

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Nintedanib (Ofev®)

Boehringer Ingelheim Pharma GmbH & Co. KG

Modul 4 A

Idiopathische Lungenfibrose

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Inhaltsverzeichnis.....	1
Tabellenverzeichnis.....	4
Abbildungsverzeichnis.....	9
Abkürzungsverzeichnis.....	12
4 Modul 4 – allgemeine Informationen.....	16
4.1 Zusammenfassung der Inhalte von Modul 4.....	17
4.2 Methodik.....	32
4.2.1 Fragestellung.....	32
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	33
4.2.3 Informationsbeschaffung.....	40
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	40
4.2.3.2 Bibliografische Literaturrecherche.....	40
4.2.3.3 Suche in Studienregistern.....	42
4.2.3.4 Selektion relevanter Studien.....	43
4.2.4 Bewertung der Aussagekraft der Nachweise.....	44
4.2.5 Informationssynthese und -analyse.....	46
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	46
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	47
4.2.5.3 Meta-Analysen.....	67
4.2.5.4 Sensitivitätsanalysen.....	70
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	71
4.2.5.6 Indirekte Vergleiche.....	73
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	75
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	75
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	75
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	75
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	79
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	82
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	84
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	87
4.3.1.2.1 Studiendesign und Studienpopulationen.....	87
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	101
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	103
4.3.1.3.1 Mortalität – RCT.....	106
4.3.1.3.1.1 Endpunkte <i>Zeit bis zum Tod</i>	107
4.3.1.3.1.2 Endpunkte <i>Forcierte Vitalkapazität</i>	116
4.3.1.3.2 Morbidität – RCT.....	131
4.3.1.3.2.1 Endpunkte <i>akute Exazerbationen</i>	132

4.3.1.3.2.2	Endpunkt <i>HRCT quantitativer Lungenfibrosescore (HRCT-QLF)</i>	146
4.3.1.3.2.3	Endpunkt <i>Veränderung der Gehstrecke im 6-Minuten-Gehtest (6 MWT)</i>	149
4.3.1.3.2.4	Endpunkt <i>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</i>	152
4.3.1.3.3	Gesundheitsbezogene Lebensqualität – RCT.....	155
4.3.1.3.3.1	Endpunkte zum <i>SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI- C und EQ-5D</i>	156
4.3.1.3.4	Unerwünschte Ereignisse – RCT	188
4.3.1.3.4.1	Endpunkt <i>Gesamtzahl UE</i>	189
4.3.1.3.4.2	Endpunkt <i>schwerwiegende UE</i>	196
4.3.1.3.4.3	Endpunkt <i>UE, die zum Therapieabbruch führten</i>	202
4.3.1.3.4.4	Endpunkt <i>UE von besonderem Interesse</i>	208
4.3.1.3.5	Subgruppenanalysen – RCT.....	231
4.3.1.3.5.1	Mortalität	234
4.3.1.3.5.1.1	Endpunkt <i>Zeit bis zum Tod, unabhängig von der Ursache, Zeit bis zum respiratorisch bedingten Tod</i>	234
4.3.1.3.5.1.2	Endpunkt <i>Forcierte Vitalkapazität</i>	235
4.3.1.3.5.2	Morbidität	236
4.3.1.3.5.2.1	Endpunkte zu <i>akuten Exazerbationen</i>	236
4.3.1.3.5.2.2	Endpunkt <i>HRCT-QLF Score</i>	237
4.3.1.3.5.2.3	Endpunkt <i>Veränderung der Gehstrecke im 6 MWT</i>	237
4.3.1.3.5.3	Gesundheitsbezogene Lebensqualität.....	237
4.3.1.3.5.3.1	Endpunkte zum <i>SGRQ, SGRQ-I, SOBQ, CASA-Q,PGI-C und EQ-5D-VAS</i>	238
4.3.1.3.5.4	Unerwünschte Ereignisse	244
4.3.1.3.5.4.1	Endpunkt <i>Gesamtzahl UE</i>	245
4.3.1.3.5.4.2	Endpunkt <i>Gesamtzahl SUE</i>	245
4.3.1.3.5.4.3	Endpunkt <i>UE, die zum Therapieabbruch führten</i>	245
4.3.1.3.5.4.4	Endpunkt <i>UE von besonderem Interesse</i>	247
4.3.1.3.6	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	250
4.3.2	Weitere Unterlagen.....	259
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	259
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	259
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	259
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	259
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	260
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	262
4.3.2.2	Nicht-randomisierte vergleichende Studien.....	263
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht-randomisierte vergleichende Studien	263
4.3.2.2.2	Charakteristika der nicht-randomisierten vergleichenden Studien	263
4.3.2.2.3	Ergebnisse aus nicht-randomisierten vergleichenden Studien	264
4.3.2.2.3.1	<Endpunkt xxx> – nicht-randomisierte vergleichende Studien	264

