0	50.0, 100.0
Brody/CF-CT Total Score			
n	20	20	40
Mean (SD)	38.29 (22.91)	43.68 (33.96)	40.98 (28.72)
Median	39.13	36.50	39.00
Min, Max	6.00, 70.25	0.00, 120.75	0.00, 120.75
Brody/CF-CT Bronchiectasis Subscore			
n	20	20	40
Mean (SD)	11.05 (7.33)	13.25 (13.05)	12.15 (10.51)
Median	11.00	9.13	10.00
Min, Max	0.00, 24.50	0.00, 48.75	0.00, 48.75
Brody/CF-CT Bronchial Wall Thickening Subscore			
n	20	20	40
Mean (SD)	7.86 (6.82)	9.05 (9.45)	8.46 (8.15)
Median	8.13	5.00	7.50
Min, Max	0.00, 19.00	0.00, 30.00	0.00, 30.00
Brody/CF-CT Mucus Plugging Subscore			
n	20	20	40
Mean (SD)	5.6 (4.2)	6.4 (5.3)	6.0 (4.7)
Median	5.0	5.5	5.0
Min, Max	0.0, 12.0	0.0, 18.0	0.0, 18.0
Brody/CF-CT Parenchymal Opacities Subscore			
n	20	20	40
Mean (SD)	2.1 (1.9)	2.9 (2.6)	2.5 (2.3)
Median	2.0	2.0	2.0
Min, Max	0.0, 6.0	0.0, 8.0	0.0, 8.0
[illegible]			
Brody/CF-CT Hyperinflation Subscore			
n	20	20	40
Mean (SD)	11.8 (5.3)	12.1 (7.4)	12.0 (6.4)
Median	11.5	14.3	13.3
Min, Max	4.0, 21.0	0.0, 27.0	0.0, 27.0

Source: Table 14.1.4

BMI: body mass index; CF-CT: Cystic Fibrosis-Computer Tomography; CFQ-R: Cystic Fibrosis Questionnaire-Revised; FEV_1: forced expiratory volume in 1 second; IVA: ivacaftor; n: size of subsample; N: total sample size; TEZ: tezacaftor

Note: Baseline was defined as the most recent non-missing measurement before the first dose of study drug.

[a] BMI = weight/(height × height), in kg/m^2.

[b] Includes medications started before the first dose of study drug and continuing during the treatment period.

Efficacy results

Absolute change from baseline of Total Brody/CF-CT score at Week 72 using LDCT

Results of the ANCOVA analysis of the primary efficacy endpoint are presented in table below.

Table 11-1 ANCOVA Analysis of Absolute Change in Total Brody/CF-CT Score Using Low-dose CT, Full Analysis Set

Parameter	TEZ/IVA N = 20	Placebo N = 21
Baseline		
n	20	20
Mean (SD)	38.29 (22.91)	43.68 (33.96)
Absolute Change from Baseline at Week 72, Within-Group		
n	20	20
LS Mean (SE)	0.90 (2.09)	2.38 (2.07)
95% CI	(−3.34, 5.14)	(−1.82, 6.58)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	−1.48 (2.96)	NA
95% CI	(−7.47, 4.52)	NA

Source: Table 14.2.1.2

ANCOVA: analysis of covariance; IVA: ivacaftor; LS Mean: least squares mean; n: size of subsample; N: total sample size; TEZ: tezacaftor

Notes: Baseline was defined as the measurement on the nominal Day 1 visit. ANCOVA model includes change in Total Brody CF-CT score at Week 72 as dependent variable, treatment group as fixed effect, and sex (male vs. female) and age (<18years vs. ≥18years) as covariates.

The change from baseline at Week 72 in the total Brody/CF-CT Score was numerically lower for the TEZ/IVA group compared with the placebo group. The least squares (LS) mean treatment difference between the TEZ/IVA and placebo groups was −1.48 (95% CI: −7.47, 4.52).

Absolute change from baseline at Week 72 in Brody/CF-CT subscores using LDCT

Subscores of the Brody/CF-CT, analysed by ANCOVA change from baseline at Week 72, are presented in table below.

Table 11-2 ANCOVA Analysis of Absolute Change in Brody CF-CT Subscores Using Low-dose CT, Full Analysis Set

Category	TEZ/IVA N = 20	Placebo N = 21
Brody CF-CT Score: Bronchiectasis		
Baseline		
n	20	20
Mean (SD)	11.05 (7.33)	13.25 (13.05)
Absolute Change from Baseline at Week 72, Within-Group		
n	20	20
LS Mean (SE)	0.74 (0.88)	1.32 (0.87)
95% CI	(−1.05, 2.53)	(−0.45, 3.09)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	−0.58 (1.25)	NA
95% CI	(−3.11, 1.95)	NA
Brody CF-CT Score: Peribronchial Thickening		
Baseline		
n	20	20
Mean (SD)	7.86 (6.82)	9.05 (9.45)
Absolute Change from Baseline at Week 72, Within-Group		
n	20	20
LS Mean (SE)	−0.03 (0.72)	−0.24 (0.72)
95% CI	(−1.50, 1.44)	(−1.70, 1.21)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	0.21 (1.02)	NA
95% CI	(−1.86, 2.29)	NA
Brody CF-CT Score: Mucus Plugging		
Baseline		
n	20	20
Mean (SD)	5.6 (4.2)	6.4 (5.3)
Absolute Change from Baseline at Week 72, Within-Group		
n	20	20
LS Mean (SE)	−0.4 (0.5)	0.4 (0.5)
95% CI	(−1.3, 0.6)	(−0.6, 1.4)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	−0.8 (0.7)	NA
95% CI	(−2.1, 0.6)	NA

Category	TEZ/IVA N = 20	Placebo N = 21
Brody CF-CT Score: Parenchymal Opacities		
Baseline		
n	20	20
Mean (SD)	2.1 (1.9)	2.9 (2.6)
Absolute Change from Baseline at Week 72, Within−Group		
n	20	20
LS Mean (SE)	0.1 (0.4)	−0.1 (0.4)
95% CI	(−0.7, 0.9)	(−0.9, 0.7)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	0.2 (0.6)	NA
95% CI	(−1.0, 1.3)	NA
Brody CF-CT Score: Hyperinflation		
Baseline		
n	20	20
Mean (SD)	11.8 (5.3)	12.1 (7.4)
Absolute Change from Baseline at Week 72, Within-Group		
n	20	20
LS Mean (SE)	0.4 (0.9)	1.0 (0.9)
95% CI	(−1.4, 2.2)	(−0.8, 2.7)
Difference vs. Placebo in Absolute Change from Baseline at Week 72		
LS Mean (SE)	−0.5 (1.3)	NA
95% CI	(−3.1, 2.0)	NA

Source: Table 14.2.1.3

ANCOVA: analysis of covariance; CF: Cystic Fibrosis; CT: Computed Topography; IVA: ivacaftor; LS Mean: least squares mean; n: size of subsample; N: total sample size; TEZ: tezacaftor

Notes: Baseline was defined as the measurement on the nominal Day 1 visit. ANCOVA model includes change in Brody CF-CT subscore at Week 72 as dependent variable, treatment group as fixed effect, and sex (male vs. female) and age (<18 years vs. ≥18 years) as covariates.

The change from baseline at Week 72 for 3 Brody subscores (bronchiectasis, hyperinflation, and mucus plugging) were numerically lower for the TEZ/IVA group compared with the placebo group. The other Brody subscores (parenchymal opacities and bronchial thickening) were numerically higher for the TEZ/IVA group compared with the placebo group.

Assessors' comments

Concerning the interpretation of the total Brody/CF-CT score and of the subdomain scores a decrease of the baseline value indicates improvement with higher scores associated to more severe structural lung disease. However, the magnitude of the baseline values and of the changes from baseline are difficult to put into context in the absence of the range of possible values of the total score and subscores (which should be provided) and without any indirect comparison with the expected values in other populations of patients with cystic fibrosis in which the same scoring system may have been used.

In study 112, the mean within-group change from baseline at Week 72 in the total Brody/CF-CT Score was numerically lower for the TEZ/IVA group (+0.90, 95% CI: −3.34, 5.14) compared with the placebo group (+2.38, 95% CI: (−1.82, 6.58). The least squares (LS) mean treatment difference between the TEZ/IVA and placebo groups was −1.48 (95% CI: −7.47, 4.52) which favours the TEZ/IVA group although statistical significance is not seen. Regarding the subdomains, the within-group change from baseline at Week 72 for 3 Brody subscores (bronchiectasis, hyperinflation, and mucus plugging) were numerically lower for the TEZ/IVA group (0.74, 0.4, and -0.4 respectively)

compared with the placebo group (1.32, 1.0, and 0.4 respectively). The other Brody subscores (parenchymal opacities and bronchial thickening) were numerically higher for the TEZ/IVA group (0.1 and -0.03 respectively) compared with the placebo group (-0.1 and -0.24). Overall, these within-group changes from baseline are of limited magnitude. Similarly, the mean differences between treatments in the change from baseline at week 72 in the subscores (which range from -0.58 to 0.21) are very modest.

The discussion of the MAH to justify the 72-week treatment duration of study 112 and to give an idea of the expected changes in the CF-CT score in patients under treatment with ivacaftor is limited to two studies. De Jong and colleagues (2004) reported in a 2-year natural history study in 48 children with CF, with a mean age of 11 years and a mean ppFEV1 of 74.3 percentage points at the start of the study, a mean change in the total Brody score of +2.2 points per year. The component CT scores for bronchiectasis and mucous plugging worsened. Based on that and other longitudinal studies (which are not discussed), the MAH estimated that the progression of disease observed in the proposed study population would be a 3.3 point increase in the total Brody score at Week 72. The magnitude of the change observed in the total score in study 112 is similar to that observed in the study by Jong et al. However, it is unclear whether the score used in this study is the CF-CT or the initial score from Brody (Brody I). Furthermore, the study population are children and adolescents.

The study by Sheikh and colleagues (2015) reported in a cohort of 10 subjects (aged 10 to 44 years) who carried the CFTR-G551D mutation a within-treatment annual improvement of 13.6 points (a decrease from 28.8 pre-ivacaftor to 15.2 after one year of ivacaftor therapy) in the total Brody/CF-CT score. Bronchiectasis score decreased by 2.7, mucous plugging decreased by 5.6 points, and airway wall thickness decreased by 5 points. However, scoring in this study was based on only four CT images and not on the whole lung CT acquisition. This may have consequences given that by scoring a large number of images more abnormalities may be detected. The decrease in the bronchiectasis score observed in this study deserves particular attention as bronchiectasis are considered irreversible (with the exception perhaps of cylindrical bronchiectasis which are considered to be early-stage disease) as opposed to peribronchial thickening and mucus plugging which are known to represent reversible CT changes during exacerbations in CF patients. Due to the above (including small sample size) the results from this study should be viewed with caution.

Overall, the clinical relevance of the results observed in the CT scoring system used in study 112 is difficult to ascertain in the absence of comparative data from other studies where the same scoring system is used in similar or different study populations in terms of age, lung function etc. and lacking details about the possible range of values of the total score and subscores and how these are expressed and considered in the statistical analysis. The MAH was requested to discuss these issues. In their response, the MAH states that there is no established minimum clinically important difference (MCID) for the CF-CT score and consequently, no firm conclusions can be drawn on the clinical relevance of the CF-CT results in Study 112. Furthermore, comparisons of Study 112 to the literature are difficult without more details of how the scoring was conducted, as studies can differ with respect to the numeric score ranges and adjudication procedures. It is concluded that based on the exploratory nature of Study 112, the variability between scorers, the difficulty of comparisons to the literature, and the lack of an established MCID for the CF-CT score, the clinical relevance of the results of study 112 are not known.

In spite of the above, data from study 112 are valuable in that they contribute to the experience on the potential use of CT scoring systems to document response to therapy (i.e., as a surrogate endpoint) which may be particularly relevant in the case of agents that aim to stop or slow progression of structural lung disease such as CFTR modulators.

Absolute change in ppFEV1

The mean (SD) absolute within-group change in ppFEV1 from baseline at Week 72 was 1.2 (8.4) points in the TEZ/IVA group and −3.6 (9.0) points in the placebo group based on data from 18 and 17 patients respectively.

Assessor's comments

Mean (SD) baseline ppFEV1 was 91.4 (16.0) and 86.6 (12.7) in the TEZ/IVA and placebo groups. The mean (SD) within-group change in ppFEV1 from baseline at Week 72 was 1.2 (8.4) points in the TEZ/IVA group and −3.6 (9.0) points in the placebo group based on data from 18 and 17 patients respectively. In the pivotal study 106 the mean (SD) within-group change from baseline at week 24 was 3.4 in the TEZ/IVA group and -0.6 in the placebo group. The mean (SD) absolute within-group change at week 24 in Study 112 was 2.2 (5.8) and -0.8 (6.2) in the TEZ/IVA and placebo groups. The study population of both studies differ in that the mean baseline ppFEV1 in study 106 was 59.6 and 60.4 in the TEZ/IVA and placebo groups respectively which may have had an impact on the magnitude of the change observed in the TEZ/IVA group of study 112. The decrease seen in the placebo group at week 72 of study 112 is higher than that observed in study 106 at week 24 which may be explained by the timing of the assessment (refer to the week 24 values of ppFEV1 above quoted for study 112). The annual rate of decline of FEV1 for young adults homozygous for F508del reported by Sawicki et al (2017) is −2.52.

