

Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Riociguat (Adempas®)

Bayer Vital GmbH

Modul 4 A

*Chronisch thromboembolische
pulmonale Hypertonie*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 24.04.2014

Inhaltsverzeichnis**Seite**

Inhaltsverzeichnis	1
Tabellenverzeichnis	4
Abbildungsverzeichnis	6
Abkürzungsverzeichnis	7
4 Modul 4 – allgemeine Informationen	10
4.1 Zusammenfassung der Inhalte von Modul 4.....	11
4.2 Methodik.....	14
4.2.1 Fragestellung.....	14
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	15
4.2.3 Informationsbeschaffung.....	15
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	15
4.2.3.2 Bibliografische Literaturrecherche.....	15
4.2.3.3 Suche in Studienregistern.....	16
4.2.3.4 Selektion relevanter Studien.....	17
4.2.4 Bewertung der Aussagekraft der Nachweise.....	17
4.2.5 Informationssynthese und -analyse.....	19
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	19
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	19
4.2.5.3 Meta-Analysen.....	20
4.2.5.4 Sensitivitätsanalysen.....	21
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	21
4.2.5.6 Indirekte Vergleiche.....	22
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	24
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	24
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	24
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	24
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	26
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	27
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	28
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	29
4.3.1.2.1 Studiendesign und Studienpopulationen.....	29
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	31
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	32
4.3.1.3.1 <Endpunkt xxx> – RCT.....	32
4.3.1.3.2 Subgruppenanalysen – RCT.....	35
4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	36
4.3.2 Weitere Unterlagen.....	36
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien.....	36
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche.....	37

4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	37
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	37
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	38
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	40
4.3.2.2	Nicht randomisierte vergleichende Studien	40
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	40
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	41
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	42
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	42
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	43
4.3.2.3	Weitere Untersuchungen.....	43
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	43
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	44
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	44
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	44
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	45
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	45
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	45
4.4.1	Beurteilung der Aussagekraft der Nachweise	45
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	46
4.4.2.1	Hintergrund/Fragestellung	46
4.4.2.2	Charakteristika der Zulassungsstudie CHEST-1	48
4.4.2.2.1	Studienendpunkte	58
4.4.2.2.1.1	Primärer Endpunkt.....	59
4.4.2.2.1.2	Sekundäre Endpunkte	60
4.4.2.2.1.3	Sicherheitsparameter	63
4.4.2.2.1.4	Operationalisierung der Studienendpunkte	63
4.4.2.2.1.5	Zusammenfassende Bewertung des Verzerrungspotentials	68
4.4.2.3	Ergebnisse der Zulassungsstudie CHEST-1	72
4.4.2.3.1	Belege für und Hinweise auf Subgruppeneffekte	72
4.4.2.3.2	Ergebnisse der Wirksamkeitsanalysen	74
4.4.2.3.2.1	Primärer Endpunkt 6MWD	84
4.4.2.3.2.2	Sekundäre Endpunkte	87
4.4.2.3.3	Ergebnisse der Sicherheitsanalysen	89
4.4.2.4	Ergebnisse einer Interimsanalyse der unkontrollierten Folgestudie CHEST-2	99
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	101
4.4.3.1	Vorbemerkung	101
4.4.3.2	Mortalität	101
4.4.3.3	Morbidität	101
4.4.3.4	Lebensqualität.....	102
4.4.3.5	Nebenwirkungen (Unerwünschte Ereignisse)	102
4.4.3.6	Ergebnissicherheit.....	103

4.4.3.7	Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen und Ausmaß des Zusatznutzens	104
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	105
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	105
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	105
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	105
4.5.4	Verwendung von Surrogatendpunkten	105
4.6	Liste der eingeschlossenen Studien.....	107
4.7	Referenzliste.....	108
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		111
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		113
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		114
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		115
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		116
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		147

Tabellenverzeichnis

	Seite
Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	25
Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	26
Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	28
Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	29
Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	30
Tabelle 4-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	31
Tabelle 4-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	31
Tabelle 4-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	32
Tabelle 4-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	32
Tabelle 4-10: Operationalisierung von <Endpunkt xxx>.....	34
Tabelle 4-11: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT mit dem zu bewertenden Arzneimittel	34
Tabelle 4-12: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel.....	35
Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	37
Tabelle 4-14: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	38
Tabelle 4-15: Operationalisierung von <Endpunkt xxx>.....	38
Tabelle 4-16: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	39
Tabelle 4-17: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	39
Tabelle 4-18: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	41
Tabelle 4-19: Operationalisierung von <Endpunkt xxx>.....	42
Tabelle 4-20: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	42
Tabelle 4-21: Operationalisierung von <Endpunkt xxx> - weitere Untersuchungen	44
Tabelle 4-22: Liste der relevanten Studien des pharmazeutischen Unternehmers.....	47

Tabelle 4-23: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden	47
Tabelle 4-24: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	48
Tabelle 4-25: Charakterisierung der Studie CHEST-1	50
Tabelle 4-26: Charakterisierung der Interventionen in der Studie CHEST-1	52
Tabelle 4-27: Charakterisierung der Studienpopulationen zu Beginn der CHEST-1 Studie (ITT- und Safety-Population).....	54
Tabelle 4-28: <i>A priori</i> definierte Subgruppen	57
Tabelle 4-29: Bewertung des Verzerrungspotenzials für die Studie CHEST-1	58
Tabelle 4-30: Matrix der patientenrelevanten Endpunkte der Studie CHEST-1	59
Tabelle 4-31: Weitere sekundäre, supportive Endpunkte in CHEST-1	59
Tabelle 4-32: Operationalisierung der Endpunkte der Studie CHEST-1	64
Tabelle 4-33: Bewertung des Verzerrungspotenzials für die Endpunkte der Studie CHEST-1	69
Tabelle 4-34: In CHEST-1 beobachtete Effektmodifikationen.....	73
Tabelle 4-35: Ergebnisse der Wirksamkeitsanalysen CHEST-1: Zusammenfassung	75
Tabelle 4-36: Ergebnisse der Wirksamkeitsanalysen der Studie CHEST-1: Primär- und Sekundäranalysen für die Riociguat-1,0-2,5 mg-Gruppe und die Placebo-Gruppe.....	77
Tabelle 4-37: Sensitivitätsanalysen zur absoluten Veränderung der 6MWD [m] bis Woche 16 im Vergleich zum Ausgangswert	85
Tabelle 4-38: Ergebnisse der Wirksamkeitsanalysen: Responderanalyse	86
Tabelle 4-39: Zusammenfassung unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse in CHEST-1.....	90
Tabelle 4-40: Ergebnisse der Sicherheitsanalysen: Statistische Gruppenvergleiche	91
Tabelle 4-41: Anzahl der Patienten mit mindestens einem UE nach Systemorganklasse und Preferred Term in RCT mit dem zu bewertenden Arzneimittel mit einer Inzidenz $\geq 5\%$	94
Tabelle 4-42: SUE nach Systemorganklasse und Preferred Term in CHEST-1	95
Tabelle 4-43: Anzahl der Patienten mit mindestens einem mit der Studienmedikation assoziierten UE nach Systemorganklasse und Preferred Term in CHEST-1	97
Tabelle 4-44: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	104
Tabelle 4-45 (Anhang): Studiendesign und -methodik für Studie CHEST-1	117
Tabelle 4-46 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie CHEST-1	148

Abbildungsverzeichnis

Seite

Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	27
Abbildung 4-2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	35
Abbildung 4-3: Patientenfluss in einem Flow-Chart gemäß CONSORT, CHEST-1	146

Abkürzungsverzeichnis

Abkürzung	Bedeutung
6MWD	6-Minuten-Gehstrecke (6 minute walk distance)
Abschn.	Abschnitt
ALT	Alanin Aminotransferase
ANCOVA	Kovarianzanalyse (Analysis of Covariance)
ARR	Absolute Risikoreduktion
AST	Aspartat Aminotransferase
BOCF	Baseline Observation Carried Forward
BPM	Schläge pro Minute (Beats per Minute)
CO	Herzzeitvolumen (Cardiac Output)
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronisch obstruktive Lungenerkrankung (Chronic Obstructive Pulmonary Disease)
CSR	Klinischer Studienbericht (Clinical Study Report)
CSR-ASA	Zusätzliche Subgruppenanalysen (Additional Subgroup Analyses)
CTEPH	Chronisch thromboembolische pulmonale Hypertonie
CTX	Carboxyterminales vernetzendes Telopeptid des Knochenkollagens
DBD	Diastolischer Blutdruck
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMC	Datenkontrollkomitee (Data Monitoring Committee)
eCRF	Elektronischer Prüfbogen (electronic Case Report Form)
EG	Europäische Gemeinschaft
eGFR	geschätzte glomeruläre Filtrationsrate (estimated Glomerular Filtration Rate)
EKG	Elektrokardiogramm
EMA	European Medicines Agency
EQ-5D	EuroQol-Fragebogen-5 Dimensionen
ERA	Endothelin-Rezeptor-Antagonist
FC	Funktionsklasse (functional class)
Fig.	Abbildung (figure)
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
hCG	Humanes Choriongonadotropin

IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-treat
i.v.	intravenös
IVRS	Interactive Voice Response System
KI	Konfidenzintervall
LOCF	Last Observation Carried Forward
LPH	Leben mit pulmonaler Hypertonie (Living with Pulmonary Hypertension)
LS	Least Squares
LTE	Folgestudie zur Langzeitbeobachtung (Long-Term Extension)
m	Meter
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal klinisch relevante Differenz (minimal important difference)
MMRM	Mixed-Effect Models for Repeated Measures
MTC	Mixed Treatment Comparison
MW	Mittelwert
NO	Stickstoffmonoxid
NT-proBNP	N-terminal prohormone of brain natriuretic peptide; N-terminales Prohormon des natriuretischen Peptids Typ B
OR	Chancenverhältnis (odds ratio)
PAH	Pulmonal arterielle Hypertonie
PAPdiast	Diastolischer pulmonal arterieller Druck
PAPm	Pulmonal-arterieller Mitteldruck (pulmonary arterial pressure mean)
PAPsyst	Systolischer pulmonal arterieller Druck
PCWP	Pulmonal-kapillärer Verschlussdruck (pulmonary capillary wedge pressure)
PEA	Pulmonale Endarteriektomie
PH	Pulmonale Hypertonie
Plc	Placebo
PRO	Patient-Reported Outcomes
PT	Preferred Term
PVR	Pulmonal vaskulärer Widerstand (pulmonary vascular resistance)
RCT	Randomisierte kontrollierte Studie (randomized controlled trial)
RR	Relatives Risiko

SAP	Statistischer Analyseplan
SBD	Systolischer Blutdruck
SC	Steuerungskomitee (steering committee)
SD	Standardabweichung (standard deviation)
SGB	Sozialgesetzbuch
SOC	Systemorganklasse (system organ class)
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
Tab.	Tabelle
tid	Ter in die; 3x täglich
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TTCW	Zeit bis zur klinischen Verschlechterung (Time To Clinical Worsening)
UE	Unerwünschtes Ereignis
WHO	Weltgesundheitsorganisation (World Health Organization)
VerfO	Verfahrensordnung

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Riociguat ist ein Arzneimittel zur Behandlung eines seltenen Leidens (Orphan Drug) entsprechend Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 (1). Gemäß § 35a SGB V Abs. 1 Satz 10 Halbs. 1 Sozialgesetzbuch V (SGB V) gilt der medizinische Zusatznutzen durch die Zulassung als belegt und Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie müssen nicht vorgelegt werden (§ 35a Abs. 1 Satz 10 Halbs. 2 SGB V).

Fragestellung

Gegenstand dieses Dossiermoduls ist der Nachweis des Ausmaßes des medizinischen Zusatznutzens von Riociguat in der Behandlung erwachsener Patienten der WHO Funktionsklassen II bis III mit inoperabler CTEPH, persistierender oder rezidivierender CTEPH nach chirurgischer Behandlung.

Datenquellen

Gemäß Absatz 1.3 der Anlage II.1 zur Verfahrensordnung des G-BA ist der Nachweis ausschließlich über die die Zulassung begründenden Studien zu erbringen. Als relevante Studien wurden der Zulassungsbehörde im Zulassungsdossier neben der Zulassungsstudie (11348, CHEST-1) zwei Phase-II-Studien (12166 und 12166 LTE) und eine unkontrollierte Folgestudie zur Langzeitbeobachtung der CHEST-1-Patienten (11349, CHEST-2) übermittelt. Die Erstellung des vorliegenden Nutzendossiers erfolgte ausschließlich auf Grundlage der randomisierten, Placebo-kontrollierten, multizentrischen, multinationalen, doppelblinden, pivotalen Parallelgruppen-Studie der Phase III, CHEST-1. Demnach finden sich alle Angaben in den Abschnitten 4.4.2 und 4.4.3. Für eine detaillierte Darstellung der Zulassungsstudie wurden zusätzlich die Anhänge 4-E und 4-F ausgefüllt.

Ein-Ausschlusskriterien für Studien

Entfällt

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Entfällt

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die in der EU zur Zulassung eingereichte Phase III-Studie CHEST-1 weist durch ihr Design als kontrollierte, doppelblinde, randomisierte Studie eine hohe Studienqualität auf, deren Verzerrungspotential auf Studienebene als niedrig eingestuft wurde. Auch das Verzerrungspotential auf Endpunktebene ist niedrig aufgrund der hohen Aussagekraft aller patientenrelevanten Endpunkte (ITT-Population, Verblindung, ergebnisunabhängige Berichterstattung der prospektiv geplanten Analysen, adäquater Umgang mit fehlenden Werten).

Riociguat zeigte in CHEST-1 statistisch signifikante und klinisch relevante Vorteile (Angaben jeweils für die ITT-Gesamtpopulation):

- Primärer Endpunkt war die Veränderung der 6-Minuten-Gehstrecke (6MWD) nach Woche 16 im Vergleich zum Ausgangswert zur Messung der Belastbarkeit. Riociguat zeigte einen signifikanten Vorteil mit einem Behandlungsunterschied zu Placebo von 45,69 m (95%-KI 24,74 m; 66,63 m). Der gezeigte statistisch signifikante Vorteil von Riociguat lag oberhalb der in der Literatur beschriebenen minimal klinisch relevante Differenz (MID) von 33 m. Dieser Vorteil wird in der a priori definierten Responderanalyse mit dem Responsekriterium „Verbesserung von mind. 40m“ bestätigt. Signifikant mehr Patienten erreichten das Responsekriterium unter Riociguat (52,6%) als unter Placebo (23,9%) (OR 3,541, 95%-KI 1,995; 6,285). Patienten unter Riociguat erfuhren somit eine „spürbare Linderung der Erkrankung“, was einer „deutlichen Verbesserung des therapielevanten Nutzens“ entspricht.
- Auch die patientenrelevanten sekundären Endpunkte zeigten klinisch relevante und konsistente Vorteile für Riociguat:
 - „Veränderung der WHO-Funktionsklasse“ zur Messung des Schweregrads der Erkrankung: Da jede Verbesserung der WHO-Funktionsklasse ein klinisch relevanter Effekt ist, der mit einer „spürbaren Linderung der Erkrankung“ einhergeht, impliziert der statistisch signifikante Vorteil von Riociguat ($p=0,0026$) für die Patienten eine „deutliche Verbesserung des therapielevanten Nutzens“.
 - „Veränderung des Werts auf der Borg-Skala“ zur Messung der Dyspnoe: Der gezeigte statistisch signifikante Vorteil von Riociguat (-1,07, 95%-KI -1,60; -0,52) lag oberhalb der in der Literatur beschriebenen MID von 1 Einheit. Patienten unter Riociguat erfuhren somit eine „spürbare Linderung der Erkrankung“, was einer „deutlichen Verbesserung des therapielevanten Nutzens“ entspricht.
 - „Veränderung des EQ-5D-Index“ zur Messung der Lebensqualität: Der gezeigte statistisch signifikante Vorteil von Riociguat (0,136, 95%-KI 0,065; 0,207) lag deutlich oberhalb der in der Literatur beschriebenen MID von 0,074. Patienten unter Riociguat erfuhren somit eine „deutliche Abschwächung schwerwiegender Symptome“, was einer „deutlichen Verbesserung des therapielevanten Nutzens“ entspricht.
 - Numerische Vorteile zeigten sich für die Endpunkte „Zeit bis zur klinischen Verschlechterung“ ($p=0,1724$) und „Veränderung des LPH-Index“, ein weiteres Instrument zur Messung der Lebensqualität (-5,76, 95%-KI -10,45; -1,06).
- Bezüglich der Häufigkeit der erfassten Nebenwirkungen gab es keine signifikanten Unterschiede zwischen Riociguat und Placebo: In der Placebo-Gruppe traten mehr Todesfälle auf, als in der Riociguat-Gruppe (3,4% vs. 1,2%). Mindestens ein unerwünschtes Ereignis (UE) trat bei 91,9% der Riociguat-Patienten und 86,4% der

Placebo-Patienten auf. Schwerwiegende unerwünschte Ereignisse (SUE) traten bei 19,7% der Riociguat-Patienten und 15,9% der Placebo-Patienten auf. Bei 2,9% der Riociguat-Patienten und 2,3% der Placebo-Patienten traten UEs auf, die zu einer Unterbrechung der Studienmedikation führten.

In der unkontrollierten Folgestudie zur Langzeitbeobachtung CHEST-2 konnte gezeigt werden, dass der gezeigte Behandlungseffekt erhalten bleibt und Riociguat auch bei längerer Gabe sicher und gut verträglich ist.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Für erwachsene Patienten der WHO Funktionsklassen II bis III mit inoperabler CTEPH, persistierender oder rezidivierender CTEPH nach einer chirurgischen Behandlung liegt ein beträchtlicher Zusatznutzen vor, da für die patientenrelevanten Morbiditätsendpunkte 6MWD, WHO-Funktionsklasse und Borg sowie den Lebensqualitätsendpunkt EQ-5D statistisch signifikante und klinisch relevante Effekte von Riociguat 1,0-2,5 mg ohne einen das Ausmaß des Zusatznutzens mindernden Schaden nachgewiesen wurden. Es wurde damit eine bisher medikamentös nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne von 5. Kapitel § 3 Absatz 1 VerfO G-BA erreicht, insbesondere eine „Abschwächung schwerwiegender Symptome“ und eine „für die Patientinnen und Patienten spürbare Linderung der Erkrankung“.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d.h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Riociguat ist ein Arzneimittel zur Behandlung eines seltenen Leidens (Orphan Drug) entsprechend Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 (1). Gemäß § 35a SGB V Abs. 1 Satz 10 Halbs. 1 Sozialgesetzbuch V (SGB V) gilt der medizinische Zusatznutzen durch die Zulassung als belegt und Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie müssen nicht vorgelegt werden (§ 35a Abs. 1 Satz 10 Halbs. 2 SGB V). Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfahrensordnung (VerfO) des Gemeinsamen Bundesausschuss (G-BA) besagt, dass der G-BA das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien bestimmt. In Modul 4 seien nur die Abschnitt 4.4.2 und 4.4.3 der Dossievorlage auszufüllen.

Demnach finden sich alle Angaben in den Abschnitten 4.4.2 und 4.4.3. Für eine detaillierte Darstellung der Zulassungsstudie wurden zusätzlich die Anhänge 4-E und 4-F ausgefüllt.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende

endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der

¹ Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.

² Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.

³ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z.B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ werden.

⁴ DerSimonian R, Laird N. *Meta-analysis in clinical trials. Control Clin Trials* 1986;7(3):177-188.

⁵ Deeks JJ, Higgins JPT, Altman DG. *Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁶ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte

⁶ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter

indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹.

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*

⁷ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

⁸ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

⁹ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹⁰ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

¹¹ B. Schöttker, D. Lühmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹² Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*
- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
<Studie 1>	ja	ja	abgeschlossen	12 Monate	Medikament A, Medikament B, Placebo

Geben Sie an, welchen Stand die Information in Tabelle 4-1 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-1 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

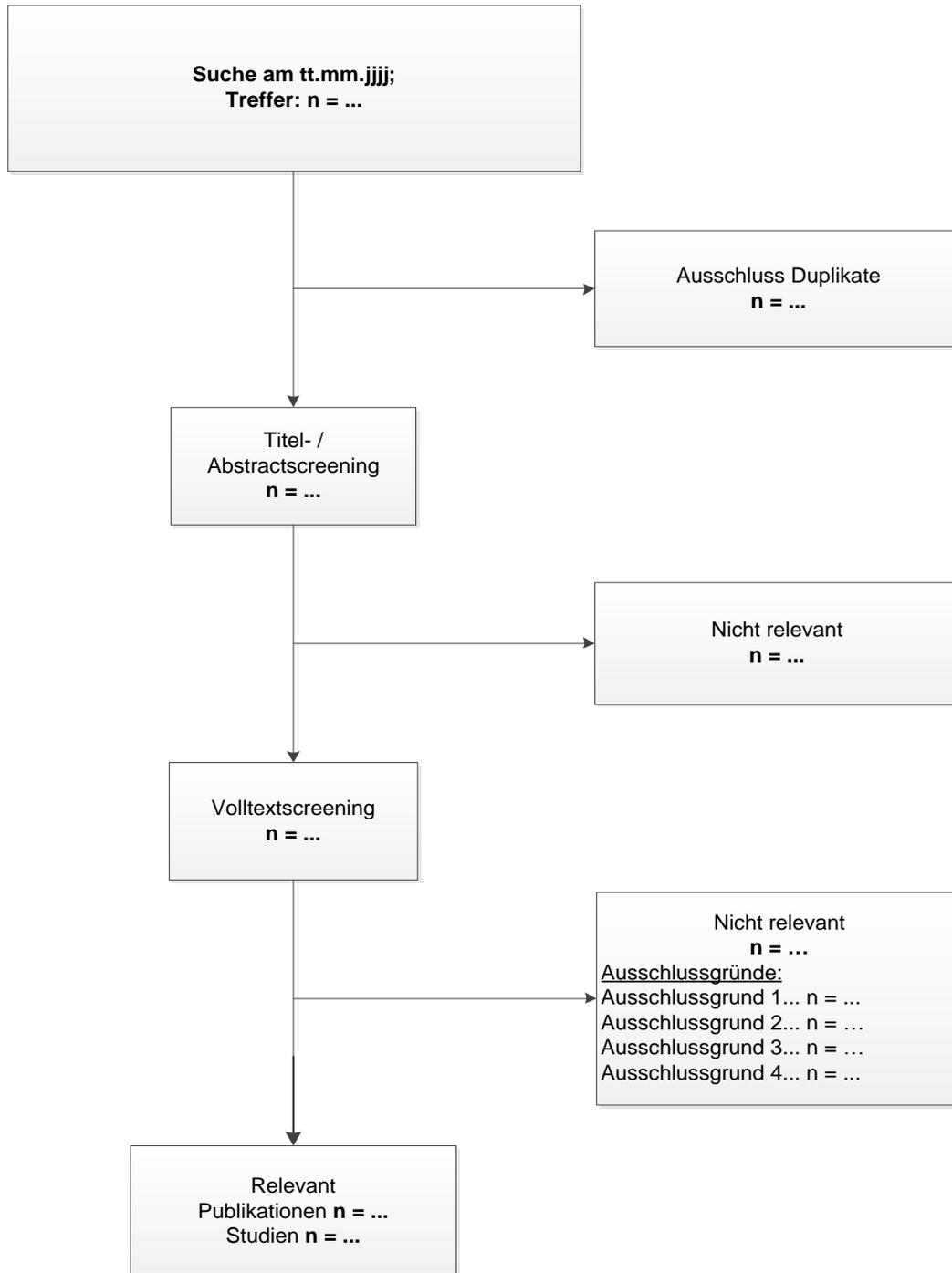


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die

Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-1) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein)	Status (abgeschlossen / abgebrochen / laufend)
<Studie 1>	clinicaltrials.gov [6, 7]	ja	nein	abgeschlossen

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

Geben Sie an, welchen Stand die Information in Tabelle 4-3 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilen-

überschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
<Studie 1>	ja	ja	nein	ja [5]	ja [6, 7]	ja [8]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 0 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
<Studie 1>	RCT, doppelblind, parallel	Jugendliche und Erwachsene, leichtes bis mittelschweres Asthma	<Gruppe 1> (n= 354) <Gruppe 2> (n= 347)	Run-in: 2 Wochen Behandlung: 6 Monate	Europa (Deutschland, Frankreich, Polen) 9/2003 – 12/2004	FEV1; Asthma-Symptome, Exazerbationen, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse

Tabelle 4-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	<Gruppe 1>	<Gruppe 2>	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
<Studie 1>	xxx 250 µg, 1 Inhalation bid + Placebo 2 Inhalationen bid	yyy 200 µg, 2 Inhalationen bid + Placebo 1 Inhalation bid	Vorbehandlung: zzz 1000 µg pro Tag, 4 Wochen vor Studienbeginn Bedarfsmedikation: aaa

Tabelle 4-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie	N	Alter (Jahre)	Geschlecht w /m (%)	<i>ggf. weitere Spalten mit Populationscharakteristika z. B. Dauer der Erkrankung, Schweregrad, weitere Basisdaten projektabhängig</i>
<Studie 1> <Gruppe 1> <Gruppe 2>				

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	<Mortalität>	<Gesundheits-bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.1.3.1 <Endpunkt xxx> – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die

Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-10: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-11: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-12: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)
<Studie 1>	

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

<Abbildung Meta-Analyse>

Abbildung 4-2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie**

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.**

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-14: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs-therapie 1>	<Vergleichs-therapie 2>	<Vergleichs-therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2> <Studie 3>	• •		• •	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-15: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-16: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-17: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei

an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung

(einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-18: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-19: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-20: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen**4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen**

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-21: Operationalisierung von <Endpunkt xxx> - weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerFO (s. Abschnitt 4.2.1).

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

4.4.2.1 Hintergrund/Fragestellung

Riociguat ist zugelassen als Arzneimittel zur Behandlung zweier seltener Leiden nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 (1) über Arzneimittel für seltene Leiden: der chronisch thromboembolischen pulmonalen Hypertonie (CTEPH, Anwendungsgebiet A) und der pulmonal arteriellen Hypertonie (PAH, Anwendungsgebiet B). Für Orphan Drugs gilt gemäß § 35a Absatz 1 Satz 10 SGB V der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Im Rahmen der frühen Nutzenbewertung ist daher lediglich das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen.

Gegenstand dieses Dossiermoduls ist der Nachweis des Ausmaßes des medizinischen Zusatznutzens von Riociguat in der Behandlung von erwachsenen Patienten der WHO Funktionsklassen II bis III mit inoperabler CTEPH, persistierender oder rezidivierender CTEPH nach chirurgischer Behandlung. Gemäß Absatz 1.3 der Anlage II.1 zur Verfahrensordnung des G-BA ist der Nachweis ausschließlich über die die Zulassung begründenden Studien zu erbringen. Demnach sind eine systematische Recherche in Studienregistern und eine bibliografische Literaturrecherche nicht erforderlich. Als relevante

Studien wurden der Zulassungsbehörde im Zulassungsdossier vier Studien übermittelt (Tabelle 4-22).

Tabelle 4-22: Liste der relevanten Studien des pharmazeutischen Unternehmers

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
11348 (CHEST-1)	ja	ja	abgeschlossen	16 Wochen	Riociguat 1,0-2,5 mg tid Placebo
11349 (CHEST-2)	ja	ja	laufend	max. 5 Jahre, 2 Monate, 14 Tage	Riociguat 1,0-2,5 mg tid
12166	ja	ja	abgeschlossen	12 Wochen	Riociguat 0,5-2,5 mg tid
12166 LTE	ja	ja	abgeschlossen	max. 7 Jahre	Riociguat 0,5-2,5 mg tid

CHEST-2 ist eine noch laufende Folgestudie von CHEST-1 zur Langzeitbeobachtung, zu der eine Interimsanalyse vorliegt. Da es sich um eine unkontrollierte Studie handelt, haben die Ergebnisse ihrer Interimsanalyse lediglich kommentierenden Charakter und werden daher nur zusammenfassend in Abschnitt 4.4.2.4 dargestellt. Bei den Studien 12166 und 12166 LTE handelt es sich um Phase-II-Studien, die als solche nicht für die Bewertung des Zusatznutzens von Riociguat herangezogen wurden (Tabelle 4-23).

Tabelle 4-23: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
11349 (CHEST-2)	Unkontrollierte Folgestudie von Studie 11348 (CHEST-1) zur Langzeitbeobachtung. Zur Darstellung der Langzeitwirkung wird die Studie in Abschnitt 4.4.2.4 trotzdem kurz beschrieben.
12166	Phase-II-Studie, nicht randomisiert, unkontrolliert, nicht verblindet, und daher zum Nachweis eines Zusatznutzens ungeeignet
12166 LTE	Phase-II-Folgestudie von Studie 12166 zur Langzeitbeobachtung

Die Herleitung des Ausmaßes des Zusatznutzens in dieser Indikation basiert folglich auf der zulassungsrelevanten Phase-III-Studie CHEST-1 „Randomized, double-blind, placebo-controlled, multicentre, multi-national study to evaluate the efficacy and safety of oral BAY 63-2521 (1 mg, 1.5 mg, 2 mg, or 2.5 mg tid) in patients with chronic thromboembolic pulmonary hypertension (CTEPH). CHEST-1 study” (Tabelle 4-24).

Tabelle 4-24: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein (Zitat))	Registereintrag ^c (ja/nein (Zitat))	Publikation (ja/nein (Zitat))
placebokontrolliert						
CHEST-1	ja	ja	nein	ja (Clinical Study Report A62508) (2)	ja (EudraCT 2007-000072-16, ClinicalTrials.gov NCT00855465)	ja (3)
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 0 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden. b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>Quelle: CSR Abschn. 2. Synopsis (2)</p>						

Im Folgenden wird zunächst die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe beschrieben (Abschnitt 4.4.2.2). Anschließend wird anhand von Studienergebnissen aufgezeigt, welche Effekte Riociguat auf CTEPH-spezifische, patientenrelevante Nutzen- und Schadenaspekte hat (Abschnitt 4.4.2.3) Im abschließenden Abschnitt 4.4.3 wird das Ausmaß des medizinischen Zusatznutzens von Riociguat von den Aussagen aus den beiden vorangehenden Kapiteln unter besonderer Berücksichtigung ihrer Relevanz für die Endpunktkategorien Mortalität, Morbidität, Lebensqualität und Nebenwirkungen abgeleitet.