4.3.2.2.3.2	Subgruppenanalysen – nicht-randomisierte vergleichende Studien	265
4.3.2.3	Weitere Untersuchungen.....	266
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	266
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	266
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	266
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	266
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	267
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	267
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	268
4.4.1	Beurteilung der Aussagekraft der Nachweise	268
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	271
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	281
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	281
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	281
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	281
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	281
4.5.4	Verwendung von Surrogatendpunkten	282
4.6	Liste der eingeschlossenen Studien.....	291
4.7	Referenzliste.....	292
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		303
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		308
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		311
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		314
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		344
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		377

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für die Nutzenbewertung des Arzneimittels Nintedanib.....	17
Tabelle 4-2: Ergebnisse für UE von besonderem Interesse (INPULSIS Studien und Studie 1199.187 gepoolt)	28
Tabelle 4-3: Ergebnisse zum medizinischen Zusatznutzen von Nintedanib.....	30
Tabelle 4-4: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für die Nutzenbewertung des Arzneimittels Nintedanib.....	37
Tabelle 4-5: Faktoren zur Charakterisierung der Studienpopulation der RCT mit dem zu bewertenden Arzneimittel	48
Tabelle 4-6: Endpunkte der RCT mit dem zu bewertenden Arzneimittel für die vorliegende Nutzenbewertung.....	49
Tabelle 4-7: Akute Exazerbationen gemäß Adjudizierungsergebnis, aufgeschlüsselt nach SUE und nicht SUE in den RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-8: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	76
Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	79
Tabelle 4-10: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	85
Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	88
Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	92
Tabelle 4-14: Charakterisierung der Studienpopulation (Ausgangswert) – RCT mit dem zu bewertenden Arzneimittel	93
Tabelle 4-15: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	102
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	103
Tabelle 4-17: Operationalisierung von <i>Zeit bis zum Tod</i>	107
Tabelle 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt <i>Zeit bis zum Tod</i> in RCT mit dem zu bewertenden Arzneimittel	108
Tabelle 4-19: Ergebnisse für <i>Zeit bis zum Tod</i> aus RCT mit dem zu bewertenden Arzneimittel.....	110
Tabelle 4-20: Operationalisierung von <i>FVC</i>	116
Tabelle 4-21: Bewertung des Verzerrungspotenzials für den Endpunkt <i>FVC</i> in RCT mit dem zu bewertenden Arzneimittel.....	118