Absolute change in additional spirometry variables

The mean (SD) absolute within-group change in FEV1 (L) from baseline at Week 72 was 0.21 (0.35) L in the TEZ/IVA group and 0.01 (0.31) L in the placebo group.

The mean (SD) absolute within-group change in FVC (L) from baseline at Week 72 was 0.19 (0.41) L in the TEZ/IVA group and 0.07 (0.29) L in the placebo group. These figures for percent predicted FVC were -0.6 (7.1) and -2.4 (5.9) points in the TEZ/IVA and placebo groups, respectively.

The mean (SD) absolute within-group change in Forced expiratory flow rate 25-75% (L/s) from baseline at Week 72 was 0.31 (0.50) in the TEZ/IVA group and -0.06 (0.61) in the placebo group.

The mean (SD) absolute within-group change in the ratio of FEV1/FVC was 0.02 (0.03) in the TEZ/IVA group and -0.02 (0.05) in the placebo group. These figures for Percent predicted ratio of FEV1 to FVC (%) were 1.9 (3.5) and -1.5 (5.7) points in the TEZ/IVA and placebo groups, respectively.

Assessor's comments

For all spirometry variables data are only available for 17 and 18 subjects at week 72 in the TEZ/IVA and placebo groups respectively. Within-group changes have been calculated regardless of the fact that at baseline data from 20 and 21 subjects were available. The impact of the lacking data at week 72 on the within-group changes is not addressed by the MAH. This is not requested to be done by the assessor as Study 112 is an exploratory study and the main interest relies on the CT scoring system for which complete set of data are available.

Cystic Fibrosis Questionnaire-Revised

Summary statistics and within-group change from baseline for the CFQ-R respiratory domain to Week 72 are provided in table below.

Table 11-5 ANCOVA Analysis of Absolute Change From Baseline in CFQ-R Respiratory Domain Score at Week 72, Full Analysis Set

Parameter	TEZ/IVA N = 20	Placebo N = 21
Baseline Score		
n	20	21
Mean (SD)	76.8 (13.3)	82.4 (11.7)
Median	80.6	83.3
Min, Max	50.0, 94.4	61.1, 100.0
Week 72		
n	20	20
Mean (SD)	75.1 (21.3)	77.8 (13.8)
Median	76.4	77.8
Min, Max	33.3, 100.0	44.4, 100.0
Absolute Change from Baseline at Week 72 (points)		
n	20	20
Mean (SD)	−1.7 (15.6)	−4.6 (15.8)
Median	0.0	0.0
Min, Max	−38.9, 22.2	−33.3, 22.2

Source: Table 14.2.3.1

ANCOVA: analysis of covariance; CFQ-R: Cystic Fibrosis Questionnaire–Revised; IVA: ivacaftor; n: size of subsample; N: total sample size; TEZ: tezacaftor

Note: Baseline was defined as the most recent non-missing measurement before the first dose of study drug.

Decreases in absolute change in mean from baseline at Week 72 in CFQ-R respiratory domain score were observed in the TEZ/IVA (−1.7 points) and placebo (−4.6 points) groups.

Assessor's comments

The Cystic Fibrosis Questionnaire-Revised (CFQ-R) is a CF–specific instrument that measures health-related quality of life (HRQOL). Three different versions were used in Study 112, i.e., subjects who were 12 and 13 years of age at the time of questionnaire completion completed the CFQ-R Child version (self-report, evaluating 8 domains), and their parents/caregivers completed the CFQ-R Parent version (evaluating 11 domains). Subjects 14 years of age and older at the time of questionnaire completion completed the CFQ-R Adolescent/Adult version of the questionnaire (self-reported). The teen/adult version assesses 12 domains: physical functioning, role, vitality, emotional functioning, social, body image, eating disturbances, treatment burden, health perceptions, weight, respiratory symptoms, and digestive symptoms. Each domain is composed of a variable number of self-report questions with four possible answers, with a total of 50 questions. The scores range from 0 to 100, with higher scores indicating a higher patient-reported quality of life with regard to the domain being evaluated.

For both the respiratory domain and for the non-respiratory domains (see below) summary statistics in this report are presented for the pooled Children Ages 12 and 13 and Adolescents and Adults Versions of the CFQ-R which may lead to the lost of relevant information as the domains assessed are not exactly the same in each version of the questionnaire. In this respect the results of 11 domains are presented under the non-respiratory domains of CFQ-R (see below) but for some of them data at baseline and at week 72 are available only for approximately 30 patients out of the 41 enrolled in Study 112 which seems to be the consequence of the specific questionnaire not asking questions about these domains for children aged 12 and 13 years old (vitality, health perceptions, role, and weight).

In Study 112, mean (SD) baseline CFQ-R respiratory domain score were 76.8 (13.3) and 82.4 (11.7) points. The mean (SD) within-group change in CFQ-R respiratory domain score from baseline at Week 72 was -1.7 (15.6) in the TEZ/IVA group and -4.6 (15.8) in the placebo group. When compared to results of the pivotal study 106, the within-group change at week 24 in the TEZ/IVA group was 5.0 points in the TEZ/IVA group and -0.1 points in the placebo group. Mean baseline values in study 106 were 70.1 (16.8) and 69.9 (16.6) in the TEZ/IVA and placebo groups. At week 24 of study 112 the within-group change was -1.9 (10.5) and -6.1 (18.7) points in the TEZ/IVA and placebo groups. It is likely again that differences in severity between study populations and the different treatment duration may partially explain the within-group changes in the respiratory domain score of CFQ-R in Study 112 which are comparatively lower than in Study 106 something to be expected at week 72 but less so at week 24.

Non-respiratory domains of CFQ-R

Summary statistics and within-group changes from baseline to week 72 in the CFQ-R non-respiratory domains are provided in table below.

Summary Statistics for non-respiratory domains CFQ-R, Children Ages 12 and 13 and Adolescents and Adults Versions, FAS

(Source: Table 14.2.3.1 CSR)

Parameter **Mean (SD)** **Median** **Min, Max**	**TEZ/IVA (N=20)**	**Placebo (BL, N=21; W72, N=20)**
CFQ-R Domain: Physical		
Baseline Score	87.2 (18.4) 93.1 29.2, 100.0	90.5 (11.1) 95.8 58.3, 100.0
Week 72 Score	86.0 (19.8) 100.0 45.8, 100.0	86.0 (15.7) 89.6 50.0, 100.0
Absolute Change from BL at W72	-1.3 (14.4) 0.0 -37.5, 16.7	-4.0 (8.7) -2.1 -27.8, 5.6
CFQ-R Domain: Vitality		
Baseline Score	60.6 (19.0) 66.7 25.0, 83.3 *(n=15)*	58.9 (18.9) 58.3 33.3, 83.3 *(n=16)*
Week 72 Score	60.0 (23.0) 66.7 25.0, 100.0	58.9 (17.4) 50.0 33.3, 91.7
Absolute Change from BL at W72	-0.6 (17.7) 0.0 -33.3, 41.7 (n=15)	-0.6 (16.2) 0.0 -41.7, 16.7 (n=15)
CFQ-R Domain: Emotion		
Baseline Score	82.5 (12.4) 86.7 54.2, 100.0	80.5 (13.5) 80.0 53.3, 100.0
Week 72 Score	79.5 (18.6) 83.3 45.8, 100.0	78.3 (16.6) 81.7 46.7, 100.0
Absolute Change from BL at W72	-2.9 (15.1) -2.1	-3.5 (13.6) 0.0

	-40.0, 26.7	-26.7, 13.3
CFQ-R Domain: Body		
Baseline Score	86.7 (14.2) 88.9 55.6, 100.0	79.4 (23.1) 88.9 22.2, 100.0
Week 72 Score	91.1 (14.2) 100.0 55.6, 100.0	75.6 (25.9) 77.8 22.2, 100.0
Absolute Change from BL at W72	4.4 (17.1) 0.0 -44.4, 33.3	-3.3 (8.9) 0.0 -22.2, 11.1
CFQ-R Domain: Eat		
Baseline Score	95.0 (12.2) 100.0 55.6, 100.0	92.6 (14.6) 100.0 44.4, 100.0
Week 72 Score	87.8 (19.7) 100.0 33.3, 100.0	96.1 (12.1) 100.0 55.6, 100.0
Absolute Change from BL at W72	-7.2 (16.6) 0.0 -66.7, 11.1	3.9 (7.5) 0.0 0.0, 22.2
CFQ-R Domain: Treatment burden		
Baseline Score	67.8 (21.9) 77.8 11.1, 100.0	64.8 (18.4) 55.6 33.3, 100.0
Week 72 Score	67.2 (24.3) 66.7 11.1, 100.0	65.0 (19.8) 66.7 22.2, 100.0
Absolute Change from BL at W72	-0.6 (17.5) 0.0 -55.6, 22.2	0.8 (20.4) 0.0 -33.3, 55.6
CFQ-R Domain: Health perception		
Baseline Score	77.8 (22.2) 77.8 33.3, 100.0 (n=15)	68.1 (22.9) 66.7 22.2, 100.0 (n=16)
Week 72 Score	79.3 (20.1) 88.9 33.3, 100.0	66.7 (17.8) 66.7 33.3, 88.9
Absolute Change from BL at W72	1.5 (18.7) 0.0 -22.2, 44.4 (n=15)	-3.0 (19.0) 0.0 -44.4, 44.4 (n=15)
CFQ-R Domain: Weight		
Baseline Score	95.6 (17.2) 100.0 33.3, 100.0 (n=15)	89.6 (26.4) 100.0 0.0, 100.0 (n=16)
Week 72 Score	93.3 (18.7) 100.0 33.3, 100.0	93.3 (18.7) 100.0 33.3, 100.0
Absolute Change from BL at W72	-2.2 (8.6) 0.0 -33.3, 0.0 (n=15)	4.4 (17.2) 0.0 -33.3, 33.3 (n=15)
CFQ-R Domain: Digestion		
Baseline Score	82.2 (22.3) 88.9 22.2, 100.0	81.0 (21.1) 88.9 33.3, 100.0
Week 72 Score	80.0 (17.5) 83.3	80.0 (18.2) 88.9

	44.4, 100.0	33.3, 100.0
Absolute Change from BL at W72	-2.2 (19.6) 0.0 -44.4, 44.4	0.0 (20.4) 0.0 -33.3, 33.3
CFQ-R Domain: Role		
Baseline Score	89.4 (8.6) 91.7 75.0, 100.0 (n=15)	85.4 (14.1) 91.7 58.3, 100.0 (n=16)
Week 72 Score	82.8 (14.9) 83.3 58.3, 100.0	82.8 (14.9) 91.7 58.3, 100.0
Absolute Change from BL at W72	-6.7 (11.0) 0.0 -33.3, 8.3 (n=15)	-1.7 (13.8) 0.0 -41.7, 16.7 (n=15)
CFQ-R Domain: Social		
Baseline Score	77.9 (15.5) 77.8 50.0, 100.0	75.0 (17.9) 72.2 33.3, 100.0
Week 72 Score	76.8 (18.3) 77.8 44.4, 100	75.0 (17.9) 77.0 38.9, 100.0
Absolute Change from BL at W72	-1.2 (10.6) 0.0 -22.2, 22.2	-0.2 (10.3) 0.0 -16.7, 23.8

Assessor's comments

Overall, the mean within-group changes in the non-respiratory domains of CFQ-R are higher in the TEZ/IVA group than in the placebo group (even if a decrease from baseline is observed for most of them). However, the following ones deserve to be commented. Regarding the "Eating" domain, the mean (SD) absolute within-group change at week 72 was -7.2 (16.6) in the TEZ/IVA group and 3.9 (7.5) points in the placebo group. Similarly, in the "Role" domain the mean (SD) change was -6.7 (11.0) in the TEZ/IVA group and -1.7 (13.8) in the placebo group. Higher mean decreases (although of less magnitude than the previous ones) are also seen in the TEZ/IVA for the domains "Weight" and "Digestion". 3.

The MAH was requested to discuss this issue in terms of whether TEZ/IVA may decrease appetite or produce other effects that may explain the results of the CFQ-R domains "Eating", "Weight", "Digestion" and "Role" domains. No further insight in this respect was provided however. It was further emphasised that study 112 was exploratory in nature and that in the pivotal study 106 no significant differences between the placebo and TEZ/IVA groups at week 24 in any of the 4 CFQ-R domains of interest (eating, weight, digestion, and role) were seen either.

Measures of nutritional status

BMI and BMI-for-age z score

The mean (SD) absolute change in BMI (kg/m^2) from baseline at Week 72 was 0.48 (1.31) in the TEZ/IVA group and 0.61 (1.06) in the placebo group. These figures for BMI-for-age z score (subjects < 20 years old) were 0.21 (0.34) and 0.09 (0.43) respectively.