4.4.2.2 Charakteristika der Zulassungsstudie CHEST-1

CHEST-1 ist eine wie geplant abgeschlossene, randomisierte, Placebo-kontrollierte, multizentrische, multinationale, doppelblinde, pivotale Parallelgruppen-Studie der Phase III mit einem zweiarmigen Design, in der die Wirksamkeit und Sicherheit der Behandlung mit dreimal täglich Riociguat in der Dosierung 1,0-2,5 mg gegenüber Placebo über eine 16-wöchige Behandlungsphase untersucht wurde (Tabelle 4-25).

Das Design von CHEST-1 entspricht den aktuellen Empfehlungen der EMA für die Durchführung confirmatorischer Studien für Arzneimittel zur Behandlung der PAH, die die EMA explizit auch als geeignet für die CTEPH bezeichnet (4). Die Studiendauer von 16 Wochen war danach ausgerichtet, eine Verbesserung der körperlichen Belastungsfähigkeit der

Patienten anhand der Veränderung der 6-Minuten-Gehstrecke (6MWD) nachzuweisen. Somit liegt gemäß dem 5. Kapitel § 5 Absatz 6 VerfO G-BA Evidenz der Stufe I b vor.

In den folgenden Tabellen 4-25, 4-26, 4-27 und 4-28 sind die wichtigsten Charakteristika von CHEST-1 zusammengestellt. Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E hinterlegt.

Tabelle 4-25: Charakterisierung der Studie CHEST-1

Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt Patientenrelevante sekundäre Endpunkte Supportive sekundäre Endpunkte Sicherheitsendpunkte
<p>RCT, doppelblind, parallel, multizentrisch, multinational</p>	<p>Erwachsene mit inoperabler CTEPH oder persistierender/rezidivierender CTEPH nach PEA</p>	<p>Riociguat-1,0-2,5 mg (N=174) Placebo (N=88)</p>	<p>Gesamtbehandlungsdauer: 16 Wochen, davon: Titrationsphase: 8 Wochen Erhaltungsphase: 8 Wochen Nachbeobachtungsphase (für Patienten, die nicht an CHEST-2, einer unkontrollierten Folgestudie zur Langzeitbeobachtung, teilnahmen): 30 Tage</p>	<p>89 Zentren in 26 Ländern: Europa (Belgien, Dänemark, Deutschland, Frankreich, Italien, Niederlande, Österreich, Polen, Portugal, Russland, Slowakei, Spanien, Schweiz, Tschechien, Türkei, Vereinigtes Königreich) Nordamerika (Kanada, USA) Südamerika (Argentinien, Brasilien, Mexiko) Asien/Pazifik (Australien, Japan, Südkorea, Taiwan) China 23.02.2009-27.06.2012</p>	<p>Primärer Endpunkt Absolute Veränderung der 6-Minuten-Gehstrecke (6MWD) bis Woche 16 im Vergleich zum Ausgangswert</p> <p>Sekundäre Endpunkte Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert Zeit bis zur klinischen Verschlechterung in Studientagen Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert</p>

Supportive sekundäre Endpunkte

Absolute Veränderung des pulmonal vaskulären Widerstands (PVR) nach 16 Wochen im Vergleich zum Ausgangswert

Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert

Sicherheit

Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse

Ereignisse von speziellem Interesse:
Vermutete UE Synkopen (MedDRA Preferred Terms: Bewusstlosigkeit, Präsynkopen, Synkopen)

Vermutete UE Hypotonie (MedDRA Preferred Terms: Hypotonie, verminderter Blutdruck)

Mortalität

Standardlaborparameter

Vitalparameter

12-Kanal Elektrokardiogramm

Blutgasanalyse

Zusätzlich Dokumentation von Schwangerschaften

Unerwünschte Ereignisse wurden als therapieabhängig erachtet, wenn sie bei erstmaliger Verabreichung der Studienmedikation bis zu zwei Tage nach Beendigung der Behandlung mit der Studienmedikation auftraten.

Quelle: CSR Abschn. 2. Synopsis; Abschn. 7.1 Study Design; Abschn. 7.4.3. Efficacy variables (2)

Tabelle 4-26: Charakterisierung der Interventionen in der Studie CHEST-1

Placebo-Gruppe	Riociguat-1,0-2,5 mg-Gruppe	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
Placebo tid	1,0-2,5 mg Riociguat tid	<p data-bbox="683 367 1385 551">Titrationsphase (8 Wochen) Beginnend mit einer Dosis von 1,0 mg Riociguat bzw. Placebo wurde in Zwei-Wochen-Abständen nach Maßgabe des am Ende eines Dosierungsintervalls gemessenen systolischen Blutdrucks (SBD) die Dosis der Studienmedikation nach folgendem Algorithmus mithilfe eines IVRS titriert:</p> <ul data-bbox="683 562 1385 831" style="list-style-type: none"> • SBD ≥ 95 mmHg: Erhöhung der Dosis um 0,5 mg tid • SBD 90-94 mmHg: Beibehaltung der Dosis • SBD < 90 mmHg ohne Hypotonie-Symptome: Senkung der Dosis um 0,5 mg tid • SBD < 90 mmHg mit Hypotonie-Symptomen wie Schwindel oder Präsynkope: Behandlung unterbrechen und nach 24 Stunden mit einer um 0,5 mg tid gesenkten Dosis neu beginnen <p data-bbox="683 842 1385 1111">Am Ende der Titrationsphase (Tag 56, Visite 5) war keine weitere Dosiserhöhung mehr möglich. Nach Maßgabe des am Ende des Dosierungsintervalls gemessenen SBD entschied der Prüfarzt, ob die Dosis beibehalten (SBD ≥ 90 mmHg) oder gesenkt (SBD < 90 mmHg) wurde. Die zum Ende der Titrationsphase am Tag 56 erreichte Dosisstärke wurde in der anschließenden 8-wöchigen Erhaltungsphase, der Nachbeobachtungsphase und bei Weiterführung auch in CHEST-2 beibehalten.</p> <p data-bbox="683 1167 863 1193">Vorbehandlung</p> <p data-bbox="683 1205 1385 1473">Prinzipiell mussten geeignete Patienten bei Studieneinschluss in Bezug auf PAH-spezifische Medikamente unbehandelt sein. In Ausnahmefällen, d.h. bei nicht akzeptablen unerwünschten Ereignissen und/oder fehlendem Nutzen einer PAH-spezifischen Medikation konnte eine Studienteilnahme erwogen werden. In diesen Fällen musste die PAH-spezifische Medikation mindestens 30 Tage vor der Rechtsherzkatheteruntersuchung zur Bestimmung der hämodynamischen Ausgangswerte abgesetzt werden.</p> <p data-bbox="683 1529 1278 1556">Erlaubte und nicht erlaubte (Begleit-)Medikationen</p> <p data-bbox="683 1568 1385 1657">Grundsätzlich sollten die Studienteilnehmer eine medizinisch erforderliche Medikation beibehalten, so sie nicht dezidiert ein Ausschlusskriterium darstellte.</p> <p data-bbox="683 1668 1385 1758">Folgende spezifische Medikationen waren im Rahmen der Behandlungsphase nicht erlaubt und führten bei Einnahme zum Ausschluss aus der Studie:</p> <ul data-bbox="683 1769 1385 1977" style="list-style-type: none"> • Endothelin-Rezeptor-Antagonisten • Prostanoid • Spezifische (z.B. Sildenafil oder Tadalafil) oder unspezifische (z.B. Dipyridamol, Theophyllin) Phosphodiesterase-Inhibitoren • NO-Donoren (z.B. Nitrate)

<p>Die Teilnahme an einem unterstützenden körperlichen Trainingsprogramm war in der Vorbeobachtungs- und Behandlungsphase nicht erlaubt.</p> <p>Die Einnahme anderer Begleitmedikationen sollte wegen möglicher Interaktionen von Riociguat mit Cytochrom P450 Isoenzym 3A4 Inhibitoren mit Vorsicht verabreicht werden (d.h. zusätzliches Blutdruck-Monitoring).</p> <p>Unspezifische Optionen für die Behandlung einer pulmonalen Hypertonie waren erlaubt, wie z.B. orale Antikoagulantien (soweit Behandlung mindestens 90 Tage vor Visite 1 begonnen); Diuretika (soweit Behandlung stabil seit mindestens 30 Tagen vor Visite 1); Digitalis; Kalziumantagonisten oder eine Langzeit-Sauerstoffbehandlung (soweit stabil seit mindestens 90 Tagen vor Visite 1).</p> <p>Modifikation oder Unterbrechung der Studienmedikation</p> <p>Im Falle einer klinisch relevanten Verschlechterung der PH lag es in der Entscheidung des Prüfarztes, auf Basis zusätzlicher hämodynamischer Messungen eine alternative Therapie zu beginnen oder eine eventuelle Begleitmedikation zu modifizieren.</p> <p>Eine Unterbrechung der Einnahme der Studienmedikation von mehr als drei Tagen in Folge (neun Dosen) im Verlauf der Studiendauer führte zum Ausschluss des Patienten.</p>
<p>Quelle: CSR Abschn. 7.1 Study Design; Abschn. 7.2.1 Inclusion criteria; Abschn. 7.2.2 Exclusion criteria; Abschn. 7.3.9 Permitted /prohibited prior and concomitant therapy; 7.2.3 Withdrawal of subjects from the study (2)</p>

Tabelle 4-27 gibt einen Überblick über die Eigenschaften der ITT-/Safety-Population zu Beginn der CHEST-1 Studie.

Tabelle 4-27: Charakterisierung der Studienpopulationen zu Beginn der CHEST-1 Studie (ITT- und Safety-Population)

Studie: CHEST-1		Placebo-Gruppe	Riociguat-1,0–2,5 mg-Gruppe
N	Randomisiert	88	174
	ITT/Safety	88	173
Alter	Durchschnitt in Jahren (SD)	59,2 (12,7)	59,3 (13,9)
	Median (Min-Max)	61,0 (26-77)	62,0 (19-80)
Geschlecht	Männlich	34 (38,6%)	55 (31,8%)
	Weiblich	54 (61,4%)	118 (68,2%)
Ethnizität	Weißer (Kaukasier)	65 (73,9%)	120 (69,4%)
	Farbige	1 (1,1%)	7 (4,0%)
	Asiaten	20 (22,7%)	37 (21,4%)
	Hispanics oder Latinos	2 (2,3%)	8 (4,6%)
	Gemischt	0 (0,0%)	1 (0,6%)
Geographische Region	Europa	53 (60,2%)	104 (60,1%)
	Nordamerika	9 (10,2%)	15 (8,7%)
	Südamerika	6 (6,8%)	15 (8,7%)
	Asien	20 (22,7%)	37 (21,4%)
	Australien	0 (0,0%)	2 (1,2%)
CTEPH Klassifikation	Inoperabel	68 (77,3%)	121 (69,9%)
	Persist./rezidivierende CTEPH nach PEA	20 (22,7%)	52 (30,1%)
WHO Funktionsklasse	I	0	3 (1,7%)
	II	25 (28,4%)	55 (31,8%)
	III	60 (68,2%)	107 (61,8%)
	IV	2 (2,3%)	8 (4,6%)
	Fehlend	1 (1,1%)	0 (0%)
6MWD	(m) (SD)	356,0 (74,7)	342,3 (81,9)
	Median (Min-Max)	372,0 (152-474)	360,0 (150-557)
PVR		(N=82)*	(N=151)*
	Durchschnitt (dyn·s·cm ⁻⁵) (SD)	779,3 (400,9)	790,7 (431,6)
	Median (Min-Max)	691,4 (258,1-2046,8)	711,1 (195,2-3942,0)
PAPm		(N=84)*	(N=156)*
	Durchschnitt (mmHg) (SD)	44,37 (10,00)	45,19 (12,82)
	Median (Min-Max)	45,67 (27,70-70,00)	43,17 (25,70-120,30)
*mITT-Population; PAPm: Mittlerer pulmonal arterieller Druck			
Quelle: CSR Fig. 8-1; Tab. 8-1, 8-5, 8-7, 8-8, 9-2; Tab. 14.2.4/5 (2)			

Es wurden weltweit in 89 Studienzentren in 26 Ländern 262 erwachsene Patienten mit diagnostizierter, inoperabler CTEPH oder persistierender/rezidivierender CTEPH nach PEA und einem 6MWD-Basiswert zwischen 150 m und 450 m randomisiert (Tabelle 4-25). Über die Hälfte der Patienten (N=157), d.h. 59,9%, wurde in Europa eingeschlossen (Tabelle 4-27).

Die Populationen in den beiden Studienarmen waren im Allgemeinen vergleichbar und wiesen keine auffälligen Unterschiede bezüglich Ethnizität, Alter oder geographischer Herkunft auf.

Eine stratifizierte Randomisierung nach der geographischen Herkunft stellte die Strukturgleichheit zwischen der Riociguat-Gruppe und der Placebo-Gruppe sicher.

Das durchschnittliche Alter lag in den beiden Studienarmen bei 59 Jahren. Über 40% der Patienten (42,8% in der Riociguat-1,0-2,5 mg-Gruppe, 40,9% in der Placebo-Gruppe) war mindestens 65 Jahre alt.

Mehr als 60% der eingeschlossenen Patienten waren Frauen (68,2% in der Riociguat-1,0-2,5 mg-Gruppe, 61,4% in der Placebo-Gruppe). Diese Verteilung spiegelt die epidemiologisch verfügbaren Daten zur geschlechterspezifischen Verteilung der Erkrankung wider (s. Modul 3A).

Die meisten Patienten waren Kaukasier (69,4% in der Riociguat-1,0-2,5 mg-Gruppe, 73,9% in der Placebo-Gruppe), durchschnittlich 22% waren Asiaten. Der hohe Anteil kaukasischer Patienten (70,9%) ist relevant für die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Die 12 beteiligten deutschen Zentren randomisierten 56 von 262 Patienten (21,4%) (Tabelle 4-27).

Die Mehrheit der eingeschlossenen Patienten litt an einer inoperablen CTEPH (69,9% in der Riociguat-1,0-2,5 mg-Gruppe, 77,3% in der Placebo-Gruppe). Der Anteil der Patienten mit einer persistierenden/rezidivierenden CTEPH nach einer PEA war in der Riociguat-1,0-2,5 mg-Gruppe etwas höher (30,1%) als in der Placebo-Gruppe (22,7%).

Zum gegenwärtigen Zeitpunkt gibt es in Deutschland keine zugelassene medikamentöse Therapie der CTEPH. Die im Rahmen von CHEST-1 erlaubten und verabreichten Begleitmedikationen sind in Deutschland zugelassen und werden von der Kölner Konsensus-Konferenz 2010 (5) empfohlen.

Mehr als 90% der Patienten wurden in die WHO-Funktionsklassen II oder III eingestuft und entsprachen damit der Zielpopulation laut Zulassung. Dabei war der Anteil der Patienten in der Funktionsklasse III in der Riociguat-1,0-2,5 mg-Gruppe (61,8%) etwas niedriger als in der Placebo-Gruppe (68,2%).

Der durchschnittliche 6MWD-Ausgangswert war in der Riociguat-1,0-2,5 mg-Gruppe mit 342,3 m nur geringfügig niedriger als in der Placebo-Gruppe (356,0 m), so dass sichergestellt war, dass die Patienten in allen Gruppen bei Studienbeginn einen vergleichbaren Status in Bezug auf den primären Endpunkt der Studie (s. 4.4.2.2.1.1) hatten.

Die durchschnittlichen Ausgangswerte für den PVR ($790,7-779,3 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$) und den PAPm ($45,19-44,37 \text{ mmHg}$) waren in den zwei Studienarmen vergleichbar und lagen deutlich über den Einschlusschwellenwerten von $300 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$ für den PVR, bzw. 25 mmHg für den PAPm. Der PAPm ist konstitutiv für die Diagnose einer pulmonalen Hypertonie (6) und kann wie der PVR als ein Verlaufsparemeter in der PAH-Therapie verwendet werden. Vergleichbare PVR-Ausgangswerte waren allein deshalb wichtig, weil die Veränderung des PVR ein sekundärer Studienendpunkt war (s. 4.4.2.2.1.2).

Die ITT- und Safety-Populationen waren identisch.

Für die Analyse der primären und sekundären Endpunkte wurden im Statistischen Analyseplan (7) die in Tabelle 4-28 beschriebenen Subgruppen präspezifiziert.

Tabelle 4-28: *A priori* definierte Subgruppen

Merkmal	Subgruppen
CTEPH-Klassifikation	Inoperabel Postoperativ, d.h. persistierende/ rezidivierende CTEPH nach einer PEA
WHO-Funktionsklasse-Ausgangswert	WHO-Funktionsklasse I/II WHO-Funktionsklasse III/IV
6MWD-Ausgangswert - Schwellenwert 320 m	6MWD-Ausgangswert < 320 m 6MWD-Ausgangswert ≥ 320 m
6MWD-Ausgangswert - Schwellenwert 380 m	6MWD-Ausgangswert < 380 m 6MWD-Ausgangswert ≥ 380 m
Geschlecht	Männer Frauen
Alter	Alter < 65 Jahre Alter ≥ 65 Jahre
Ethnizität	Weißer (Kaukasier) Farbige (Nicht-Kaukasier) Asiaten (Nicht-Kaukasier) Hispanics oder Latinos (Nicht-Kaukasier) Gemischt (Nicht-Kaukasier)
Region	Nordamerika Südamerika Europa China Asien-Pazifik
Quelle: CSR, Abschn. 7.6.4.4 (2)	

Wesentlich für den Nachweis der Ergebnissicherheit der abschließend getroffenen Aussagen zum Zusatznutzen sind das Verzerrungspotenzial auf Studienebene (Tabelle 4-29), sowie die Operationalisierung (Tabelle 4-32) und die Verzerrungspotenziale (Tabelle 4-33) der einzelnen Studienendpunkte. Details zu den Bewertungen der Verzerrungspotenziale sind im Anhang 4-F dokumentiert.

Tabelle 4-29: Bewertung des Verzerrungspotenzials für die Studie CHEST-1

Adäquate Erzeugung der Randomisierungs- sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
		Patient	Behandelnde Personen			
ja	ja	ja	ja	ja	ja	Niedrig
Quelle: CSR Abschn. 7.3.3 Treatment assignment; Abschn. 7.3.5 Blinding (2); Anhang 4-F						

Die Studie wurde als qualitativ hochwertige, randomisierte, kontrollierte, multizentrische, doppelblinde Phase-III-Studie durchgeführt. Die Studie wurde in 89 Studienzentren in 26 Ländern weltweit durchgeführt, von denen jedes zwischen 1-24 Patienten rekrutierte, so dass nicht ein Zentrum die Ergebnisse dominiert. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wurde das Verzerrungspotenzial auf Studienebene als niedrig eingestuft. Details zu den Bewertungen der Verzerrungspotenziale auf Studienebene sind im Anhang 4-F (Tabelle 4-46) dokumentiert.

4.4.2.2.1 Studienendpunkte

Ein Überblick über die patientenrelevanten Endpunkte, auf denen die Bewertung des Ausmaßes des Zusatznutzens beruht, findet sich in Tabelle 4-30. Weitere sekundäre, supportive Wirksamkeitsendpunkte sind in Tabelle 4-31 genannt. Die Operationalisierung der Endpunkte findet sich in Tabelle 4-32.

Tabelle 4-30: Matrix der patientenrelevanten Endpunkte der Studie CHEST-1

Mortalität	Morbidität				Gesundheits- bezogene Lebensqualität	Sicher- heit	
Gesamt mortalität*	Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert	Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert	Zeit bis zur klinischen Verschlechterung in Studientagen	Veränderung der Dyspnoe und des Ermüdungs-empfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert	Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert	Veränderung der gesundheitsbezogenen Lebensqualität gemäß LPH-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert	Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse
ja	ja	ja	ja	ja	ja	ja	Ja
*CHEST-1 war wie von der EMA beschrieben (4) auf den Nachweis von Effekten von Riociguat auf die Belastbarkeit der Patienten über die Studiendauer ausgelegt, aber nicht auf den Nachweis von Effekten auf die Mortalität. Die Gesamtmortalität wurde als Sicherheitsparameter erfasst. Quelle: CSR Abschn. 7.4.3 Efficacy variables; Abschn. 7.4.4 Safety variables (2)							

Tabelle 4-31: Weitere sekundäre, supportive Endpunkte in CHEST-1

Absolute Veränderung des PVR nach 16 Wochen im Vergleich zum Ausgangswert	Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert
Quelle: CSR Abschn.7.4.3 Efficacy variables; Abschn. 7.4.4 Safety variables (2)	

4.4.2.2.1 Primärer Endpunkt

Als primärer Endpunkt wurde die absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert definiert (Tabelle 4-32). Dieser Endpunkt wird von der EMA als einziger empfohlener Endpunkt zur Belastbarkeit der Patienten angesehen, solange das untersuchte Arzneimittel keinen negativen Effekt auf die Mortalität hat (4). In Deutschland ist die 6MWD ein im Versorgungsalltag gebräuchlicher Parameter zur Unterscheidung zwischen stabilen und instabilen Patienten als Grundlage für die Definition von Therapiezielen (8). Die

6MWD wird in Übereinstimmung mit den europäischen Leitlinien (9) als Verlaufsparemeter empfohlen (8). Der G-BA stuft die 6MWD als patientenrelevantes Maß für die Belastbarkeit eines Patienten ein (10).

4.4.2.2.1.2 Sekundäre Endpunkte

In CHEST-1 wurden weitere patientenrelevante, sekundäre Wirksamkeitsendpunkte untersucht.

Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert

Dieser Endpunkt wird explizit von der EMA für klinische Studien empfohlen (4). In Deutschland ist die WHO-Funktionsklasse ein im Versorgungsalltag gebräuchlicher Parameter zur Unterscheidung zwischen stabilen und instabilen Patienten als Grundlage für die Definition von Therapiezielen (8). Die WHO-Funktionsklassifizierung (11) wird in Übereinstimmung mit den europäischen Leitlinien (9) als Verlaufsparemeter empfohlen (8). Wie aus Tabelle 4-32 ersichtlich, bilden die WHO-Funktionsklassen die Fähigkeit des Patienten ab, seine Aktivitäten wahrzunehmen, was im Rahmen von Nutzenbewertungen als patientenrelevant zu verstehen ist (12).

Zeit bis zur klinischen Verschlechterung in Studientagen

Der kombinierte Endpunkt Zeit bis zur klinischen Verschlechterung in Studientagen (TTCW) setzt sich aus Endpunkten verschiedener Endpunktkategorien (Mortalität und Morbidität) zusammen (Tabelle 4-32):

Mortalität

1. Tod (Gesamt mortalität)

Morbidität

2. Herz- oder Lungentransplantation
3. PEA als Ultima Ratio wegen anhaltender Verschlechterung der WHO-Funktionsklasse (für weitere Details s. Tabelle 4-32)
4. Krankenhauseinweisung wegen anhaltender Verschlechterung der PH
5. Beginn einer neuen PH-spezifischen Therapie wegen Verschlechterung der PH
6. Anhaltende Verringerung der 6MWD um mehr als 15% gegenüber dem Ausgangswert oder um mehr als 30% verglichen mit der letzten Studien-bedingten Messung der 6MWD wegen Verschlechterung der PH. Die Verschlechterung musste durch eine zweite Messung nach 14 Tagen bestätigt werden
7. Anhaltende Verschlechterung der WHO-Funktionsklasse wegen Verschlechterung der PH

Dieser kombinierte Endpunkt wird ebenfalls von der EMA für klinische Studien empfohlen, wobei explizit die unter 1, 4, 6 und 7 genannten Endpunktkomponenten vorgeschlagen werden (4).

Die Patientenrelevanz der Gesamtmortalität ergibt sich unmittelbar aus dem 5. Kapitel § 3 Absatz 1 Verfo des G-BA.

Herz- und Lungentransplantationen sowie PEA sind aufwändige, mit erheblichen Risiken behaftete, operative Eingriffe und als solche patientenrelevant.

Eine pulmonale Hypertonie hat eine ähnliche Symptomatik wie eine Herzinsuffizienz bei koronarer Herzkrankheit (13). In seinen Empfehlungen für „Anforderungen an ein Modul ‚Chronische Herzinsuffizienz‘ für strukturierte Behandlungsprogramme für Koronare Herzkrankheit“ konstatiert der G-BA, dass eine hohe Hospitalisationsrate die Lebensqualität der Patienten vermindert und die Vermeidung von Hospitalisationen daher ein Therapieziel ist (14). Hieraus ist zu folgern, dass eine Krankenhauseinweisung ein patientenrelevantes Ereignis ist.

Obwohl vor der Einführung von Riociguat bisher kein Medikament für die Behandlung der CTEPH zugelassen war, bestand in Deutschland der Konsens, dass spezifische PH-Medikamente für ausgewählte CTEPH-Patienten in drei verschiedenen Situationen erwogen werden konnten: 1. Wenn Patienten nicht als Operationskandidaten beurteilt werden; 2. Wenn die präoperative Behandlung zur Verbesserung der Hämodynamik angemessen erscheint und 3. Wenn Patienten eine symptomatische residuale PH nach einer chirurgischen Behandlung aufweisen oder sich wieder eine PH nach erfolgter chirurgischer Behandlung entwickelt (5). Der Beginn einer neuen PH-spezifischen Therapie ist damit ein Indikator für eine weitergehende Behandlungsbedürftigkeit, die als solche patientenrelevant ist.

Die Validität und Patientenrelevanz der Änderungen der 6MWD und der WHO-Funktionsklasse wurden bereits unter 4.4.2.2.1.1, bzw. 4.4.2.2.1.2 erörtert.

Somit wird die Patientenrelevanz des kombinierten Endpunktes Zeit bis zur klinischen Verschlechterung durch die Patientenrelevanz seiner einzelnen Komponenten unterstrichen.

Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert

In den auch für die CTEPH relevanten Empfehlungen für die Durchführung klinischer PAH-Studien zählt die EMA explizit die Dyspnoe zu den Symptomen, die mit fortschreitender Morbidität der Patienten auftreten können (4). Nach den Vorgaben der American Thoracic Society (15) ist nach jedem 6MWD-Test die Dyspnoe und das Ermüdungsempfinden anhand einer Borg-Skala zu bewerten (Tabelle 4-32). Die Patientenrelevanz einer Dyspnoe hat der G-BA bereits an anderer Stelle bestätigt (16).

Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert

Die Veränderung der gesundheitsbezogenen Lebensqualität wird explizit als sekundärer Studienendpunkt von der EMA empfohlen (4). Die Patientenrelevanz der Lebensqualität ergibt sich unmittelbar aus dem 5. Kapitel § 3 Absatz 1 Verfo des G-BA. Der EQ-5D wird

vom G-BA „als weitverbreitetes und valides, jedoch generisches Erhebungsinstrument der Lebensqualität gesehen“ (17).

Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert

Die Veränderung der gesundheitsbezogenen Lebensqualität wird explizit als sekundärer Studienendpunkt von der EMA empfohlen (4). Die Patientenrelevanz der Lebensqualität ergibt sich unmittelbar aus dem 5. Kapitel § 3 Absatz 1 VerfO des G-BA. Der LPH wurde für Patienten mit pulmonal arterieller Hypertonie entwickelt, einer anderen klinischen Gruppe der pulmonalen Hypertonie mit einer sehr ähnlichen Symptomatik wie bei einer CTEPH, und lässt daher krankheitsspezifische Aussagen zur Veränderung der gesundheitsbezogenen Lebensqualität zu (Tabelle 4-32).

Absolute Veränderung des PVR nach 16 Wochen im Vergleich zum Ausgangswert

Eine pulmonale Hypertonie ist mit einem fortschreitenden Anstieg des pulmonal vaskulären Widerstands (PVR) verbunden, der zu Rechtsherzversagen und Tod führen kann. Dieser Endpunkt wird von der EMA für klinische Studien insbesondere zu Arzneimitteln aus neuen Wirkstoffklassen empfohlen (4), wurde bisher aber nicht als Surrogatendpunkt validiert. Die Patientenrelevanz dieses Endpunkts ist also fraglich, weswegen er nicht zur Bewertung des Zusatznutzens von Riociguat herangezogen wird. Aus dem Verständnis der Pathogenese der CTEPH und dem Wirkmechanismus von Riociguat musste jedoch eine Senkung des PVR durch eine Behandlung mit Riociguat erwartet werden. Für eine Verlaufskontrolle im Versorgungsalltag ist der PVR nicht geeignet, da er indirekt und mit vergleichsweise hohem Aufwand durch invasive Druckmessungen bei einer Rechtsherzkatheteruntersuchung ermittelt wird (Tabelle 4-32).

Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert

Eine hohe Serumkonzentration (≥ 1.400 pg/mL) des N-terminalen Prohormons des natriuretischen Peptids Typ B (NT-proBNP) korreliert mit einer schlechten Langzeitprognose für Patienten mit einer pulmonalen Hypertonie (18). Wenngleich die Veränderung des NT-proBNP noch nicht als Surrogatendpunkt validiert worden ist, sieht ihn die EMA als supportiven Wirksamkeitsendpunkt an (4). Die Patientenrelevanz dieses Endpunkts ist also noch fraglich, weswegen er nicht zur Bewertung des Zusatznutzens von Riociguat herangezogen wird. In Anlehnung an die europäischen Leitlinien (9) wird auch für den Versorgungsalltag in Deutschland empfohlen, die NT-proBNP-Serumkonzentration als Parameter zur Unterscheidung zwischen stabilen und instabilen Patienten, als Grundlage für die Definition von Therapiezielen und als Verlaufparameter zu nutzen (8), zumal dieses Prohormon relativ unaufwändig und für den Patienten wenig belastend nachgewiesen werden kann (Tabelle 4-32).

4.4.2.2.1.3 Sicherheitsparameter

Die Untersuchung der Sicherheit von Riociguat erfolgte anhand verschiedener Parameter in der Safety-Population, die in der Placebo-Gruppe 88 Patienten und in der Riociguat-1,0-2,5 mg-Gruppe 173 Patienten umfasste, und der ITT-Population entsprach.

Mortalität

Jeder im Studienverlauf aufgetretene Todesfall wurde unter Angabe des Studientages und der Todesursache sowie der Einschätzung des Prüfarztes dokumentiert, ob der Todesfall mit der Studienmedikation assoziiert war. Die Patientenrelevanz der Gesamtmortalität ergibt sich unmittelbar aus dem 5. Kapitel § 3 Absatz 1 Verfo des G-BA.