Tabelle 4-22: Ergebnisse für <i>jährliche FVC-Abnahme</i> (ml) aus RCT mit dem zu bewertenden Arzneimittel	120
Tabelle 4-23: Ergebnisse für <i>FVC-Responder zu Woche 52</i> bezogen auf die Schwellenwerte für die absolute Verschlechterung gegenüber dem Ausgangswert um 10%-Punkte bzw. 5%-Punkte aus RCT mit dem zu bewertenden Arzneimittel	124
Tabelle 4-24: Ergebnisse für <i>FVC-Responder zu Woche 24</i> bezogen auf die Schwellenwerte für die absolute Verschlechterung gegenüber dem Ausgangswert um 10%-Punkte bzw. 5%-Punkte aus RCT mit dem zu bewertenden Arzneimittel	125
Tabelle 4-25: Operationalisierung von <i>akute Exazerbationen</i>	132
Tabelle 4-26: Bewertung des Verzerrungspotenzials für den Endpunkt <i>akute Exazerbationen</i> in RCT mit dem zu bewertenden Arzneimittel	133
Tabelle 4-27: Ergebnisse für <i>Zeit bis zur ersten adjudizierten akuten Exazerbation</i> und <i>Zeit bis zur ersten akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel	136
Tabelle 4-28: Sensitivitätsanalyse - Ergebnisse für <i>Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation</i> und <i>Zeit bis zur ersten schwerwiegenden akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel	140
Tabelle 4-29: <i>Akute Exazerbationen</i> in den gepoolten INPULSIS Studien und der Extensionsstudie INPULSIS-ON	145
Tabelle 4-30: Operationalisierung von <i>HRCT-QLF Score</i>	146
Tabelle 4-31: Bewertung des Verzerrungspotenzials für den Endpunkt <i>HRCT-QLF Score</i> in RCT mit dem zu bewertenden Arzneimittel	146
Tabelle 4-32: Ergebnisse für <i>Veränderung des HRCT-QLF Scores bis Woche 24</i> aus RCT mit dem zu bewertenden Arzneimittel	148
Tabelle 4-33: Operationalisierung von <i>6 MWT</i>	149
Tabelle 4-34: Bewertung des Verzerrungspotenzials für den Endpunkt <i>6 MWT</i> in RCT mit dem zu bewertenden Arzneimittel	150
Tabelle 4-35: Ergebnisse für <i>Veränderung der Gehstrecke im 6 MWT bis Woche 24</i> aus RCT mit dem zu bewertenden Arzneimittel	151
Tabelle 4-36: Operationalisierung von <i>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</i>	152
Tabelle 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt <i>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</i> in RCT mit dem zu bewertenden Arzneimittel ..	153
Tabelle 4-38: Ergebnisse für <i>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</i> mit dem zu bewertenden Arzneimittel	154
Tabelle 4-39: Operationalisierung zu <i>SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI-C</i> und <i>EQ-5D</i>	156
Tabelle 4-40: Bewertung des Verzerrungspotenzials für die Endpunkte <i>SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI-C</i> und <i>EQ-5D</i> in RCT mit dem zu bewertenden Arzneimittel	159
Tabelle 4-41: Ergebnisse für <i>Veränderung SGRQ-Gesamtscore</i> und <i>Veränderung SGRQ-I-Gesamtscore bis Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel	162

Tabelle 4-42: Ergebnisse für <i>Veränderung SGRQ-Gesamtscore</i> und <i>Veränderung SGRQ-I-Gesamtscore bis Woche 24</i> aus RCT mit dem zu bewertenden Arzneimittel	165
Tabelle 4-43: Ergebnisse für <i>SGRQ-(≤ 4 Punkte)-Responder</i> und <i>SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel.....	167
Tabelle 4-44: Ergebnisse für <i>SGRQ-(≤ 4 Punkte)-Responder</i> und <i>SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 24</i> aus RCT mit dem zu bewertenden Arzneimittel	169
Tabelle 4-45: Ergebnisse für <i>Veränderung SOBQ-Gesamtscore bis Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel	170
Tabelle 4-46: Ergebnisse für <i>Veränderung SOBQ-Gesamtscore bis Woche 24</i> aus RCT mit dem zu bewertenden Arzneimittel	172
Tabelle 4-47: Ergebnisse für <i>Veränderung CASA-Q (Husten-Symptome bzw. Husten-Impact) bis Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel	173
Tabelle 4-48: Ergebnisse für <i>PGI-C-Responder zu Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel	175
Tabelle 4-49: Ergebnisse für <i>Veränderung EQ-5D-VAS bis Woche 52</i> aus RCT mit dem zu bewertenden Arzneimittel	177
Tabelle 4-50: Operationalisierung von <i>Gesamtzahl UE</i>	189
Tabelle 4-51: Bewertung des Verzerrungspotenzials für den Endpunkt <i>Gesamtzahl UE</i> in RCT mit dem zu bewertenden Arzneimittel	190
Tabelle 4-52: Ergebnisse für die <i>Gesamtzahl UE</i> aus RCT mit dem zu bewertenden Arzneimittel.....	191
Tabelle 4-53: Sensitivitätsanalyse: Ergebnisse für die <i>Gesamtzahl UE ohne Berücksichtigung von adjudizierten Exazerbationen</i> aus RCT mit dem zu bewertenden Arzneimittel.....	192
Tabelle 4-54: Häufige UE ($\geq 5\%$ in mindestens einem Studienarm) nach MedDRA SOC und PT aus RCT mit dem zu bewertenden Arzneimittel – INPULSIS Studien und Studie 1199.187 (gepoolt).....	193
Tabelle 4-55: Operationalisierung von <i>SUE</i>	196
Tabelle 4-56: Bewertung des Verzerrungspotenzials für die Endpunkte zu den <i>SUE</i> in RCT mit dem zu bewertenden Arzneimittel	197
Tabelle 4-57: Ergebnisse für die <i>SUE</i> aus RCT mit dem zu bewertenden Arzneimittel	198
Tabelle 4-58: Sensitivitätsanalyse: Ergebnisse für die <i>Gesamtzahl SUE ohne Berücksichtigung von adjudizierten Exazerbationen</i> aus RCT mit dem zu bewertenden Arzneimittel.....	199
Tabelle 4-59: Häufige <i>SUE</i> ($\geq 1\%$ in mindestens einem Studienarm) nach MedDRA SOC und PT aus RCT mit dem zu bewertenden Arzneimittel – INPULSIS Studien und Studie 1199.187 (gepoolt).....	200
Tabelle 4-60: Operationalisierung von <i>UE, die zum Therapieabbruch führten</i> in RCT mit dem zu bewertenden Arzneimittel.....	202
Tabelle 4-61: Bewertung des Verzerrungspotenzials für den Endpunkt <i>UE, die zum Therapieabbruch führten</i> , in RCT mit dem zu bewertenden Arzneimittel	202