Weight and weight-for-age z score

The mean (SD) absolute change in weight (kg) from baseline at Week 72 was 2.40 (5.18) in the TEZ/IVA group and 3.00 (3.92) in the placebo group. These figures for weight-for-age z score (subjects < 20 years old) were 0.15 (0.41) and 0.06 (0.42) respectively.

Height and height-for-age z score

The mean (SD) absolute within-group change in height (cm) from baseline at Week 72 was 3.08 (3.94) in the TEZ/IVA group and 2.44 (2.85) in the placebo group. These figures for height-for-age z score (subjects < 20 years old) were −0.17 (0.50) and −0.04 (0.40) respectively.

Assessor's comments

In terms of weight, height, BMI and their z scores, the mean within-groups changes in both groups are very modest. The pivotal study 106 did not show any meaningful effect of TEZ/IVA on body weight or BMI either.

Data at week 72 for height are only available for 12 and 18 subjects in the TEZ/IVA and placebo groups respectively which limit the reliability of the calculated change from baseline at week 72 for this parameter. If height data were not available at week 72 for a number of subjects is unclear how BMI could be calculated for 20 subjects in each group at week 72. The MAH was requested to clarify this issue. In response, it has been clarified that height values were measured for all subjects at screening, and at most visits thereafter (excluding the Day 15 visit due to the short interval) for as long as the subject was ≤21 years old. Per the protocol, once a subject was >21 years of age, height was no longer collected because subjects will have stopped growing by age 21. For subjects >21 years old at the Week 72 visit, a height value collected at an earlier visit was used to calculate body mass index (BMI) explaining why all subjects had a height measurement to calculate BMI at Week 72.

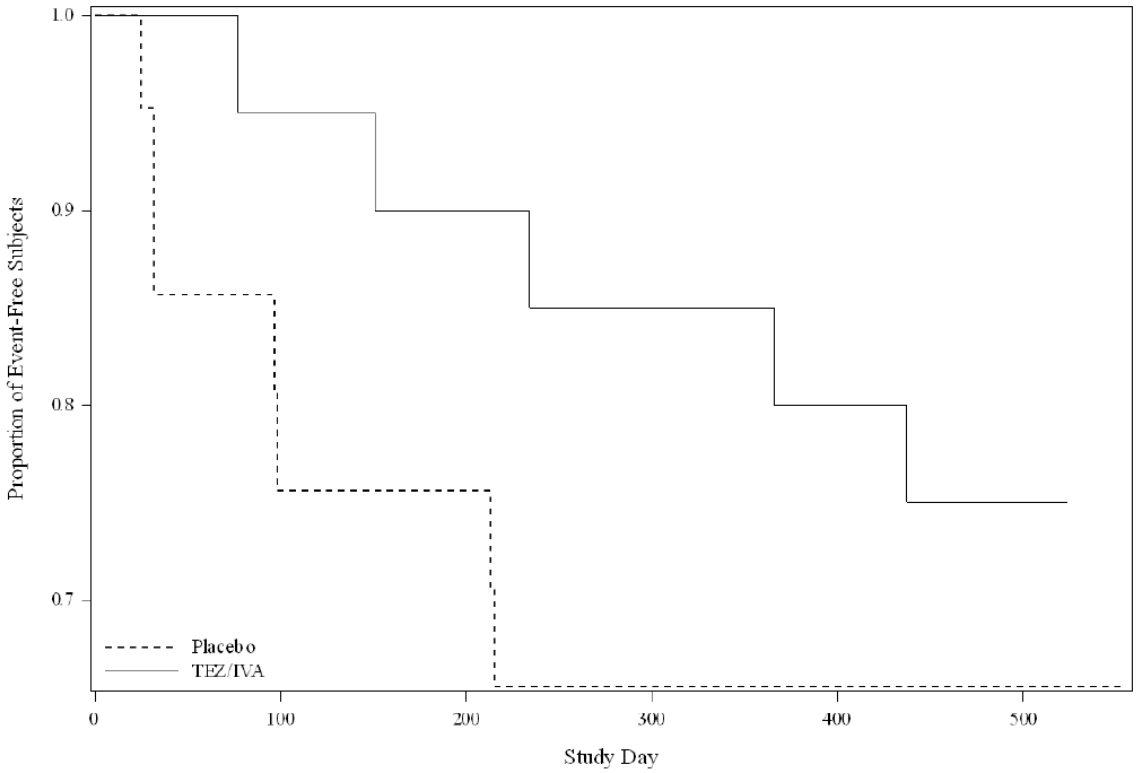
At baseline, z scores were calculated for 10 subjects in the TEZ/IVA group and 13 in the placebo group (less than 20 years old). At week 72, data were available for 8 and 11 subjects respectively.

Pulmonary exacerbations and hospitalization

The observed rate of PEx was 0.59 events per year in the TEZ/IVA group and 0.75 events per year in the placebo group. The rate of PEx requiring hospitalization or IV antibiotics was 0.17 events per year in the TEZ/IVA group and 0.52 events per year in the placebo group.

Kaplan-Meier plot for time-to-first IV antibiotics or hospitalization for pulmonary exacerbation is shown below.

Figure 11-2: Kaplan-Meier Plot for Time-to-First IV Antibiotics or Hospitalization for Pulmonary Exacerbation, Full Analysis Set



Source: Figure 14.2.5.5.1
Note: Pulmonary Exacerbation = new or change in antibiotic therapy for ≥4 sinopulmonary signs/symptoms.

The observed rate of unplanned hospitalisations was 0.17 events per year in the TEZ/IVA group and 0.23 events per year in the placebo group.

Assessor's comments

The event rate is calculated as the Total number of events x 336 / Total number of days on study; (48 weeks = 336 days). To obtain an annualised rate, the total number of years is obtained by dividing the total number of days by 336 days (which corresponds to 12 months of 28 days). An alternative approach would have been dividing the period in days by 365.25 days which is how it is usually done for long periods. However, for short periods (as it is the case of the present study) it is preferable that the period used is the one corresponding to the study.

SF-36

Summary statistics and within-group changes from baseline to week 72 in the SF-36 (Version 2) domains are provided in table below.

Summary statistics for SF-36 domains (norm-based scores), FAS

(Source: Table 14.2.7 CSR)

Parameter **Mean (SD)** **Median** **Min, Max**	**TEZ/IVA (N=20)**	**Placebo (BL, N=21; W72, N=20)**
SF-36 Domain: *Physical functioning*		
Baseline Score	55.5 (3.7) 57.6 42.2, 57.6	55.6 (3.4) 57.6 46.0, 57.6
Week 72 Score	55.7 (3.6) 57.6 46.0, 57.6	55.5 (3.1) 56.6 48.0, 57.6
Absolute Change from BL at W72	0.2 (3.0) 0.0 -7.7, 5.8	0.0 (2.1) 0.0 -5.8, 3.9
SF-36 Domain: *Role physical scale*		
Baseline Score	52.6 (5.6) 54.9 41.7, 57.1	53.9 (6.4) 57.1 28.5, 57.1
Week 72 Score	51.7 (7.4) 56.0 32.9, 57.1	53.9 (4.5) 56.0 43.9, 57.1
Absolute Change from BL at W72	-0.9 (6.8) 0.0 -15.4, 15.4	0.2 (5.5) 0.0 -13.2, 15.4
SF-36 Domain: *Bodily pain scale*		
Baseline Score	53.8 (6.5) 54.6 45.9, 60.9	54.3 (6.0) 54.6 45.5, 60.9
Week 72 Score	54.3 (7.5) 57.7 41.5, 60.9	52.8 (7.5) 54.6 37.6, 60.9
Absolute Change from BL at W72	0.4 (8.2) 0.0 -19.3, 15.0	-1.5 (8.1) 0.0 -19.3, 9.1
SF-36 Domain: *General health scale*		
Baseline Score	47.0 (10.2) 47.5 28.8, 63.2	47.2 (9.8) 48.6 28.8, 63.2
Week 72 Score	47.8 (9.3) 48.6 27.9, 61.0	44.6 (9.5) 44.2 27.9, 65.4
Absolute Change from BL at W72	0.8 (8.3) 1.1 -16.8, 23.4	-2.7 (6.2) -2.2 -17.6, 7.9
SF-36 Domain: *Vitality scale*		
Baseline Score	52.5 (10.2) 54.2 33.8, 69.2	51.0 (9.3) 50.1 31.1, 69.2
Week 72 Score	52.4 (12.5) 55.5 31.1, 69.2	50.5 (10.8) 50.1 31.1, 69.2
Absolute Change from BL at W72	-0.1 (7.4) -1.4 -13.6, 13.6	-0.7 (8.3) 0.0 -16.3, 19.1
SF-36 Domain: *Social functioning scale*		
Baseline Score	53.3 (4.8) 56.7	52.0 (8.3) 56.7

	41.9, 56.7	27.1, 56.7
Week 72 Score	49.3 (10.3) 56.7 27.1, 56.7	52.5 (7.2) 56.7 32.0, 56.7
Absolute Change from BL at W72	-4.0 (8.4) 0.0 -24.7, 9.9	0.0 (7.9) 0.0 -24.7, 14.8
SF-36 Domain: *Role emotional scale*		
Baseline Score	52.2 (5.9) 55.6 36.6, 55.6	51.5 (7.3) 55.6 28.9, 55.6
Week 72 Score	51.6 (7.6) 55.6 25.1, 55.6	49.0 (10.6) 55.6 21.3, 55.6
Absolute Change from BL at W72	-0.6 (5.3) 0.0 -11.5, 11.5	-3.6 (9.1) 0.0 -30.5, 3.8
SF-36 Domain: *Mental health scale*		
Baseline Score	53.6 (7.8) 55.2 35.4, 62.7	51.9 (8.0) 55.2 32.9, 62.7
Week 72 Score	53.4 (10.1) 56.5 30.5, 62.7	51.5 (9.5) 52.8 28.0, 62.7
Absolute Change from BL at W72	-0.2 (7.7) 0.0 -17.4, 19.8	-1.4 (6.5) -2.5 -12.4, 9.9

Assessor's comments

The SF-36 questionnaire is a generic tool (i.e., not disease-specific) which includes 36 questions divided into 8 categories: Physical Functioning (PF), Role Physical (RF), Bodily Pain (BP), Social functioning (SF), Mental Health (MH), Role Emotional (RE), Vitality (VT) and General Health Perceptions (GH). Physical Component Summary (PCS) is assessed by grouping all physical components (PF, RP, BP and VT) together; similarly, the Mental Component Summary (MCS) encompasses mental components, such as SF, RE, MH and GH. Each answer is assigned a certain number of points; the score obtained in a given category may equal from 0 to 100 points. It is widely accepted that the lower the score, the worse the quality of life.

In this report norm-based scores are presented as it is generally accepted that their interpretation is simpler. Norm-based scoring equates all scores, so scores above 50 are better than the general population average for all scales and summary measures, while scores below 50 are worse. Normative data from the Australian population are available. Because health status is generally related to age and gender (with females usually reporting poorer health status), normative values are often presented separately for specific age and gender groups. This is not done in the present study. Taking into account the main objective of Study 112 providing mean changes in SF-36 scores sorted out by sex and age will not be requested.

Safety results

A total of 41 subjects received at least 1 dose of study drug during the Treatment Emergent (TE) Period. The mean duration of exposure was 72.4 weeks in the TEZ/IVA group and 69.4 weeks in the

placebo group. The majority of subjects received >72 weeks of treatment, including 11 (55.0%) subjects in the TEZ/IVA group and 13 (61.9%) subjects in the placebo group.

Table below provides summary statistics for study drug exposure.

Table 12-1: Summary of Exposure, Safety Set

	TEZ/IVA N = 20	Placebo N = 21	Total N = 41
Total exposure (patient years[a])	27.8	27.9	55.7
Exposure duration (years)[b]			
n	20	21	41
Mean (SD)	1.4 (0.0)	1.3 (0.3)	1.4 (0.2)
Median	1.4	1.4	1.4
Min, Max	1.3, 1.6	0.2, 1.5	0.2, 1.6
Exposure duration (weeks)			
n	20	21	41
Mean (SD)	72.4 (2.2)	69.4 (13.9)	70.9 (10.0)
Median	72.1	72.4	72.3
Min, Max	69.3, 81.1	9.9, 79.0	9.9, 81.1
Exposure duration category (weeks) n (%)			
>0 and ≤2	0	0	0
>2 and ≤4	0	0	0
>4 and ≤12	0	1 (4.8)	1 (2.4)
>12 and ≤24	0	0	0
>24 and ≤36	0	0	0
>36 and ≤48	0	0	0
>48 and ≤60	0	0	0
>60 and ≤72	9 (45.0)	7 (33.3)	16 (39.0)
>72	11 (55.0)	13 (61.9)	24 (58.5)

Source: Table 14.1.10

IVA: ivacaftor; n: size of subsample; N: total sample size; TEZ: tezacaftor

[a] Patient years = Total of Duration of study drug exposure (days) of all subjects/365.25.