Unerwünschte Ereignisse

Wesentliche Bestandteile der Sicherheitsanalyse waren die Beobachtung und Dokumentation der Art und Häufigkeit unerwünschter Ereignisse (UE) und schwerwiegender unerwünschter Ereignisse (SUE) (Tabelle 4-32). Die Patientenrelevanz unerwünschter Ereignisse ergibt sich unmittelbar aus dem 5. Kapitel § 3 Absatz 1 Verfo des G-BA.

Wegen der gefäßerweiternden Wirkung von Riociguat waren folgende UE von besonderem Interesse:

- Vermutete UE Synkopen (MedDRA Preferred Terms: Bewusstlosigkeit, Präsynkopen, Synkopen)
Synkopen gehören zu den unspezifischen Symptomen einer pulmonalen Hypertonie (11). Synkopen wurden in der Phase II-Studie für Riociguat bei 4 von 75 Patienten beobachtet, waren aber nicht mit der Studienmedikation assoziiert (19).
- Vermutete UE Hypotonie (MedDRA Preferred Terms: Hypotonie, verminderter Blutdruck)
In der Phase II-Studie für Riociguat wurde Hypotonie bei 11 von 75 Patienten beobachtet, führte aber nicht zum Studienabbruch (19). Bei neun Patienten normalisierte sich der Blutdruck ohne und bei zwei Patienten mit Dosisänderung.

Weitere Sicherheitsparameter

Zusätzlich wurden Standardlaborparameter (hämatologische Parameter, Koagulationstests, klinische Chemie), Vitalparameter (Blutdruck- und Herzfrequenzmessungen), 12-Kanal Elektrokardiogramme, Blutgasanalysen (SaO₂ (%), PaO₂ (mmHg), PaCO₂ (mmHg)) erhoben und Schwangerschaften dokumentiert. Schwangerschaften weiblicher Studienteilnehmer und der Partnerin männlicher Studienteilnehmer wurden festgehalten und zur Erfassung möglicher Auffälligkeiten über die gesamte Zeit begleitet.

4.4.2.2.1.4 Operationalisierung der Studienendpunkte

Informationen zur Operationalisierung der Endpunkte der CHEST-1 Studie finden sich in Tabelle 4-32.

Tabelle 4-32: Operationalisierung der Endpunkte der Studie CHEST-1

Endpunkt „Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert“
<p>Der 6-Minuten-Gehstrecken-Test (6MWD-Test) wurde bei jeder Visite, d.h. in zwei-wöchentlichem Turnus, nach den Vorgaben der American Thoracic Society (15) durchgeführt. So wurde die 6MWD innerhalb eines Gebäudes, in einem vorzugsweise 30 m, minimal 25 m langen, ebenen, geraden, abgeschlossenen, wenig genutzten Gang mit festem Boden, mit Längenmarkierungen und Wendemarken ermittelt.</p> <p>Die Patienten wurden angehalten, innerhalb von 6 Minuten so viel Wegstrecke wie möglich durch Gehen von einem Ende zum anderen des Ganges zurückzulegen. Sie sollten dabei ihre Gehgeschwindigkeit selber bestimmen. Es war ihnen gestattet, stehen zu bleiben und sich gegen die Wand zu lehnen, sie sollten aber weitergehen, sobald sie sich dazu wieder in der Lage fühlten. Eine längere Wegstrecke entspricht also einer höheren Belastbarkeit eines Patienten.</p> <p>Eine Aufwärmphase sollte es nicht geben. Die Patienten wurden dazu aufgefordert, sich vor Testbeginn mindestens 10 Minuten auf einem Stuhl nahe der Startposition auszuruhen.</p> <p>Die Beobachter durften die Patienten nicht begleiten und nur standardisierte, einheitliche Testanweisungen geben, bzw. nur standardisierte, einheitliche Aussagen zur Aufmunterung im Minutentakt machen.</p> <p>In der Vorbeobachtungsphase war ein 6MWD-Test zur Überprüfung der Studieneignung und, ggf. ein 6MWD-Test zur Messung des 6MWD-Ausgangswertes, einschließlich der Ermittlung des Borg-Skalenwertes, durchzuführen. Beide 6MWD-Tests konnten am selben Tag durchgeführt werden, wenn der Patient bereits mit dem 6MWD-Test vertraut war. Andernfalls musste der Patient vor Überprüfung seiner Studieneignung zunächst einen 6MWD-Test zur Eingewöhnung machen, der nicht am selben Tag wie der Eignungstest stattfinden durfte.</p> <p>Zur Beschränkung der intraindividuellen Variabilität der Messungen an verschiedenen Messzeitpunkten waren die Rahmenbedingungen der Messung des Ausgangswertes für alle weiteren Messungen vorgeschrieben:</p> <ul style="list-style-type: none"> • Falls der Patient bereits zum Zeitpunkt der Erhebung des 6MWD-Ausgangswertes eine Langzeit-Sauerstofftherapie erhielt, waren für alle nachfolgenden 6MWD-Tests derselbe Sauerstofffluss, derselbe Applikationsweg und dieselbe Trageweise des mobilen Sauerstoffgerätes wie bei der Erhebung des 6MWD-Ausgangswertes vorgeschrieben. • Wenn eine Langzeit-Sauerstofftherapie erst nach Erhebung des 6MWD-Ausgangswertes begonnen, bzw. eine bestehende Therapie modifiziert wurde, war es dennoch nicht gestattet, die nachfolgenden 6MWD-Tests unter anderen als den Bedingungen bei der Messung des 6MWD-Ausgangswertes durchzuführen. • Nutzte der Patient eine Gehhilfe wie einen Gehstock oder einen Gehwagen bei der Erhebung des 6MWD-Ausgangswertes, hatte er diese Gehhilfen auch bei allen nachfolgenden 6MWD-Test zu verwenden. <p>Ab Visite 2 musste der 6MWD-Test durch einen zweiten Beobachter durchgeführt werden, der nicht in den Titrationsprozess eingebunden war und der nicht etwaige Änderungen des Blutdrucks und des Herzschlags als unmittelbare Reaktion des Patienten auf die Verabreichung der Studienmedikation beobachten konnte. Dieser zweite Beobachter notierte auf einem separaten Arbeitsblatt die Testergebnisse, die von einer anderen Person in den eCRF übertragen wurden.</p> <p>Zur Bewertung der klinischen Relevanz der beobachteten Effekte wurde zusätzlich eine <i>a priori</i> definierte Responderanalyse durchgeführt. Prädefiniertes Responsekriterium war eine Verbesserung der 6MWD um mindestens 40 m gegenüber dem Ausgangswert.</p>

Endpunkt „Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert“

Als weiteres Maß für die Morbidität der Patienten wurde bei jeder Visite, d.h. in zwei-wöchentlichem Turnus, die WHO-Funktionsklasse (WHO FC) (11) mit folgenden Spezifikationen bestimmt:

- WHO FC I: Keine Einschränkung der körperlichen Aktivität. Normale körperliche Belastungen führen nicht zu vermehrter Dyspnoe oder Müdigkeit, thorakalen Schmerzen oder Schwächeanfällen (asymptomatisch).
- WHO FC II: Leichte Einschränkung der körperlichen Aktivität. Keine Beschwerden in Ruhe. Normale körperliche Aktivität führt zu vermehrter Dyspnoe oder Müdigkeit, thorakalen Schmerzen oder Schwächeanfällen.
- WHO FC III: Deutliche Einschränkung der körperlichen Aktivität. Keine Beschwerden in Ruhe. Bereits leichtere als normale Belastungen führen zu Dyspnoe oder Müdigkeit, thorakalen Schmerzen oder Schwächeanfällen.
- WHO FC IV: Unfähigkeit, irgendwelche körperliche Belastung ohne Beschwerden auszuführen. Zeichen der manifesten Rechtsherzinsuffizienz. Dyspnoe und/oder Müdigkeit können bereits in Ruhe vorhanden sein. Bei geringster Aktivität werden die Beschwerden verstärkt.

Eine niedrigere WHO-Funktionsklasse entspricht also einer geringeren Morbidität eines Patienten. Ab Visite 2 musste die Bestimmung der WHO-Funktionsklasse durch einen zweiten Arzt durchgeführt werden, der nicht in den Titrationsprozess eingebunden war und der nicht etwaige Änderungen des Blutdrucks und des Herzschlags als unmittelbare Reaktion des Patienten auf die Verabreichung der Studienmedikation beobachten konnte. Dieser zweite Arzt notierte auf einem separaten Arbeitsblatt die Testergebnisse, die von einer anderen Person in den eCRF übertragen wurden.

Endpunkt „Zeit bis zur klinischen Verschlechterung in Studientagen“

Der kombinierte Endpunkt „Zeit bis zur klinischen Verschlechterung in Studientagen“ reflektiert das erstmalige Auftreten von besonders interessierenden Ereignissen, die unter eine der folgenden Kategorien fielen:

- Tod (Gesamtsterblichkeit).
- Herz- oder Lungentransplantation.
- PEA als Ultima Ratio wegen anhaltender Verschlechterung der PH. Obwohl nur Patienten mit inoperabler CTEPH oder persistierender/rezidivierender CTEPH nach einer PEA eingeschlossen werden durften, konnte es in Ausnahmefällen nicht ausgeschlossen werden, dass bei einer beträchtlichen Verschlechterung eine PEA als Ultima Ratio notwendig wurde.
- Krankenhauseinweisung wegen anhaltender Verschlechterung der PH. Eine Verschlechterung wurde nicht gezählt, wenn eine Entlassung innerhalb von 48 Stunden möglich war.
- Beginn einer neuen PH-spezifischen Therapie, d.h. mit Endothelin-Rezeptor-Antagonisten, Prostanoiden oder Phosphodiesterase-5-Inhibitoren wegen Verschlechterung der PH.
- Anhaltende Verringerung der 6MWD um mehr als 15% gegenüber dem Ausgangswert oder um mehr als 30% verglichen mit der letzten Studien-bedingten Messung der 6MWD wegen Verschlechterung der PH. Hierzu musste eine zweite, 14 Tage später stattfindende Messung bestätigen, dass es sich um eine anhaltende Verringerung der 6MWD handelt. Lagen zwischen der ersten Beobachtung der entsprechenden Gehstrecken-Verringerung und der Visite 7, bzw. der Abschlussvisite weniger als 14 Tage, musste bei Visite 7, bzw. der Abschlussvisite bestätigt werden, dass es sich um eine anhaltende Verringerung der 6MWD handelt.
- Anhaltende Verschlechterung der WHO-Funktionsklasse wegen Verschlechterung der PH. Bei Patienten, die sich von WHO-Funktionsklasse II oder III in die WHO-Funktionsklasse IV verschlechterten, musste analog zur Verringerung der 6MWD im Nachgang zur Erstbeobachtung bestätigt werden, dass es sich um eine anhaltende Verschlechterung handelt.

Je später also eines der oben genannten Ereignisse eintrat, desto langsamer verschlechterte sich der klinische Zustand eines Patienten. Im Falle einer klinischen Verschlechterung musste der Prüfarzt eingehend untersuchen, ob die Verschlechterung tatsächlich PH-bedingt war oder eine andere Ursache hatte.

Patienten, bei denen irgendeines der beschriebenen, besonders interessierenden Ereignisse im Sinne einer anhaltenden, klinischen Verschlechterung beobachtet wurde, mussten aus der Studie ausgeschlossen werden.

Endpunkt „Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert“

Borg-Skalenwerte wurden nach allen 6MWD-Tests ermittelt. Hierzu wurde die Borg CR10-Skala, bzw. die modifizierte Borg Dyspnoe Skala verwendet; letztere nur für Patienten, die vor der nationalen Genehmigung der dritten nachträglichen Protokolländerung (*amendment 3*) in die Studie eingeschlossen wurden.

Mit Fragebögen und Anleitungen in Landessprache wurde den Patienten die Erhebung des Borg-Skalenwertes unmittelbar vor einem 6MWD-Test erklärt. Der Patient wurde gebeten, unmittelbar nach dem 6MWD-Test sein Ermüdungsempfinden und seine Atemnot mithilfe der für ihn relevanten Borg-Skala zu beschreiben. Die folgende Gegenüberstellung zeigt die Gemeinsamkeiten und Unterschiede der beiden verwendeten Borg-Skalen-Versionen.

Skalenwert	Borg CR10-Skala	Modifizierte Borg Dyspnoe Skala
0	Gar nichts	Gar nichts
0,3		
0,5	Extrem schwach	Kaum merkbar
0,7		
1	Sehr schwach	Sehr schwach
1,5		
2	Schwach	Leicht
2,5		
3	Mäßig	Mäßig
4		
5	Stark	Schwer
6		
7	Sehr stark	Sehr schwer
8		
9		Sehr, sehr schwer
10	Extrem stark	Maximal
11		
•	Absolutes Maximum (12 oder sogar mehr)	

Die Skalierungen als auch die begrifflichen Auflösungen der Skalenwerte beider verwendeter Borg-Skalen stimmen weitestgehend überein, so dass eine parallele Nutzung dieser Skalen möglich war. Ein niedrigerer Skalenwert entspricht also einem geringeren Ausmaß an Dyspnoe und Ermüdungsempfinden.

Ab Visite 2 musste die Befragung durch einen zweiten Beobachter durchgeführt werden, der nicht in den Titrationsprozess eingebunden war und der nicht etwaige Änderungen des Blutdrucks und des Herzschlags als unmittelbare Reaktion des Patienten auf die Verabreichung der Studienmedikation beobachten konnte. Dieser zweite Beobachter notierte auf einem separaten Arbeitsblatt die Testergebnisse, die von einer anderen Person in den eCRF übertragen wurden.

Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert“

Der generische Fragebogen EQ-5D zur Bestimmung der gesundheitsbezogenen Lebensqualität wurde den Patienten in der 3-Level Version in Landessprache ausgehändigt und von ihnen selbständig zur Bestimmung des Ausgangswertes und des Wertes nach 16 Wochen (Visite 7) ausgefüllt. Sollte ein Patient jedoch Probleme mit dem Ausfüllen des Fragebogens haben, wurde versucht, dem Patienten auf neutrale Weise die Fragen zu erklären, ohne ihn zu beeinflussen. Ein höherer EQ-5D-Index entspricht einer höheren Lebensqualität.

Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert“

Zur Bestimmung der gesundheitsbezogenen Lebensqualität mit einem krankheitsspezifischen Instrument wurde der validierte Living with Pulmonary Hypertension-Fragebogen (LPH-Fragebogen) eingesetzt (13). Der LPH-Fragebogen ist eine Adaptation des Minnesota Living with Heart Failure Questionnaire (MLHFQ), der bereits in mehreren PAH-Studien eingesetzt wurde ((13); (20)).

Zur Messung der Auswirkungen seiner PH-Symptome, funktionellen Einschränkungen und der psychologischen Belastung auf seine Lebensqualität sollte der Patient auf einer 6-stufigen Likert-Skala ankreuzen, wie sehr jeder einzelne von 21 Aspekten ihn daran hinderte, ein Leben zu führen, wie er es sich wünschte. Auf dieser Skala bedeutete 0: nein und 5: sehr stark, so dass der Summenscore einen Wert von 0 bis 105 annehmen konnte und ein höherer Wert eine stärkere Beeinträchtigung des Patienten bedeutete.

Der LPH-Fragebogen wurde den Patienten in Landessprache ausgehändigt und von ihnen selbständig zur Bestimmung des Ausgangswertes und nach 16 Wochen (Visite 7) ausgefüllt. Sollte ein Patient jedoch Probleme mit dem Ausfüllen des Fragebogens haben, wurde versucht, dem Patienten auf neutrale Weise die Fragen zu erklären, ohne ihn zu beeinflussen. Der Fragebogen hatte in der deutschen Version folgenden Wortlaut:

„FRAGEBOGEN ZUM LEBEN MIT PULMONALER HYPERTONIE

In den folgenden Fragen geht es darum, wie sehr Ihre pulmonale Hypertonie (Bluthochdruck im Lungenkreislauf) sich in den letzten 7 Tagen auf Ihr Leben ausgewirkt hat. Kreuzen Sie bei jeder Frage die 0, 1, 2, 3, 4 oder 5 an, um anzugeben, wie sehr Ihr Leben davon betroffen war. Wenn eine Frage auf Sie nicht zutrifft, kreuzen Sie bitte die 0 an.

Hat Ihre pulmonale Hypertonie Sie in den letzten 7 Tagen an der von Ihnen gewünschten Lebensweise gehindert, dadurch dass

1. Schwellungen Ihrer Knöchel oder Beine auftraten?
2. Sie sich tagsüber hinlegen oder hinsetzen mussten, um sich auszuruhen?
3. Sie beim Gehen oder Treppensteigen Schwierigkeiten hatten?
4. Sie bei der Haus- oder Gartenarbeit Schwierigkeiten hatten?
5. Sie Schwierigkeiten hatten, außer Haus unterwegs zu sein?
6. Sie Schwierigkeiten hatten, nachts gut zu schlafen?
7. Sie Schwierigkeiten hatten, mit Familie oder Freunden Kontakte zu pflegen oder gemeinsame Unternehmungen durchzuführen?
8. Sie Schwierigkeiten hatten, Ihren Lebensunterhalt zu verdienen?
9. Sie bei Freizeitbeschäftigungen, Sport oder Hobbys Schwierigkeiten hatten?
10. Sie in Ihrem Sexualleben beeinträchtigt waren?
11. Sie weniger von dem essen konnten, was Sie mögen?
12. Sie unter Kurzatmigkeit litten?
13. Sie müde, erschöpft oder energielos waren?
14. Sie im Krankenhaus bleiben mussten?
15. Sie Geld für Ihre medizinische Versorgung bezahlen mussten?
16. Sie unter Nebenwirkungen Ihrer Behandlung litten?
17. Sie sich als Belastung für Ihre Familie oder Freunde empfanden?
18. Sie das Gefühl hatten, weniger Kontrolle über Ihr Leben zu haben?
19. Sie sich Sorgen machten?
20. Sie Schwierigkeiten hatten, sich zu konzentrieren oder sich an etwas zu erinnern?
21. Sie sich deprimiert fühlten?“

Hierbei bilden die Fragen 2–7 sowie 12 und 13 die physische Subdomäne des LPH-Fragebogens, die Fragen 17–21 die emotionale Subdomäne. Für beide Subdomänen wurde jeweils ein LPH-Teilscore berechnet.

Endpunkt „Absolute Veränderung des pulmonal vaskulären Widerstands (PVR) nach 16 Wochen im Vergleich zum Ausgangswert“

Der PVR wurde innerhalb eines Zeitfensters von 8 Wochen vor Behandlungsbeginn (Ausgangswert) und nach 16 Wochen (Visite 7) bestimmt. Eine Senkung des PVR entspricht einer Verbesserung.

Die hämodynamischen Parameter systolischer pulmonal arterieller Druck (PAPsyst), diastolischer pulmonal arterieller Druck (PAPdiast) und pulmonal kapillärer Verschlussdruck (PCWP) wurden durch invasive Druckmessungen bei einer Rechtsherzkatheteruntersuchung in Rückenlage mit einem Swan-Ganz-Katheter standardisiert gemessen, und dabei das Herzzeitvolumen (CO) mit der Thermodilutionsmethode bestimmt.

Der PVR wurde nach folgender Formel berechnet:

$PVR = 80 \cdot (PAPm - PCWP) / CO$ [dyn·sec·cm⁻⁵], mit

$PAPm = (PAPsyst + (PAPdiast \cdot 2)) / 3$ [mmHg]

Endpunkt „Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert“

Blutproben für die Bestimmung des N-terminalen Prohormons des natriuretischen Peptids Typ B (NT-proBNP) wurden im zweiwöchigen Turnus, d.h. bei jeder Visite zeitgleich mit den Proben zur Bestimmung der sicherheitsrelevanten Laborwerte genommen. Eine Senkung der NT-proBNP-Serumkonzentration entspricht einer Verbesserung. Die Proben wurden in einem Zentrallabor analysiert.

Um die Verblindung aufrecht zu erhalten, wurden den Prüfarzten erst nach dem Schließen der Studiendatenbank die Ergebnisse der NT-proBNP-Bestimmung mitgeteilt.

Endpunkt „Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse“

Die Erhebung unerwünschter Ereignisse (UE) und schwerwiegender unerwünschter Ereignisse (SUE) erfolgte gemäß des Medical Dictionary for Regulatory Activities (MedDRA, Version 15.0) unter Verwendung der Systemorganklasse (SOC) und dem Preferred Term (PT). Explizit wurden dabei auch UE von besonderem Interesse (Synkopen, Hypotonie), UE, die zu einer Unterbrechung der Behandlung mit der Studienmedikation führten, sowie alle Todesfälle erfasst. Es wurden alle (S)UE im eCRF dokumentiert, die im Zeitraum zwischen der Einwilligungserklärung und 30 Tagen nach letztmaliger Verabreichung der Studienmedikation beobachtet wurden. Da nur solche Ereignisse als therapiebedingt in Betracht kamen, die im Zeitraum von der erstmaligen Verabreichung der Studienmedikation bis zwei Tage nach ihrer letztmaligen Verabreichung auftraten, wurden nur diese Ereignisse eingehend analysiert. Patienten, bei denen ein Ereignis mehrmals vorkam, oder bei denen ein Ereignis kontinuierlich auftrat, wurden nur einmal gezählt. Nur der jeweils höchste Schweregrad eines Ereignisses wurde erfasst. Der Prüfarzt beurteilte, ob eine UE einen kausalen Zusammenhang mit der Studienmedikation aufwies und dann als mit der Studienmedikation assoziierte UE einzustufen war.

Der Schweregrad einer UE wurde vom Prüfarzt wie folgt klassifiziert: leicht – meist vorübergehender Natur und grundsätzlich ohne Beeinträchtigung gewöhnlicher Aktivitäten; mittelschwer – unangenehm genug, um gewöhnliche Aktivitäten zu beeinträchtigen; schwer – verhindert gewöhnliche Aktivitäten.

Quelle: CSR Abschn. 7.4.3 Efficacy variables; Abschn. 7.4.4 Safety variables (2)

4.4.2.2.1.5 Zusammenfassende Bewertung des Verzerrungspotentials

Die Bewertung des Verzerrungspotenzials für die Endpunkte der Studie CHEST-1 wird in Tabelle 4-33 dargestellt, die die Angaben aus Tabelle 4-46 zusammenfasst.

Für den primären Endpunkt 6MWD und die sekundären Endpunkte WHO FC, TTCW, Borg, EQ-5D, LPH und UE/SUE wurde das ITT-Prinzip umgesetzt. Laut Studienbericht wurden die Analysen in Übereinstimmung mit der ICH-Guideline E9 (21) zu „Statistical Principles for Clinical Trials“ durchgeführt. In das „Full Analysis Set“ einer ITT-Analyse sollten möglichst

alle Patienten eingeschlossen werden, die randomisiert wurden. Drei Ausnahmen von dieser Regel sind zulässig. Eine davon, der Ausschluss eines Patienten, der keine einzige Dosis der Studienmedikation erhalten hat, traf auf die Analyse der oben genannten Endpunkte (6MWD, WHO FC, TTCW, Borg, EQ-5D, LPH und UE/SUE) zu und ist Guideline-konform. In der Riociguat-1,0-2,5 mg-Gruppe traf dies auf einen Patienten von 174 Patienten (0,6%) zu (Randomisierungsfehler). In der Placebo-Gruppe wurden ausnahmslos alle randomisierten Patienten eingeschlossen. Das „Full Analysis Set“ umfasst in der Riociguat-1,0-2,5 mg-Gruppe 173 und in der Placebo-Gruppe 88 Patienten.

Tabelle 4-33: Bewertung des Verzerrungspotenzials für die Endpunkte der Studie CHEST-1

Endpunkt	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
6MWD	niedrig	ja	ja	ja	ja	niedrig
WHO FC	niedrig	ja	ja	ja	ja	niedrig
TTCW	niedrig	ja	ja	ja	ja	niedrig
Borg	niedrig	ja	ja	ja	ja	niedrig
EQ-5D	niedrig	ja	ja	ja	ja	niedrig
LPH	niedrig	ja	ja	ja	ja	niedrig
PVR	niedrig	ja	nein	ja	ja	hoch
NT-proBNP	niedrig	ja	nein	ja	ja	hoch
UE/SUE	niedrig	ja	ja	ja	ja	Niedrig
<p>6MWD: Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert WHO FC: Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert TTCW: Zeit bis zur klinischen Verschlechterung in Studientagen Borg: Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert EQ-5D: Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert LPH: Veränderung der gesundheitsbezogenen Lebensqualität gemäß LPH-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert PVR: Absolute Veränderung des PVR nach 16 Wochen im Vergleich zum Ausgangswert NT-proBNP: Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert UE/SUE: Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse Quelle: CSR Abschn. 7.3.3 Treatment assignment; Abschn. 7.3.5 Blinding (2)</p>						

Die Analyse des primären Endpunktes „Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert“ wurde in der ITT-Population durchgeführt. Alle Daten, die

zum Studienende (Woche 16) vorlagen, wurden verwendet. Fehlende Daten wurden adäquat durch Imputation ersetzt. Im Falle einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangsvs. verwendet. Lagen zum Studienende keine Daten vor, wurde der schlechtestmögliche Wert, d.h. 0 m, eingesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

Die Analyse des sekundären Endpunktes „Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert“ umfasste in der Riociguat-1,0-2,5 mg-Gruppe 173/173 (100%) und in der Placebo-Gruppe 87/88 (98,9%) (fehlende WHO FC bei Einschluss) der Patienten der ITT-Population. Eine Abweichung von < 5% von der ITT-Population wird als adäquat angesehen. Alle Daten, die zum Studienende (Woche 16) vorlagen, wurden verwendet. Fehlende Daten wurden adäquat durch Imputation ersetzt indem der schlechtestmögliche Wert, d.h. „IV“ bei klinischer Verschlechterung oder „V“ bei Tod, eingesetzt wurde. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

Die Analyse der sekundären Endpunkte „Zeit bis zur klinischen Verschlechterung“ (TTCW) und „Veränderung des Ermüdungsempfindens und der Dyspnoe gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert“ (Borg) wurde in der ITT-Population durchgeführt. Fehlende Daten wurden adäquat durch die Time-to-Event-Analyse mittels Logrank-Tests (TTCW) einbezogen, indem Todesfälle und die Beendigung der Studie wegen klinischer Verschlechterung als Ereignis gezählt wurden oder Imputation (Borg), indem der schlechtestmögliche Wert, d.h. „10“, eingesetzt wurde. Im Falle einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangsvs. verwendet. Bei Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

Die Analyse des sekundären Endpunktes „Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert“ wurde für den EQ-5D-Index in der ITT-Population durchgeführt, die für die Riociguat-1,0-2,5 mg-Gruppe 172/173 (99,4%) (fehlende EQ-5D-Messung bei Einschluss) Patienten und für die Placebo-Gruppe 87/88 (98,9%) (fehlende EQ-5D-Messung bei Einschluss) Patienten umfasste. Damit lag die Abweichung von der ITT-Population < 5% und wird als adäquat angesehen. Fehlende Daten wurden adäquat durch Imputation ersetzt, indem der schlechtestmögliche EQ-5D-Wert, d.h. -0,594, eingesetzt wurde. Im Falle von

Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

Die Analyse des sekundären Endpunktes „Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert“ wurde in einer ITT-Population durchgeführt, die für die Riociguat-1,0-2,5 mg-Gruppe 170/173 (98,3%) (fehlende LPH-Messungen bei Einschluss) und für die Placebo-Gruppe 86/88 (97,7%) (fehlende LPH-Messungen bei Einschluss) Patienten umfasste. Damit lag die Abweichung von der ITT-Population < 5% und wird als adäquat angesehen. Fehlende Daten wurden adäquat durch Imputation ersetzt, indem der schlechtestmögliche LPH-Wert, d.h. 105 eingesetzt wurde. Bei einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangsvsiste verwendet. Lagen zum Studienende wegen eines Todesfalls oder der Beendigung der Studie wegen klinischer Verschlechterung keine Daten vor, wurde der schlechtestmögliche LPH-Wert eingesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

Die Analyse des supportiven sekundären Endpunktes „Absolute Veränderung des pulmonal vaskulären Widerstands (PVR) nach 16 Wochen im Vergleich zum Ausgangswert“ wurde in einer modifizierten Intention-to-Treat (mITT)-Population durchgeführt, die für die Riociguat-1,0-2,5 mg-Gruppe 151/173 (87,3%) Patienten der ITT-Population und in der Placebo-Gruppe 82/88 (93,2%) der ITT-Population umfasste.

Die Analyse des supportiven sekundären Endpunktes „Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert“ wurde in einer mITT-Population durchgeführt, die für Riociguat 1,0-2,5 mg 150/173 (86,7%) Patienten der ITT-Population und in der Placebo-Gruppe 73/88 (83,0%) Patienten der ITT-Population umfasste.

Fehlende Daten wurden für die beiden supportiven sekundären Endpunkte im Fall von Tod oder fehlenden Post-Ausgangswerten nicht durch Imputation ergänzt. Im Falle einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangsvsiste verwendet. Es wurden keine weiterführenden Sensitivitätsanalysen zur Untersuchung des Effektes der fehlenden Daten durchgeführt. Aufgrund dieser Abweichung wird das Verzerrungspotenzial für diese zwei Endpunkte als hoch eingestuft.

Die Analyse des Sicherheitsendpunktes „Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse“ wurde in der Safety-Population durchgeführt, die mit der ITT-Population übereinstimmte.

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor (Tabelle 4-29). Auf Endpunktebene lagen für alle Endpunkte eine Verblindung des Endpunkterhebers und eine ergebnisunabhängige Berichterstattung vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen jeweils vollständig ab. Für die patientenrelevanten Endpunkte 6MWD, WHO FC, TTCW, Borg, EQ-5D, LPH und UE/SUE zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diese Endpunkte durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial für alle patientenrelevanten Endpunkte, d.h. 6MWD, WHO FC, TTCW, Borg, EQ-5D, LPH und UE/SUE als niedrig eingestuft.