Tabelle 4-62: Ergebnisse für <i>UE, die zum Therapieabbruch führten</i> aus RCT mit dem zu bewertenden Arzneimittel	204
Tabelle 4-63: Sensitivitätsanalyse: Ergebnisse für <i>UE, die zum Therapieabbruch führten</i> ohne Berücksichtigung von adjudizierten Exazerbationen aus RCT mit dem zu bewertenden Arzneimittel	205
Tabelle 4-64: Ergebnisse für <i>UE, die zum Therapieabbruch führten</i> mit Inzidenz $\geq 1\%$ aus RCT mit dem zu bewertenden Arzneimittel - INPULSIS Studien und Studie 1199.187 (gepoolt)	206
Tabelle 4-65: Operationalisierung von <i>UE von besonderem Interesse</i>	208
Tabelle 4-66: Bewertung des Verzerrungspotenzials für den Endpunkt <i>UE von besonderem Interesse</i> in RCT mit dem zu bewertenden Arzneimittel	209
Tabelle 4-67: Ergebnisse für <i>UE – gastrointestinale Ereignisse</i> aus RCT mit dem zu bewertenden Arzneimittel	211
Tabelle 4-68: Ergebnisse für <i>UE – Leber- und Gallenerkrankungen</i> (Leberenzym-, Bilirubin-, AP-Erhöhung) aus RCT mit dem zu bewertenden Arzneimittel	218
Tabelle 4-69: Ergebnisse für <i>UE - kardiale Ereignisse</i> aus RCT mit dem zu bewertenden Arzneimittel	222
Tabelle 4-70: Ergebnisse für <i>UE - respiratorische Ereignisse</i> aus RCT mit dem zu bewertenden Arzneimittel	225
Tabelle 4-71: Ergebnisse für <i>UE - Blutungsereignisse</i> aus RCT mit dem zu bewertenden Arzneimittel	228
Tabelle 4-72: Subgruppenmerkmale, die als potenzielle Effektmodifikatoren untersucht wurden	232
Tabelle 4-73: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur Mortalität basierend auf der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187	234
Tabelle 4-74: Subgruppenergebnisse für <i>FVC-Responder</i> (Schwellenwert 5% des Sollwertes) aus RCT für die gepoolte Auswertung der INPULSIS Studien	235
Tabelle 4-75: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur Morbidität	236
Tabelle 4-76: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur gesundheitsbezogenen Lebensqualität	237
Tabelle 4-77: Subgruppenergebnisse für die Veränderung des SGRQ-I-Gesamtscores – basierend auf der gepoolten Analyse der Studien INPULSIS 1 und INPULSIS 2	238
Tabelle 4-78: Subgruppenergebnisse für <i>SGRQ-Responder</i> und <i>SGRQ-I-Responder</i> zu Woche 52 basierend auf der gepoolten Analyse der INPULSIS Studien	239
Tabelle 4-79: Subgruppenergebnisse für die Veränderung des CASA-Q Husten-Symptom Scores und die Veränderung auf der EQ-5D-VAS – basierend auf der gepoolten Analyse der INPULSIS Studien	242
Tabelle 4-80: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zu unerwünschten Ereignissen – basierend auf der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187	244