[b] Duration of study drug exposure (years) = (last dose date – first dose date + 1)/365.25, regardless of any study drug interruption.

Treatment-emergent adverse events (AE)

Treatment-emergent AEs were defined as any AE that increased in severity or that was newly developed at or after the first dose of study drug through the end of TE Period.

Table below summarizes the percentage of subjects with AEs.

Table 12-2 Overview of Adverse Events, Safety Set

	TEZ/IVA N = 20 n (%)	Placebo N = 21 n (%)	Total N = 41 n (%)
Number of AEs (Total)	118	221	339
Subjects with any AEs	20 (100.0)	21 (100.0)	41 (100.0)
Subjects with AEs by strongest relationship			
Related	0	0	0
Possibly related	7 (35.0)	7 (33.3)	14 (34.1)
Unlikely related	7 (35.0)	12 (57.1)	19 (46.3)
Not related	6 (30.0)	2 (9.5)	8 (19.5)
Subjects with AEs by maximum severity			
Mild	5 (25.0)	2 (9.5)	7 (17.1)
Moderate	12 (60.0)	11 (52.4)	23 (56.1)
Severe	3 (15.0)	8 (38.1)	11 (26.8)
Life-threatening	0	0	0
Subjects with AE leading to treatment discontinuation	0	1 (4.8)	1 (2.4)
Subjects with AE leading to treatment interruption	2 (10.0)	4 (19.0)	6 (14.6)
Subjects with grade 3/4 AEs	3 (15.0)	8 (38.1)	11 (26.8)
Subjects with related AEs	7 (35.0)	7 (33.3)	14 (34.1)
Subjects with serious AEs	8 (40.0)	13 (61.9)	21 (51.2)
Subjects with related serious AEs	2 (10.0)	2 (9.5)	4 (9.8)
Subjects with AE leading to death	0	0	0

Source: Table 14.3.1.1

AE: adverse event; IVA: ivacaftor; MedDRA: Medical Dictionary for Regulatory Activities; n: size of subsample; N: total sample size; PT: preferred term; SAE: serious AE; TEZ: tezacaftor

Notes: AEs were coded using MedDRA version 21.0. When summarizing number of events, a subject with multiple events within a category is counted multiple times in that category. When summarizing number and % of subjects, a subject with multiple events within a category is counted only once in that category.

AEs were reported for all subjects. The majority of subjects had AEs that were considered either mild (17.1%) or moderate (56.1%) in severity. A total of 3 (15.0%) subjects in the TEZ/IVA group and 8 (38.1%) subjects in the placebo group had severe AEs. No subject had a life-threatening (i.e., Grade 4) event. There were no deaths.

The most common AEs by SOC (occurring in ≥30% of subjects in any treatment group) were infections and infestations (85.4%); respiratory, thoracic, and mediastinal disorders (70.7%); investigations (43.9%), and gastrointestinal disorders (31.7%), and were consistent with the expected manifestations of CF disease.

The most common AEs by PT (occurring in ≥20% of subjects overall) were infective PEx of CF (51.2%) and cough (41.5%) and were consistent with the expected manifestations of CF disease.

Assessor's comments

A total of 41 subjects received at least 1 dose of study drug. Nine subjects (45.0%) in the TEZ/IVA group and 7 (33.3%) in the placebo group were exposed for >60 and ≤72 weeks. Eleven (55.0%) and 13 (61.9%) in the TEZ/IVA and placebo groups were exposed for >72 weeks up to a maximum of 81.1 weeks.

All subjects reported at least an adverse event but more subjects in the placebo group reported an adverse event (221 events) than in the TEZ/IVA group (118).

Very common adverse events (≥1/10) in the TEZ/IVA group (n=20) were infective pulmonary exacerbations of cystic fibrosis (9 subjects, 45.0%), followed by cough (8 subjects, 40.0%). Haemoptysis, productive cough, and upper respiratory tract infection were reported at a frequency of 20.0% each (4 subjects each) while the frequency of reporting of lower respiratory tract infection bacterial, lung infection pseudomonal, nasopharyngitis, and pyrexia was 15% each (3 subjects each). Gastroenteritis, influenza, migraine, oropharyngeal pain, pharyngitis, sputum increased, and abdominal pain were reported at 10% each (2 subjects each).

In the placebo group, the most frequently reported adverse events were infective pulmonary exacerbations of cystic fibrosis (12 subjects, 57.1%) followed by cough (9, 42.9%), bacterial test positive (5, 23.8%), productive cough, nausea, vomiting and headache (4, 19.0% each), and sunburn (3, 14.3%). Haemoptysis, upper respiratory tract infection, pyrexia, oropharyngeal pain, abdominal pain, fatigue, fungal test positive, rhinorrhea, upper respiratory tract congestion, blood alkaline phosphatase increased, chest pain, constipation, and epistaxis were reported at a frequency of 9.5% (2 subjects).

From a quantitative point of view it would appear that preferred terms that are related to the same event are split diluting the frequency of reporting. As an example, a single subject reported a headache in the TEZ/IVA group (5.0%) while 4 subjectsreported in the placebo group (19%). However, in the TEZ/IVA group three additional subjects reported tension headache (5.0%) and migraine (2 events, 10%). If these 4 subjects were counted together the frequency of reporting would have been 20% in the TEZ/IVA group (instead of 5%). Regarding abdominal pain (which is listed in section 4.8 of the SmPC of Symkevi with a frequency of very common), two preferred terms were used, i.e., "abdominal pain" and "abdominal pain upper". Counted together the frequency of abdominal pain (any) would have been 10% in the TEZ/IVA group (instead of 5%) and 14.3% in the placebo group (instead of 9.5%). The MAH was requested to justify the strategy used for the analysis of adverse events i.e. splitting up the preferred terms that are related to the same AE (e.g. headache, abdominal pain), and provide the incident rates for AEs of those as harmonised with the section 4.8 of the SmPC of Symkevi. In this respect, the MAH states that a standard approach for the analysis and tabulation of adverse events (AEs) in clinical studies is used which includes presentation of AE incidences in one table by Preferred Term (PT) and in another table by System Organ Class (SOC) and PT. The SOCs and PTs were coded using MedDRA Version 21.0, the most recent version available during the analysis of Study 112. This approach is consistent with how the AEs are presented in the Symkevi Summary of Product Characteristics (SmPC) Section 4.8 where frequency is presented by individual PT ("adverse reaction" column), which is equivalent to how the incidence data in Study 112 are presented. While the response provided may be acceptable in that no changes in the SmPC are proposed or requested (taking into account the sample size of study 112), it remains at the discretion of assessors to ask for the frequency of reporting of preferred terms that are related and/or to request Standardised MedDRA Queries (SMQs) to gather further information on certain events (e.g. acute pancreatitis based on the clustering of various preferred terms).

Within the SOC "Infections and infestations", more events were reported in the TEZ/IVA group (19, 95%) than in the placebo group (16, 76.2%). Nine events (45.0%) of infective pulmonary exacerbations of cystic fibrosis occurred in the TEZ/IVA group and 12 (57.1%) in the placebo group. Other adverse events such as upper and lower respiratory tract infection, lung infection pseudomonal, nasopharyngitis, and pharyngitis were observed more frequently in the TEZ/IVA group.

Adverse events related to the increase of transaminases and of other liver function tests (within the SOC "Investigations") are described by several preferred terms such as aspartate aminotransferase increased, transaminases increased, alanine aminotransferase increased, blood bilirubin increased,

blood alkaline phosphatase increased, and gamma-glutamyltransferase increased. In the TEZ/IVA group events of aspartate aminotransferase increased, transaminases increased, and blood bilirubin increased were reported (5.0% each). In the placebo group, single events of aspartate and alanine aminotransferase increased and gamma-glutamyltransferase increased were reported (4.8% each) as well as 2 events of blood alkaline phosphatase increased (9.5%). The discussion on liver function tests (refer to laboratory abnormalities) shows that 8 patients in the TEZ/IVA group had ALT or AST elevations that were >ULN to ≤5 x ULN in spite of which only 2 events of aspartate aminotransferase increased and transaminases increased are described. For additional clarification in relation to this issue, please refer to Laboratory abnormalities further below.

According to severity, Grade 3 AEs occurred in 11 subjects (3 subjects in the TEZ/IVA group and 8 subjects in the placebo group). The most common Grade 3 AE was infective PEx of CF, which occurred in 2 (10.0%) subjects in the TEZ/IVA group and 2 (9.5%) subjects in the placebo group. All other Grade 3 AEs occurred in the placebo group, except for 1 AE of testicular torsion in the TEZ/IVA group.

Assessor's comments

The majority of subjects had AEs that were considered either mild (17.1%) or moderate (56.1%) in severity. A total of 3 (15.0%) subjects in the TEZ/IVA group and 8 (38.1%) subjects in the placebo group had severe AEs. Severe adverse events reported in the TEZ/IVA group were infective pulmonary exacerbation of cystic fibrosis (2 events, 10%) and testicular torsion (an event, 5%). In the placebo group, severe adverse events were as follows: gastroenteritis (an event, 4.8%), haemoptysis (4.8%), alanine aminotransferase increased (4.8%), gamma-glutamyltransferase increased (4.8%), abdominal pain (4.8%), nausea (4.8%), vomiting (4.8%), hyperglycemia (4.8%). No subject had a life-threatening (i.e., Grade 4) event. There were no deaths in the study.

Related adverse events

The number (%) of subjects with any related AEs (including related, possibly related, and missing AEs) was 7 (35.0) and 7 (33.3) in the TEZ/IVA and placebo groups respectively. These subjects reported 12 and 11 events respectively. Within the SOC "Investigations", 4 events were reported in the TEZ/IVA group which included aspartate aminotransferase increased, transaminase increased, ultrasound liver abnormal, and blood triglycerides increased (each reported at a frequency of 5%) while in the placebo group three related events of blood alkaline phosphatase increased (2 events, 9.5%) and electrocardiogram PR shortened (4.8%) were reported.

Two events (one per group) of infective pulmonary exacerbation of cystic fibrosis (5% and 4.8% in the TEZ/IVA and placebo groups respectively) were considered treatment-related. Other related adverse events reported in the TEZ/IVA group were as follows: atrioventricular block first degree (5%), sunburn (5%), dizziness (5%), testicular torsion (5%), and haemoptysis (5%).

Serious adverse events (SAEs)

A total of 21 (51.2%) subjects had SAEs during the TE period: 8 (40.0%) subjects in the TEZ/IVA group, and 13 (61.9%) subjects in the placebo group. Ten serious adverse events were reported in the TEZ/IVA group while 17 were reported in the placebo group.

By PT, the most common SAE overall was infective PEx of CF, which occurred in 5 (25.0%) subjects in the TEZ/IVA group and 6 (28.6%) subjects in the placebo group. The only other SAE that occurred in more than 1 subject in any treatment group was lung infection pseudomonal (2 subjects in the

TEZ/IVA group and no subjects in the placebo group). All other SAEs occurred in 1 subject only. The majority of SAEs had an outcome of recovered/resolved.

Overall, a total of 4 (9.8%) subjects had SAEs that were considered by the investigator to be related or possibly related to study drugs, including 2 (10.0%) subjects in the TEZ/IVA group and 2 (9.5%) subjects in the placebo group. In the TEZ/IVA group, 1 subject had a first-degree atrioventricular block and 1 subject had testicular torsion. In the placebo group, 1 subject had hepatic cirrhosis and 1 subject had infective pulmonary exacerbation of CF.

Adverse events leading to study drug discontinuation or interruption

No subject in the TEZ/IVA group had an AE that led to treatment discontinuation. The subject in the placebo group that had hepatic cirrhosis had treatment discontinued but completed all study visits. A total of 6 (14.6%) subjects had AEs that led to treatment interruption: 2 (10.0%) subjects in the TEZ/IVA group and 4 (19.0%) subjects in the placebo group. The adverse events in the placebo group that led to treatment interruption were alanine aminotransferase increased (4.8%), gamma-glutamyltransferase increased (4.8%), nausea (4.8%), vomiting (4.8%), gastroenteritis (4.8%), dyspnea (4.8%), and respiration abnormal (4.8%). In the TEZ/IVA group, the adverse events that led to treatment interruption were atrioventricular block first degree (5%) and blood bilirubin increased (5%).

Assessor's comments

Study drug was interrupted in 2 (10%) subjects in the TEZ/IVA group due to atrioventricular block first degree (a subject, 5%) and blood bilirubin increased (5%). All other study drug interruptions (4 subjects, 19.0%) occurred in the placebo group. Of note, blood bilirubin increased is not reported as a related adverse event.