4.4.2.3 Ergebnisse der Zulassungsstudie CHEST-1

4.4.2.3.1 Belege für und Hinweise auf Subgruppeneffekte

Die Wirksamkeits- (Tabellen 4-35 und 4-36) und Sicherheitsanalysen (Tabelle 4-39) anhand patientenrelevanter Endpunkte und Sicherheitsparameter lieferten Belege für und Hinweise auf Effektmodifikationen (Tabelle 4-34) durch Merkmale mit präspezifizierten Subgruppen (Tabelle 4-28). Aufgrund ihres gemischt kategorialen Charakters waren Interaktionstests für die patientenrelevanten Endpunkte WHO FC und TTCW nicht möglich. Die Endpunkte PVR und NTproBNP werden aufgrund der fraglichen Patientenrelevanz nicht dargestellt.

Tabelle 4-34: In CHEST-1 beobachtete Effektmodifikationen

Endpunkt / Parameter	6MWD Primäranalyse	6MWD Responderanalyse	Borg	EQ-5D	LPH	UE	SUE	Unterbrechung Studienmedik. wg. UE	UE von bes. Interesse – Hypotonie	UE von bes. Interesse - Hämorrhagie	UE von bes. Interesse - Synkope	Tod
CTEPH-Klassifikation		H				B						
WHO FC-Ausgangswert			H			B						
6MWD 320					H							
6MWD 380												
Geschlecht		H		H			B					
Alter		H										
Ethnizität		H		H								
Region				H								
B: Beleg für eine Effektmodifikation (Interaktionstest P -Wert $< 0,05$) H: Hinweis auf eine Effektmodifikation (Interaktionstest $0,05 \leq P$ -Wert $< 0,2$) □: Weder Beleg für noch Hinweis auf eine Effektmodifikation (Interaktionstest P -Wert $\geq 0,2$) 6MWD 320: 6MWD-Ausgangswert - Schwellenwert 320 m 6MWD 380: 6MWD-Ausgangswert - Schwellenwert 380 m Quelle: CSR-ASA												

Belege für Subgruppeneffekte gab es nur in Bezug auf die Sicherheitsparameter UE und SUE durch die Merkmale CTEPH-Klassifikation, WHO-Funktionsklasse-Ausgangswert und Geschlecht (Tabelle 4-34). Die beobachteten Effektunterschiede zwischen den jeweiligen Subgruppen waren größtenteils auf deutliche Unterschiede in den absoluten Inzidenzen der UE, bzw. SUE in den Placebo-Subgruppen zurückzuführen (Tabelle 4-40). Da in Bezug auf andere Sicherheitsparameter keine weiteren Signale für Effektmodifikationen gefunden wurden, erscheinen diese singulären Belege nicht sinnvoll interpretierbar.

Ebenso wenig erwiesen sich die vereinzelt Hinweise für Subgruppeneffekte auf die Wirksamkeitsendpunkte 6MWD (Responderanalyse), Borg, EQ-5D und LPH als relevant für

die Identifikation von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht.

Daher werden im Folgenden die jeweiligen Subgruppenergebnisse für die betreffenden Merkmale in den Tabellen 4-36, 4-38 und 4-40 wiedergegeben, aber nicht tiefergehend diskutiert.

4.4.2.3.2 Ergebnisse der Wirksamkeitsanalysen

Die Studie erreichte ihr Ziel, die Überlegenheit von Riociguat 1,0-2,5 mg vs. Placebo in der Behandlung erwachsener Patienten mit inoperabler CTEPH oder persistierender/rezidivierender CTEPH nach einer PEA („postoperativer CTEPH“) für den primären Endpunkt 6MWD nachzuweisen. Gleichgerichtet und ebenso statistisch signifikant waren die Effekte von Riociguat 1,0-2,5 mg auf die sekundären Endpunkte WHO-Funktionsklasse, Borg, EQ-5D, PVR und NT-proBNP. Für die Endpunkte Zeit bis zur klinischen Verschlechterung und LPH wurde ein gleichgerichteter Effekt mit numerischem Vorteil von Riociguat 1,0-2,5 mg gegenüber Placebo, ermittelt (Tabelle 4-35).

Tabelle 4-35: Ergebnisse der Wirksamkeitsanalysen CHEST-1: Zusammenfassung

Endpunkt	Population	Riociguat 1,0-2,5 mg vs. Placebo LS MW-Differenz [95%-KI]	P-Wert	Positiver Effekt von Riociguat
Primärer Endpunkt				
6MWD	ITT	45,69 m [24,74; 66,63 m]	<0,0001 ^W	ja
Sekundäre patientenrelevante Wirksamkeitsendpunkte				
WHO FC	ITT	*	0,0026 ^W	ja
TTCW	ITT	*	0,1724 ^L	ja
Borg	ITT	-1,07 [-1,60; -0,52]	0,0035 ^W	ja
EQ-5D	ITT	0,136 [0,065; 0,207]	<0,0001 ^W	ja
LPH	ITT	-5,76 [-10,45; -1,06]	0,1220 ^W	ja
Sekundäre, supportive Wirksamkeitsendpunkte				
PVR	mITT	-246,4 dyn×sec×cm ⁻⁵ [-303,3; -189,5 dyn×sec×cm ⁻⁵]	<0,0001 ^W	ja
NT-proBNP	mITT	-444,0 pg/mL [-843,0; -45,0 pg/mL]	<0,0001 ^W	ja
<p>6MWD: Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert WHO FC: Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert TTCW: Zeit bis zur klinischen Verschlechterung in Studientagen Borg: Veränderung des Ermüdungsempfindens und der Dyspnoe gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert EQ-5D: Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert LPH: Veränderung der gesundheitsbezogenen Lebensqualität gemäß LPH-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert PVR: Absolute Veränderung des PVR nach 16 Wochen im Vergleich zum Ausgangswert NT-proBNP: Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert mITT: Modifizierte ITT-Population LS MW: Least Squares Mittelwert KI: Konfidenzintervall, berechnet auf Grundlage des nachfolgend beschriebenen ANCOVA-Modells Berechnung der P-Werte für die Unterschiede zwischen Riociguat und Placebo nach 16 Wochen im Vergleich zum Ausgangswert: ANCOVA-Modell mit dem Ausgangswert als Kovariate und dem Behandlungsarm sowie der Region (d.h. Nordamerika, Südamerika, Europa, China, und Asien/Pazifikregion) und dem Stratum als feste Effekte. Bei Signifikanz des ggf. nachfolgenden Shapiro-Wilk-Tests der Residuen auf Normalverteilung Durchführung eines stratifizierten Wilcoxon-Tests (s. W) zur formalen Bestimmung der statistischen Signifikanz. L: Stratifizierter Logrank-Test mit Kaplan-Meier-Schätzern der Anteile der Patienten mit dem entsprechenden Ereignis W: Stratifizierter Wilcoxon-Test nach Region und Stratum</p>				

*Gemischter kategorialer Endpunkt, für den keine Konfidenzintervalle berechnet werden können

Quelle: CSR; 6MWD Tab. 9-2; WHO Tab. 9-10; TTCW Tab. 9-11; Borg Tab. 9-12; EQ-5D Tab. 9-13; LPH Tab. 9-15; PVR Tab. 9-7; NTproBNP Tab. 9-9; 95%-KI-Werte finden sich in den Subgruppenanalysen (2)

Im Detail sind die Ergebnisse der prädefinierten Primär- und Sekundäranalysen für die Riociguat 1,0-2,5 mg-Gruppe und die Placebo-Gruppe in Tabelle 4-36 dargestellt. Es sind Ergebnisse nur für diejenigen prädefinierten Subgruppen zusammengestellt, für die es nach Maßgabe eines Interaktionstests einen Beleg ($P < 0,05$) für oder einen Hinweis ($0,05 \leq P < 0,2$) auf einen Subgruppeneffekt gibt.

Tabelle 4-36: Ergebnisse der Wirksamkeitsanalysen der Studie CHEST-1: Primär- und Sekundäranalysen für die Riociguat-1,0-2,5 mg-Gruppe und die Placebo-Gruppe

Population	Placebo			Riociguat			Riociguat vs. Plc		
Endpunkt "Absolute Veränderung der 6MWD [m] bis Woche 16 im Vergleich zum Ausgangswert"									
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	LS MW- Differenz [95%-KI] P-Wert
ITT	88	356,0 (74,7)	350,4 (122,2)	-5,5	173	342,3 (81,9)	381,2 (119,2)	38,9	45,69 [24,74; 66,63] <0,0001 ^A <0,0001 ^W
Endpunkt "Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert"									
	N	Ausgangswert Pat. mit FC I II III IV N (%)	Letzte Visite Pat. mit FC I II III IV V (Tod) N (%)	Pat. mit Δ FC -2 FC -1 FC ±0 FC +1 FC +2 FC +3 FC N (%)	N	Ausgangswert Pat. mit FC I II III IV N (%)	Letzte Visite Pat. mit FC I II III IV V (Tod) N (%)	Pat. mit Δ FC -2 FC -1 FC ±0 FC +1 FC +2 FC +3 FC N (%)	* P-Wert
ITT	87	0 (0,0%) 25 (28,7%) 60 (69,0%) 2 (2,3%)	3 (3,4%) 30 (34,5%) 48 (55,2%) 3 (3,4%) 3 (3,4%)	0 (0,0%) 13 (14,9%) 68 (78,2%) 3 (3,4%) 3 (3,4%) 0 (0,0%)	173	3 (1,7%) 55 (31,8%) 107 (61,8%) 8 (4,6%)	14 (8,1%) 85 (49,1%) 65 (37,6%) 7 (4,0%) 2 (1,2%)	4 (2,3%) 53 (30,6%) 107 (61,8%) 7 (4,0%) 1 (0,6%) 1 (0,6%)	* 0,0026 ^W
*Für den gemischten kategorialen Endpunkt „Änderung der WHO-Funktionsklasse“ ist keine LS MW-Differenz berechenbar.									

Population	Placebo	Riociguat	Riociguat vs. Plc	
Endpunkt: "Zeit bis zur klinischen Verschlechterung in Studientagen"				
	Patienten mit ... A0. mindestens einem relevanten Ereignis A1. Krankenhauseinweisung wegen der PH A2. neuer PH-spezifischer Therapie A3. Anhaltende Verringerung der 6MWD wegen der PH A4. Anhaltende Verschlechterung der FC wegen der PH A5. Tod N (%) [Zeit bis zum relevanten Ereignis in Studientagen]	Patienten mit ... A0. mindestens einem relevanten Ereignis A1. Krankenhauseinweisung wegen der PH A2. neuer PH-spezifischer Therapie A3. Anhaltende Verringerung der 6MWD wegen der PH A4. Anhaltende Verschlechterung der FC wegen der PH A5. Tod N (%) [Zeit bis zum relevanten Ereignis in Studientagen]	*	P-Wert
ITT	88 A0. 5 (5,7%) A1. 1 (1,1%) [68] A2. 1 (1,1%) [87] A3. 2 (2,3%) [41, 57] A4. 1 (1,1%) [41] A5. 3 (3,4%) [19, 70, 71]	173 A0. 4 (2,3%) A1. 0 (0,0%) A2. 2 (1,2%) [58, 60] A3. 1 (0,6%) [44] A4. 0 (0,0%) A5. 2 (1,2%) [64, 88]	*	0,1724 ^L 0,2180 ^M
*Für den gemischten kategorialen Endpunkt „Zeit bis zur klinischen Verschlechterung“ ist keine LS MW-Differenz berechenbar.				

Population	Placebo				Riociguat				Riociguat vs. Plc	
Endpunkt „Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert“										
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
ITT	88	4,44 (2,20)	4,61 (2,56)	0,17	173	4,29 (2,25)	3,47 (2,34)	-0,83	-1,07 [-1,60; -0,52]	0,0035 ^w
Endpunkt „Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert“ – Subgruppenergebnisse										
Subgruppen	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	LS MW- Differenz [95%-KI]	
WHO-Funktionsklasse-Ausgangswert (<i>P</i> -Wert Interaktionstest = 0,0851)										
WHO FC I/II	25	4,16 (2,01)	3,28 (2,09)	-0,88	58	3,29 (2,11)	2,56 (2,10)	-0,73	-0,22 [-1,06; 0,62]	
WHO FC III/IV	62	4,61 (2,24)	5,21 (2,50)	0,6	115	4,80 (2,15)	3,92 (2,33)	-0,87	-1,36 [-2,06; -0,67]	

Population	Placebo				Riociguat				Riociguat vs. Plc	
Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D-Index nach 16 Wochen im Vergleich zum Ausgangswert“										
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
ITT	87	0,658 (0,246)	0,576 (0,349)	-0,082	172	0,645 (0,244)	0,706 (0,266)	0,062	0,136 [0,065; 0,207]	0,0002 ^A <0,0001 ^W
Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D-Index nach 16 Wochen im Vergleich zum Ausgangswert“ – Subgruppenergebnisse										
Subgruppen	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	LS MW- Differenz [95%-KI]	
Geschlecht (<i>P</i> -Wert Interaktionstest = 0,1595)										
Männlich	34	0,669 (0,272)	0,693 (0,222)	0,024	55	0,651 (0,219)	0,756 (0,177)	0,105	0,072 [0,003; 0,142]	
Weiblich	53	0,651 (0,230)	0,502 (0,395)	-0,150	117	0,642 (0,256)	0,682 (0,296)	0,041	0,185 [0,084; 0,286]	
Ethnizität (<i>P</i> -Wert Interaktionstest = 0,1963)										
Kaukasier	64	0,649 (0,260)	0,575 (0,352)	-0,074	119	0,649 (0,243)	0,677 (0,283)	0,028	0,102 [0,016; 0,187]	
Nicht-Kaukasier	23	0,684 (0,204)	0,581 (0,349)	-0,103	53	0,635 (0,249)	0,772 (0,209)	0,138	0,207 [0,083; 0,332]	

Population	Placebo			Riociguat				Riociguat vs. Plc	
Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D-Index nach 16 Wochen im Vergleich zum Ausgangswert“ – Subgruppenergebnisse (Forts.)									
Region (<i>P</i> -Wert Interaktionstest = 0,1987)									
Nordamerika	8	0,612 (0,306)	0,639 (0,271)	0,028	15	0,634 (0,264)	0,608 (0,339)	-0,026	-0,047 [-0,274; 0,179]
Südamerika	6	0,780 (0,177)	0,754 (0,129)	-0,027	15	0,572 (0,226)	0,755 (0,213)	0,184	0,081 [-0,122; 0,284]
Europa	53	0,640 (0,272)	0,556 (0,371)	-0,084	103	0,651 (0,243)	0,683 (0,280)	0,032	0,122 [0,026; 0,219]
China	11	0,621 (0,113)	0,499 (0,421)	-0,122	21	0,539 (0,270)	0,787 (0,182)	0,248	0,316 [0,100; 0,531]
Asien-Pazifik	9	0,770 (0,149)	0,618 (0,279)	-0,152	18	0,801 (0,132)	0,783 (0,200)	-0,018	0,147 [-0,039; 0,333]

Population	Placebo				Riociguat				Riociguat vs. Plc	
Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß LPH-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert“										
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
ITT	86	46,03 (22,58)	43,94 (26,56)	-2,09	170	41,48 (21,69)	34,76 (22,36)	-6,72	-5,76 [-10,45; -1,06]	0,0165 ^A 0,1220 ^W
Endpunkt „Veränderung der gesundheitsbezogenen Lebensqualität gemäß LPH-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert“ - Subgruppenergebnis										
Subgruppen	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
6MWD-Ausgangswert - Schwellenwert 320 m (P-Wert Interaktionstest = 0,1419)										
6MWD < 320 m	24	49,69 (19,91)	57,45 (25,10)	7,76	59	47,47 (22,12)	44,04 (24,94)	-3,43	-12,05 [-22,36; -1,74]	
6MWD ≥ 320 m	62	44,61 (23,53)	38,71 (25,41)	-5,90	111	38,30 (20,87)	29,82 (19,21)	-8,47	-4,33 [-9,02; 0,37]	
Endpunkt „Absolute Veränderung des PVR [dyn·sec·cm⁻⁵] nach 16 Wochen im Vergleich zum Ausgangswert“										
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
mITT	82	779,3 (400,9)	802,4 (415,2)	32,0	151	790,7 (431,6)	565,0 (312,3)	-214,5	-246,4 [-303,3; -189,5]	<0,0001 ^A <0,0001 ^W

Population	Placebo				Riociguat				Riociguat vs. Plc	
Endpunkt „Absolute Veränderung des PVR [dyn·sec·cm⁻⁵] nach 16 Wochen im Vergleich zum Ausgangswert“ - Subgruppenergebnisse										
Subgruppen	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
CTEPH-Klassifikation (<i>P</i> -Wert Interaktionstest = 0,0089)										
Inoperable CTEPH	63	851,5 (422,7)	884,8 (414,9)	33,4	105	866,5 (471,2)	609,4 (333,2)	-257,1	-284,9 [-357,1; -212,6]	
Postoperative CTEPH	19	540,2 (171,6)	529,1 (282,8)	-11,0	46	617,5 (252,3)	463,6 (230,7)	-153,9	-131,0 [-214,0; -48,0]	
Endpunkt „Absolute Veränderung des Biomarkers NT-proBNP (pg/mL) bis Woche 16 im Vergleich zum Ausgangswert“										
	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	N	Ausgangswert MW (SD)	Letzte Visite MW (SD)	LS MW Absolute Veränderung	LS MW- Differenz [95%-KI]	P-Wert
mITT	73	1.705,8 (2.567,2)	1.782,2 (2.263,5)	-25,7	150	1.508,3 (2.337,8)	1.217,6 (2.062,5)	-469,7	-444,0 [-843,0; -45,0]	0,0293 ^A <0,0001 ^W
<p>Berechnung der 95%-KI auf Grundlage des nachfolgend beschriebenen ANCOVA-Modells</p> <p>Berechnung der <i>P</i>-Werte für die Unterschiede zwischen Riociguat und Placebo nach 16 Wochen im Vergleich zum Ausgangswert:</p> <p>A: ANCOVA-Modell mit dem Ausgangswert als Kovariate und dem Behandlungsarm sowie der Region (d.h. Nordamerika, Südamerika, Europa, China, und Asien/Pazifikregion) und dem Stratum als feste Effekte. Bei Signifikanz des ggf. nachfolgenden Shapiro-Wilk-Tests der Residuen auf Normalverteilung Durchführung eines stratifizierten Wilcoxon-Tests (s. W) zur formalen Bestimmung der statistischen Signifikanz.</p> <p>L: Stratifizierter Logrank-Test mit Kaplan-Meier-Schätzern der Anteile der Patienten mit dem entsprechenden Ereignis</p> <p>M: Stratifizierter Mantel-Haenszel-Test nach Region und Stratum</p> <p>W: Stratifizierter Wilcoxon-Test nach Region und Stratum</p> <p>Für Subgruppen wurden keine <i>P</i>-Werte berechnet.</p> <p>Plc: Placebo</p> <p>Quelle: 6MWD: CSR Tab. 9-2; 14.2.2/1; 14.2.2/2 // WHO: CSR Tab. 9-10; 14.2.3/21 // TTCW: CSR Tab. 9-11; 14.2.3/26 // Borg: CSR Tab. 9-12; 14.2.3/35; 14.2.5/7 (2); CSR-ASA: Tab. 17/4, Fig. 17/4 (22) // EQ-5D: CSR Tab. 9-13; 9-14; 14.2.3/39; 14.2.4/59; 14.2.5/8 (2); CSR-ASA: Tab. 17/5 ; Fig. 17/5 (22) // LPH: CSR Tab. 9-15; 14.2.3/54; 14.2.5/9 (2); CSR-ASA Tab. 17/6; Fig. 17/6 (22) // PVR: CSR Tab. 9-7; 14.2.3/1; 14.2.3/2; 14.2.5/3 (2); CSR-ASA Tab. 17/3; Fig. 17/3 (22) // NTproBNP: CSR Tab. 9-9; 14.2.3/11; 14.2.3/12 (2)</p>										

4.4.2.3.2.1 Primärer Endpunkt 6MWD

Bis Woche 16 zeigte sich für die 6MWD in der Riociguat-1,0-2,5 mg-Gruppe eine Verbesserung von durchschnittlich 38,9 m (Least Squares, LS) gegenüber einer Verschlechterung von -5,5 m (LS) im Vergleich zum Ausgangswert in der Placebo-Gruppe. Der geschätzte Behandlungsunterschied zwischen diesen Gruppen betrug 45,69 m (LS MW-Differenz, 95%-KI 24,74; 66,63 m) und war statistisch signifikant ($P < 0,0001$) (Tabelle 4-36). Für CTEPH-Patienten wurde bisher keine Minimal Important Difference (MID) für die 6MWD publiziert, für PAH-Patienten liegt jedoch eine durch eine Kombination von verteilungs- und ankerbasierten Methoden berechnete MID von 33 m vor (23). Es gibt keinen medizinischen Grund, warum sich PAH- und CTEPH-Patienten in Bezug auf die MID unterscheiden sollten, da es sich um Erkrankungen mit vergleichbarer Symptomatik handelt. Nach dieser Maßgabe wird der in CHEST-1 beobachtete, positive Effekt von Riociguat 1,0-2,5 mg als klinisch relevante, „für die Patientinnen und Patienten spürbare Linderung der Erkrankung“ angesehen und als deutliche Verbesserung des therapielevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 VerfO G-BA gewertet.

Zur Bewertung des Einflusses fehlender Daten und der Robustheit der primären Analyse wurden auf Basis aktueller Empfehlungen der EMA (24) Sensitivitätsanalysen unter Anwendung weiterer Verfahren für den Umgang mit fehlenden Daten durchgeführt. Dazu zählten ein Mixed-Effect Models for Repeated Measures (MMRM)-Ansatz unter Zuhilfenahme diverser Varianz-Kovarianzmatrizen sowie multiple Imputationen, bei denen mit festen oder zunehmenden Abzügen vom Gruppendurchschnitt der Erkenntnis beim „Blinded Review“ Rechnung getragen wurde, dass Studienabbrecher tendenziell schlechtere Ergebnisse erzielten als Patienten, die die Studie planmäßig beendeten. Zur adäquaten Behandlung von Ausreißern wurde zudem eine robuste ANCOVA mit Huber-Schätzung durchgeführt (Tabelle 4-37).

Alle Sensitivitätsberechnungen bestätigen die Ergebnisse der primären Analysen, so dass eine mögliche Verzerrung des Effektes unwahrscheinlich ist. Die unteren Limits der 95%-KI lagen immer deutlich über 0 m, was als Beleg für einen methodenunabhängigen Behandlungseffekt zu werten ist.

Multiple Imputationen mit festen oder zunehmenden Abzügen in gleicher Höhe für beide Behandlungsgruppen ergaben geschätzte Behandlungsunterschiede von 43,69 m (95%-KI 26,25; 61,13 m), bzw. 41,81 m (95%-KI 24,05; 59,58 m) zugunsten von Riociguat 1,0-2,5 mg. Selbst mit sehr konservativen Modellen, in denen es lediglich für die Riociguat-1,0-2,5 mg-Gruppe Abzüge gab, wurden deutlich positive Effekte für Riociguat 1,0-2,5 mg berechnet (40,07 m, 95%-KI 22,94; 57,21 m, bzw. 38,71 m, 95%-KI 21,27; 56,15 m).

Nach dem MMRM-Ansatz ergab sich ein Behandlungseffekt von 44,40 m (95%-KI 27,94; 60,85 m), der ungefähr dem Ergebnis der Primäranalyse entspricht. Nach der robusten ANCOVA mit Huber-Schätzung ergab sich im Vergleich mit der primären Analyse ein etwas geringerer Behandlungseffekt von 40,31 m (95%-KI 25,86; 54,75 m). Dieser Befund ist dadurch zu erklären, dass bei diesem Verfahren *worst case* Imputationen nicht so sehr ins

Gewicht fallen wie bei den anderen Verfahren, und es in der Placebo-Gruppe mehr Todesfälle und Ereignisse im Sinne einer klinischen Verschlechterung gab.

Tabelle 4-37: Sensitivitätsanalysen zur absoluten Veränderung der 6MWD [m] bis Woche 16 im Vergleich zum Ausgangswert

Analyse	Geschätzter Behandlungsunterschied Riociguat 1,0-2,5 mg - Placebo	95% KI
Mixed-Effect Models for Repeated Measures (MMRM)	44,40	[27,94; 60,85]
Multiple Imputation – Fixed penalty: Riociguat 1,0-2,5 mg –60 m und Placebo –60 m	43,69	[26,25; 61,13]
Multiple Imputation – Decreasing slope: Riociguat 1,0-2,5 mg –20 m und Placebo –20 m pro Visite	41,81	[24,05; 59,58]
Multiple Imputation – Fixed penalty: Riociguat 1,0-2,5 mg –60 m und Placebo –0 m	40,07	[22,94; 57,21]
Multiple Imputation – Decreasing slope: Riociguat 1,0-2,5 mg –20 m und Placebo –0 m pro Visite	38,71	[21,27; 56,15]
Robuste Regression	40,31	[25,86; 54,75]
Quelle: CSR Tab. 9-4 (2)		

Insgesamt wichen alle Ergebnisse nur geringfügig von den Ergebnissen der Primäranalyse (LS MW-Differenz 45,69 m, 95%-KI 24,74; 66,63) ab. Dies kann mit der geringen Anzahl von Studienabbruchern in beiden Gruppen (Placebo-Gruppe 5,7%, Riociguat-1,0-2,5 mg-Gruppe 8,0%) erklärt werden.

Ergebnisse einer *a priori* definierten Responderanalyse bestätigten die klinische Relevanz der beobachteten Effekte. Prädefiniertes Responsekriterium war eine Verbesserung der 6MWD um mindestens 40 m gegenüber dem Ausgangswert. Dieser Wert ist nahezu identisch mit der lediglich verteilungsbasiert ermittelten klinisch relevanten Änderung (MID) (25) von 41 m und deutlich höher als die durch eine Kombination von verteilungs- und ankerbasierten Methoden berechnete MID von 33 m (23), so dass das Responsekriterium dem aktuellen Stand der Wissenschaft entspricht.

Die Analyse für die ITT-Population ergab, dass in der Riociguat-1,0-2,5 mg-Gruppe statistisch signifikant ($P < 0,0001$) mehr Responder waren (52,6%) als in der Placebo-Gruppe (23,9%; OR 3,541, 95%-KI 1,995; 6,285; RR 2,204, 95%-KI 1,479; 3,285; ARR 0,287, 95%-KI 0,171; 0,403) (Tabelle 4-38). Die Responderanalysen bestätigen somit die „für die Patientinnen und Patienten spürbare Linderung der Erkrankung“.

Tabelle 4-38: Ergebnisse der Wirksamkeitsanalysen: Responderanalyse

Population Subgruppe	Placebo		Riociguat		Riociguat vs. Placebo			
	Total N	Verbessert N (%)	Total N	Verbessert N (%)	RR [95%-KI]	ARR [95%-KI]	OR [95%-KI]	OR P-Wert
Patienten mit einer Verbesserung der 6MWD um mindestens 40 m im Vergleich von Ausgangswert zu letzter Visite								
ITT	88	21 (23,9%)	173	91 (52,6%)	2,204 [1,479; 3,285]	0,287 [0,171; 0,403]	3,541 [1,995; 6,285]	< 0,0001
Patienten mit einer Verbesserung der 6MWD um mindestens 40 m im Vergleich von Ausgangswert zu letzter Visite – Subgruppenergebnisse								
CTEPH-Klassifikation (<i>P</i> -Wert Interaktionstest = 0,0970)								
Inoperable CTEPH	68	16 (23,5%)	121	73 (60,3%)			4,943 [2,534; 9,643]	< 0,0001
Postoperative CTEPH	20	5 (25,0%)	52	18 (34,6%)			1,588 [0,497; 5,077]	0,4266
Geschlecht (<i>P</i> -Wert Interaktionstest = 0,1990)								
Männlich	34	12 (35,3%)	55	31 (56,4%)			2,368 [0,980; 5,723]	0,0520
Weiblich	54	9 (16,7%)	118	60 (50,8%)			5,172 [2,320; 11,528]	< 0,0001
Alter (<i>P</i> -Wert Interaktionstest = 0,1046)								
Alter < 65 Jahre	52	18 (34,6%)	99	58 (58,6%)			2,672 [1,330; 5,366]	0,0049
Alter ≥ 65 Jahre	36	3 (8,3%)	74	33 (44,6%)			8,854 [2,492; 31,453]	< 0,0001
Ethnizität (<i>P</i> -Wert Interaktionstest = 0,0622)								
Kaukasier	65	17 (26,2%)	120	56 (46,7%)			2,471 [1,278; 4,777]	0,0057
Nicht-Kaukasier	23	4 (17,4%)	53	35 (66,0%)			9,235 [2,730; 31,246]	< 0,0001
Rel. Ereignis: Relevantes Ereignis, OR: Odds Ratio, RR: Relatives Risiko, ARR: Absolute Risikoreduktion; RR und ARR wurden für Subgruppen nicht berechnet Dargestellt sind nur Ergebnisse für diejenigen Subgruppen, für die es nach Maßgabe des Interaktionstests einen Beleg für (<i>P</i> -Wert < 0,05) oder einen Hinweis (0,05 ≤ <i>P</i> -Wert < 0,2) auf einen Subgruppeneffekt gibt. Quelle: CSR-ASA Tab. 17/2; Fig 17/2 (22)								

4.4.2.3.2 Sekundäre Endpunkte

Konsistent mit den Ergebnissen für den primären Endpunkt 6MWD zeigte die Behandlung mit Riociguat 1,0-2,5 mg ausnahmslos gleichgerichtete, positive Effekte auf alle prädefinierten sekundären Endpunkte. In der ITT-Population wurde für die patientenrelevanten Endpunkte WHO-Funktionsklasse, Borg und EQ-5D ein statistisch signifikanter Unterschied zu Placebo zugunsten von Riociguat ermittelt (Tabelle 4-35). In Bezug auf die Endpunkte Zeit bis zur klinischen Verschlechterung und LPH hatten Patienten unter Riociguat 1,0-2,5 mg einen numerischen Vorteil gegenüber Patienten unter Placebo. Anhand der supportiven, sekundären Endpunkte PVR und NT-proBNP konnten ebenfalls, wengleich in einer mITT-Population, statistisch signifikante Vorteile von Riociguat 1,0-2,5 mg demonstriert werden (Tabelle 4-35).

WHO-Funktionsklasse

Für die WHO-Funktionsklasse wurde ein statistisch signifikanter Behandlungsunterschied ($P = 0,0026$) zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe zugunsten von Riociguat ermittelt (Tabelle 4-36).