Tabelle 4-81: Subgruppenergebnisse für die <i>UE, die zum Therapieabbruch führten</i> für die gepoolte Auswertung der INPULSIS Studien und der Studie 1199.187	245
Tabelle 4-82: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zu <i>UE von besonderem Interesse</i>	247
Tabelle 4-83: Subgruppenergebnisse für die <i>UE von besonderem Interesse</i> für die gepoolte Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187	248
Tabelle 4-84: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	260
Tabelle 4-85: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	260
Tabelle 4-86: Operationalisierung von <Endpunkt xxx>.....	261
Tabelle 4-87: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	261
Tabelle 4-88: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	262
Tabelle 4-89: Verzerrungsaspekte auf Studienebene – nicht-randomisierte vergleichende Interventionsstudien	264
Tabelle 4-90: Operationalisierung von <Endpunkt xxx>.....	264
Tabelle 4-91: Verzerrungsaspekte für <Endpunkt xxx> – nicht-randomisierte vergleichende Studien	265
Tabelle 4-92: Operationalisierung von <Endpunkt xxx> - weitere Untersuchungen	267
Tabelle 4-93: Bewertung des Zusatznutzens mit Erläuterung - Mortalität	273
Tabelle 4-94: Bewertung des Zusatznutzens mit Erläuterung - Morbidität	275
Tabelle 4-95: Bewertung des Zusatznutzens mit Erläuterung - gesundheitsbezogene Lebensqualität	277
Tabelle 4-96: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	281
Tabelle 4-97: Liste der eingeschlossenen Studien	292
Tabelle 4-98: Übersicht ausgeschlossene Studienregistereinträge.....	314
Tabelle 4-99: Ausgeschlossene Studienregistereinträge	314
Tabelle 4-100 (Anhang): Studiendesign und -methodik für Studie INPULSIS 1.....	344
Tabelle 4-101 (Anhang): Studiendesign und -methodik für Studie INPULSIS 2.....	354
Tabelle 4-102 (Anhang): Studiendesign und -methodik für Studie 1199.187	364
Tabelle 4-103 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INPULSIS 1	378
Tabelle 4-104 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INPULSIS 2	396
Tabelle 4-105 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 1199.187	414