Laboratory abnormalities (Liver Function Tests)

AEs associated with LFTs included: AST increased (TEZ/IVA: 1 subject; placebo: 1 subject); ALT increased (TEZ/IVA: no subjects; placebo: 1 subject); blood bilirubin increased (TEZ/IVA: 1 subject; placebo: no subjects); transaminases increased (TEZ/IVA: 1 subject; placebo: no subjects); blood alkaline phosphatase increased (TEZ/IVA: no subjects; placebo: 2 subjects); GGT increased (TEZ/IVA: no subjects; placebo: 1 subject); and abnormal ultrasound of the liver (TEZ/IVA: 1 subject; placebo: no subjects). All of the AEs associated with LFTs were classified as mild or moderate, with the exception of 1 subject in the placebo group with 2 AEs associated with LFTs classified as severe. Most of these AEs resolved, with the exception of 1 TEZ/IVA subject with increased transaminase, the TEZ/IVA subject with an abnormal liver ultrasound, and 1 placebo subject with blood alkaline phosphatase increase. None of them led to study drug discontinuation. One subject in the TEZ/IVA group with an increase in blood bilirubin, and 1 subject in the placebo group with 2 AEs associated with LFTs, both Grade 3 or higher, had treatment interrupted.

Liver Function Test results that met threshold criteria are summarized in table below.

Table 12-9 Threshold Analysis of LFTs During the Treatment-Emergent Period, Safety Set

Parameter Threshold Analysis Criteria n/N1 (%)	TEZ/IVA N = 20	Placebo N = 21	Total N = 41
ALT			
>ULN to ≤3 × ULN	6/20 (30.0)	8/21 (38.1)	14/41 (34.1)
>3 to ≤5 × ULN	0/20	3/21 (14.3)	3/41 (7.3)
>5 to ≤8 × ULN	0/20	0/21	0/41
>8 to ≤20 × ULN	0/20	0/21	0/41
>20 × ULN	0/20	0/21	0/41
AST			
>ULN to ≤3 × ULN	5/20 (25.0)	7/21 (33.3)	12/41 (29.3)
>3 to ≤5 × ULN	1/20 (5.0)	1/21 (4.8)	2/41 (4.9)
>5 to ≤8 × ULN	0/20	0/21	0/41
>8 to ≤20 × ULN	0/20	0/21	0/41
>20 × ULN	0/20	0/21	0/41
ALT or AST			
>ULN to ≤3 × ULN	7/20 (35.0)	9/21 (42.9)	16/41 (39.0)
>3 to ≤5 × ULN	1/20 (5.0)	3/21 (14.3)	4/41 (9.8)
>5 to ≤8 × ULN	0/20	0/21	0/41
>8 to ≤20 × ULN	0/20	0/21	0/41
>20 × ULN	0/20	0/21	0/41
ALP			
>ULN to ≤1.5 × ULN	4/20 (20.0)	4/21 (19.0)	8/41 (19.5)
>1.5 to ≤2.5 × ULN	0/20	2/21 (9.5)	2/41 (4.9)
>2.5 to ≤5 × ULN	0/20	1/21 (4.8)	1/41 (2.4)
>5 to ≤20 × ULN	0/20	1/21 (4.8)	1/41 (2.4)
>20 × ULN	0/20	0/21	0/41
Total bilirubin			
>ULN to ≤1.5 × ULN	0/20	1/21 (4.8)	1/41 (2.4)
>1.5 to ≤2 × ULN	0/20	0/21	0/41
>2 to ≤3 × ULN	1/20 (5.0)	0/21	1/41 (2.4)
>3 to ≤10 × ULN	0/20	0/21	0/41
>10 × ULN	0/20	0/21	0/41
ALT and total bilirubin			
ALT >3 × ULN and total bilirubin >2 × ULN	0/20	0/21	0/41
AST and total bilirubin			
AST >3 × ULN and total bilirubin >2 × ULN	0/20	0/21	0/41
ALT or AST and total bilirubin			
AST or AST >3 × ULN and total bilirubin >2 × ULN	0/20	0/21	0/41

Source: Table 14.3.4.12

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; IVA: ivacaftor; LFT: liver function test; n: size of subsample; N: total sample size; N1: number of subjects with at least 1 non-missing measurement during the TE Period; TE: treatment-emergent; TEZ: tezacaftor: ULN: upper limit of normal

Notes: Treatment-emergent period is from the first dose of study drug within a treatment period through the corresponding Safety Follow-up Visit or 28 days after the last dose of study drug within the same period for subjects who did not have a Safety Follow-up Visit. For n, subject is only counted once in the worst category of all assessments during the TE period. Criteria involving 2 LFT parameters may be determined by assessments at different visits.

No subjects had an ALT or AST elevation >5 × ULN or with an ALT or AST elevation >3 × ULN with a concurrent total bilirubin elevation >2 × ULN.

Assessor's comments

Adverse events related to the increase of transaminases and of other liver function tests (within the SOC "Investigations") are described by several preferred terms as previously mentioned. In the

TEZ/IVA group events of aspartate aminotransferase increased, transaminases increased, and blood bilirubin increased were reported (5.0% each). In the placebo group, events of aspartate and alanine aminotransferase increased and gamma-glutamyltransferase increased were reported (4.8% each) as well as 2 events of blood alkaline phosphatase increased (9.5%). The analysis of the liver function test results that met threshold criteria shows that 7 subjects (35.0%) in the TEZ/IVA group and 9/21 (42.9%) in the placebo group had ALT or AST elevations that were >ULN to ≤3 x ULN. Four additional subjects had ALT or AST elevations that were >3 to ≤5 x ULN (one [5%] in the TEZ/IVA group and 3 [14.3%] in the placebo group). No subjects had an ALT or AST elevation >5 × ULN or with an ALT or AST elevation >3 × ULN with a concurrent total bilirubin elevation >2 × ULN.

There seems to be some inconsistencies between the frequency of adverse events related to liver function tests and the analysis of subjects who reported such events. Eight subjects in the TEZ/IVA group presented ALT or AST elevations that were >ULN to ≤5 x ULN in spite of which only 2 adverse events of aspartate aminotransferase increased and transaminases increased are described. While no events of "Alanine aminotransferase increased" were reported in the TEZ/IVA group, the analysis of liver function test results that met threshold criteria shows that ALT increase >ULN to ≤3 x ULN occurred in 6 (30.0%) subjects in the TEZ/IVA group. The MAH was requested to clarify the strategy used for the analysis of liver functions tests results as well as the difference between the preferred terms used to describe adverse events related to the liver (in particular those related to the increase of transaminases either isolated or combined). The reason why ALT increase was not considered as an adverse event in subjects of the TEZ/IVA group in Study 112 should also be clarified. In response, it has been clarified that liver function tests are analysed as adverse events and reported based on the investigators' clinical determination, as well as laboratory data (i.e., using summary statistics [including mean post-baseline values and changes from baseline] and threshold analyses (i.e., the number and percentage of subjects meeting prespecified threshold criteria). Transaminase (alanine transaminase [ALT] and aspartate transaminase) laboratory elevations are common in patients with CF, and as such, not all transaminase elevations would be considered clinically significant and reported as adverse events, which explains why blood bilirubin increased and ALT increase observed in the TEZ/IVA group were not considered (related) adverse events. The MAH response is acknowledged. However, excluding patients with abnormal liver function tests at screening (defined as any 2 or more of the following: ≥3 x upper limit of normal (ULN) aspartate aminotransferase (AST), ≥3 x ULN alanine aminotransferase (ALT), ≥3 x ULN gamma-glutamyl transpeptidase (GGT), ≥3 x ULN alkaline phosphatase (ALP), or ≥2 x ULN total bilirubin OR abnormal liver function defined as any increase of ≥5 x ULN in AST or ALT) is not endorsed. Moderate liver abnormalities are very frequent in patients with cystic fibrosis. By excluding them from clinical trials, this prevents the generation of safety data in these patients while in clinical practice it is unlikely that this is the case.

Overall, no clinically meaningful adverse trends attributable to TEZ/IVA treatment were identified from serum chemistry, vitamins and lipids, hematology, coagulation, or urinalysis results. Similarly, there were no safety concerns identified in ECGs or vital signs and no clinically meaningful adverse trends attributable to TEZ/IVA treatment were identified ophthalmological examinations, or pulse oximetry results.

2.3.3. Discussion on clinical aspects

This Article 46 procedure of Regulation (EC) No1901/2006, as amended concerns the submission of a stand-alone study which is Study VX15-661-112 (Study 112), a phase 2, randomized, placebo-controlled, double-blind study to evaluate the effect of VX-661 in combination with ivacaftor on chest

imaging endpoints in subjects aged 12 years and older with cystic fibrosis, homozygous for the *F508del-CFTR* mutation.

The same study is being submitted to meet regulatory requirements for both Kalydeco and Symkevi given that both medicinal products are approved in combination for the treatment of adolescents and adults subjects homozygous for *F508del*.

Design and conduct of clinical study

Study 112 was performed in 9 centres in Australia. It included a Screening Period (28-days), a 72-week Treatment Period, and a Safety Follow-up Visit (approximately 28 days after the last dose). Study subjects were administered TEZ 100-mg once daily (qd)/IVA 150-mg every 12 hours (q12h), the approved dose of Symkevi. Subjects randomized to placebo received an inactive matching regimen.

The main objective of the study was to evaluate the treatment effect of TEZ/IVA on chest imaging endpoints as evaluated using low-dose computed tomography (LDCT) at Week 72 in subjects with CF 12 years of age and older who are homozygous for the *F508del-CFTR* mutation. This was done by assessing the CT images with the Brody/CF-CT scoring system. Magnetic resonance imaging (MRI) was planned, but the equipment needed to conduct ultra-short echo time chest MRI was unavailable at the start of enrolment, and consequently, MRI scans were not performed on any subject during the study.

The primary endpoint was the absolute change in Total Brody/CF-CT score from baseline at Week 72 using LDCT. Secondary and other endpoints were safety assessments as well as the absolute change from baseline at week 72 in Brody/CF-CT subscores (bronchiectasis, bronchial wall thickening, mucus plugging, parenchyma, and hyperinflation). Other endpoints and assessments were the absolute change from baseline at week 72 in percent predicted FEV1 (ppFEV1), absolute change in the respiratory domain score of CFQ-R, absolute change in anthropometric parameters, number of pulmonary exacerbations etc. which are the usual endpoints in studies of CFTR modulators.

Treatment duration of 72 weeks is considered appropriate for the main objective of assessing chest imaging endpoints based on review of the natural history progression of radiographic changes demonstrated on longitudinal studies seen on CT scans (de Jong PA. et al, 2004).

Chest imaging is recognized as essential in the assessment of respiratory disease of children and adults with CF. Chest CT offers advantages over other pulmonary endpoints because it has greater sensitivity to detect early structural lung disease even while the commonly used endpoint, ppFEV1, is within normal range and respiratory symptoms are absent. Additionally, CT can demonstrate the structural abnormalities that are specific to CF lung disease and can be performed similarly in all age groups. However, concerns about radiation exposure and need for sedation/anaesthesia (particularly in young children) apparently limited a general recommendation for routine use of CT in the follow-up of patients with cystic fibrosis in the EU (Castellani C., Duff AJA., Bell SC. et al, 2018).

In Study 112 the CT scans were optimized for low radiation dose and high image quality which is endorsed due to the risk of radiation. Inspiratory and expiratory (the latter needed to quantify air trapping) LDCT scans were obtained at Day 1 and week 72 and the images were read by 2 blinded, independent, qualified readers which is essential for the use of CT scoring systems as endpoints in clinical trials with medicinal products to track response to therapy.

Quantitative evaluation of chest CT images requires application of a scoring system to derive numerical values. Several CT scoring systems for CF have been developed for adults and children over the age of 6. The Brody's scoring system is a lobar scoring system, which semi-quantitatively scores the degree of structural lung disease by assigning a score to each lobe separately (including the lingula). It requires a complex method of calculation of the extent at different lung zones/segments and inclusion

of weighting factors. To improve standardization and training, the CF-CT scoring system, based on the Brody II system, was developed which consists of a large training module and 7 training sets that were scored by Brody and de Jong (the most experienced observers at that time) to define the 'gold standard' ratings (Szczesniak R, 2017). The score provides both localization and quantification of 5 abnormalities characteristic of CF lung disease: bronchiectasis, air trapping, mucus plugging, bronchial wall thickening, and parenchymal changes (includes parenchymal opacities, ground glass opacities, and cysts/bullae). It has been used in multiple studies to validate chest CT as an outcome and therefore it is considered an acceptable choice within the available scoring systems although the clinical value of the numbers generated is difficult to understand. The Brody scoring system has been reported as a total score with a maximal possible value of 207 and as a score representing the average severity of each of the six lobes, including the lingula as a separate lobe, with a maximum of 40.5 (Sanders DB. et al, 2015).

The MAH has further clarified that in Study 112 the subscore ranges were bronchiectasis 0 to 12; mucus plugging 0 to 6; peribronchial thickening 0 to 9; parenchymal opacities 0 to 5; and air trapping 0 to 4.5. The maximum possible score for a single lobe is thus 36.5 (12 + 6 + 9 + 5 + 4.5). Each lobe is scored separately, with the lingula considered a lobe (6 total), resulting in a range for the total score of 0 to 219. Some publications express the score on a percentage scale of 0 to 100, which allows for comparisons to scoring systems with different ranges. For the purposes of Study 112, in which changes from baseline were compared, the absolute score was chosen as the primary outcome. Two readers scored each lobe, and any subdomain score differing by more than 2 points was adjudicated by a third independent, expert reader who chose 1 of the scores to report. The expert reader noted that there was a greater than usual disagreement between the readers in Study 112, which added variability to the results.