Zu Behandlungsbeginn wurde für mehr als 90% der Patienten der ITT-Population in jedem Studienarm die WHO-Funktionsklasse II oder III zugeordnet. In der Riociguat-1,0-2,5 mg-Gruppe waren 61,8% der Patienten in WHO FC III und 31,8% in WHO FC II. In der Placebo-Gruppe war der Anteil der Patienten in der WHO FC III höher (69,0%), der in der WHO FC II etwas niedriger (28,7%). Insgesamt zeigte die Analyse, dass in der Riociguat-1,0-2,5 mg-Gruppe deutlich mehr Patienten (32,9%) eine klinisch relevante Verbesserung um mindestens eine Funktionsklasse erfuhren als in der Placebo-Gruppe (14,9%) (Tabelle 4-36). Da jede Änderung der WHO-Funktionsklasse ein klinisch relevanter Effekt ist, der mit einer „für die Patientinnen und Patienten spürbaren Linderung der Erkrankung“ einhergeht, wird dieser Befund als deutliche Verbesserung des therapierlevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 VerfO G-BA gewertet.

Zeit bis zur klinischen Verschlechterung

In Bezug auf die Zeit bis zur klinischen Verschlechterung gab es keinen statistisch signifikanten Behandlungseffekt von Riociguat-1,0-2,5 mg ($P = 0,1724$). Ebenso war der Unterschied in der Inzidenz der relevanten Ereignisse zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe statistisch nicht signifikant ($P = 0,2180$) (Tabelle 4-36).

Trotz der relativ kurzen Studiendauer und der geringen Ereignisrate zeigte sich jedoch ein positiver Trend zugunsten von Riociguat 1,0-2,5 mg. Anzahl und Anteil von Patienten mit mindestens einem relevanten Ereignis waren mit vier Patienten (2,3%) in der Riociguat-1,0-2,5 mg-Gruppe niedriger als in der Placebo-Gruppe mit fünf Patienten (5,7%), wobei in keinem Fall eine Herztransplantation, Lungentransplantation oder PEA vorgenommen wurde. Lediglich in Bezug auf die nach Relevanz und Schwere geringste Endpunktkomponente, den Beginn einer neuen PH-spezifischen Therapie, gab es keinen Vorteil, aber auch keinen Nachteil für die Patienten unter Riociguat (1,1 % der Patienten unter Placebo vs. 1,2% der Patienten unter Riociguat 1,0-2,5 mg). Alle anderen relevanten Ereignisse traten nicht nur

relativ, sondern auch absolut häufiger in der Placebo-Gruppe auf: Krankenhauseinweisung 1 (Placebo) zu 0 (Riociguat), Verringerung der 6MWD 2 zu 1 und Verschlechterung der WHO-Funktionsklasse 1 zu 0 Patienten (Tabelle 4-36). Dies galt auch für die Gesamtmortalität mit 3 Todesfällen in der Placebo-Gruppe gegenüber 2 Todesfällen in der Riociguat-1,0-2,5 mg-Gruppe. Mit Ausnahme eines Todesfalls in der Riociguat-1,0-2,5 mg-Gruppe hatten alle verstorbenen Patienten eine inoperable CTEPH. Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Borg

Bis Woche 16 zeigte sich für den Borg-Wert ein statistisch signifikanter Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe (LS MW-Differenz -1,07, 95%-KI -1,60; -0,52, $P = 0,0035$) zugunsten von Riociguat (Tabelle 4-36).

Bei nahezu identischen Ausgangsmittelwerten sank, d.h. verbesserte sich bis Woche 16 der durchschnittliche Borg-Wert in der Riociguat-1,0-2,5 mg-Gruppe um 0,83, wohingegen er in der Placebo-Gruppe um 0,17 anstieg.

Eine MID für die Veränderung des Borg-Skalenwertes wurde bisher nicht spezifisch für die CTEPH, wohl aber für die COPD ermittelt (26). Danach kann eine Änderung um eine Borg-Skaleneinheit als klinisch relevant angesehen werden. Ein Effekt dieser Größenordnung wurde in der Riociguat-1,0-2,5 mg-Gruppe erzielt und damit die klinische Relevanz des Effekts von Riociguat belegt. Dieser Effekt wird als deutliche Verbesserung des therapielevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 VerFO G-BA gewertet, da er eine „für die Patientinnen und Patienten spürbare Linderung der Erkrankung“ ist.

EQ-5D

Für den EQ-5D-Index zeigte sich in der ITT-Population bis Woche 16 bei Rücklaufquoten von 91,9% für die Riociguat-1,0-2,5 mg-Gruppe und 94,3% für die Placebo-Gruppe ein statistisch signifikanter Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe ($P < 0,0001$).

Die Ausgangswerte waren nahezu identisch. Bis Woche 16 stieg der EQ-5D-Index in der Riociguat-1,0-2,5 mg-Gruppe durchschnittlich um 0,062 im Vergleich zum Ausgangswert, während er in der Placebo-Gruppe um 0,082 sank. Dies entsprach einer LS MW-Differenz von 0,136 (95%-KI 0,065; 0,207) (Tabelle 4-36).

Auch für den EQ-5D wurde noch keine CTEPH-spezifische MID publiziert. Jedoch wurde durch Sekundäranalyse von acht longitudinalen Studien mit elf verschiedenen Patientengruppen, u.a. COPD-Patienten, eine mittlere MID von 0,074 für den EQ-5D berechnet (27). In der Riociguat-1,0-2,5 mg-Gruppe wurde diese MID deutlich überschritten, was für eine klinisch relevante Verbesserung der gesundheitsbezogenen Lebensqualität durch Riociguat 1,0-2,5 mg spricht. Dies wird als deutliche Verbesserung des therapielevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 VerFO G-BA durch eine „Abschwächung schwerwiegender Symptome“ angesehen.

LPH

Bis Woche 16 zeigte sich für den LPH-Gesamtscore bei Rücklaufquoten von 94,1% für die Riociguat-1,0-2,5 mg-Gruppe und 95,3% für die Placebo-Gruppe kein statistisch signifikanter Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe ($P = 0,1220$) (Tabelle 4-36).

Bei recht unterschiedlichen Ausgangswerten (46,03 in der Placebo-Gruppe, 41,48 in der Riociguat-1,0-2,5 mg-Gruppe) sank der LPH-Gesamtscore bis Woche 16 um durchschnittlich -6,72 in der Riociguat-1,0-2,5 mg-Gruppe und um -2,09 in der Placebo-Gruppe. Die LS MW-Differenz von -5,76 (95%-KI -10,45; -1,06) bedeutete eine stärkere Verbesserung der krankheitsspezifischen Lebensqualität in der Riociguat-1,0-2,5 mg-Gruppe. Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

PVR

Auch für den supportiven, sekundären Endpunkt PVR wurde bis Woche 16, hier jedoch in einer mITT-Population, ein statistisch signifikanter Behandlungsunterschied zugunsten von Riociguat 1,0-2,5 mg ($P < 0,0001$).

Bei vergleichbaren Ausgangswerten nahm der PVR-Wert bis Woche 16 in der Riociguat-1,0-2,5 mg-Gruppe deutlich um $-214,5 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$ ab, stieg dagegen in der Placebo-Gruppe um $32,0 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$. Dies entsprach einer LS MW-Differenz von $-246,43 \text{ dyn}\cdot\text{s}\cdot\text{cm}^{-5}$ (95%-KI -303,3; -189,5 $\text{dyn}\cdot\text{s}\cdot\text{cm}^{-5}$) (Tabelle 4-36).

Da die Patientenrelevanz des PVR zumindest fraglich, und er zudem mit einem hohen Verzerrungspotential behaftet ist (Tabelle 4-33), werden seine Ergebnisse im Folgenden nicht weiter diskutiert.

NT-proBNP

Der zweite supportive, sekundäre Endpunkt NT-proBNP fiel bis Woche 16 ebenfalls statistisch signifikant ($P < 0,0001$) zugunsten von Riociguat 1,0-2,5 mg aus. In dieser Zeit sank die Serumkonzentration des NT-ProBNP in der Riociguat-1,0-2,5 mg-Gruppe deutlich um durchschnittlich $-469,7 \text{ pg/mL}$, aber nur um $-25,7 \text{ pg/mL}$ in der Placebo-Gruppe (jeweils LS MW absolute Änderung). Das entsprach einer LS MW-Differenz von $-444,0 \text{ pg/mL}$ (95%-KI -843,0; -45,0 pg/mL) (Tabelle 4-36).

Da die Patientenrelevanz des NT-proBNP zumindest fraglich, und er zudem mit einem hohen Verzerrungspotential behaftet ist (Tabelle 4-33), werden auch seine Ergebnisse im Folgenden nicht weiter diskutiert.

4.4.2.3.3 Ergebnisse der Sicherheitsanalysen

Tabelle 4-39 fasst die Inzidenz der Todesfälle und anderer unerwünschter und schwerwiegender unerwünschter Ereignisse in der CHEST-1 Studie zusammen.

Tabelle 4-39: Zusammenfassung unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse in CHEST-1

CHEST-1 (Safety-Population)				
Patienten mit ...	Placebo		Riociguat 1,0-2,5 mg	
	N=88	(100%)	N=173	(100%)
mindestens 1 unerwünschtem Ereignis (UE)	76	(86,4%)	159	(91,9%)
mindestens 1 mit der Studienmedikation assoziierten UE	36	(40,9%)	103	(59,5%)
mindestens 1 schweren UE	10	(11,4%)	19	(11,0%)
mindestens 1 mit der Studienmedikation assoziierten schweren UE	2	(2,3%)	4	(2,3%)
mindestens 1 schwerwiegenden UE (SUE)	14	(15,9%)	34	(19,7%)
mindestens 1 mit der Studienmedikation assoziierten SUE	1	(1,1%)	6	(3,5%)
zu einer Unterbrechung der Behandlung mit der Studienmedikation führendem UE	2	(2,3%)	5	(2,9%)
mit der Studienmedikation assoziiertem, zu einer Unterbrechung der Behandlung mit der Studienmedikation führendem UE	0	(0%)	1	(0,6%)
zum Tode führendem UE	3	(3,4%)	2	(1,2%)
Quelle: CSR Tab. 10-4 (2)				

Zur Überprüfung, ob es analog zu den Wirksamkeitsendpunkten auch in Bezug auf die Sicherheitsparameter statistisch signifikante und klinische relevante Unterschiede zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe gibt, wurden *post hoc* auch für die Sicherheitsparameter statistische Gruppenvergleiche durchgeführt (Tabelle 4-40).

Tabelle 4-40: Ergebnisse der Sicherheitsanalysen: Statistische Gruppenvergleiche

Population	Placebo		Riociguat		Riociguat vs. Placebo							
	Total N	Rel. Ereignis N (%)	Total N	Rel. Ereignis N (%)	RR [95%-KI]	ARR [95%-KI]	OR [95%-KI]	OR P-Wert				
Patienten mit unerwünschtem Ereignis												
Safety	88	76 (86,4%)	173	159 (91,9%)	1,064 [0,969; 1,169]	0,055 [-0,027; 0,138]	1,793 [0,791; 4,064]	0,1661				
Patienten mit unerwünschtem Ereignis - Subgruppenergebnisse												
CTEPH-Klassifikation (<i>P</i> -Wert Interaktionstest = 0,0199)												
Inoperable CTEPH	68	61 (89,7%)	121	108 (89,3%)					0,953 [0,361; 2,517]	0,9230		
Postoperative CTEPH	20	15 (75,0%)	52	51 (98,1%)					16,998 [1,841; 156,940]	0,0028		
WHO-Funktionsklasse-Ausgangswert (<i>P</i> -Wert Interaktionstest = 0,0247)												
WHO FC I/II	25	19 (76,0%)	58	56 (96,6%)					8,842 [1,643; 47,575]	0,0056		
WHO FC III/IV	62	56 (90,3%)	115	103 (89,6%)	0,920 [0,327; 2,583]	0,8732						
Patienten mit schwerwiegendem unerwünschtem Ereignis												
Safety	88	14 (15,9%)	173	34 (19,7%)	1,235 [0,701; 2,178]	0,037 [-0,059; 0,134]	1,293 [0,653; 2,560]	0,4562				
Patienten mit schwerwiegendem unerwünschtem Ereignis - Subgruppenergebnisse												
Geschlecht (<i>P</i> -Wert Interaktionstest = 0,0078)												
Männlich	34	1 (2,9%)	55	15 (27,3%)					12,375 [1,552; 98,665]	0,0013		
Weiblich	54	13 (24,1%)	118	19 (16,1%)					0,605 [0,274; 1,339]	0,2202		
Patienten mit einer Unterbrechung der Behandlung mit der Studienmedikation wegen unerwünschtem Ereignis												
Safety	88	2 (2,3%)	173	5 (2,9%)	1,272 [0,252; 6,423]	0,006 [-0,034; 0,046]	1,280 [0,243; 6,732]	0,7673				

Population	Placebo		Riociguat		Riociguat vs. Placebo			
	Total N	Rel. Ereignis N (%)	Total N	Rel. Ereignis N (%)	RR [95%-KI]	ARR [95%-KI]	OR [95%-KI]	OR P-Wert
Patienten mit Hypotonie oder Blutdrucksenkung								
Safety	88	4 (4,5%)	173	19 (11,0%)	2,416 [0,848; 6,886]	0,064 [0,001; 0,128]	2,591 [0,853; 7,865]	0,0679
Patienten mit Hämorrhagie								
Safety	88	10 (11,4%)	173	23 (13,3%)	1,170 [0,583; 2,348]	0,019 [-0,064; 0,103]	1,196 [0,542; 2,638]	0,6547
Patienten mit Bewusstlosigkeit, Präsynkope oder Synkope								
Safety	88	3 (3,4%)	173	6 (3,5%)	1,017 [0,261; 3,971]	0,001 [-0,046; 0,047]	1,018 [0,248; 4,170]	0,9802
Patienten mit Tod (Gesamtmortalität)								
Safety	88	3 (3,4%)	173	2 (1,2%)	0,339 [0,058; 1,992]	-0,023 [-0,064; 0,019]	0,331 [0,054; 2,020]	0,2256
Rel. Ereignis: Relevantes Ereignis, OR: Odds Ratio, RR: Relatives Risiko, ARR: Absolute Risikoreduktion Dargestellt sind nur Ergebnisse für diejenigen Subgruppen, für die es nach Maßgabe des Interaktionstests einen Beleg für (P -Wert $< 0,05$) oder einen Hinweis ($0,05 \leq P$ -Wert $< 0,2$) auf einen Subgruppeneffekt gibt. Quelle: CSR-ASA Tab. 17/7-17/13; Fig. 17/7-17/13 (22)								

Mortalität

In der Placebo-Gruppe traten mehr Todesfälle auf (3/88 Patienten, 3,4%) als in der Riociguat-1,0-2,5 mg-Gruppe (2/173 Patienten, 1,2%) (Tabelle 4-39). Todesursachen in der Placebo-Gruppe waren in einem Fall ein Herz-Lungen-Versagen und in zwei Fällen Herzstillstand. Todesursachen in der Riociguat-1,0-2,5 mg-Gruppe waren in einem Fall Herzversagen, im anderen ein akutes Nierenversagen mit nachfolgender Anämie und Blutungen an der Kathetereinstichstelle. Hierbei wurde das akute Nierenversagen vom Prüfarzt als mit der Studienmedikation assoziiert klassifiziert. Alle anderen Todesfälle wurden nicht mit der Studienmedikation assoziiert.

Für die Gesamtmortalität nach 16 Wochen wurden eine statistisch nicht signifikante OR von 0,331 (95%-KI 0,054; 2,020, $P = 0,2256$), ein RR von 0,339 (95%-KI 0,058; 1,992) und eine ARR von -0,023 (95%-KI -0,064; 0,019) zugunsten von Riociguat 1,0-2,5 mg ermittelt (Tabelle 4-40). Es ließ sich keine Aussage zum Zusatznutzen von Riociguat ableiten.

Unerwünschte Ereignisse

Mindestens ein UE trat bei 91,9% der Patienten der Riociguat-1,0-2,5 mg-Gruppe, in der Placebo-Gruppe waren es mit 86,4% geringfügig weniger Patienten (Tabelle 4-39). Dieser Unterschied war statistisch nicht signifikant (OR 1,793, 95%-KI 0,791; 4,064, $P = 0,1661$, RR 1,064, 95%-KI 0,969; 1,169, ARR 0,055, 95%-KI -0,027; 0,138) (Tabelle 4-40). Es wurde kein geringerer oder größerer Schaden von Riociguat gefunden, der den Zusatznutzen verringert.

Tabelle 4-41 zeigt die Inzidenz aller unerwünschten Ereignisse in CHEST-1 nach Systemorganklasse und Preferred Term und einer Inzidenz $\geq 5\%$ bezogen auf die Systemorganklasse. Die häufigsten UE waren demnach Erkrankungen des Nervensystems, wovon in der Riociguat-1,0-2,5 mg-Gruppe mit 49,1% mehr Patienten betroffen waren als in der Placebo-Gruppe mit 30,7%. Ebenfalls häufiger traten gastrointestinale Erkrankungen in der Riociguat-1,0-2,5 mg-Gruppe (48,0% der Patienten) als in der Placebo-Gruppe (30,7% der Patienten) auf. Infektionen oder parasitäre Erkrankungen (Riociguat-1,0-2,5 mg-Gruppe 37,6%, Placebo-Gruppe 33,0% der Patienten) und allgemeine Erkrankungen und Beschwerden am Verabreichungsort (Riociguat-1,0-2,5 mg-Gruppe 31,8%, Placebo-Gruppe 35,2% der Patienten) wurden für ungefähr ein Drittel der Patienten berichtet.

Dafür traten in den drei Klassen „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ sowie „Untersuchungen“ in der Placebo-Gruppe (35,2%, 36,4% sowie 25,0%) häufiger Beschwerden auf als in der Riociguat-1,0-2,5 mg-Gruppe (31,8%, 24,9% sowie 17,9%).

Das häufigste, einzelne Symptom (Preferred Term) in der Riociguat-1,0-2,5 mg-Gruppe waren Kopfschmerzen (24,9%), gefolgt von Schwindel (22,5%). In der Placebo-Gruppe waren dies mit 20,5% periphere Ödeme, gefolgt von Husten (18,2%).

Tabelle 4-41: Anzahl der Patienten mit mindestens einem UE nach Systemorganklasse und Preferred Term in RCT mit dem zu bewertenden Arzneimittel mit einer Inzidenz $\geq 5\%$

Studie CHEST-1 (Safety-Population)				
Systemorganklasse Preferred Term	Placebo		Riociguat 1,0-2,5 mg	
MedDRA, Version 15.0	N=88	(100%)	N=173	(100%)
Gesamtzahl Patienten mit irgendeinem UE	76	(86,4%)	159	(91,9%)
Gastrointestinale Erkrankungen	27	(30,7%)	83	(48,0%)
Obstipation	1	(1,1%)	10	(5,8%)
Diarrhöe	4	(4,5%)	17	(9,8%)
Dyspepsie	7	(8,0%)	31	(17,9%)
Nausea	7	(8,0%)	19	(11,0%)
Erbrechen	3	(3,4%)	17	(9,8%)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	31	(35,2%)	55	(31,8%)
Peripheres Ödem	18	(20,5%)	27	(15,6%)
Infektionen und parasitäre Erkrankungen	29	(33,0%)	65	(37,6%)
Nasopharyngitis	8	(9,1%)	26	(15,0%)
Infektion der oberen Atemwege	4	(4,5%)	10	(5,8%)
Untersuchungen	22	(25,0%)	31	(17,9%)
Erhöhung des Kreatininspiegels	5	(5,7%)	3	(1,7%)
Erhöhte International Normalized Ratio	4	(4,5%)	10	(5,8%)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	19	(21,6%)	34	(19,7%)
Rückenschmerzen	5	(5,7%)	7	(4,0%)
Extremitätenschmerzen	5	(5,7%)	3	(1,7%)
Erkrankungen des Nervensystems	27	(30,7%)	85	(49,1%)
Schwindel	11	(12,5%)	39	(22,5%)
Kopfschmerzen	12	(13,6%)	43	(24,9%)
Psychiatrische Störungen	8	(9,1%)	8	(4,6%)
Schlaflosigkeit	6	(6,8%)	4	(2,3%)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	32	(36,4%)	43	(24,9%)
Husten	16	(18,2%)	9	(5,2%)
Dyspnoe	12	(13,6%)	8	(4,6%)
Gefäßerkrankungen	13	(14,8%)	29	(16,8%)
Hypotonie	3	(3,4%)	16	(9,2%)

Quelle: CSR Tab. 10-7 (2)

In der Riociguat-1,0-2,5 mg-Gruppe traten etwas häufiger schwerwiegende UE (Tabelle 4-42) als in der Placebo-Gruppe (19,7% vs. 15,9% der Patienten) auf, wobei auch dieser Unterschied statistisch nicht signifikant war (OR 1,293, 95%-KI 0,653; 2,560, $P = 0,4562$, RR 1,235, 95%-KI 0,701; 2,178, ARR 0,037, 95%-KI -0,059; 0,134) (Tabelle 4-40). In Bezug auf SUE gab es einen Beleg (P -Wert Interaktionstest = 0,0078) für eine Effektmodifikation durch das Merkmal Geschlecht. Da die Interaktionstests bei keinem anderen Sicherheitsparameter einen Beleg für oder einen Hinweis auf einen Subgruppeneffekt durch das Merkmal Geschlecht lieferten (22), ist dieser Befund singular und lässt sich nicht sinnvoll interpretieren. Es wurde kein geringerer oder größerer Schaden von Riociguat gefunden, der den Zusatznutzen verringert.

Tabelle 4-42: SUE nach Systemorganklasse und Preferred Term in CHEST-1

Studie CHEST-1 (Safety-Population) Systemorganklasse Preferred Term	Placebo		Riociguat 1,0-2,5 mg	
	N=88	(100%)	N=173	(100%)
MedDRA, Version 15.0				
Gesamtanzahl Patienten mit irgendeinem SUE	14	(15,9%)	34	(19,7%)
Herzerkrankungen	6	(6,8%)	10	(5,8%)
Herzstillstand	2	(2,3%)	0	(0,0%)
Rechtsherzversagen	3	(3,4%)	6	(3,5%)
Gastrointestinale Erkrankungen	1	(1,1%)	4	(2,3%)
Gastritis	0	(0,0%)	2	(1,2%)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	0	(0,0%)	4	(2,3%)
Brustschmerzen	0	(0,0%)	2	(1,2%)
Erkrankungen des Nervensystems	3	(3,4%)	4	(2,3%)
Synkope	3	(3,4%)	4	(2,3%)
Erkrankungen der Nieren und Harnwege	1	(1,1%)	4	(2,3%)
Akutes Nierenversagen	0	(0,0%)	2	(1,2%)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	2	(2,3%)	11	(6,4%)
Hämoptyse	0	(0,0%)	3	(1,7%)
Pulmonal arterielle Hypertonie	0	(0,0%)	2	(1,2%)
Respiratorische Insuffizienz	0	(0,0%)	2	(1,2%)

Dargestellt sind nur diejenigen Preferred Terms, die mindestens zweimal in irgendeinem Behandlungsarm zu finden sind.

Quelle: CSR Tab. 10-12. (2)

Bis Woche 16 unterbrochen in der Riociguat-1,0-2,5 mg-Gruppe 5 von 173 Patienten (2,9%), in der Placebo-Gruppe 2 von 88 Patienten (2,3%) die Behandlung mit der Studienmedikation wegen eines unerwünschten Ereignisses. Dieser Unterschied zugunsten von Placebo war statistisch nicht signifikant (OR 1,280, 95%-KI 0,243; 6,732, $P = 0,7673$) (Tabelle 4-40). Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Wie aufgrund der gefäßerweiternden Wirkung von Riociguat (19) zu erwarten war, wurde eine Hypotonie in der Riociguat-1,0-2,5 mg-Gruppe (9,2%) häufiger beobachtet als in der Placebo-Gruppe (3,4%) (Tabelle 4-41). Eine Blutdrucksenkung trat ebenfalls in der Riociguat-1,0-2,5 mg-Gruppe häufiger auf als in der Placebogruppe (1,7% vs. 1,1% der Patienten). Der Unterschied im gemeinsam betrachteten Auftreten beider UE war jedoch statistisch nicht signifikant (OR 2,591, 95%-KI 0,853; 7,865, $P = 0,0679$, RR 2,416, 95%-KI 0,848; 6,886, ARR 0,064, 95%-KI 0,001; 0,128) (Tabelle 4-40). Mit nur einer Ausnahme wurden diese UE unter Riociguat jedoch von den Prüfern als leicht oder mittelschwer bewertet. Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Weil Anämien für Patienten der Riociguat-1,0-2,5 mg-Gruppe häufiger gemeldet wurden (6 Patienten, 3,5%) als für Patienten der Placebo-Gruppe (1 Patient, 1,1%) wurde retrospektiv gezielt nach Blutungen (Hämorrhagie) in der klinischen Datenbank gesucht. Hämorrhagien traten in den beiden Behandlungsgruppen ungefähr gleich häufig auf (Riociguat-1,0-2,5 mg-

Gruppe: 23 Patienten, 13,3% vs. Placebo-Gruppe: 10 Patienten, 11,4%), der statistische Gruppenvergleich erbrachte keine statistisch signifikanten Unterschiede (OR 1,196, 95%-KI 0,542; 2,638, $P = 0,6547$, RR 1,170, 95%-KI 0,583; 2,348, ARR 0,019, 95%-KI -0,064; 0,103) (Tabelle 4-40). Das ausschließliche Auftreten schwerwiegender Hämoptysen in der Riociguat-1,0-2,5 mg-Gruppe (3 Patienten, 1,7%, Tabelle 4-42) kann jedoch ein Hinweis auf eine Erhöhung des Blutungsrisikos durch Riociguat sein. Weitere Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Synkopen traten in beiden Gruppen mit geringer Häufigkeit auf, d.h. bei 4 Patienten (2,3%) in der Riociguat-1,0-2,5 mg-Gruppe und 3 Patienten (3,4%) in der Placebo-Gruppe (Tabelle 4-42). Wurden auch die nicht als SUE eingestuft Synkopen sowie Präsynkopen und Bewusstlosigkeiten in den Gruppenvergleich einbezogen, erhöhte sich der Anteil der betroffenen Patienten in der Riociguat-1,0-2,5 mg-Gruppe auf 3,5%. Es gab keinen statistisch signifikanten Unterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe (Tabelle 4-40). Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Ein Vergleich aller nach Meinung des Prüfarztes mit der Studienmedikation assoziierten unerwünschten Ereignisse (Tabelle 4-43) unter Riociguat 1,0-2,5 mg und Placebo weist Riociguat als Arzneimittel mit einem guten Nebenwirkungsprofil aus, das außer den erwartbaren, in der Regel nicht schweren Hypotonien und Dyspepsien, sowie Schwindel und Kopfschmerzen keine auffälligen Unterschiede zu Placebo aufweist. Damit ist für Riociguat 1,0-2,5 mg kein Schaden festzustellen, der das Ausmaß des Zusatznutzens mindert.