Abbildungsverzeichnis**Seite**

Abbildung 1: Studiendesign und für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte der INPULSIS Studien und Studie 1199.187	20
Abbildung 2: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zum Tod, unabhängig von der Ursache</i> für die gepoolte Analyse der INPULSIS Studien und der Studie 1199.187 (kontrollierte Phase)	21
Abbildung 3: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zum respiratorisch bedingten Tod</i> für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	22
Abbildung 4: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten adjudizierten akuten Exazerbation</i> für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	24
Abbildung 5: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten akuten Exazerbation</i> für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	25
Abbildung 6: Studiendesign und für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte der INPULSIS Studien und Studie 1199.187	69
Abbildung 7: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	81
Abbildung 8: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zum Tod, unabhängig von der Ursache</i> für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	112
Abbildung 9: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zum respiratorisch bedingten Tod</i> für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	113
Abbildung 10: Forest-Plot zur Meta-Analyse für Endpunkt <i>Zeit bis zum Tod, unabhängig von der Ursache</i> - RCT mit dem zu bewertenden Arzneimittel	114
Abbildung 11: Forest-Plot zur Meta-Analyse für Endpunkt <i>Zeit bis zum respiratorisch bedingten Tod</i> - RCT mit dem zu bewertenden Arzneimittel	115
Abbildung 12: Mittlere <i>FVC-Abnahme</i> (Standardfehler des Mittelwertes, SEM) gegenüber dem Ausgangswert (ml) im Zeitverlauf - aus RCT mit dem zu bewertenden Arzneimittel (gepoolte Analyse der INPULSIS Studien und Studie 1199.187 [kontrollierte Phase])	122
Abbildung 13: Forest-Plot zur Meta-Analyse für Endpunkt <i>jährliche FVC-Abnahme</i> - RCT mit dem zu bewertenden Arzneimittel	123
Abbildung 14: Forest-Plot zur Meta-Analyse für Endpunkt <i>FVC-Responder zu Woche 52 (Schwellenwert 10 Prozentpunkte)</i> - RCT mit dem zu bewertenden Arzneimittel	127
Abbildung 15: Forest-Plot zur Meta-Analyse für Endpunkt <i>FVC-Responder zu Woche 52 (Schwellenwert 5 Prozentpunkte)</i> - RCT mit dem zu bewertenden Arzneimittel	128
Abbildung 16: Studiendesign der INPULSIS Studien und der Studie INPULSIS-ON	129

Abbildung 17: <i>Jährliche Abnahme der FVC (ml)</i> in den gepoolten INPULSIS Studien über 52 Wochen (links) und in der Extensionsstudie INPULSIS-ON über 192 Wochen (rechts). Fehlerbalken stellen Standardfehler dar.....	130
Abbildung 18: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten adjudizierten akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel; gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	138
Abbildung 19: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel; gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	139
Abbildung 20: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel (Sensitivitätsanalyse); gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	141
Abbildung 21: Kaplan-Meier-Kurven für Endpunkt <i>Zeit bis zur ersten schwerwiegenden akuten Exazerbation</i> aus RCT mit dem zu bewertenden Arzneimittel (Sensitivitätsanalyse); gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)	142
Abbildung 22: Forest-Plot zur Meta-Analyse für Endpunkt <i>Zeit bis zur ersten adjudizierten akuten Exazerbation</i> - RCT mit dem zu bewertenden Arzneimittel	143
Abbildung 23: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Zeit bis zur ersten akuten Exazerbation</i> - RCT mit dem zu bewertenden Arzneimittel.....	144
Abbildung 24: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung des SGRQ-Gesamtscores</i> - RCT mit dem zu bewertenden Arzneimittel.....	179
Abbildung 25: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung des SGRQ-I-Gesamtscores</i> - RCT mit dem zu bewertenden Arzneimittel.....	180
Abbildung 26: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung des SOBQ-Gesamtscores</i> - RCT mit dem zu bewertenden Arzneimittel.....	181
Abbildung 27: Forest-Plot zur Meta-Analyse für den Endpunkt <i>SGRQ-(≤ 4 Punkte)-Responder zu Woche 52</i> - RCT mit dem zu bewertenden Arzneimittel.....	182
Abbildung 28: Forest-Plot zur Meta-Analyse für den Endpunkt <i>SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 52</i> - RCT mit dem zu bewertenden Arzneimittel.....	183
Abbildung 29: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung des CASA-Q Husten-Symptom-Scores</i> - RCT mit dem zu bewertenden Arzneimittel.....	184
Abbildung 30: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung des CASA-Q Husten-Impact-Scores</i> - RCT mit dem zu bewertenden Arzneimittel	185
Abbildung 31: Forest-Plot zur Meta-Analyse für den Endpunkt <i>PGI-C-Responder zu Woche 52</i> - RCT mit dem zu bewertenden Arzneimittel	186
Abbildung 32: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Veränderung EQ-5D-VAS</i> - RCT mit dem zu bewertenden Arzneimittel	187
Abbildung 33: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Gesamtzahl UE</i> - RCT mit dem zu bewertenden Arzneimittel.....	195