Study 112 was exploratory in nature and the sample size was not based on statistical power. The analysis of the primary efficacy variable (total Brody/CF-CT score) and of the sub-scores was based on an analysis of the covariance. For these variables the estimated between-group difference in treatment effects and 95% CI was provided. For the remaining efficacy variables, only within-group changes were shown. No estimation of the difference between treatments was provided. Although the table showing the results of the absolute change from baseline in the respiratory domain of the CFQ-R indicates that the results of an ANCOVA analysis are presented, the MAH has clarified that the title for Table 11-5 in the clinical study report (and therefore in this report) is mislabelled, i.e., absolute within-group changes are presented.

Inclusion and exclusion criteria are similar to those used in the pivotal study 106 in subjects homozygous for *F508del*. No specific criteria are related to the primary endpoint of the study other than trying to ensure that subjects have stable disease at the time of screening/inclusion into the study and an exclusion criterion related to the presence of contraindications to undergoing LCDT.

Demographics and patient characteristics

A total of 41 subjects were randomized: 20 subjects in the TEZ/IVA group and 21 subjects in the placebo group. Out of these 41 who received at least 1 dose of study drug, 40 (97.6%) subjects completed the study and 39 (95.1%) completed study drug dosing.

The mean (SD) age of subjects included in the Full Analysis Set (n=41) was 20.2 (8.4) years. Eight subjects (40.0%) in the TEZ/IVA group and 11 (52.4%) in the placebo group were adolescents. Fifty-five percent of subjects in the TEZ/IVA group and 47.6% in the placebo group were females. Their mean (SD) BMI was 21.93 (3.51) and 21.10 (3.40) kg/m2 in the TEZ/IVA and placebo groups respectively. Mean (SD) ppFEV1 was 91.4 (16.0) and 86.6 (12.7) percentage points in the TEZ/IVA

and placebo groups, respectively , with 50% of subjects in the TEZ/IVA group having a ppFEV1 above 90 percentage points. This figure was 47.6% in the placebo group. When compared to the baseline data of the study population of the pivotal study 106, the population of Study 112 is less severely affected, particularly in regard to ppFEV1.

The total Brody/CF-CT score at baseline was higher in the placebo group which suggests higher structural lung disease (mean [SD] value was 38.29 [22.91+ in the TEZ/Iva group and 43.68 [33.96] in the placebo group).

Discussion on clinical efficacy

The mean within-group change from baseline at Week 72 in the total Brody/CF-CT Score was numerically lower for the TEZ/IVA group (+0.90, 95% CI: −3.34, 5.14) compared with the placebo group (+2.38, 95% CI: (−1.82, 6.58). The least squares (LS) mean treatment difference between the TEZ/IVA and placebo groups was −1.48 (95% CI: −7.47, 4.52) which favours the TEZ/IVA group although statistical significance is not seen. Regarding the subdomains, the within-group change from baseline at Week 72 for 3 Brody subscores (bronchiectasis, hyperinflation, and mucus plugging) were numerically lower for the TEZ/IVA group (0.74, 0.4, and -0.4 respectively) compared with the placebo group (1.32, 1.0, and 0.4 respectively). The other Brody subscores (parenchymal opacities and bronchial thickening) were numerically higher for the TEZ/IVA group (0.1 and -0.03 respectively) compared with the placebo group (-0.1 and -0.24).

Overall, these within-group changes from baseline are of limited magnitude. Similarly, the mean differences between treatments in the change from baseline at week 72 in the subscores (which range from -0.58 to 0.21) seem modest.

The study by Sheikh and colleagues (2015) reported in a cohort of 10 subjects (aged 10 to 44 years) who carried the *CFTR-G551D* mutation a within-treatment annual improvement of 13.6 points (a decrease from 28.8 pre-ivacaftor to 15.2 after one year of ivacaftor therapy) in the total Brody/CF-CT score. Bronchiectasis score decreased by 2.7, mucous plugging decreased by 5.6 points, and airway wall thickness decreased by 5 points. However, scoring in this study was based on only four CT images and not on the whole lung CT acquisition. This may have consequences given that by scoring a large number of images more abnormalities may be detected. The decrease in the bronchiectasis score observed in this study deserves particular attention as bronchiectasis are considered irreversible (with the exception perhaps of cylindrical bronchiectasis which are considered to be early-stage disease) as opposed to peribronchial thickening and mucus plugging which are known to represent reversible CT changes during exacerbations in CF patients. Due to the above (including small sample size) the results from this study should be viewed with some caution.

The clinical relevance of the results observed in the CT scoring system used in study 112 is difficult to ascertain in the absence of comparative data from other studies where the same scoring system is used in similar or different study populations in terms of age, lung function etc. and lacking details about the possible range of values of the total score and subscores and how these are expressed and considered in the statistical analysis. In this respect, the MAH discussed that based on the exploratory nature of Study 112, the variability between scorers, the difficulty of comparisons to the literature, and the lack of an established MCID for the CF-CT score, the clinical relevance of the results of study 112 are not known. Comparisons of Study 112 to the literature are difficult without more details of how the scoring was conducted, as studies can differ with respect to the numeric score ranges and adjudication procedures. Overall, there are a number of factors beyond the study population (e.g. age) which difficult any attempts to compare results between studies using as an endpoint CT scores such as the CT score itself, the selection of the CT protocol and image acquisition technique, the variation in the

experience, training and skill of readers which limits the reliability of scoring systems because of inter-observer variability and bias, and the variation in the measure of variability itself, with various statistical methods employed to measure reproducibility (Calder AD. e al 2014).

In spite of the above, data from Study 112 are valuable in that they contribute to the experience on the potential use of CT scoring systems to document response to therapy (i.e., as a surrogate endpoint) which may be particularly relevant in the case of agents that aim to stop or slow progression of structural lung disease such as CFTR modulators.

Mean (SD) baseline ppFEV1 was 91.4 (16.0) and 86.6 (12.7) in the TEZ/IVA and placebo groups, respectively. The mean (SD) within-group change in ppFEV1 from baseline at Week 72 was 1.2 (8.4) points in the TEZ/IVA group and −3.6 (9.0) points in the placebo group based on data from 18 and 17 patients respectively. In the pivotal study 106 the mean (SD) within-group change from baseline at week 24 was 3.4 in the TEZ/IVA group and -0.6 in the placebo group which are similar values to those observed in Study 112 at week 24.

The population of both studies differ in that the mean baseline ppFEV1 in study 106 was 59.6 and 60.4 in the TEZ/IVA and placebo groups respectively which may have had an impact on the magnitude of the change observed in the TEZ/IVA group of study 112. The decrease seen in the placebo group at week 72 of study 112 is higher than that observed in study 106 at week 24 which may be explained by the timing of the assessment as the annual rate of decline of FEV1 for young adults homozygous for *F508del* reported by Sawicki et al (2017) is −2.52.

For all spirometry variables data are only available for 17 and 18 subjects at week 72 in the TEZ/IVA and placebo groups respectively. Within-group changes have been calculated regardless of the fact that at baseline data from 20 and 21 subjects were available. The impact of the lacking data at week 72 on the within-group changes is not addressed by the MAH. This is not requested to be done as Study 112 is an exploratory study and the main interest relies on the CT scoring system for which complete set of data are available. Therefore, no recalculation of the mean changes in spirometry variables or on other efficacy variables based on the available number of subjects with values at both baseline and week 72 is requested.

In Study 112, mean (SD) baseline CFQ-R respiratory domain score were 76.8 (13.3) and 82.4 (11.7) points. The mean (SD) within-group change in CFQ-R respiratory domain score from baseline at Week 72 was -1.7 (15.6) in the TEZ/IVA group and -4.6 (15.8) in the placebo group. When compared to results of the pivotal study 106, the within-group change at week 24 in the TEZ/IVA group was 5.0 points in the TEZ/IVA group and -0.1 points in the placebo group. Mean baseline values in study 106 were 70.1 (16.8) and 69.9 (16.6) in the TEZ/IVA and placebo groups. At week 24 of study 112 the within-group change was -1.9 (10.5) and -6.1 (18.7) points in the TEZ/IVA and placebo groups. It is likely again that differences in severity between study populations and the different treatment duration may partially explain the within-group changes in the respiratory domain score of CFQ-R in Study 112 which are comparatively lower than in Study 106 something to be expected at week 72 but less so at week 24 of study 112.

Three different versions of the CFQ-R questionnaire were used in Study 112, i.e., subjects who were 12 and 13 years of age at the time of questionnaire completion completed the CFQ-R Child version (self-report, evaluating 8 domains), and their parents/caregivers completed the CFQ-R Parent version (evaluating 11 domains). Subjects 14 years of age and older at the time of questionnaire completion completed the CFQ-R Adolescent/Adult version of the questionnaire (self-reported). The teen/adult version assesses 12 domains: physical functioning, role, vitality, emotional functioning, social, body image, eating disturbances, treatment burden, health perceptions, weight, respiratory symptoms, and

digestive symptoms. Each domain is composed of a variable number of self-report questions with four possible answers, with a total of 50 questions. The scores range from 0 to 100, with higher scores indicating a higher patient-reported quality of life with regard to the domain being evaluated.

For both the respiratory domain and for the non-respiratory domains (see below) summary statistics in this report are presented for the pooled Children Ages 12 and 13 and Adolescents and Adults Versions of the CFQ-R which may lead to the lost of relevant information as the domains assessed are not exactly the same in each version of the questionnaire. In this respect the results of 11 domains are discussed under the non-respiratory domains of CFQ-R but for some of them data at baseline and at week 72 are available only for approximately 30 patients out of the 41 enrolled in Study 112 which seems to be the consequence of the specific questionnaire not asking questions about some domains (vitality, health perceptions, role, and weight) for children aged 12 and 13 years old.

Overall, the mean within-group changes in the non-respiratory domains of CFQ-R are higher in the TEZ/IVA group than in the placebo group (even if a decrease from baseline is observed for most of them). However, the following ones deserve to be commented. Regarding the "Eating" domain, the mean (SD) absolute within-group change at week 72 was -7.2 (16.6) in the TEZ/IVA group and 3.9 (7.5) points in the placebo group. Similarly, in the "Role" domain the mean (SD) change was -6.7 (11.0) in the TEZ/IVA group and -1.7 (13.8) in the placebo group. Higher mean decreases (although of less magnitude than the previous ones) are also seen in the TEZ/IVA for the domains "Weight" and "Digestion". The MAH was requested to discuss this issue in terms of whether TEZ/IVA may decrease appetite or produce other effects that may explain the results of the CFQ-R domains "Eating", "Weight" and "Digestion". No further insight was provided, however. It was clarified that in the pivotal study 106 no significant differences between the placebo and TEZ/IVA groups at week 24 in any of the 4 CFQ-R domains of interest (eating, weight, digestion, and role) were seen either.

In terms of weight, height, BMI and their z scores, the mean within-groups changes in both groups are very modest. The pivotal study 106 did not show any meaningful effect of TEZ/IVA on body weight or BMI either.

The observed rate of PEx was 0.59 events per year in the TEZ/IVA group and 0.75 events per year in the placebo group. The rate of PEx requiring hospitalization or IV antibiotics was 0.17 events per year in the TEZ/IVA group and 0.52 events per year in the placebo group. The observed rate of unplanned hospitalisations was 0.17 events per year in the TEZ/IVA group and 0.23 events per year in the placebo group.

Regarding the within-group changes in the domains of the SF-36 questionnaire, the magnitude of the changes observed is limited. In this report norm-based scores are presented as it is generally accepted that their interpretation is simpler. Norm-based scoring equates all scores, so scores above 50 are better than the general population average for all scales and summary measures, while scores below 50 are worse. At week 72 most scores of the SF-36 were close to or above 50 points. Because health status is generally related to age and gender, normative values are often presented separately for specific age and gender groups. This is not done in the present study. Taking into account the main objective of Study 112 providing mean changes in SF-36 scores sorted out by sex and age will not be requested.

Discussion on safety

A total of 41 subjects received at least 1 dose of study drug. Nine subjects (45.0%) in the TEZ/IVA group and 7 (33.3%) in the placebo group were exposed for >60 and ≤72 weeks. Eleven (55.0%) and 13 (61.9%) in the TEZ/IVA and placebo groups were exposed for >72 weeks up to a maximum of 81.1 weeks.

All subjects reported at least an adverse event but more subjects in the placebo group reported an adverse event (221 events) than in the TEZ/IVA group (118).