Tabelle 4-43: Anzahl der Patienten mit mindestens einem mit der Studienmedikation assoziierten UE nach Systemorganklasse und Preferred Term in CHEST-1

Studie CHEST-1 (Safety-Population)				
Systemorganklasse	Placebo		Riociguat	
Preferred Term				
MedDRA, Version 15.0	N=88	(100%)	N=173	(100%)
Anzahl Patienten mit mindestens einem mit der Studienmedikation assoziierten UE	36	(40,9%)	103	(59,5%)
Herzerkrankungen	4	(4,5%)	6	(3,5%)
Bradykardie	1	(1,1%)	0	(0,0%)
Palpitationen	2	(2,3%)	4	(2,3%)
Sinus-Tachykardie	0	(0,0%)	1	(0,6%)
Supraventrikuläre Extrasystolen	1	(1,1%)	0	(0,0%)
Tachykardie	0	(0,0%)	1	(0,6%)
Erkrankungen des Ohrs und des Labyrinths	1	(1,1%)	1	(0,6%)
Vertigo	1	(1,1%)	1	(0,6%)
Augenerkrankungen	1	(1,1%)	7	(4,0%)
Konjunktivale Hyperämie	0	(0,0%)	2	(1,2%)
Augenschwellung	1	(1,1%)	1	(0,6%)
Lidödeme	0	(0,0%)	2	(1,2%)
Okuläre Hyperämie	0	(0,0%)	1	(0,6%)
Nebelsehen	0	(0,0%)	1	(0,6%)
Gastrointestinale Erkrankungen	15	(17,0%)	53	(30,6%)
Abdominelle Beschwerden	1	(1,1%)	6	(3,5%)
Abdominelle Blähung	0	(0,0%)	1	(0,6%)
Abdominalschmerzen	0	(0,0%)	2	(1,2%)
Abdominalschmerzen (untere)	0	(0,0%)	1	(0,6%)
Abdominalschmerzen (obere)	2	(2,3%)	3	(1,7%)
Konstipation	0	(0,0%)	4	(2,3%)
Diarrhöe	3	(3,4%)	7	(4,0%)
Mundtrockenheit	1	(1,1%)	3	(1,7%)
Dyspepsie	6	(6,8%)	21	(12,1%)
Dysphagia	0	(0,0%)	3	(1,7%)
Gastritis	0	(0,0%)	3	(1,7%)
Gastroösophageale Refluxkrankheit	0	(0,0%)	5	(2,9%)
Glossodynie	0	(0,0%)	1	(0,6%)
Nausea	5	(5,7%)	11	(6,4%)
Ösophageale Beschwerden	0	(0,0%)	1	(0,6%)
Erbrechen	2	(2,3%)	8	(4,6%)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	11	(12,5%)	15	(8,7%)
Asthenie	1	(1,1%)	1	(0,6%)
Brustbeschwerden	0	(0,0%)	1	(0,6%)
Brustschmerzen	0	(0,0%)	1	(0,6%)
Gesichtsödem	0	(0,0%)	2	(1,2%)
Müdigkeit	1	(1,1%)	2	(1,2%)
Kältegefühl	0	(0,0%)	1	(0,6%)
Ödem	1	(1,1%)	2	(1,2%)
Peripheres Ödem	8	(9,1%)	5	(2,9%)
Pyrexie	0	(0,0%)	1	(0,6%)
Fremdkörpergefühl	0	(0,0%)	1	(0,6%)
Wirbelsäulenschmerzen	0	(0,0%)	1	(0,6%)
Affektionen der Leber und Gallenblase	1	(1,1%)	0	(0,0%)
Abnormale Leberfunktion	1	(1,1%)	0	(0,0%)

Infektionen und parasitäre Erkrankungen	0	(0,0%)	3	(1,7%)
Chronische Sinusitis	0	(0,0%)	1	(0,6%)
Pneumonie	0	(0,0%)	1	(0,6%)
Infektionen der oberen Atemwege	0	(0,0%)	1	(0,6%)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	1	(1,1%)	1	(0,6%)
Verletzung	1	(1,1%)	0	(0,0%)
Postprozedurale Beschwerden	0	(0,0%)	1	(0,6%)
Untersuchungen	5	(5,7%)	10	(5,8%)
Erhöhung der alkalischen Phosphatase	0	(0,0%)	1	(0,6%)
Erhöhung der Creatin-Phosphokinase	1	(1,1%)	1	(0,6%)
Erhöhung des Kreatininspiegels	1	(1,1%)	0	(0,0%)
Niedriger Blutdruck	1	(1,1%)	3	(1,7%)
Erhöhter Harnstoffwert	1	(1,1%)	1	(0,6%)
Eingeschränkte Kreatinin-Clearance	1	(1,1%)	0	(0,0%)
Erhöhung der Gamma-Glutamyltransferase	1	(1,1%)	1	(0,6%)
Vermindertes Hämoglobin	0	(0,0%)	1	(0,6%)
Erhöhung der Leberenzyme	1	(1,1%)	0	(0,0%)
Erhöhte International Normalized Ratio	0	(0,0%)	1	(0,6%)
Verringerung der Leukozytenzahl	0	(0,0%)	2	(1,2%)
Stoffwechsel- und Ernährungsstörungen	1	(1,1%)	1	(0,6%)
Verringerter Appetit	1	(1,1%)	1	(0,6%)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	4	(4,5%)	3	(1,7%)
Hämarthrose	0	(0,0%)	1	(0,6%)
Gelenkschwellung	1	(1,1%)	0	(0,0%)
Muskelkrämpfe	2	(2,3%)	1	(0,6%)
Muskuloskeletale Steifigkeit	0	(0,0%)	1	(0,6%)
Extremitätenschmerzen	1	(1,1%)	0	(0,6%)
Gutartige, bösartige und unspezifische Neubildungen (einschl. Zysten und Polypen)	0	(0,0%)	2	(1,2%)
Seborrhoische Keratose	0	(0,0%)	1	(0,6%)
Schilddrüsenneoplasma	0	(0,0%)	1	(0,6%)
Erkrankungen des Nervensystems	12	(13,6%)	53	(30,6%)
Gleichgewichtsstörungen	1	(1,1%)	0	(0,0%)
Brennendes Gefühl	0	(0,0%)	2	(1,2%)
Schwindel	3	(3,4%)	26	(15,0%)
Kopfschmerzen	7	(8,0%)	27	(15,6%)
Migräne	0	(0,0%)	1	(0,6%)
Parästhesie	2	(2,3%)	2	(1,2%)
Präsynkope	0	(0,0%)	1	(0,6%)
Synkope	1	(1,1%)	3	(1,7%)
Tremor	1	(1,1%)	1	(0,6%)
Psychiatrische Erkrankungen	2	(2,3%)	0	(0,0%)
Schlaflosigkeit	1	(1,1%)	0	(0,0%)
Verminderte Libido	1	(1,1%)	0	(0,0%)
Erkrankungen der Nieren und Harnwege	0	(0,0%)	2	(1,2%)
Akutes Nierenversagen	0	(0,0%)	1	(0,6%)
Abnormaler Uringeruch	0	(0,0%)	1	(0,6%)
Erkrankungen der Geschlechtsorgane und der Brustdrüse	0	(0,0%)	1	(0,6%)
Erektile Dysfunktion	0	(0,0%)	1	(0,6%)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	3	(3,4%)	10	(5,8%)
Asthma	0	(0,0%)	1	(0,6%)
Würgen	0	(0,0%)	1	(0,6%)
Husten	1	(1,1%)	1	(0,6%)

Dyspnoe	0	(0,0%)	1	(0,6%)
Hämoptyse	0	(0,0%)	1	(0,6%)
Nasale Obstruktion	2	(2,3%)	3	(1,7%)
Nasentrockenheit	0	(0,0%)	1	(0,6%)
Oropharyngeale Beschwerden	0	(0,0%)	1	(0,6%)
Pharyngeales Erythem	0	(0,0%)	1	(0,6%)
Erkrankungen der Haut und des Unterhautzellgewebes	4	(4,5%)	4	(2,3%)
Dermatitis	0	(0,0%)	1	(0,6%)
Ekzem	0	(0,0%)	1	(0,6%)
Erythem	3	(3,4%)	0	(0,0%)
Hyperhidrose	1	(1,1%)	0	(0,0%)
Pruritus	0	(0,0%)	1	(0,6%)
Exfoliation	0	(0,0%)	1	(0,6%)
Gefäßerkrankungen	3	(3,4%)	22	(12,7%)
Erröten	2	(2,3%)	7	(4,0%)
Hitzewallung	1	(1,1%)	1	(0,6%)
Hypotonie	0	(0,0%)	14	(8,1%)
Orthostatische Hypotonie	0	(0,0%)	1	(0,6%)
Vasodilatation	0	(0,0%)	1	(0,6%)
Quelle: CSR Tab. 14.3.1/10 (2)				

4.4.2.4 Ergebnisse einer Interimsanalyse der unkontrollierten Folgestudie CHEST-2

Von den 243 Patienten, die die Behandlung in CHEST-1 abgeschlossen hatten (Abbildung 3), traten 237 (97,5%) in CHEST-2 ein. Für Patienten aus der ehemaligen Placebo-Gruppe von CHEST-1 wurde in einer analog zu CHEST-1 durchgeführten (Tabelle 4-26) Titrationsphase die Riociguat-Erhaltungsdosis bestimmt. Patienten aus der ehemaligen Riociguat-1,0-2,5 mg-Gruppe von CHEST-1 erhielten weiterhin ihre in CHEST-1 ermittelte Riociguat-Erhaltungsdosis, wobei zur Aufrechterhaltung der Verblindung bei diesen Patienten eine Scheintitration durchgeführt wurde. Nach der 8-wöchigen Titrationsphase wurde die Studie entblindet.

Es wurde eine explorative, deskriptive Interimsanalyse zur Wirksamkeit und Sicherheit von Riociguat 1,0-2,5 mg in den ersten 12 Wochen nach Studienbeginn von CHEST-2 durchgeführt (28). Sie umfasste Daten von 194 Patienten, von denen 129 Patienten ehemals in der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe und 65 Patienten ehemals in der CHEST-1-Placebo-Gruppe waren.

Im Vergleich zum CHEST-1-Ausgangswert verbesserte sich die 6MWD in der ehemaligen CHEST-1-Riociguat-1,0-2,5 mg-Gruppe um durchschnittlich 63,3 m (SD 64,4 m), ein Wert, der wiederum höher lag als die Verbesserung dieser Gruppe von durchschnittlich 51,2 m (SD 61,8 m) bei Abschlussvisite von CHEST-1 im Vergleich zum Ausgangswert. In der ehemaligen CHEST-1-Placebo-Gruppe ergab sich eine Verbesserung um durchschnittlich 35,3 m (SD 74,8 m) gegenüber dem CHEST-1-Ausgangswert und lag somit höher als die Verbesserung dieser Gruppe von durchschnittlich 4,1 m (SD 66,2 m) bei Abschlussvisite von CHEST-1 im Vergleich zum Ausgangswert.

Konsistente Ergebnisse in Bezug auf die Wirksamkeitsparameter WHO-Funktionsklasse, Borg, EQ-5D und LPH deuten an, dass der positive Behandlungseffekt von Riociguat 1,0-2,5 mg bei Patienten aus der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe erhalten bleibt, und Patienten aus der CHEST-1-Placebo-Gruppe von Riociguat 1,0-2,5 mg profitieren.

In CHEST-2 traten 5 (2,6%) Todesfälle im Vergleich zu 2 (1,2%) Todesfällen in der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe (Tabelle 4-39) auf, wobei 3 der in CHEST-2 verstorbenen Patienten der ehemaligen CHEST-1-Riociguat-1,0-2,5 mg-Gruppe angehörten (2,3%), 2 der ehemaligen CHEST-1-Placebo-Gruppe (3,1%). Keiner der Todesfälle in CHEST-2 wurde vom Prüfarzt als mit der Studienmedikation assoziiert klassifiziert. Auch in Bezug auf andere Sicherheitsparameter zeigten sich keine auffälligen Unterschiede zwischen CHEST-1 und CHEST-2. In CHEST-2 hatten 3 (1,5%) Patienten ein UE, das zu einer Unterbrechung der Behandlung mit der Studienmedikation führte, im Vergleich zu 5 (2,9%) Patienten in der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe. Anzahl und Anteil der Patienten mit mindestens einem mit der Studienmedikation assoziierten SUE waren in CHEST-2 mit 8 Fällen (4,1%) etwa so hoch wie in der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe mit 6 Fällen (3,5%). Ebenso waren Anzahl und Anteil der Patienten mit mindestens einem UE in CHEST-2 mit 173 Fällen (89,2%) vergleichbar mit der CHEST-1-Riociguat-1,0-2,5 mg-Gruppe mit 159 Fällen (91,9%). Insgesamt deuten die Ergebnisse dieser Sicherheitsanalyse an, dass Riociguat 1,0-2,5 mg auch bei längerer Gabe sicher und gut verträglich ist.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

4.4.3.1 Vorbemerkung

Riociguat ist die einzige zugelassene Möglichkeit, CTEPH medikamentös zu behandeln. Mit der Zulassungsstudie CHEST-1 wurde das Ziel erreicht, die Überlegenheit von Riociguat 1,0-2,5 mg gegenüber Placebo in der Behandlung erwachsener Patienten sowohl mit inoperabler CTEPH als auch mit persistierender/rezidivierender CTEPH nach einer chirurgischen Behandlung für den primären Endpunkt 6MWD nachzuweisen. Konsistent mit den Ergebnissen für den primären Endpunkt 6MWD hatte die Behandlung mit Riociguat 1,0-2,5 mg ausnahmslos gleichgerichtete, positive Effekte auf alle prädefinierten sekundären Endpunkte. Die Ergebnisse einer Interimsanalyse der unkontrollierten Folgestudie CHEST-2 deuten auf einen positiven Langzeiteffekt von Riociguat 1,0-2,5 mg hin.

4.4.3.2 Mortalität

In CHEST-1 wurden insgesamt 5 Todesfälle beobachtet, davon 3 (3,4%) in der Placebo-Gruppe und 2 (1,2%) in der Riociguat-1,0-2,5 mg-Gruppe. Einer der Todesfälle aus der Riociguat-1,0-2,5 mg-Gruppe war nach Einschätzung der Prüfarzte mit der Studienmedikation assoziiert. Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

4.4.3.3 Morbidität

Für den primären Endpunkt 6MWD lag ein statistisch signifikanter Unterschied bezüglich der Mittelwertdifferenzen zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe (Tabelle 4-36) vor, der deutlich größer als die relevante MID war. Ein statistisch signifikant höherer Anteil an Respondern in der Riociguat-1,0-2,5 mg-Gruppe gegenüber der Placebo-Gruppe bestätigte die klinische Relevanz des beobachteten Effekts von Riociguat (Tabelle 4-38). Damit erfuhren Patienten unter Riociguat eine für sie „spürbare Linderung der Erkrankung“, was als deutliche Verbesserung des therapierelevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 VerfO G-BA gewertet wird.

Für die WHO-Funktionsklasse wurde ein statistisch signifikanter Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe zugunsten von Riociguat ermittelt (Tabelle 4-36). Die WHO-Funktionsklassen bilden die Fähigkeit des Patienten ab,

seine körperlichen Aktivitäten wahrnehmen zu können, was im Rahmen von Nutzenbewertungen als patientenrelevant zu verstehen ist. Daher war jede Änderung der WHO-Funktionsklasse ein klinisch relevanter Effekt, der wiederum als „eine für die Patientinnen und Patienten spürbare Linderung der Erkrankung“, mithin als deutliche Verbesserung des therapielevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 Verfo G-BA gewertet wird.

In Bezug auf die Zeit bis zur klinischen Verschlechterung gab es einen positiven, statistisch nicht signifikanten Behandlungseffekt von Riociguat 1,0-2,5 mg (Tabelle 4-36), so dass sich von diesem Endpunkt keine Aussage zum Zusatznutzen von Riociguat ableiten lässt.

Für den Borg-Wert lag ein statistisch signifikanter Behandlungsunterschied in Höhe der für COPD ermittelten MID zugunsten der Riociguat-1,0-2,5 mg-Gruppe gegenüber der Placebo-Gruppe vor (Tabelle 4-36), der daher als klinisch relevant angesehen wird. Dies bedeutete für die Patienten unter Riociguat eine „spürbare Linderung der Erkrankung“, was als deutliche Verbesserung des therapielevanten Nutzens im Sinne des 5. Kapitel § 3 Absatz 1 Verfo G-BA gewertet wird.

4.4.3.4 Lebensqualität

Für den generischen EQ-5D-Index wurde ein statistisch signifikanter Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe zugunsten von Riociguat ermittelt (Tabelle 4-36). Dieser Effekt hatte fast die doppelte Höhe der für vergleichbare Patienten ermittelten MID und wird daher als klinisch relevant und als deutliche Verbesserung des therapielevanten Nutzens durch eine „Abschwächung schwerwiegender Symptome“ im Sinne des 5. Kapitel § 3 Absatz 1 Verfo G-BA angesehen.

Für den krankheitsspezifischen LPH zeigte sich ein numerischer Behandlungsunterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe zugunsten von Riociguat (Tabelle 4-36), so dass sich von diesem Endpunkt keine Aussage zum Zusatznutzen von Riociguat ableiten lässt.

4.4.3.5 Nebenwirkungen (Unerwünschte Ereignisse)

Mindestens ein UE trat bei 91,9% der Patienten der Riociguat-1,0-2,5 mg-Gruppe, in der Placebo-Gruppe waren es mit 86,4% geringfügig weniger Patienten (Tabelle 4-39). Dieser Unterschied war statistisch nicht signifikant (OR 95%-KI 0,791; 4,064) (Tabelle 4-40). Aussagen zum Zusatznutzen waren nicht ableitbar.

In der Riociguat-1,0-2,5 mg-Gruppe traten etwas häufiger schwerwiegende UE (Tabelle 4-42) als in der Placebo-Gruppe (19,7% vs. 15,9% der Patienten) auf. Auch dieser Unterschied war statistisch nicht signifikant (OR 95%-KI 0,653; 2,560) (Tabelle 4-40), so dass keine Aussage zum Zusatznutzen ableitbar war.

Ebenso wenig gab es einen statistisch signifikanten Unterschied zwischen der Riociguat-1,0-2,5 mg-Gruppe und der Placebo-Gruppe in Bezug auf den Anteil an Patienten, die die Behandlung mit der Studienmedikation wegen eines unerwünschten Ereignisses unterbrachen (2,9% vs. 2,3% der Patienten, OR 95%-KI 0,243; 6,732) (Tabelle 4-40). Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Synkopen, Präsynkopen und Bewusstlosigkeiten traten in beiden Behandlungsgruppen gleich häufig auf (OR 95%-KI 0,248; 4,170) (Tabelle 4-40), d.h. in der Riociguat-1,0-2,5 mg-Gruppe bei 3,5% der Patienten, in der Placebo-Gruppe bei 3,4% der Patienten. Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Hämorrhagien traten in den beiden Behandlungsgruppen ungefähr gleich häufig auf (Riociguat-1,0-2,5 mg-Gruppe: 13,3% vs. Placebo-Gruppe: 11,4%). Der statistische Gruppenvergleich erbrachte keinen statistisch signifikanten Unterschied (OR 95%-KI 0,542; 2,638) (Tabelle 4-40). Das ausschließliche Auftreten schwerwiegender Hämoptysen in der Riociguat-1,0-2,5 mg-Gruppe (3 Patienten, 1,7%, Tabelle 4-42) kann jedoch ein Hinweis auf eine Erhöhung des Blutungsrisikos durch Riociguat sein. Weitere Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Hypotonien und Blutdrucksenkungen wurden erwartungsgemäß in der Riociguat-1,0-2,5 mg-Gruppe (11,0%) häufiger beobachtet als in der Placebo-Gruppe (4,5%), dieser Unterschied war allerdings nicht statistisch signifikant (OR 95%-KI 0,853; 7,865) (Tabelle 4-40). Aussagen zum Zusatznutzen von Riociguat waren nicht ableitbar.

Für Riociguat 1,0-2,5 mg ist kein Schaden festzustellen, der das Ausmaß des Zusatznutzens mindert.

4.4.3.6 Ergebnissicherheit

Mit der randomisierten Studie CHEST-1 wurde eine Datenlage mit der Evidenzstufe I b geschaffen. Die herangezogenen Endpunkte sowie die übrigen Studiencharakteristika entsprechen den Empfehlungen der EMA. Der hohe Anteil der nach Herkunft und Ethnizität der deutschen Bevölkerung entsprechenden Studienteilnehmer, die hohe Anzahl an deutschen Studienzentren, die hiesige Verfügbarkeit und Akzeptanz aller in CHEST-1 genutzten diagnostischen und therapeutischen Optionen sprechen für die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Auf Studienebene liegt ein niedriges Verzerrungspotential vor (Tabelle 4-29). Auf Endpunktebene lagen für alle Endpunkte eine Verblindung des Endpunkterhebers und eine ergebnisunabhängige Berichterstattung vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen jeweils vollständig ab. Für die patientenrelevanten Endpunkte 6MWD, WHO-Funktionsklasse, Zeit bis zur klinischen Verschlechterung, Borg, EQ-5D, LPH und UE/SUE zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diese

Endpunkte durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial für alle patientenrelevanten Endpunkte, d.h. 6MWD, WHO-Funktionsklasse, Zeit bis zur klinischen Verschlechterung, Borg, EQ-5D, LPH und UE/SUE, als niedrig eingestuft.

Vor diesem Hintergrund wird die Ergebnissicherheit der Aussagen zum Zusatznutzen als hoch bewertet.

4.4.3.7 Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen und Ausmaß des Zusatznutzens

Für erwachsene Patienten der WHO Funktionsklassen II bis III mit inoperabler CTEPH und erwachsene Patienten der WHO Funktionsklassen II bis III mit persistierender/rezidivierender CTEPH nach einer chirurgischen Behandlung liegt ein beträchtlicher Zusatznutzen vor, da für die patientenrelevanten Morbiditätsendpunkte 6MWD, WHO-Funktionsklasse und Borg sowie den Lebensqualitätsendpunkt EQ-5D statistisch signifikante und klinisch relevante Effekte von Riociguat 1,0-2,5 mg ohne einen das Ausmaß des Zusatznutzens mindernden Schaden nachgewiesen wurden. Es wurde damit eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne von 5. Kapitel § 3 Absatz 1 Verfo G-BA erreicht, insbesondere eine „Abschwächung schwerwiegender Symptome“ und eine „für die Patientinnen und Patienten spürbare Linderung der Erkrankung“.

Die Quintessenz der bis hierhin getroffenen Aussagen unter Berücksichtigung der Schwere der Erkrankung gibt Tabelle 4-44 wieder.

Tabelle 4-44: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten der WHO Funktionsklassen II bis III mit inoperabler CTEPH und erwachsene Patienten der WHO Funktionsklassen II bis III mit persistierender/ rezidivierender CTEPH nach einer chirurgischen Behandlung	Beträchtlich

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹³,

¹³ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁵) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁶) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der VerfO (s. Abschnitt 4.2.1).

¹⁴ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. *Stat Methods Med Res* 2010; 19(3): 205-236.

¹⁵ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* 2006; 5(3): 173-186.

¹⁶ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; 25(2): 183-203.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

4.7 Referenzliste

1. Europäisches Parlament. Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden: Amtsblatt der Europäischen Gemeinschaften; 1999 [Zugriff am 10.02.2014]. URL:<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:018:0001:0005:de:PDF>.
2. Bayer HealthCare AG. Clinical Study Report (CSR) No. A62508: Randomized, double-blind, placebo-controlled, multicentre, multi-national study to evaluate the efficacy and safety of oral BAY 63-2521 (1 mg, 1.5 mg, 2 mg, or 2.5 mg tid) in patients with chronic thromboembolic pulmonary hypertension (CTEPH). CHEST-1 Study; 2012.
3. Ghofrani HA, D'Armini AM, Grimminger F, Hoeper MM, Jansa P, Kim NH et al. Riociguat for the Treatment of Chronic Thromboembolic Pulmonary Hypertension. *N Engl J Med* 2013; 369(4):319–29.
4. EMA. Guideline on Pulmonary Arterial Hypertension: COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP); 2009 [Zugriff am 10.02.2014]. URL:http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/12/WC500016686.pdf.
5. Wilkens H, Lang I, Behr J, Berghaus T, Grohe C, Guth S et al. Chronisch thromboembolische pulmonale Hypertonie. *Dtsch med Wochenschr* 2010; 135(S 03):S125–S130.
6. Hoeper MM, Ghofrani HA, Gorenflo M, Grünig E, Rosenkranz S, Schranz D. Diagnostik und Therapie der pulmonalen Hypertonie. *Der Kardiologe* 2010; 4(3):189–207.
7. Bayer HealthCare AG. CHEST-1 - Statistical methods and database errata and notes to file; 2012.
8. Opitz CF, Blindt R, Blumberg F, Borst MM, Bruch L, Leuchte H et al. Pulmonary hypertension: hemodynamic evaluation: hemodynamic evaluation - recommendations of the Cologne Consensus Conference 2010. *Dtsch Med Wochenschr* 2010; 135 Suppl 3:S78-86.
9. Galiè N, Hoeper MM, Humbert M, Torbicki A, Vachiery JL, Barbera JA et al. Guidelines for the diagnosis and treatment of pulmonary hypertension: The Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). *European Heart Journal* 2009; 30(20):2493–537.
10. G-BA. Zusammenfassende Dokumentation Stand: 9. Mai 2012 über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Pirfenidon; 2012 [Zugriff am 10.02.2014]. URL:http://www.g-ba.de/downloads/40-268-1915/2012-03-15_AM-RL-XII_Pirfenidon_ZD.pdf.
11. Rosenkranz S. Pulmonale Hypertonie: In: *Klinische Kardiologie*. Erdmann E, Hrsg. Springer Medizin Verlag Heidelberg 2011:309–34.
12. IQWiG. Allgemeine Methoden 4.1; 2013 [Zugriff am 10.02.2014]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf
13. Bonner N, Abetz L, Meunier J, Sikirica M, Mathai SC. Development and validation of the living with pulmonary hypertension questionnaire in pulmonary arterial hypertension patients. *Health and Quality of Life Outcomes* 2013, 11:116.
14. G-BA. Anforderungen an ein Modul „Chronische Herzinsuffizienz“ für strukturierte Behandlungsprogramme für Koronare Herzkrankheit (KHK): Empfehlungen des Gemeinsamen Bundesausschusses gemäß § 137f Abs. 2 SGB V für die Rechtsverordnung

- nach § 266 Abs. 7 SGB V; 2008 [Zugriff am 10.02.2014]. URL:http://www.g-ba.de/downloads/40-268-634/2008-06-20_DMP_KHK_Modul-HI.pdf.
15. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166(1):111–7.
 16. G-BA. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Ticagrelor; 2011 [Zugriff am 10.02.2014]. URL:http://www.g-ba.de/downloads/40-268-1828/2011-12-15_AM-RL-XII_Ticagrelor_TrG.pdf.
 17. G-BA. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO, Wirkstoff: Ivacaftor; 2012 [Zugriff am 10.02.2014]. URL:http://www.g-ba.de/downloads/40-268-2263/2013-02-07_AM-RL-XII_Ivacaftor_ZD.pdf.
 18. Fijalkowska A, Kurzyna M, Torbicki A, Szewczyk G, Florczyk M, Pruszczyk P et al. Serum N-terminal brain natriuretic peptide as a prognostic parameter in patients with pulmonary hypertension. *Chest* 2006; 129(5):1313–21.
 19. Ghofrani HA, Hoeper MM, Halank M, Meyer FJ, Staehler G, Behr J et al. Riociguat for chronic thromboembolic pulmonary hypertension and pulmonary arterial hypertension: a phase II study. *European Respiratory Journal* 2010; 36(4):792–9.
 20. Chen H, Taichman DB, Doyle RL. Health-related Quality of Life and Patient-reported Outcomes in Pulmonary Arterial Hypertension. *Proceedings of the American Thoracic Society* 2008; 5(5):623–30.
 21. ICH. E9 Statistical Principles for Clinical Trials: Guidance for Industry [U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER)]; 1998 [Zugriff am 10.02.2014]. URL:http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
 22. Bayer HealthCare AG. CHEST-1 Clinical Study Report - Additional Subgroup Analyses (CSR-ASA); 2013.
 23. Mathai SC, Puhan MA, Lam D, Wise RA. The Minimal Important Difference in the 6-Minute Walk Test for Patients with Pulmonary Arterial Hypertension. *Am J Respir Crit Care Med* 2012; 186(5):428–33.
 24. EMA. Guideline on Missing Data in Confirmatory Clinical Trials; 2010 [Zugriff am 10.02.2014]. URL:http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf.
 25. Gilbert C, Brown MCJ, Cappelleri JC et al. Estimating a Minimally Important Difference in Pulmonary Arterial Hypertension Following Treatment With Sildenafil. *Chest* 2009; 135(1):137.
 26. Ries AL. Minimally Clinically Important Difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD* 2005; 2(1):105–10.
 27. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005; 14(6):1523–32.
 28. Bayer HealthCare AG. Long-term extension, multi-centre, multi-national study to evaluate the safety and tolerability of oral BAY 63 2521 (1mg, 1.5 mg, 2 mg, or 2.5 mg tid) in patients with Chronic Thromboembolic Pulmonary Hypertension (CTEPH). CHEST-2 Study; 2012.

29. ICH. E6 Good Clinical Practice: Consolidated Guidance: Guidance for Industry [U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER)]; 1996 [Zugriff am 10.02.2014]. URL: <http://www.fda.gov/downloads/Drugs/Guidances/ucm073122.pdf>.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 (Quelle ¹⁷) – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁷ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) (ALL-FIELDS) AND ("Phase II" OR "Phase III" OR "Phase IV") (PHASE)
Treffer	23

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht auszufüllen gemäß Absatz 1.3 der Anlage II.1 zum 5. Kapitel der Verfo (s. Abschnitt 4.2.1).