Abbildung 34: Forest-Plot zur Meta-Analyse für den Endpunkt <i>Gesamtzahl SUE</i> - RCT mit dem zu bewertenden Arzneimittel	201
Abbildung 35: Forest-Plot zur Meta-Analyse für den Endpunkt <i>UE, die zum Therapieabbruch führten</i> - RCT mit dem zu bewertenden Arzneimittel.....	207
Abbildung 36: Ergebnisse zur FVC-Abnahme und Gesamtmortalität der klinischen Studien zu Nintedanib und Pirfenidon	285
Abbildung 37: Grafische Darstellung der berichteten Behandlungsdifferenz in der Abnahme der FVC (ml) zwischen Vergleichstherapie und aktiver Behandlung und dem HR für die Zeit bis zum Tod.	287
Abbildung 38: 4-Jahres-Überleben in Abhängigkeit vom Ausmaß der FVC-Abnahme in % des Sollwertes über sechs Monate bei Patienten mit IPF (n=84).....	289
Abbildung 39: Patientenfluss INPULSIS 1 Studie	353
Abbildung 40: Patientenfluss INPULSIS 2 Studie	364
Abbildung 41: Patientenfluss Studie 1199.187	376

Abkürzungsverzeichnis

Abkürzung	Bedeutung
6 MWT	6-Minute-Walk-Test (6-Minuten-Gehtest)
AE	Adverse Event (Unerwünschtes Ereignis)
ALAT	Latin American Thoracic Association
ALT	Alanin-Aminotransferase
AP	Alkalische Phosphatase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Kovarianzanalyse
AST	Aspartat-Aminotransferase
ATS	American Thoracic Society
BI	Boehringer Ingelheim Pharma GmbH & Co. KG
BMI	Body-Mass-Index
BSC	Best Supportive Care
CASA-Q	Cough and Sputum Assessment Questionnaire
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease (Chronisch obstruktive Lungenerkrankung)
CRF	Case Report Form
CTP	Clinical Trial Protocol (Studienprotokoll)
DGP	Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DL _{CO}	Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität)
EC	European Commission (Europäische Kommission)
eCRF	Electronic Case Report Form
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	European Medicines Agency (Europäische Arzneimittel-Agentur)
EQ-5D	European Quality of Life-5 Dimensions

ERS	European Respiratory Society
EU	Europäische Union
EU-CTR	European Union Clinical Trials Register
FDA	U. S. Food and Drug Administration
FEM	Fixed Effects Model (Modell mit festen Effekten)
FEV ₁	Forced Expiratory Volume in one Second (Forcierte expiratorische Einsekunden-Kapazität)
FVC	Forced Vital Capacity (Forcierte Vitalkapazität)
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
GFR	Glomeruläre Filtrationsrate
GKV	Gesetzliche Krankenversicherung
HR	Hazard Ratio
HRCT	High Resolution Computed Tomography (Hochauflösende Computer-Tomografie)
HRQoL	Health-Related Quality of Life (Gesundheitsbezogene Lebensqualität)
IAC	Independent Adjudication Committee (Unabhängiges Bewertungskomitee)
ICTRP	International Clinical Trials Registry Platform
INR	International Normalized Ratio
IPD	Individual Participant Data (Individuelle Patientendaten)
IPF	Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interactive Response Technology System
ITT	Intention-to-Treat
IXRS	Interaktives Telefon/Internet-Response-System
JRS	Japanese Respiratory Society
k. A.	keine Angabe
KI	Konfidenzintervall
MCID	Minimal Clinically Important Difference (Minimaler Schwellenwert für eine klinisch relevante Verbesserung)

MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Effect Model for Repeated Measures
MTC	Mixed Treatment Comparison
MW	Mittelwert
MWD	Mittelwertdifferenz
NAC	N-Acetylcystein
n. b.	nicht berichtet
n. e.	nicht erhoben
NICE	National Institute for Health and Care Excellence
OC	Observed Case (Auswertung nur der tatsächlich beobachteten Werte)
OR	Odds Ratio
PF-ILD	Progressive Fibrosing Interstitial Lung Disease (Progressive Fibrotisierende Interstitielle Lungenerkrankung)
PGI-C	Patient's Global Impression of Change
PRO	Patient Reported Outcome (Patientenberichteter Endpunkt)
PT	Preferred Term
PTT	Partielle Thromboplastinzeit
QLF	Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore)
RCT	Randomized Controlled Trial (randomisierte kontrollierte Studie)
RD	Risk Difference (Absolute Risikodifferenz)
RR	Risk Ratio (Relatives Risiko)
SAP	Statistischer Analyseplan
SD	Standard Deviation (Standardabweichung)
SE	Standard Error (Standardfehler)
SEM	Standard Error of the Mean (Standardfehler des Mittelwertes)
SGB	Sozialgesetzbuch
SGRQ	St. George's Respiratory Questionnaire
SGRQ-I	St. George's Respiratory Questionnaire IPF-specific version
SMQ	Standardised MedDRA Queries