Very common adverse events (≥1/10) in the TEZ/IVA group (n=20) were infective pulmonary exacerbations of cystic fibrosis (9 subjects, 45.0%), followed by cough (8 subjects, 40.0%). Haemoptysis, productive cough, and upper respiratory tract infection were reported at a frequency of 20.0% each (4 subjects each) while the frequency of reporting of lower respiratory tract infection bacterial, lung infection pseudomonal, nasopharyngitis, and pyrexia was 15% each (3 subjects each). Gastroenteritis, influenza, migraine, oropharyngeal pain, pharyngitis, sputum increased, and abdominal pain were reported at 10% each (2 subjects each).

In the placebo group, the most frequently reported adverse events were infective pulmonary exacerbations of cystic fibrosis (12 subjects, 57.1%) followed by cough (9, 42.9%), bacterial test positive (5, 23.8%), productive cough, nausea, vomiting and headache (4, 19.0% each), and sunburn (3, 14.3%). Haemoptysis, upper respiratory tract infection, pyrexia, oropharyngeal pain, abdominal pain, fatigue, fungal test positive, rhinorrhea, upper respiratory tract congestion, blood alkaline phosphatase increased, chest pain, constipation, and epistaxis were reported at a frequency of 9.5% (2 subjects).

From a quantitative point of view it would appear that preferred terms that are related to the same event are split diluting the frequency of reporting. As an example, a single event of headache was reported in the TEZ/IVA group (5.0%) while 4 events were reported in the placebo group (19%). However, in the TEZ/IVA group three additional events were tension headache (5.0%) and migraine (2 events, 10%). If these 4 events were counted together the frequency of reporting would have been 20% in the TEZ/IVA group (instead of 5%). Regarding abdominal pain (which is listed in section 4.8 of the SmPC of Symkevi with a frequency of very common), two preferred terms were used, i.e., "abdominal pain" and "abdominal pain upper". Counted together the frequency of abdominal pain (any) would have been 10% in the TEZ/IVA group (instead of 5%) and 14.3% in the placebo group (instead of 9.5%). The MAH was requested to justify the strategy used for the analysis of adverse events, and provide the incident rates for AEs of those as harmonised with the section 4.8 of the SmPC of Symkevi.

In this respect, the MAH states that a standard approach for the analysis and tabulation of adverse events (AEs) in clinical studies is used which includes presentation of AE incidences in one table by Preferred Term (PT) and in another table by System Organ Class (SOC) and PT. This approach is consistent with how the AEs are presented in the Symkevi Summary of Product Characteristics (SmPC) Section 4.8 where frequency is presented by individual PT ("adverse reaction" column), which is equivalent to how the incidence data in Study 112 are presented. While the response provided may be acceptable in that no changes in the SmPC are proposed or requested (taking into account the sample size of study 112), it remains at the discretion of assessors to ask for the frequency of reporting of preferred terms that are related and/or to request Standardised MedDRA Queries (SMQs) to gather further information on certain events (e.g. acute pancreatitis based on the clustering of various preferred terms).

Within the SOC "Infections and infestations", more events were reported in the TEZ/IVA group (19, 95%) than in the placebo group (16, 76.2%). Nine events (45.0%) of infective pulmonary exacerbations of cystic fibrosis occurred in the TEZ/IVA group and 12 (57.1%) in the placebo group. Other adverse events such as upper and lower respiratory tract infection, lung infection pseudomonal, nasopharyngitis, and pharyngitis were observed more frequently in the TEZ/IVA group.

Adverse events related to the increase of transaminases and of other liver function tests (within the SOC "Investigations") are described by several preferred terms such as aspartate aminotransferase

increased, transaminases increased, alanine aminotransferase increased, blood bilirubin increased, blood alkaline phosphatase increased, and gamma-glutamyltransferase increased. In the TEZ/IVA group events of aspartate aminotransferase increased, transaminases increased, and blood bilirubin increased were reported (5.0% each). In the placebo group, events of aspartate and alanine aminotransferase increased and gamma-glutamyltransferase increased were reported (4.8% each) as well as 2 events of blood alkaline phosphatase increased (9.5%). Please also refer to the summary of laboratory abnormalities (liver function tests) further below.

The majority of subjects had AEs that were considered either mild (17.1%) or moderate (56.1%) in severity. No subject had a life-threatening (i.e., Grade 4) event. There were no deaths in the study.

A total of 3 (15.0%) subjects in the TEZ/IVA group and 8 (38.1%) subjects in the placebo group had severe AEs. Severe adverse events reported in the TEZ/IVA group were infective pulmonary exacerbation of cystic fibrosis (2 events, 10%) and testicular torsion (an event, 5%).

The number (%) of subjects with any related AEs (including related, possibly related, and missing AEs) was 7 (35.0) and 7 (33.3) in the TEZ/IVA and placebo groups respectively. These subjects reported 12 and 11 events respectively. Within the SOC "Investigations", 4 events were reported in the TEZ/IVA group which included aspartate aminotransferase increased, transaminase increased, ultrasound liver abnormal, and blood triglycerides increased (each reported at a frequency of 5%) while in the placebo group three related events of blood alkaline phosphatase increased (2 events, 9.5%) and electrocardiogram PR shortened (4.8%) were reported. Two events (one per group) of infective pulmonary exacerbation of cystic fibrosis (5% and 4.8% in the TEZ/IVA and placebo groups respectively) were considered treatment-related. Other related adverse events reported in the TEZ/IVA group were as follows: atrioventricular block first degree (5%), sunburn (5%), dizziness (5%), testicular torsion (5%), and haemoptysis (5%).

Eight subjects (40.0%) in the TEZ/IVA group reported 10 serious adverse events (SAE) versus 13 patients (61.9%) in the placebo group who reported 17 SAEs. By Preferred Term, the most common SAE overall was infective PEx of CF, which occurred in 5 (25.0%) subjects in the TEZ/IVA group and 6 (28.6%) subjects in the placebo group. The only other SAE that occurred in more than 1 subject in any treatment group was lung infection pseudomonal (2 subjects in the TEZ/IVA group and no subjects in the placebo group). All other SAEs occurred in 1 subject only. The majority of SAEs had an outcome of recovered/resolved.

Overall, a total of 4 (9.8%) subjects had SAEs that were considered by the investigator to be related or possibly related to study drugs. In the TEZ/IVA group, 1 subject had a first-degree atrioventricular block and 1 subject had testicular torsion. In the placebo group, 1 subject had hepatic cirrhosis and 1 subject had infective pulmonary exacerbation of CF.

Study drug was interrupted in 2 (10%) subjects in the TEZ/IVA group due to atrioventricular block first degree (a subject, 5%) and blood bilirubin increased (5%). All other study drug interruptions (4 subjects, 19.0%) occurred in the placebo group. Of note, blood bilirubin increased is not reported as a related adverse event in the TEZ/IVA group.

Adverse events related to the increase of transaminases and of other liver function tests (within the SOC "Investigations") are described by several preferred terms as previously mentioned. In the TEZ/IVA group events of aspartate aminotransferase increased, transaminases increased, and blood bilirubin increased were reported (5.0% each). In the placebo group, events of aspartate and alanine aminotransferase increased and gamma-glutamyltransferase increased were reported (4.8% each) as well as 2 events of blood alkaline phosphatase increased (9.5%). The analysis of the liver function test results that met threshold criteria shows that 7 subjects (35.0%) in the TEZ/IVA group and 9 (42.9%)

in the placebo group had ALT or AST elevations that were >ULN to ≤3 x ULN. Four additional subjects had ALT or AST elevations that were >3 to ≤5 x ULN (one [5.0%] in the TEZ/IVA group and 3 [14.3%] in the placebo group). No subjects had an ALT or AST elevation >5 × ULN or with an ALT or AST elevation >3 × ULN with a concurrent total bilirubin elevation >2 × ULN.

Eight subjects in the TEZ/IVA group presented ALT or AST elevations that were >ULN to ≤5 x ULN in spite of which only 2 adverse events of aspartate aminotransferase increased and transaminases increased are described. While no events of "Alanine aminotransferase increased" were reported in the TEZ/IVA group, the analysis of liver function test results that met threshold criteria shows that ALT increase >ULN to ≤3 x ULN occurred in 6 (30.0%) subjects in the TEZ/IVA group. The MAH was requested to clarify the strategy used for the analysis of liver functions tests results as well as the difference between the preferred terms used to describe adverse events related to the liver (in particular those related to the increase of transaminases either isolated or combined). Clarification was also requested on the reason why ALT increase was not considered as an adverse event in subjects of the TEZ/IVA group in Study 112. In response, it has been clarified that liver function tests are analysed as adverse events and reported based on the investigators' clinical determination, as well as laboratory data (i.e., using summary statistics [including mean post-baseline values and changes from baseline] and threshold analyses (i.e., the number and percentage of subjects meeting prespecified threshold criteria). Transaminase (alanine transaminase [ALT] and aspartate transaminase) laboratory elevations are common in patients with CF, and as such, not all transaminase elevations would be considered clinically significant and reported as adverse events, which explains why blood bilirubin increased and ALT increase observed in the TEZ/IVA group were not considered (related) adverse events. The MAH response is acknowledged. However, excluding patients with abnormal liver function tests at screening (defined as any 2 or more of the following: ≥3 x upper limit of normal (ULN) aspartate aminotransferase (AST), ≥3 x ULN alanine aminotransferase (ALT), ≥3 x ULN gamma-glutamyl transpeptidase (GGT), ≥3 x ULN alkaline phosphatase (ALP), or ≥2 x ULN total bilirubin OR abnormal liver function defined as any increase of ≥5 x ULN in AST or ALT) is not endorsed. Moderate liver abnormalities are very frequent in patients with cystic fibrosis. By excluding them from clinical trials, this prevents the generation of safety data in these patients while in clinical practice it is unlikely that this is the case.

3. Rapporteur's overall conclusion and recommendation

Study 112 was designed to explore the treatment effect of TEZ/IVA on chest imaging endpoints during 72 weeks of treatment. Low-dose computed tomography was used for chest imaging and the images were evaluated using the Brody/CF-CT scoring system. Magnetic resonance imaging (MRI) was planned, but not performed due to the unavailability of the equipment needed. The study is exploratory in nature, of relatively small sample size but randomised, double-blind, and placebo controlled. Treatment duration (72 weeks), the use of CT protocols which ensure low-dose radiation and quality of the images, the selected scoring system, and the interpretation of the CT images by blinded readers are considered appropriate although more details should have been provided for the CT protocol as it appears that each participating centre used its own equipment.

The main interest of the present study relies on the fact that it expands the experience with the use of CT scoring systems as endpoints in clinical trials of medicinal products for the treatment of subjects with cystic fibrosis to document response to therapy (i.e., as a surrogate endpoint) which may be particularly relevant in the case of agents that aim to stop or slow progression of structural lung disease such as CFTR modulators.

The clinical relevance of the results observed in the CT scoring system used in study 112 is difficult to ascertain in the absence of comparative data from other studies where the same scoring system is used in similar or different study populations in terms of age, lung function etc. and lacking details about the possible range of values of the total score and subscores and how these are expressed and considered in the statistical analysis. Upon request, the MAH further discussed that comparisons of Study 112 to the literature are difficult without more details of how the scoring was conducted, as studies can differ with respect to the numeric score ranges and adjudication procedures. Furthermore, additional factors that need to be considered are the study population (e.g. age), the selection of the CT protocol and image acquisition technique, the variation in the experience, training and skill of readers which limits the reliability of scoring systems because of inter-observer variability and bias, and the variation in the measure of variability itself, with various statistical methods employed to measure reproducibility. Additional clarifications were requested in relation to other endpoints which have been sufficiently addressed by the responses provided by the MAH.

From a safety perspective, the safety profile in Study 112 did not reveal unknown adverse events (i.e., the safety profile is overall consistent with that described for Symkevi at the time of MA). Clarifications were requested regarding the strategy for the analysis of adverse events which are closely related within a particular SOC as well as that of the analysis of adverse events related to liver function tests and laboratory abnormalities. These have been sufficiently addressed in the responses provided taking into account the size of study 112 and that no amendments of the SmPC are proposed or requested based on safety data from this study.

☒ **Fulfilled:**

☐ **Not fulfilled:**

Based on the data submitted, the MAH should provide additional clarifications regarding the efficacy endpoints of Study 112 as part of this procedure (see section "Additional clarification requested").

4. Additional clarification requested

Based on the data submitted, the MAH should address the following questions as part of this procedure:

1. The MAH is requested to discuss the clinical relevance of the results of Study 112 in terms of the scoring system (total score and subscores). Even though the limitations of indirect comparison are acknowledged, this discussion can be accomplished by comparing these results with the ones of other studies where the same score has been used and giving appropriate consideration to the potential differences in study populations in terms of age, ling function etc. In addition, the measuring range of the CF-CT scoring system for the total score and subscores and how the values derived from it were considered for the statistical analysis should be clarified.
2. The MAH is requested to clarify the strategy of analysis for the absolute change in the respiratory domain of the CFQ-R from baseline at week 72 as apparently an ANCOVA analysis has been performed while only summary statistics were planned in the statistical analysis plan.
3. Regarding non-respiratory subdomains of CFQ-R, the mean absolute within-group change at week 72 in the "Eating" domain (particularly) as well as in "Weight" and "Digestion" domains

were worse in the TEZ/IVA group than in the placebo group. This was also the case for the "Role" domain. The MAH is requested to discuss this issue in terms of whether TEZ/IVA may decrease appetite or produce other effects that may explain the results seen in the "Eating", "Weight", "Digestion" and "Role" domains of the CFQ-R.