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-45 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-45 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-45 (Anhang): Studiendesign und -methodik für Studie CHEST-1

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Ziel</p> <p>Das Ziel der vorliegenden Studie war es, die Wirksamkeit und Sicherheit einer Behandlung mit Riociguat, einem Stimulator der löslichen Guanylatcyclase, für 16 Wochen bei erwachsenen Patienten mit einer inoperablen chronisch thromboembolischen pulmonalen Hypertonie (CTEPH) oder einer persistierenden oder rezidivierenden CTEPH nach einer chirurgischen Behandlung.</p> <p>Fragestellung</p> <p>Ist die Behandlung von erwachsenen Patienten mit einer inoperablen CTEPH oder einer persistierenden oder rezidivierenden CTEPH nach einer PEA mit Riociguat 1,0-2,5 mg für 16 Wochen einer Behandlung mit Placebo überlegen?</p> <p>Hypothese</p> <p>Nullhypothese: Änderung für den primären Endpunkt vom Ausgangswert ist für die Riociguat-Gruppe und die Placebo-Gruppe gleich.</p> <p>Alternativhypothese: Änderung für den primären Endpunkt vom Ausgangswert ist für die Riociguat-Gruppe größer als für die Placebo-Gruppe.</p> <p>Im Rahmen der Überlegenheitstestung kann die Nullhypothese abgelehnt werden, wenn das zweiseitige Signifikanzniveau gleich oder kleiner 0,05 ist.</p>
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Randomisierte, Placebo-kontrollierte, multizentrische, multinationale, doppelblinde, pivotale Parallelgruppen-Studie der Phase III</p> <p>Die an einer inoperablen CTEPH oder einer persistierenden oder rezidivierenden CTEPH nach einer PEA leidenden Patienten wurden den beiden Studienarmen im Verhältnis 2:1 zugeordnet:</p> <p>Studienarm 1 (aktive Behandlung mit Riociguat 1,0-2,5 mg)</p>

Item ^a	Charakteristikum	Studieninformation
		<p data-bbox="906 275 1361 338">Orale Gabe von Riociguat 1,0-2,5 mg tid über einen Zeitraum von 16 Wochen</p> <p data-bbox="906 371 1230 400">Studienarm 2 (Placeboarm)</p> <p data-bbox="906 403 1334 465">Orale Gabe von Placebo tid über einen Zeitraum von 16 Wochen</p> <p data-bbox="906 499 1385 887">In einer vierwöchigen Vorbeobachtungsphase wurde neben der Charakterisierung der Patienten die (In)Operabilität entsprechend internationalen Leitlinien durch einen erfahrenen Chirurgen beurteilt. Dazu war mindestens ein pulmonares Angiogramm, ergänzt um eine Lungenperfusions-/Ventilationsszintigraphie, oder alternativ ein Pulmonalis CTA ergänzt um eine Lungenperfusions-/Ventilationsszintigraphie als bildgebendes Verfahren erforderlich.</p> <p data-bbox="906 889 1385 1227">Nach der Vorbeobachtungsphase wurden die Patienten randomisiert. Die sich anschließende 16-wöchige Behandlungsphase bestand aus einer 8-wöchigen Titrationsphase und einer 8-wöchigen Erhaltungsphase. Patienten, die nicht in CHEST-2 (NCT00910429 bei ClinicalTrials.gov), einer unkontrollierten Folgestudie zur Langzeitbeobachtung, eintraten oder die die Studie vorzeitig verließen, wurden 30 Tage nachbeobachtet.</p> <p data-bbox="906 1261 1385 1413">Die Studie wurde in Übereinstimmung mit den ethischen Prinzipien der Deklaration von Helsinki und der ICH-Guideline E6: Good Clinical Practice (GCP) durchgeführt (29).</p> <p data-bbox="906 1447 1334 1509">Die Studie ist unter NCT00855465 bei ClinicalTrials.gov eingetragen.</p>

Item ^a	Charakteristikum	Studieninformation
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Protokolländerungen</p> <p>Zum Originalstudienprotokoll vom 10.09.2008 gab es sieben Änderungen, davon waren vier globale und drei nationale Änderungen. Die Protokolländerungen 1 vom 10.10.2008 und 2 vom 04.03.2009 betrafen nur Studienzentren in Japan, Protokolländerung 7 vom 28.02.2011 nur Studienzentren in Kanada und den USA. Diese sind nicht relevant für Deutschland und werden deshalb hier nicht beschrieben.</p> <p>Protokolländerung Nr. 3 vom 16.06.2009, für alle Studienzentren</p> <p>Die wichtigsten Änderungen waren:</p> <ul style="list-style-type: none"> • Klarstellungen in Bezug auf einige Einschluss- und Ausschlusskriterien • Spezifikation des 6MWD-Tests • Wechsel der Borg Dyspnoe-Skala • Sammlung von Informationen zu Ressourcen • Hinzufügen der pharmakogenetischen Probenahme (optional) • Hinzufügen der Definition des körperlichen Trainingsprogrammes • Spezifikation des Zeitplans für die Dosierung der Studienmedikation • Verwendung lokaler Laborergebnisse für den Einschluss von Patienten • Hinzufügen der Methodik zur Blutdruckmessung • Hinzufügen der unerwünschten Wirkungen Schwindel und Synkope • Hinzufügen der Beschreibung des zentralen Prozesses der Bewertung der Operabilität • Hinzufügen der detaillierten Aufgaben des Steuerungskomitees (SC; übergeordnete wissenschaftliche Verantwortung für die Studie) und des Datenkontrollkomitees (DMC; Verantwortung für die Sicherheit der Patienten) <p>Begründung:</p> <ul style="list-style-type: none"> • Umsetzung von Empfehlungen aus mehreren globalen Prüfarzte-Treffen <p>Protokolländerung Nr. 4 vom</p>

Item ^a	Charakteristikum	Studieninformation
		<p>24.03.2010, für alle Studienzentren</p> <p>Die wichtigsten Änderungen waren:</p> <ul style="list-style-type: none"> • Abschaffung der obligatorischen Übernachtungen bei der ersten Untersuchung • Anpassung der oberen Altersgrenze von 75 auf 80 Jahre mit strengeren PCWP Kriterien zum Ausschluss signifikanter Linksherzinsuffizienz • Klarstellung der Verhütungsmethoden in den Ausschlusskriterien • Klarstellung des Schwangerschaftsnachweises • Klarstellung des 6MWD-Tests für die Einschlusskriterien und andere Ausschlusskriterien • Streichung eines Ausschlusskriteriums zu Allergien • Änderungen von Beobachtungszeiträumen • Verwendung der modifizierten Borg Dyspnoe-Skala • Erfassung des Raucherstatus • Aufnahme von Rauchen als Interaktion • Hinzufügen von Erbrechen und Gastritis als unerwünschte Ereignisse • Erweiterung des Zeitfensters für den Folgebesuch auf 30 +5 Tage • Definition „erfahrener Chirurg“ für das Komitee zur Beurteilung der Operabilität; hinterlegt im Deklarationsformular • Berichtigung von Fehlern, die durch Protokolländerung Nr. 3 in den Ausschlusskriterien eingeführt wurden <p>Begründung:</p> <ul style="list-style-type: none"> • Empfehlungen mehrerer globaler Prüfarzte-Treffen und des SC <p>Protokolländerung Nr. 5 vom 11.10.2010, für alle Studienzentren</p> <p>Die wichtigsten Änderungen waren:</p> <ul style="list-style-type: none"> • Reduktion des für den Studieneinschluss erforderlichen Basis-PVR-Wertes von $>480 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$ auf $>300 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$ • Weitere Spezifizierung und Harmonisierung der Beurteilung der

Item ^a	Charakteristikum	Studieninformation
		<p>Inoperabilität auf Basis von studien-spezifischen Handbüchern</p> <p>Begründung: Spezifizierung der Modalitäten für die bildgebenden Verfahren zur Feststellung der Inoperabilität, der Einschlusskriterien (PVR), der Studienmodalitäten und Korrektur von Inkonsistenzen vorheriger Protokollversionen.</p> <p>Protokolländerung Nr. 6 vom 15.02.2011, für alle Studienzentren</p> <p>Die wichtigsten Änderungen waren:</p> <ul style="list-style-type: none"> • Hinzufügen der Messung von Calcium und Phosphat (bei jeder Visite) und Calcitriol (bei Visiten 1, 2 und 7). Untersuchung der schon vorhandenen Proben, um herauszufinden, ob ein klinisch relevanter Effekt von Riociguat auf die Calcium-Homöostase bei Patienten mit CTEPH im Vergleich zu Placebo vorhanden war • Änderung des Zeitraumes, in dem unerwünschte Ereignisse als potentiell therapiebedingt erachtet wurden von 7 auf 2 Tage nach Beendigung der Behandlung mit der Studienmedikation. Grundlage ist eine Empfehlung, dass das Zeitfenster für therapiebedingte unerwünschte Ereignisse dem Fünffachen der Halbwertszeit der Studienmedikation entspricht • Separate Erfassung von unerwünschten Ereignissen, die nach drei oder mehr Tagen auftraten <p>Begründung: Spezifizierung zusätzlicher Labormessungen für die Sicherheitsuntersuchungen, Modifikation der Definition des Behandlungsaustrittes bei unerwünschten Ereignissen und Korrektur von Inkonsistenzen vorheriger Protokollversionen.</p> <p>Änderungen und Ergänzungen des statistischen Analyseplans Update des statistischen Analyseplans, Version 3.0, 14.12.2010</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Die wichtigsten Änderungen waren:</p> <ul style="list-style-type: none"> • Ergänzung der Per-Protocol-Validierungskriterien: <ul style="list-style-type: none"> ○ 6MWD strikt weniger als 450 m und größer als 150 m ○ Rechtsherzkatheteruntersuchung mehr als 8 Wochen vor Visite 1 ○ Ausweitung der Adhärenz-Regeln ○ Alter zwischen 18 und 80 Jahren bei Visite 1 ○ Klarstellung des Zeitraums für nicht erlaubte Begleitmedikationen ○ $PVR \geq 300 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$ bei Visite 1 ○ Änderung der Anforderungen an bildgebende Verfahren für die CTEPH-Diagnose und Feststellung der Operabilität • Austausch der statistischen Methoden für den Endpunkt Zeit bis zur klinischen Verschlechterung (Log-Rank als Haupttest, Cochran-Mantel-Haenszel Gewichtung supportiv) • Wechsel des Statistikers • Aktualisierung der regionalen Gruppierung • Hinzufügen zusätzlicher deskriptiver Subgruppen (inoperabel/persistierende oder rezidivierende CTEPH nach einer PEA) • Aktualisierung spezifischer Begleitmedikations-Gruppen • Hinzufügen des Raucherstatus und -historie zur Demographie • Änderung der Imputationsregeln für die Borg-Skala • Aktualisierung von Labormessungen und Vitalzeichen • Hinzufügen eines Absatzes zur möglichen Entblindung wegen Open-Label Behandlung in CHEST-2 <p>Ergänzung des Statistikanalyseplans, Version 2.0, 15.03.2012 und Version 3.0, 11.07.2012</p> <ul style="list-style-type: none"> • Hinzufügen zusätzlicher Analysen zu hämodynamischen Parametern

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Hinzufügen zusätzlicher Analysen zu PRO (EQ-5D, LPH) • Kategoriale Analysen von Wirksamkeitsvariablen • Sensitivitätsanalyse der Zeit bis zur klinischen Verschlechterung • Sensitivitätsanalyse der 6MWD (Handhabung fehlender Daten) • Robuste Regressionsanalyse der 6MWD • Hinzufügen von Analysen zur Darstellung der Beziehung zwischen Outcomes <p>Update des Statistikanalyseplans, Version 4.0, 11.07.2012</p> <ul style="list-style-type: none"> • Änderung des Per-Protocol-Validierungskriteriums: Änderung der Erhebung der 6MWD bei der Abschlussvisite bei Ausscheiden eines Patienten (nicht auf Tod oder klinischer Verschlechterung beruhend) • Kein statistisches Testen der Basis-Demographie, da nicht länger von den internen Bayer-Standards unterstützt • Klarstellung der Bedingungen für das Ereignis klinische Verschlechterung • Ergänzung des Behandlungsunterschieds bei Subgruppen um Konfidenzintervalle • Hinzufügen zusätzlicher Subgruppen: Region, 6MWD-Ausgangswert <380 m, 6MWD-Ausgangswert ≥380 m • Hinzufügen der eGFR als Laborparameter

Item ^a	Charakteristikum	Studieninformation
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien der Patienten</p> <ul style="list-style-type: none"> • 18 bis 80 Jahre bei Visite 1 (Berücksichtigung der altersbedingten Restriktionen in Hinblick auf PCWP, die untere Altersgrenze kann in Abhängigkeit von den Vorschriften des jeweiligen Landes höher liegen) • Männer und Frauen mit diagnostizierter CTEPH, einer Studieneignung und einem Basiswert zwischen 150 m und 450 m für den 6MWD-Test entweder definiert als: <ul style="list-style-type: none"> ○ Inoperabel, mit PVR >300 dyn·sec·cm⁻⁵, mindestens 90 Tage nach Beginn der vollständigen Antikoagulation gemessen und einem PAPm > 25 mmHg (Bestätigung der Inoperabilität vor der Randomisierung durch einen erfahrenen Chirurgen) oder ○ Mit persistierender oder rezidivierender CTEPH nach einer PEA (PVR >300 dyn·sec·cm⁻⁵, mindestens 180 Tage nach der Operation gemessen) • Erlaubnis unspezifischer Behandlung mit oralen Antikoagulantien, Diuretika, Digitalis, Kalziumkanalblockern oder Sauerstoffgabe, die auch für die Therapie der PH eingesetzt werden. Die Behandlung mit Antikoagulantien musste mindestens 90 Tage vor Visite 1 begonnen werden und die mit Diuretika mindestens 30 Tage vor Visite 1 stabil sein • Zusätzliche Langzeit-Sauerstofftherapie, wenn die Sauerstoffkonzentration mindestens 90 Tage vor Visite 1 konstant war • Ergebnisse der Rechtsherzkatheteruntersuchung nicht älter als acht Wochen vor Visite 1, nach mindestens 90 Tagen vollständiger Antikoagulation und in Zusammenarbeit mit dem teilnehmenden Zentrum unter Standardbedingungen gemessen. Andernfalls in der Vorbeobachtungsphase nach Unterzeichnung der Einverständniserklärung oder bei Visite

Item ^a	Charakteristikum	Studieninformation
		<p>1 vor der Randomisierung zu messen</p> <ul style="list-style-type: none"> • Frauen, die eines der folgenden Kriterien erfüllen: <ul style="list-style-type: none"> ○ Postmenopause seit mindestens zwei Jahren ○ Bilaterale Tubenligatur ○ Bilaterale Ovariectomie ○ Hysterektomie ○ Negativer serologischer Schwangerschaftstest und eine Kombination sicherer Verhütungsmethoden während der Studiendauer • Verstehen und Befolgen von Instruktionen und Teilnahme über die gesamte Laufzeit der Studie • Schriftliche Einverständniserklärung nach vorheriger Information und vor Beginn jeglicher Studien-spezifischer Prozeduren <p>Ausschlusskriterien der Patienten</p> <p>Genereller Ausschluss</p> <ul style="list-style-type: none"> • Gleichzeitige Teilnahme an einer anderen klinischen Studie während dieser Studie oder innerhalb von 30 Tagen vor Visite 1 • Schwangere Frauen (d.h. positiver Serum β-hCG-Test oder andere Anzeichen für eine Schwangerschaft), stillende oder gebärfähige Frauen, die keine Kombination sicherer Verhütungsmethoden anwenden • Jegliche Erkrankungen oder Gesundheitszustände in der Patientenhistorie, die nach Meinung des Prüfarztes die Studienteilnahme oder –beendigung beeinträchtigen können • Drogenmissbrauch (z.B. von Alkohol oder illegalen Drogen) innerhalb von 180 Tagen vor Visite 1 • Zugrunde liegende Erkrankung mit einer angenommenen Lebenserwartung von weniger als zwei Jahren (z.B. aktive Krebserkrankung mit lokalisiertem und/oder metastasiertem Tumorgewebe) • Hypersensitivität gegenüber der Testsubstanz oder einem der Zusatzstoffe in der Medikamentenzusammensetzung • Unfähigkeit, einen validen 6MWD-Test durchzuführen (z.B. Patienten mit schwerer arterieller

Item ^a	Charakteristikum	Studieninformation
		<p>Verschlusskrankheit). Hinweis: Patienten mit Gehhilfe können durch den Prüfarzt eingeschlossen werden, wenn die Gehstrecke nicht beeinträchtigt wird</p> <ul style="list-style-type: none"> • Relative Differenz (d.h. absolute Differenz/Mittelwert) von mehr als 15% zwischen dem 6MWD-Test zur Überprüfung der Studieneignung und dem 6MWD-Test zur Messung des 6MWD-Ausgangswertes <p>Medikations- oder behandlungsbedingte Ausschlüsse</p> <p>Screening von Patienten in Hinblick auf eine mögliche Studienteilnahme sollte nicht zum Absetzen einer notwendigen Therapie führen. Widersprachen diese Therapien den Studienkriterien, durften die betroffenen Patienten nicht eingeschlossen werden. Folgende spezifischen Medikationen waren im Rahmen einer Vorbehandlung innerhalb der letzten 90 Tage vor Visite 1 nicht erlaubt:</p> <ul style="list-style-type: none"> • NO-Donoren (z.B. Nitrate) innerhalb der letzten 90 Tagen vor Visite 1 • PAH-spezifische Medikationen: <ul style="list-style-type: none"> ○ Endothelin-Rezeptor-Antagonist ○ Prostanoid ○ Spezifische (z.B. Sildenafil oder Tadalafil) oder unspezifische (z.B. Dipyridamol, Theophyllin) Phosphodiesterase-Inhibitoren <p>Prinzipiell mussten geeignete Patienten bei Studieneinschluss in Bezug auf PAH-spezifische Medikamente unbehandelt sein. In Ausnahmefällen, d.h. bei nicht akzeptablen unerwünschten Ereignissen und/oder fehlendem Nutzen einer PAH-spezifischen Medikation konnte eine Studienteilnahme erwogen werden. In diesen Fällen musste die PAH-spezifische Medikation mindestens 30 Tage vor der Rechtsherzkatheteruntersuchung zur Bestimmung der hämodynamischen Ausgangswerte abgesetzt werden.</p> <p>In Abhängigkeit vom Zeitpunkt der Rechtsherzkatheter-Untersuchung zur Bestimmung der Ausgangswerte variiert das minimale PAH-medikationsfreie Intervall zwischen 30 (Visite 1) und 86 Tagen (8 Wochen vor Visite 1).</p>

Item ^a	Charakteristikum	Studieninformation
		<p data-bbox="906 282 1310 338">Ausschlüsse in Zusammenhang mit Lungenerkrankungen</p> <ul data-bbox="906 371 1393 741" style="list-style-type: none"> • Alle Typen von PH mit Ausnahme der Subtypen 4.1 und 4.2 (Venice Clinical Classification of PH) • Mittelschwere bis schwere obstruktive Lungenerkrankung (forciertes Expirations-Einsekundenvolumen (FEV 1) <60% des prädiktiven Wertes) • Schwere restriktive Lungenerkrankung (totale Lungenkapazität <70% des prädiktiven Wertes) • Schwere angeborene Missbildungen von Lunge, Thorax und Zwerchfell <p data-bbox="906 775 1377 864">Ausschlüsse im Zusammenhang mit Blutgasanomalien (kapillär oder arteriell im Ruhezustand)</p> <ul data-bbox="906 898 1369 1144" style="list-style-type: none"> • Sauerstoff-Sättigung (SaO₂) <88% bei Visite 0 trotz zusätzlicher Sauerstoff-Therapie • Arterieller Sauerstoff-Partialdruck (PaO₂) <55 mmHg bei Visite 0 trotz zusätzlicher Sauerstoff-Therapie • Arterieller CO₂-Partialdruck (PaCO₂) >45 mmHg bei Visite 0 <p data-bbox="906 1178 1246 1211">Kardiovaskuläre Ausschlüsse</p> <ul data-bbox="906 1245 1393 2007" style="list-style-type: none"> • Historie einer unkontrollierten arteriellen Hypertonie innerhalb der letzten 90 Tage vor Visite 1 und/oder <ul data-bbox="986 1357 1393 1581" style="list-style-type: none"> ○ SBD >180 mmHg und/oder DBD >110 mmHg bei Visite 0 und/oder Visite 1 vor Randomisierung oder ○ SBD <95 mmHg bei Visite 0 und/oder bei Visite 1 vor Randomisierung • Ruhepuls im Wachzustand <50 Schläge pro Minute oder >105 Schläge pro Minute bei Visite 0 und/oder Visite 1 vor Randomisierung • Vorhofflimmern/Vorhofflattern innerhalb der letzten 90 Tage vor Visite 1 • Linksherzinsuffizienz mit einer Ejektionsfraktion von weniger als 40% innerhalb der letzten 90 Tage vor Visite 1 • Pulmonale venöse Hypertonie mit einem PCWP-Ausgangswert >15 mmHg (wenn Alter bei Visite 1

Item ^a	Charakteristikum	Studieninformation
		<p>zwischen 18-75 Jahren) oder >12 mmHg (wenn Alter >75 Jahre bei Visite 1)</p> <ul style="list-style-type: none"> • Hypertrophe obstruktive Kardiomyopathie • Schwere nachgewiesene oder vermutete koronare Herzkrankheit (Patienten mit Klasse 2-4 nach der Canadian Cardiovascular Society Angina Classification, und/oder Bedarf an Nitraten, und/oder Myokardinfarkt innerhalb der letzten 90 Tage vor Visite 1) • Klinische Evidenz einer symptomatischen atherosklerotischen Erkrankung (z.B. periphere arterielle Verschlusskrankheit mit reduzierter Gehstrecke, Historie eines Schlaganfalles mit persistierendem neurologischem Defizit etc.) • Angeborene oder erworbene Herzklappen- oder Herzmuskel-erkrankung, wenn klinisch signifikant abgesehen von Trikuspidal-klappeninsuffizienz aufgrund von PH • Klinische Evidenz für rezidivierende Thromboembolien trotz ausreichender (dokumentierter) oraler Antikoagulation - auch wenn Lungenarterien nicht betroffen sind <p>Ausschlüsse in Zusammenhang mit Organfunktionsstörungen</p> <ul style="list-style-type: none"> • Klinisch relevante hepatische Dysfunktion: <ul style="list-style-type: none"> ○ Bilirubin >2-fache des oberen Normwertes bei Visite 0 und/oder ○ ALT oder AST >3-fache des oberen Normwertes bei Visite 0 und/oder ○ Zeichen einer schweren Leberinsuffizienz (z.B. eingeschränkte Albumin-Synthese mit Albumin <32 g/L, Leberenzephalopathie > Grad 1) bei Visite 0 • Schwere Niereninsuffizienz mit glomerulärer Filtrationsrate <30 mL/min bei Visite 0 <p>Gründe für den Ausschluss von Patienten aus der Studie</p> <ul style="list-style-type: none"> • Eigener Wunsch • Wunsch des Sponsors

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none">• Auftreten von unerwünschten Ereignissen oder interkurrenter Erkrankungen, die nach Meinung des Prüfarztes eine Fortführung nicht erlaubten• Auftreten von Arzneimittelnebenwirkungen, die nach Meinung des Prüfarztes das Risiko-Nutzen-Verhältnis des Patienten negativ beeinflussten• Relevante Nicht-Adhärenz• Unterbrechung der Einnahme der Studienmedikation von mehr als drei Tagen in Folge (neun Dosen)• Schwangerschaft oder Stillen• Teilnahme an einer anderen Studie• Auftreten eines oder mehrerer Kriterien des Endpunktes „Zeit bis zur klinischen Verschlechterung“• Weitere Dosisreduktion nicht möglich und Patient tolerierte die niedrigste mögliche Dosis (0,5 mg tid) nicht

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	<p>Die Studie wurde in 89 Studienzentren in 26 Ländern weltweit durchgeführt (Anzahl Zentren je Land):</p> <p>Argentinien (1), Australien (1), Österreich (2), Belgien (2), Brasilien (3), Kanada (5), China (5), Tschechische Republik (1), Dänemark (1), Frankreich (8), Deutschland (12), Italien (1), Japan (15), Mexiko (4), Niederlande (1), Polen (2), Portugal (1), Russland (2), Slowakei (1), Südkorea (2), Spanien (2), Schweiz (1), Taiwan (1), Türkei (3), Großbritannien (3), USA (9)</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Die Studie umfasste eine 4-wöchige Vorbeobachtungsphase und eine 16-wöchige Behandlungsphase, bestehend aus einer 8-wöchigen Titrationsphase und einer 8-wöchigen Erhaltungsphase. Riociguat oder Placebo wurden dreimal täglich oral (tid) als Filmtabletten mit oder ohne Nahrung verabreicht. Die jeweiligen einzelnen Tagesdosen sollten im Abstand von 6-8 Stunden eingenommen werden.</p> <p>4-wöchige Vorbeobachtungsphase</p> <ul style="list-style-type: none"> • Einverständniserklärung • Erhebung demographischer Basisdaten • Dokumentation der CTEPH-Klassifikation (inoperable/persistierende oder rezidivierende CTEPH nach einer PEA) • Erhebung der Patientenhistorie • Beurteilung der Operabilität durch einen erfahrenen Chirurgen oder ein zentrales Adjudizierungskomitee • Visite 0 <p>16-wöchige Behandlungsphase</p> <ul style="list-style-type: none"> • 8-wöchige individuelle Titrationsphase (Visiten 1-5) • 8-wöchige Erhaltungsphase (Visiten 5-7) <p>8-wöchige Titrationsphase Riociguat-1,0-2,5 mg-Gruppe (geplant: 180 Patienten)</p> <ul style="list-style-type: none"> • Therapiestart mit 1,0 mg Riociguat tid • Individuelle Titration mit 0,5 mg Riociguat tid alle zwei Wochen • Höchste Dosis 2,5 mg Riociguat tid <p>Placebo-Gruppe (geplant: 90 Patienten)</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Scheintitration analog Riociguat-1,0-2,5 mg-Gruppe</p> <p>Für beide Studienarme galt:</p> <ul style="list-style-type: none"> • Änderungen der Dosierung durch den Prüfarzt über das IVRS durften $\pm 0,5$ mg nicht übersteigen • Dosissteigerungen waren nur zu den Visiten 2, 3 und 4 erlaubt • Dosisenkung im Fall von mit der Studienmedikation assoziierten Nebenwirkungen auf 0,5 mg tid erlaubt • Durchführen der (Schein-)Titration in Abhängigkeit vom peripheren systolischen Blutdruck (SBD) am Ende des Dosierungsintervalles, der vor der Einnahme der morgendlichen Dosis gemessen wurde: <ul style="list-style-type: none"> ○ SBD ≥ 95 mmHg: Erhöhung der Dosis um 0,5 mg 3x täglich ○ SBD 90 – 94 mmHg: Beibehaltung der Dosis ○ SBD < 90 mmHg ohne Hypotonie-Symptome: Senkung der Dosis um 0,5 mg 3x täglich ○ SBD < 90 mmHg mit Hypotonie-Symptomen wie Schwindel oder Präsynkope: Behandlung unterbrechen und nach 24 Stunden mit einer um 0,5 mg 3x täglich gesenkten Dosis neu beginnen <p>8-wöchige Erhaltungsphase</p> <ul style="list-style-type: none"> • Beibehalten der Dosisstärke am Ende der Titrationsphase von Tag 56 bei Visite 5 • Dosisreduktionen aus Sicherheitsgründen erlaubt, nicht aber ein erneuter Dosisanstieg <p>Patienten, die nicht in CHEST-2, einer unkontrollierten Folgestudie zur Langzeitbeobachtung, eintraten oder die die Studie vorzeitig verließen, wurden 30 Tage nachbeobachtet.</p>

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Endpunkt</p> <p>Absolute Veränderung der 6MWD bis Woche 16 im Vergleich zum Ausgangswert</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert • Zeit bis zur klinischen Verschlechterung in Studientagen • Veränderung des Ermüdungsempfindens und der Dyspnoe gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert • Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert • Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert <p>Supportive sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Absolute Veränderung des Pulmonal vaskulären Widerstandes (PVR) nach 16 Wochen im Vergleich zum Ausgangswert • Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert <p>Sicherheitsendpunkte</p> <ul style="list-style-type: none"> • Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse • Ereignisse von speziellem Interesse: Vermutete UE Synkopen (MedDRA Preferred Terms: Bewusstlosigkeit, Präsynkopen, Synkopen) Vermutete UE Hypotonie (MedDRA Preferred Terms: Hypotonie, verminderter Blutdruck) • Weitere signifikante unerwünschte Ereignisse

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Mortalität • Standardlaborparameter • Vitalparameter • 12-Kanal Elektrokardiogramm • Blutgasanalyse <p style="margin-left: 20px;">Zusätzlich Dokumentation von Schwangerschaften</p> <p>Sonstige Variablen</p> <ul style="list-style-type: none"> • Änderung zusätzlicher hämodynamischer Parameter nach 16 Wochen • Einsatz von Ressourcen für gesundheitsökonomische Kalkulationen nach 16 Wochen: <ul style="list-style-type: none"> ○ Besuche bei Angehörigen der Gesundheitsberufe ○ Hospitalisierungen und Rehabilitation ○ Diagnostische und therapeutische Verfahren • Explorative Biomarker: <ul style="list-style-type: none"> ○ Carboxyterminales vernetzendes Telozeptid des Knochenkollagens (CTX) ○ Osteopontin • Pharmakokinetik: Konzentration von Riociguat und seinem Hauptmetaboliten BAY 60-4552 im Plasma • Pharmakogenetik: Untersuchung möglicher Gene, die potentiell in CTEPH involviert sind, Medikamentenabsorption, Ansprechen auf die Behandlung, AEs <p>Training der Prüfarzte</p> <ul style="list-style-type: none"> • GCP-geschulte Prüfarzte • Prüfarzttreffen zur Vorbereitung auf die Studienteilnahme: <ul style="list-style-type: none"> ○ Schulung zur Substanz ○ Schulung zur Indikation • Initiierung der Zentren: <ul style="list-style-type: none"> ○ Besprechung offener und besonders kritischer Punkte • Training des DSM Boards und der adjudizierenden Chirurgen in CHEST bezüglich der Studieranforderungen (Data on file)

Item ^a	Charakteristikum	Studieninformation
		<p>Validierung des LPH-Fragebogens</p> <ul style="list-style-type: none"> • Qualitative Validierung des Fragebogeninhalts in Zusammenarbeit mit 38 PH-Patienten aus USA, Frankreich und Deutschland • Psychometrische Validierung durch Verwendung doppelblinder Daten aus der 12-wöchigen klinischen Phase-III-Studie PATENT-1 (Studie 12934) bei Patienten mit PAH • Ergebnis: Der LPH-Fragebogen ist ein valides, zuverlässiges und empfindliches Instrument zur Messung der Auswirkungen von PH auf Domänen der gesundheitsbezogenen Lebensqualität
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderungen der Zielkriterien nach Studienbeginn
7	Fallzahl	<p>Von den eingeschlossenen 446 Patienten wurden insgesamt 262 Patienten randomisiert:</p> <ul style="list-style-type: none"> • 174 Patienten in die Riociguat-1,0-2,5 mg-Gruppe • 88 Patienten in die Placebo-Gruppe
7a	Wie wurden die Fallzahlen bestimmt?	<p>Der primäre Endpunkt war die absolute Veränderung der 6MWD zum Ausgangswert bis Woche 16. Unter der Annahme einer SD von 70 m wurde berechnet, dass 261 Patienten für die ITT-Auswertung benötigt wurden (174 in der Riociguat-1,0-2,5 mg-Gruppe, 87 in der Placebo-Gruppe, 2:1 Randomisierung), um eine Placebo-adjustierte Differenz von 30 m in der 6MWD zu erfassen (Power von 90% und zweiseitiges Signifikanzniveau von 5%).</p> <p>Unter Berücksichtigung einer Ungültigkeitsquote von 3% war die Randomisierung von insgesamt 270 Patienten erforderlich.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Die Studie wurde wie geplant abgeschlossen.</p> <p>Sicherheitsreviews durch das DMC</p> <ul style="list-style-type: none"> • Überprüfen akkumulierender, unverblindeter Studiendaten-zusammenfassungen auf sicherheitsrelevante Aspekte • Einbindung eines externen Statistik-Instituts zur Sicherstellung, dass entblindete Daten Dritten nicht zur

Item ^a	Charakteristikum	Studieninformation
		Verfügung standen Kriterien für einen vorzeitigen Studienabbruch: <ul style="list-style-type: none"><li data-bbox="906 405 1385 521">• Empfehlung zum Studienabbruch durch das DMC bei unakzeptablem Sicherheitsprofil, finale Entscheidung durch SC und Sponsor<li data-bbox="906 533 1334 551">• Studienabbruch durch den Sponsor

Item ^a	Charakteristikum	Studieninformation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Gruppenzuteilung erfolgte rein zufällig durch eine zentrale, unabhängige Einheit (IVRS oder IWRS; Bracklet, London, UK). Die Zuteilungssequenz wurde mittels computergenerierter Zufalls-codes erzeugt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden in Blöcken der Größe 6 randomisiert, das 2:1 Verhältnis widerspiegelnd. Für die Wirksamkeitsanalysen fand vor dem Schließen der Datenbank eine geographische Clusterbildung in fünf Regionen (Nordamerika, Südamerika, Europa, China und Asien/Pazifik) mit jeweils mindestens 30 Patienten statt. Die statistischen Auswertungen wurden in Bezug auf die geographischen Regionen adjustiert und die Konsistenz der Behandlungseffekte überprüft.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Patienten wurden im Verhältnis 2:1 randomisiert. Die Zuteilung erfolgte durch eine zentrale, unabhängige Einheit (IVRS oder IWRS). Die Geheimhaltung war bis zur Zuteilung gewährleistet.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierungsliste enthielt die Randomisierungscodes, Patientenidentifizierungsnummern und die zugeordneten Behandlungen. Die Patienten wurden von den Prüfarzten in die Studie aufgenommen und den Gruppen zugeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Die Studie war doppelblind. a) Patienten waren verblindet b) Prüfarzte waren verblindet c) Personal von Auftragsinstituten und dem Sponsor waren verblindet Durchführung der Verblindung <ul style="list-style-type: none"> • Aussehen der Tabletten identisch • Verpacken der Studienmedikation in identischen Flaschen aus hochverdichtetem Polyäthylen mit jeweils 54 Tabletten Riociguat 0,5 mg, 1 mg, 1,5 mg, 2 mg, 2,5 mg oder Placebo • Zuteilung zu den Behandlungsgruppen und Dosistitration mit Hilfe der IVRS • Scheintitration bei Patienten der