4. Data at week 72 for height are only available for 12 and 18 subjects in the TEZ/IVA and placebo groups. If height data are not available at week 72 for a number of subjects is unclear how BMI could be calculated for 20 subjects in each group at week 72. The MAH is requested to clarify this issue.

5. The MAH is requested to justify the strategy used for the analysis of adverse events i.e. splitting up the preferred terms that are related to the same AE (e.g. headache, abdominal pain), and provide the incident rates for AEs of those as harmonised with the section 4.8 of the SmPC of Symkevi.

6. The MAH is requested to clarify the strategy used for the analysis of liver functions tests results as well as to clarify the difference between the preferred terms used to describe adverse events related to the liver (in particular those related to the increase of transaminases either isolated or combined). The reason why ALT increase is not considered as an adverse event in subjects of the TEZ/IVA group in Study 112 should also be clarified.

The timetable is a 30 day response timetable with clock stop.

5. MAH responses to Request for supplementary information

1. The MAH is requested to discuss the clinical relevance of the results of Study 112 in terms of the scoring system (total score and subscores). Even though the limitations of indirect comparison are acknowledged, this discussion can be accomplished by comparing these results with the ones of other studies where the same score has been used and giving appropriate consideration to the potential differences in study populations in terms of age, lung function etc. In addition, the measuring range of the CF-CT scoring system for the total score and subscores and how the values derived from it were considered for the statistical analysis should be clarified.

MAH's response

Study VX15-661-112 was an exploratory study to evaluate the effect of tezacaftor/ivacaftor (TEZ/IVA) on chest imaging endpoints in subjects 12 years of age and older who are homozygous for *F508del* using the Cystic Fibrosis-Computed Tomography (CF-CT) scoring system, and to explore the use of the CF-CT score itself. There is no established minimum clinically important difference (MCID) for the CF-CT score. Therefore, no firm conclusions can be drawn on the clinical relevance of the CF-CT results in Study 112.

As mentioned by the assessor, the calculation of the CF-CT score is complex and may vary in different studies. The subscore ranges used in Study 112 were bronchiectasis 0 to 12; mucus plugging 0 to 6; peribronchial thickening 0 to 9; parenchymal opacities 0 to 5; and air trapping 0 to 4.5. The maximum possible score for a single lobe is thus 36.5 (12 + 6 + 9 + 5 + 4.5). Two readers scored each lobe, and any subdomain score differing by more than 2 points was adjudicated by a third independent, expert reader who chose 1 of the scores to report. The expert reader noted that there was a greater than usual disagreement between the readers in Study 112, which added variability to the results.

Each lobe is scored separately, with the lingula considered a lobe (6 total), resulting in a range for the total score of 0 to 219. As pointed out by the assessor, some publications express the score on a percentage scale of 0 to 100, which allows for comparisons to scoring systems with different ranges. For the purposes of Study 112, in which changes from baseline were compared, the absolute score was chosen as the primary outcome.

In Study 112, the mean changes in CF-CT total score and subscores were evaluated over 72 weeks. A change in total score of 2.2% per year in children with CF (mean age: 11 years) was reported by de Jong.[1] Subscore changes in children with CF (median age: 12.6 years; range: 6 to 19 years) were reported by Tepper[2] (total score not reported). Study 112 included older subjects (mean age: 20 years; range: 12 to 43 years) who had relatively preserved lung function (percent predicted forced expiratory volume in 1 second [ppFEV1] ≥70 at baseline). Comparisons of Study 112 to the literature are difficult without more details of how the scoring was conducted, as studies can differ with respect to the numeric score ranges and adjudication procedures.

In summary, based on the exploratory nature of Study 112, the variability between scorers, the difficulty of comparisons to the literature, and the lack of an established MCID for the CF-CT score, the clinical relevance of these results is not known. However, these data may contribute to an increased understanding of the application of the CF-CT scoring system in clinical studies of CF subjects.

[1] *de Jong PA, Nakano Y, Lequin MH, Mayo JR, Woods R, Paré PD, Tiddens HA. Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis. Eur Respir J. 2004;23(1):93-7.*

[2] *Tepper LA, Caudri D, Utens EM, van der Wiel EC, Quittner AL, Tiddens HA. Tracking CF disease progression with CT and respiratory symptoms in a cohort of children aged 6-19 years. Pediatr Pulmonol. 2014;49(12):1182-9.*

Assessor's comments

The issue has been sufficiently clarified. Overall, there are a number of factors beyond the study population (e.g. age) which difficult any attempts to compare results between studies using as an endpoint CT scores such as the CT score itself, the selection of the CT protocol and image acquisition technique, the variation in the experience, training and skill of readers which limits the reliability of scoring systems because of inter-observer variability and bias, and the variation in the measure of variability itself, with various statistical methods employed to measure reproducibility (Calder AD. et al 2014).

Issue solved.

Calder AD, Bush A, Brody AS, Owens CM. Scoring of chest CT in children with cystic fibrosis: state of the art. Pediatr Radiol. 2014;44(12):1496-506

2. The MAH is requested to clarify the strategy of analysis for the absolute change in the respiratory domain of the CFQ-R from baseline at week 72 as apparently an ANCOVA analysis has been performed while only summary statistics were planned in the statistical analysis plan.

MAH's response

Vertex would like to clarify that no analysis of covariance (ANCOVA) was performed for the Cystic Fibrosis Questionnaire-Revised (CFQ-R) results in Study 112. Consistent with the statistical analysis

plan and prespecified tables, only summary statistics were performed, based on the CFQ-R results at each visit. The title for Table 11-5 in the clinical study report (CSR) is mislabeled; the results presented in the referenced source table (Table 14.2.3.1, "Summary Statistics for CFQ-R, Children Ages 12 and 13 and Adolescents and Adults Version, Full Analysis Set") and in the CSR text of Section 11.2.2.4 are summary statistics, not ANCOVA results.

Assessor's comments

The issue has been adequately clarified. ***Issue solved.***

3. Regarding non-respiratory subdomains of CFQ-R, the mean absolute within-group change at week 72 in the "Eating" domain (particularly) as well as in "Weight" and "Digestion" domains were worse in the TEZ/IVA group than in the placebo group. This was also the case for the "Role" domain. The MAH is requested to discuss this issue in terms of whether TEZ/IVA may decrease appetite or produce other effects that may explain the results seen in the "Eating", "Weight", "Digestion" and "Role" domains of the CFQ-R.

MAH's response

Study 112 was an exploratory imaging study with 41 subjects, and was not powered to test for changes in total score or any individual domain scores of the CFQ-R. Outcomes for CFQ-R components are exploratory and do not provide precise information regarding the effect of TEZ/IVA on any of these domains. In the pivotal study of TEZ/IVA in subjects 12 years of age and older, Study VX14-661-106 (N = 504, Full Analysis Set), there were no significant differences between the placebo and TEZ/IVA groups in any of the 4 CFQ-R domains noted by the assessor (eating, weight, digestion, and role), evaluated as changes from baseline at Week 24.

Assessor's comments

No further insight has been provided by the MAH. In the pivotal study 106 no significant differences between the placebo and TEZ/IVA groups at week 24 in any of the 4 CFQ-R domains of interest (eating, weight, digestion, and role) were seen either.

Issue solved.

4. Data at week 72 for height are only available for 12 and 18 subjects in the TEZ/IVA and placebo groups. If height data are not available at week 72 for a number of subjects is unclear how BMI could be calculated for 20 subjects in each group at week 72. The MAH is requested to clarify this issue.

MAH's response

Height values were measured for all subjects at screening, and at most visits thereafter (excluding the Day 15 visit due to the short interval) for as long as the subject was ≤21 years old. Per the protocol, once a subject was >21 years of age, height was no longer collected because subjects will have stopped growing by age 21. For subjects >21 years old at the Week 72 visit, a height value collected at an earlier visit was used to calculate body mass index (BMI). Thus, all subjects had a valid height measurement to calculate BMI at Week 72.

Assessor's comments

The issue has been adequately clarified. ***Issue solved.***

5. **The MAH is requested to justify the strategy used for the analysis of adverse events i.e. splitting up the preferred terms that are related to the same AE (e.g. headache, abdominal pain), and provide the incident rates for AEs of those as harmonised with the section 4.8 of the SmPC of Symkevi.**

MAH's response

Vertex uses a standard approach for the analysis and tabulation of adverse events (AEs) in clinical studies. This approach includes presentation of AE incidences in one table by Preferred Term (PT) and in another table by System Organ Class (SOC) and PT. The SOCs and PTs were coded using MedDRA Version 21.0, the most recent version available during the analysis of Study 112.

This approach is consistent with how the AEs are presented in the Symkevi Summary of Product Characteristics (SmPC) Section 4.8. Vertex would like to clarify that in the SmPC, frequency is presented by individual PT ("adverse reaction" column), which is equivalent to how the incidence data in Study 112 are presented.

The analyses in Study 112 revealed no new safety concerns for TEZ/IVA.

Assessor's comments

In the response, the MAH states that a standard approach for the analysis and tabulation of adverse events (AEs) in clinical studies is used which includes presentation of AE incidences in one table by Preferred Term (PT) and in another table by System Organ Class (SOC) and PT. The SOCs and PTs were coded using MedDRA Version 21.0, the most recent version available during the analysis of Study 112. This approach is consistent with how the AEs are presented in the Symkevi Summary of Product Characteristics (SmPC) Section 4.8 where frequency is presented by individual PT ("adverse reaction" column), which is equivalent to how the incidence data in Study 112 are presented. While the response provided may be acceptable in that no changes in the SmPC are proposed or requested (taking into account the sample size of study 112), it remains at the discretion of assessors to ask for the incident rates of preferred terms that are related and/or to request Standardised MedDRA Queries (SQMs) to gather further information on certain events (e.g. acute pancreatitis based on the clustering of various preferred terms).

Issue solved.

6. **The MAH is requested to clarify the strategy used for the analysis of liver functions tests results as well as to clarify the difference between the preferred terms used to describe adverse events related to the liver (in particular those related to the increase of transaminases either isolated or combined). The reason why ALT increase is not considered as an adverse event in subjects of the TEZ/IVA group in Study 112 should also be clarified**

MAH's response

The approach to the analysis of liver function tests (LFTs; laboratory data and AEs) in Study 112 was comprehensive and consistent with the approach in other clinical studies. LFT laboratory data were collected based on the schedule of assessments and analyzed using summary statistics (including

mean post-baseline values and changes from baseline), as well as threshold analyses (i.e., the number and percentage of subjects meeting prespecified threshold criteria).

AEs are defined as any newly developed unfavorable finding or worsening from baseline following study drug treatment, and are reported based on the investigators' clinical determination. During the study, the investigators reviewed any laboratory abnormalities, including elevated transaminases, and assessed whether these should be reported as AEs based their medical judgment of clinical significance. Transaminase (alanine transaminase [ALT] and aspartate transaminase) laboratory elevations are common in patients with CF, and as such, not all transaminase elevations would be considered clinically significant and reported as AEs.

No subject in the TEZ/IVA group had an AE of ALT increased in Study 112, reflecting the fact that no investigator reported an AE of ALT increased for any subject in this treatment group (per source Table 14.3.1.2).

Assessor's comments

It has been clarified by the MAH that liver function tests are analysed as adverse events and reported based on the investigators' clinical determination, as well as laboratory data (i.e., using summary statistics [including mean post-baseline values and changes from baseline] and threshold analyses (i.e., the number and percentage of subjects meeting prespecified threshold criteria). Transaminase (alanine transaminase [ALT] and aspartate transaminase) laboratory elevations are common in patients with CF, and as such, not all transaminase elevations would be considered clinically significant and reported as adverse events, which explains why blood bilirubin increased and ALT increase observed in the TEZ/IVA group were not considered (related) adverse events. The MAH response is acknowledged. However, excluding patients with abnormal liver function tests at screening (defined as any 2 or more of the following: ≥3 x upper limit of normal (ULN) aspartate aminotransferase (AST), ≥3 x ULN alanine aminotransferase (ALT), ≥3 x ULN gamma-glutamyl transpeptidase (GGT), ≥3 x ULN alkaline phosphatase (ALP), or ≥2 x ULN total bilirubin OR abnormal liver function defined as any increase of ≥5 x ULN in AST or ALT) is not endorsed. Moderate liver abnormalities are very frequent in patients with cystic fibrosis. By excluding them from clinical trials, this prevents the generation of safety data in these patients while in clinical practice it is unlikely that this is the case.

Issue solved.

Annex. Line listing of all the studies included in the development program

A line listing has not been provided by the MAH because Study VX15-661-112 is not part of a development program for either product.