Item ^a	Charakteristikum	Studieninformation
		<p>Placebo-Gruppe ab Visite 1</p> <ul style="list-style-type: none">• Zuweisung der verblindeten Placebo-Medikation stets durch IVRS• Durchführung des 6MWD-Tests, der Bewertung der Borg-Skala und der Evaluation der WHO-Funktionsklasse von einem zweiten, nicht in den Dosistitrationsprozess involvierten Arzt oder medizinischem Fachpersonal ab Visite 2• Weiterleitung der Ergebnisse der NT-proBNP Bestimmungen an die Prüfarzte erst nach Schließen der Studiendatenbank• Aufrechterhaltung der Verblindung des Bioanalysten für die Analyse der Pharmakokinetik• Berücksichtigung eines gewissen Niveaus von Entblindungen in den statistischen Auswertungen durch Entblindung einzelner Fälle für Berichte an Behörden durch den Sponsor

Item ^a	Charakteristikum	Studieninformation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	<p>Riociguat</p> <p>Filmtablette zur oralen Einnahme in den Dosierungen 0,5 mg, 1,0 mg, 1,5 mg, 2,0 mg, und 2,5 mg tid mit identischem Aussehen.</p> <p>Placebo</p> <p>Filmtabletten zur oralen Einnahme, die nicht von den Riociguat-Tabletten unterscheidbar waren.</p>
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Intention-to-treat (ITT) Population</p> <p>Alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten.</p> <p>Laut Studienbericht wurden die Analysen in Übereinstimmung mit der ICH-Guideline E9 zu „Statistical Principles for Clinical Trials“ durchgeführt. In das „Full Analysis Set“ einer ITT-Analyse sollten möglichst alle Patienten eingeschlossen werden, die randomisiert wurden. Drei Ausnahmen von dieser Regel sind zulässig. Eine davon, der Ausschluss eines Patienten, der keine einzige Dosis der Studienmedikation erhalten hat, trifft auf die Analyse des primären Endpunktes 6MWD zu und ist Guideline-konform. In der Riociguat-1,0-2,5 mg Gruppe traf dies auf einen Patienten (610018003) von 174 Patienten (0,6%) zu. Das „Full Analysis Set“ umfasst in der Riociguat-1,0-2,5 mg-Gruppe 173 und in der Placebo-Gruppe 88 Patienten.</p> <p>Safety-Population</p> <p>Alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten.</p> <p>Per-Protocol (PP)-Population</p> <p>Alle Patienten,</p> <ul style="list-style-type: none"> • die für die Safety- / ITT-Population geeignet waren • die einen adäquaten 6MWD-Test zur Bestimmung des Ausgangswertes hatten • die die Studie beendet hatten und einen adäquaten 6MWD-Test in Woche 16 aufwiesen

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • die keine schwerwiegenden Protokollverletzungen aufwiesen, die die Wirksamkeit hätten beeinflussen können. Dazu zählen: <ul style="list-style-type: none"> ○ CTEPH-Definition nicht erfüllt ○ Patient aus Sicht des Prüfarztes nicht geeignet ○ PVR-Ausgangswert $\leq 300 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$ ○ PAPm-Ausgangswert $\leq 25 \text{ mmHg}$ ○ PVR-Ausgangswert bei inoperablen Patienten gemessen, bevor diese 90 Tage vollständige Antikoagulation erhalten hatten oder weniger als 180 Tage nach PEA und/oder Rechtsherzkatheteruntersuchungsergebnisse älter als 8 Wochen bei Visite 1 ○ Eignungs- und/oder Ausgangs-6MWD geringer als 150 m oder größer als 450 m ○ Relative Differenz zwischen Eignungs- und Ausgangs-6MWD-Tests von mehr als 15% ○ Jünger als 18 Jahre oder älter als 80 Jahre bei Visite 1 ○ Administration der Studienmedikation nicht in Einklang mit den Vorgaben des Studienprotokolls ○ Einnahme nicht erlaubter Begleitmedikationen ab 90 Tage vor Visite 1 bis zur Woche 16 für NO-Donoren (z.B. Nitrate) und ab 30 Tage vor Aufnahme des Rechtsherzkatheterausgangswertes bis zu Woche 16 für Endothelin-Rezeptor-Antagonisten, Prostanoiden, spezifische (z.B. Sildenafil oder Tadalafil) oder unspezifische Phosphodiesterase-Inhibitoren (z. B. Dipyridamol, Theophyllin) <p>Statistischer Analyseplan</p> <p>Der statistische Analyseplan (SAP), Version 2.0, wurde am 30. Sep. 2009 erstellt, Änderungen in Versionen 3.0 (14. Dez. 2010) und Version 4.0 (11. Jul. 2012) sowie Anhängen zu Versionen 2.0 (15. März 2012) und 3.0 (11. Jul. 2012) vorgenommen. Die letzten Versionen aller SAPs wurden jeweils</p>

Item ^a	Charakteristikum	Studieninformation
		<p>vor Entblindung der Daten erstellt.</p> <p>Die Ergebnisse wurden für die Behandlungsgruppen zusammengefasst. Sofern nicht anders angegeben, wurden alle statistischen Tests zweiseitig mit einem Fehler 1. Art von $\alpha = 5\%$ durchgeführt.</p> <p>Die statistischen Analysen wurden mit dem Softwarepaket SAS Version 9.2 (SAS Institute Inc., Cary, NC, USA) durchgeführt.</p> <p>Primäre Wirksamkeitsanalyse</p> <ul style="list-style-type: none"> • Veränderung der 6MWD zum Ausgangswert bis Woche 16 (letzte Beobachtung bis Woche 16) in der ITT-Gruppe; Imputation fehlender Werte für Patienten, die vorzeitig die Studie verließen oder vor Woche 16 verstarben • Vergleich der Riociguat-1,0-2,5 mg- und Placebo-Gruppe mittels Kovarianz-Analyse (ANCOVA), mit 6MWD-Ausgangswert als einer Kovariate und Behandlungsgruppe und Region als Haupteffekte • Berechnung der geschätzten Randmittelwerte (Least squares mean, LS MW) und der 95%-KI für die Behandlungsunterschiede mittels Kovarianzanalyse (ANCOVA) • Untersuchung der Normalverteilung der Residuen von ANCOVA mittels Shapiro-Wilk-Test. Wenn dieser auf dem 5%-Niveau signifikant war, wurde statt ANCOVA der nach Region und Stratifizierungsgruppe stratifizierte Wilcoxon-Test verwendet • Überlegenheit der Riociguat-1,0-2,5 mg-Gruppe über die Placebo-Gruppe, wenn das zweiseitige Signifikanzniveau geringer oder gleich 0,05 war <p>Sekundäre Wirksamkeitsanalysen</p> <ul style="list-style-type: none"> • Durchführung der Analyse der sekundären Endpunkte <ol style="list-style-type: none"> 1. PVR 2. NT-proBNP 3. WHO-Funktionsklasse 4. Zeit bis zur klinischen Verschlechterung 5. Borg

Item ^a	Charakteristikum	Studieninformation
		<p>6. EQ-5D 7. LPH</p> <p>in der ITT-Population (sequentielles Testverfahren strikt in der oben aufgeführten Reihenfolge)</p> <ul style="list-style-type: none"> • Durchführung einer supportiven PP-Analyse • Analyse der Veränderungen von PVR, NT-proBNP, EQ-5D und LPH zum Ausgangswert bis Woche 16 wie 6MWD • Berechnung der Veränderungen der WHO-Funktionsklasse zum Ausgangswert bis Woche 16 mittels eines Scores: <ul style="list-style-type: none"> ○ von -3 (Ausgangswert Klasse IV, Woche 16 Klasse I) bis +4 (Ausgangswert Klasse I, Woche 16 Tod), statistische Analyse mit stratifiziertem Wilcoxon-Test • Analyse der Veränderungen der Borg-Skala zum Ausgangswert bis Woche 16 mit stratifiziertem Wilcoxon-Test • Analyse der Zeit bis zur klinischen Verschlechterung mit <ul style="list-style-type: none"> ○ stratifiziertem Log-Rank-Test mit Kaplan-Meier Schätzungen für den Anteil der Patienten mit einem Ereignis ○ Mantel-Haenszel Gewichtung mit Stratifizierung nach Region <p>Imputation fehlender Werte</p> <ul style="list-style-type: none"> • Verwendung der Daten, die bis zum Studienende (Woche 16) vorlagen • Verwendung der Werte der Abschluss-Visite oder der letzten Visite im Fall von vorzeitiger Beendigung der Studienmedikation • Ersetzen fehlender Werte im Fall von Tod, vorzeitigem Ausscheiden wegen klinischer Verschlechterung oder fehlender Messwerte bei der Abschluss-Visite: <ul style="list-style-type: none"> ○ 6MWD schlechtestmöglicher Wert, d.h. 0 m ○ Modifizierte Borg-Skala schlechtestmöglicher Wert, d.h. 10 ○ EQ-5D und LPH jeweils schlechtestmögliche Werte ○ WHO-Funktionsklasse schlechtestmöglicher Wert, d.h. IV bei fehlender Abschluss-Visite bzw. V bei Tod

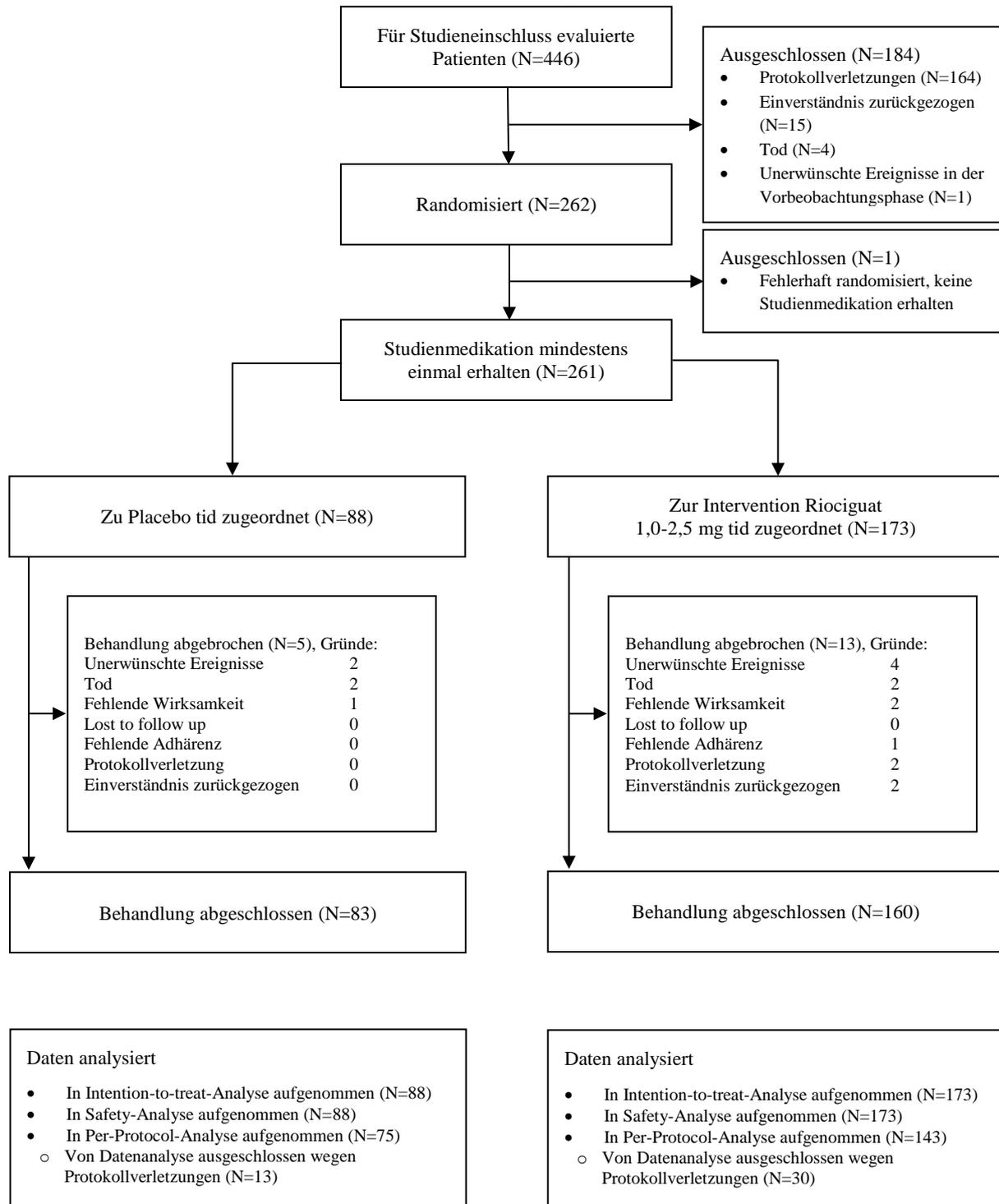
Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none">○ PVR ohne Imputation○ NT-proBNP ohne Imputation• Verwendung der Ausgangswerte bei vorzeitiger Beendigung aus anderen Gründen• Verwendung der letzten Werte, die nach den Ausgangswerten bestimmt wurden bei regulärer Beendigung der Studie und fehlenden Wirksamkeitsmessungen in Woche 16

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p data-bbox="906 282 1385 311">Deskriptive Subgruppenanalysen</p> <p data-bbox="906 340 1385 459">Die primären und sekundären Wirksamkeitsvariablen wurden in der Safety-/ITT-Population für die folgenden Subgruppen deskriptiv zusammengefasst:</p> <ul data-bbox="906 488 1385 1556" style="list-style-type: none"> <li data-bbox="906 488 1385 517">• Inoperabel <li data-bbox="906 519 1385 577">• Persistierende oder rezidivierende CTEPH nach einer PEA <li data-bbox="906 607 1385 665">• WHO-Funktionsklasse I/II als Ausgangswert <li data-bbox="906 667 1385 725">• WHO-Funktionsklasse III/IV als Ausgangswert <li data-bbox="906 763 1385 792">• 6MWD-Ausgangswert <320 m <li data-bbox="906 795 1385 824">• 6MWD-Ausgangswert ≥320 m <li data-bbox="906 862 1385 891">• 6MWD-Ausgangswert <380 m <li data-bbox="906 893 1385 922">• 6MWD-Ausgangswert ≥380 m <li data-bbox="906 960 1385 990">• Männer <li data-bbox="906 992 1385 1021">• Frauen <li data-bbox="906 1059 1385 1088">• Alter <65 Jahre <li data-bbox="906 1090 1385 1120">• Alter ≥65 Jahre <li data-bbox="906 1158 1385 1187">• Weiße (Kaukasier) <li data-bbox="906 1189 1385 1218">• Farbige (Nicht-Kaukasier) <li data-bbox="906 1220 1385 1249">• Asiaten (Nicht-Kaukasier) <li data-bbox="906 1252 1385 1310">• Hispanics oder Latinos (Nicht-Kaukasier) <li data-bbox="906 1312 1385 1364">• Gemischt (Nicht-Kaukasier) <li data-bbox="906 1366 1385 1395">• <li data-bbox="906 1411 1385 1440">• Nordamerika <li data-bbox="906 1442 1385 1471">• Südamerika <li data-bbox="906 1473 1385 1503">• Europa <li data-bbox="906 1505 1385 1534">• China <li data-bbox="906 1536 1385 1565">• Asien-Pazifik <li data-bbox="906 1568 1385 1597">• <p data-bbox="906 1579 1385 1608">Explorative Wirksamkeitsanalysen</p> <ul data-bbox="906 1659 1385 1955" style="list-style-type: none"> <li data-bbox="906 1659 1385 1718">• Zusätzliche Analysen hämodynamischer Parameter <li data-bbox="906 1720 1385 1778">• Zusätzliche Analysen der PRO EQ-5D und LPH <li data-bbox="906 1794 1385 1852">• Kategoriale Analysen von Wirksamkeitsvariablen <li data-bbox="906 1868 1385 1926">• Sensitivitätsanalyse von Zeit bis zur klinischen Verschlechterung <li data-bbox="906 1942 1385 1955">• Sensitivitätsanalyse der 6MWD

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Robuste Regression der 6MWD <p>Zusätzliche Analysen</p> <ul style="list-style-type: none"> • Beziehung zwischen Outcomes • Ressourcenverbrauch <p>Safety-Analysen</p> <ul style="list-style-type: none"> • Tabellarische Darstellung <ul style="list-style-type: none"> ○ Von mit der Studienmedikation assoziierten unerwünschten Ereignissen (Definition: Start oder Verschlechterung bis zu zwei Tage nach Beendigung der Behandlung mit der Studienmedikation) ○ Von schweren und/oder mit der Studienmedikation assoziierten unerwünschten Ereignissen ○ Von unerwünschten Ereignissen, die in der Nachbeobachtungsphase auftraten • Deskriptive Analyse der Mortalität <p>Sonstige Variablen</p> <ul style="list-style-type: none"> • Explorative Biomarker: <ul style="list-style-type: none"> ○ Analyse von Osteopontin und CTX wie NT-proBNP • Pharmakokinetik: <ul style="list-style-type: none"> ○ Bestimmung der Konzentration von Riociguat und seinem Hauptmetaboliten BAY 60-4552 im Plasma bei jeder Visite und Auswertung mittels deskriptiver Statistik
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Insgesamt wurden in 89 Studienzentren 446 Patienten rekrutiert. a) Randomisierung Riociguat-1,0-2,5 mg-Gruppe: 174 Patienten Placebo-Gruppe: 88 Patienten b) Erhalt der Intervention Riociguat-1,0-2,5 mg-Gruppe: 173 Patienten Placebo-Gruppe: 88 Patienten

Item ^a	Charakteristikum	Studieninformation
		<p>c) Analyse des primären Zielkriteriums</p> <p>Riociguat-1,0-2,5 mg-Gruppe: 173 Patienten</p> <p>Placebo-Gruppe: 88 Patienten</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>243 (92,7%) von 262 randomisierten Patienten beendeten die Behandlungsphase. Die verbleibenden 19 Patienten brachen die Studie vorzeitig ab (18 Patienten) oder erhielten keine Studienmedikation (1 Patient).</p> <p>Der Anteil war in der Riociguat-1,0-2,5 mg-Gruppe (8,0%) etwas höher als in der Placebo-Gruppe (5,7%).</p> <p>Der Unterschied war nicht auf Sicherheitsaspekte zurückzuführen, der Anteil an unerwünschten Ereignissen war in beiden Gruppen ähnlich.</p> <p>Die Gründe sind im CONSORT Flussdiagramm am Ende dieser Tabelle aufgelistet (Abbildung 3).</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Beginn der Studie: 23. Februar 2009 (Erster Patient zur ersten Visite)</p> <p>Ende der Studie: 27. Juni 2012 (Letzter Patient zur letzten Visite)</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete planmäßig nach der letzten abgeschlossenen Visite am 27. Juni 2012.
a: nach CONSORT 2010.		

Abbildung 4-3: Patientenfluss in einem Flow-Chart gemäß CONSORT, **CHEST-1**



Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-46 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie CHEST-1

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Bayer HealthCare Studienbericht Nr. A62508, Studiennummer 11348, Version 1.0 “Randomized, double-blind, placebo-controlled, multicentre, multinational study to evaluate the efficacy and safety of oral BAY 63-2521 (1 mg, 1.5 mg, 2 mg, or 2.5 mg tid) in patients with chronic thromboembolic pulmonary hypertension (CTEPH). CHEST-1 study”	A

A: Verzerrungsaspekte auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

In der Studie wurden die Behandlungen den Patienten randomisiert zugeteilt.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gruppenzuteilung erfolgte rein zufällig, die Erzeugung der Zuteilungssequenz mittels computergenerierter Zufallscodes ist beschrieben und geeignet.

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Zuteilung erfolgte durch eine zentrale, unabhängige Einheit (IVRS oder IWRS).

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten waren verblindet.

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die behandelnden Personen waren bzgl. der Behandlung verblindet.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Die Studie wurde als qualitativ hochwertige, randomisierte, kontrollierte, multizentrische, doppelblinde Phase-III-Studie durchgeführt. Die Studie wurde in 89 Studienzentren in 26 Ländern weltweit durchgeführt, von denen jedes zwischen 1-24 Patienten rekrutierte, so dass nicht ein Zentrum die Ergebnisse dominiert. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial auf Studienebene als niedrig eingestuft.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt

Endpunkt: Absolute Veränderung der 6-Minuten-Gehstrecke (6MWD) bis Woche 16 im Vergleich zum Ausgangswert (primärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Laut Studienbericht wurden die Analysen in Übereinstimmung mit der ICH-Guideline E9 zu „Statistical Principles for Clinical Trials“ durchgeführt. In das „Full Analysis Set“ einer ITT-Analyse sollten möglichst alle Patienten eingeschlossen werden, die randomisiert wurden. Drei Ausnahmen von dieser Regel sind zulässig. Eine davon, der Ausschluss eines Patienten, der keine einzige Dosis der Studienmedikation erhalten hat, trifft auf die Analyse des primären Endpunktes 6MWD zu und ist Guideline-konform. In der Riociguat-1,0-2,5 mg-Gruppe traf dies auf einen Patienten (610018003) von 174 Patienten (0,6%) zu. Das „Full Analysis Set“ umfasst in der Riociguat-1,0-2,5 mg-Gruppe 173 und in der Placebo-Gruppe 88 Patienten.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat durch Imputation ersetzt. Lagen wegen eines Todesfalls oder der Beendigung der Studie wegen klinischer Verschlechterung keine Daten vor, so wurde die 6MWD auf den schlechtestmöglichen Wert, d.h. 0 m, gesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem primären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Änderung der WHO-Funktionsklasse bis Woche 16 im Vergleich zum Ausgangswert (sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Patienten und Studienpersonal waren gegenüber der Intervention verblindet. Zusätzlich musste die WHO-Funktionsklasse ab Visite 2 von einem zweiten Arzt bestimmt werden, der nicht in die Dosisitration involviert war und dem der Blutdruck und die Herzfrequenz des Patienten nach Einnahme der Studienmedikation nicht bekannt waren. Dieser Arzt hat den Befund auf einem eigenen Arbeitsblatt dokumentiert, von dem eine dritte Person das Ergebnis in die Datenbank übertrug.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Die Analyse der WHO-Funktionsklasse umfasste in der Riociguat-1,0-2,5 mg-Gruppe 173/173 (100%) und in der Placebo-Gruppe 87/88 (98,9%) der Patienten der ITT-Population. Eine Abweichung von <5% von der ITT-Population wird als adäquat angesehen.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat durch Imputation ersetzt indem der schlechtestmögliche Wert, d.h. IV bei klinischer Verschlechterung oder V bei Tod, eingesetzt wurde. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Zeit bis zur klinischen Verschlechterung in Studientagen (sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Die Analyse des Endpunktes wurde in der ITT-Population durchgeführt, mit 173/173 Patienten (100%) in der Riociguat-1,0-2,5 mg-Gruppe und 88/88 Patienten (100%) in der Placebo-Gruppe.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt und beschrieben worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat durch die Time-to-Event Analyse mittels Log-Rank-Tests einbezogen. Todesfälle und die Beendigung der Studie wegen klinischer Verschlechterung wurden dabei als Ereignis gezählt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder ist auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich

keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Veränderung der Dyspnoe und des Ermüdungsempfindens gemäß Borg-Skala bis Woche 16 im Vergleich zum Ausgangswert (sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Die Analyse des Endpunktes wurde in der ITT-Population durchgeführt, mit 173/173 Patienten (100%) in der Riociguat-1,0-2,5 mg-Gruppe und 88/88 Patienten (100%) in der Placebo-Gruppe.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt und beschrieben worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat ersetzt. Daten, die zum Studienende (Woche 16) vorlagen, wurden verwendet. Lagen zum Studienende wegen eines Todesfalles oder der Beendigung der Studie wegen klinischer Verschlechterung keine Daten vor, wurde der BORG-Wert auf den schlechtestmöglichen Wert, d.h. 10, gesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Veränderung der gesundheitsbezogenen Lebensqualität gemäß EQ-5D nach 16 Wochen im Vergleich zum Ausgangswert (sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Die Analyse des EQ-5D-Index wurde in einer ITT-Population durchgeführt, die für die Riociguat-1,0-2,5 mg-Gruppe 172/173 (99,4%) und für die Placebo-Gruppe 87/88 (98,9%) Patienten umfasste.
- Damit lag die Abweichung von der ITT-Population unter 5% und wird als adäquat angesehen.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt und beschrieben worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat ersetzt. Daten, die zum Studienende (Woche 16) vorlagen, wurden verwendet. Lagen zum Studienende wegen eines Todesfalles oder der Beendigung der Studie wegen klinischer Verschlechterung keine Daten vor, wurde der schlechtestmögliche EQ-5D Wert eingesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Veränderung der gesundheitsbezogenen Lebensqualität gemäß Living with Pulmonary Hypertension (LPH)-Fragebogen nach 16 Wochen im Vergleich zum Ausgangswert (sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

- Die Analyse des LPH wurde in einer ITT-Population durchgeführt, die in der Riociguat-1,0-2,5 mg-Gruppe 170/173 (98,3%) Patienten und in der Placebo-Gruppe 86/88 (97,7%) Patienten umfasste. Eine Abweichung von <5% von der ITT-Population wird als adäquat angesehen.
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt und beschrieben worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung). Fehlende Daten wurden adäquat ersetzt. Daten, die zum Studienende (Woche 16) vorlagen, wurden verwendet. Lagen zum Studienende wegen eines Todesfalles oder der Beendigung der Studie wegen klinischer Verschlechterung keine Daten vor, wurde der schlechtestmögliche LPH Wert eingesetzt. Im Falle von Studienabbrüchen aus anderen Gründen ohne eine weitere Messung nach Bestimmung des Ausgangswertes, wurde der Ausgangswert eingesetzt (BOCF). Falls ein Patient die Studie wie geplant beendete, aber es keine Messung am Ende der Studie gab, wurde der letzte, nach Bestimmung des Ausgangswertes gemessene Wert verwendet (LOCF).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde umgesetzt, wobei fehlende Daten adäquat durch Imputation ersetzt wurden und es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Endpunkt: Absolute Veränderung des pulmonal vaskulären Widerstands (PVR) nach 16 Wochen im Vergleich zum Ausgangswert (supportiver sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse wurde in einer mITT-Population durchgeführt, die in der Riociguat-1,0-2,5 mg-Gruppe 151/173 (87,3%) Patienten der ITT-Population und in der Placebo-Gruppe 82/88 (93,2%) der ITT-Population umfasste.

Fehlende Daten wurden im Fall von Tod oder fehlenden Post-Ausgangswerten nicht durch Imputation ergänzt. Im Falle einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangsvs. verwendet. Es wurden keine weiterführenden Sensitivitätsanalysen zur Untersuchung des Effektes der fehlenden Daten durchgeführt. Aufgrund dieser Abweichung wird das Verzerrungspotenzial für diese zwei Endpunkte als hoch eingestuft.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig hoch

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde nicht umgesetzt und stattdessen eine mITT-Population analysiert, die weniger als 95% der ITT-Population repräsentierte. Es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten

Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Aufgrund der Abweichung von der ITT-Population wird das Verzerrungspotenzial für diesen Endpunkt als hoch eingestuft.

Endpunkt: Absolute Veränderung des Biomarkers NT-proBNP bis Woche 16 im Vergleich zum Ausgangswert (supportiver sekundärer Endpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse wurde in einer mITT-Population durchgeführt, die in der Riociguat-1,0-2,5 mg-Gruppe 150/173 (86,7%) Patienten der ITT-Population und in der Placebo-Gruppe 73/88 (83,0%) Patienten der ITT-Population umfasste.

Fehlende Daten wurden im Fall von Tod oder fehlenden Post-Ausgangswerten nicht durch Imputation ergänzt. Im Falle einer vorzeitigen Beendigung der Studienmedikation wurden die Daten der Abschlussvisite oder der letzten Post-Ausgangswerte verwendet. Es wurden keine weiterführenden Sensitivitätsanalysen zur Untersuchung des Effektes der fehlenden Daten durchgeführt. Aufgrund dieser Abweichung wird das Verzerrungspotenzial für diese zwei Endpunkte als hoch eingestuft.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem sekundären Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig hoch

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor, das ITT-Prinzip wurde nicht umgesetzt und stattdessen eine mITT-Population

analysiert, die weniger als 95% der ITT-Population repräsentierte. Es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können. Aufgrund der Abweichung von der ITT-Population wird das Verzerrungspotenzial für diesen Endpunkt als hoch eingestuft.

Endpunkt: Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse (Sicherheitsendpunkt)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Der Endpunkt wurde verblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Safety-Population stimmt mit der ITT-Population überein.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Darstellung von Ergebnissen zu dem Endpunkt ist nicht unterlassen worden oder auch nicht in einer von der Planung abweichenden Weise erfolgt. Eine ergebnisgesteuerte Berichterstattung ist daher unwahrscheinlich.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Auf Studienebene liegt ein niedriges Verzerrungspotenzial vor. Auf Endpunktebene lag eine Verblindung des Endpunkterhebers vor. Die Safety-Population stimmt mit der ITT-Population überein. Es lag eine ergebnisunabhängige Berichterstattung dieses Endpunktes vor. Der klinische Studienbericht bildet die prospektiv geplanten Analysen vollständig ab. Es zeigen sich keine sonstigen Aspekte, die zu Verzerrungen führen können.

Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind. Insgesamt wird das Verzerrungspotenzial als niedrig eingestuft.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Verzerrungsaspekte auf Studienebene:**Einstufung als randomisierte Studie**

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient**

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven

Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
 - Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