SOBQ	Shortness of Breath Questionnaire
SOC	System Organ Class (Systemorganklasse)
SpO ₂	Sauerstoffsättigung
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UCSD-SOBQ	University of California San Diego Shortness of Breath Questionnaire
UE	Unerwünschtes Ereignis
UIP	Usual Interstitial Pneumonia
ULN	Upper Limits of Normal (Oberer Normgrenzwert)
VAS	Visual Analog Scale
VEGFR	Vascular Endothelial Growth Factor Receptor (Vaskulärer endothelialer Wachstumsfaktorrezeptor)
VerfO	Verfahrensordnung
vs.	versus
WHO	World Health Organization (Weltgesundheitsorganisation)
ZNS	Zentrales Nervensystem
zVT	zweckmäßige Vergleichstherapie

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 Codierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Für Arzneimittel, die zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 zugelassen sind, müssen gemäß §35b SGB V keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen (ZN) vorgelegt werden, solange der Umsatz des Arzneimittels zu Lasten der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten zwölf Monaten einen Betrag von 50 Millionen Euro nicht übersteigt. Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, müssen jedoch vorgelegt werden. In diesem Fall müssen in Modul 4 nur die Abschnitte 4.2 und 4.4 ausgefüllt werden. Weitere Abschnitte von Modul 4 können zur strukturierten Darstellung der Ergebnisse ausgefüllt werden. In dem vorliegenden Dossier wurden daher zusätzlich weitere Abschnitte von Modul 4 entsprechend den Vorgaben der Dokumentenvorlage bearbeitet.

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.

Fragestellung

Ausgewiesen als Arzneimittel zur Behandlung seltener Leiden erhielt Nintedanib (Ofev®) am 15.01.2015 von der Europäischen Kommission (EC) die Zulassung im Anwendungsgebiet der Idiopathischen Lungenfibrose (IPF) bei erwachsenen Patienten. In der Folge wurde im Beschluss zum Nutzenbewertungsverfahren vom 03.09.2015 für Nintedanib (im Folgenden auch als Erstbewertung bezeichnet) durch den Gemeinsamen Bundesausschuss (G-BA) ein geringer Zusatznutzen vergeben (Gemeinsamer Bundesausschuss, 2015a).

Nach § 35a Abs. 1 Satz 10 SGB V (Sozialgesetzbuch) hat der pharmazeutische Unternehmer bei Überschreiten eines Umsatzes von 50 Mio. Euro des Arzneimittels zur Behandlung seltener Leiden innerhalb der letzten 12 Kalendermonate erneut ein Dossier zur Nutzenbewertung beim G-BA einzureichen. Eine entsprechende Aufforderung durch den G-BA an die Boehringer Ingelheim Pharma GmbH & Co. KG (BI) zur erneuten Einreichung eines Dossiers erging am 11.01.2019.

Ziel des vorliegenden Dossiers ist die erneute Bewertung des medizinischen Zusatznutzens im Sinne des § 35a SGB V von Nintedanib bei der Behandlung von erwachsenen Patienten mit IPF. Diese erfolgt gegenüber der vom G-BA festgelegten zweckmäßigen Vergleichstherapie (zVT) (Gemeinsamer Bundesausschuss, 2016b). Hierzu wählt BI aus den vom G-BA benannten Therapiealternativen, Best Supportive Care (BSC) oder Pirfenidon, die zVT BSC.

Ein-/Ausschlusskriterien für Studien

Die Kriterien zum Ein- und Ausschluss von Studien zur Nutzenbewertung von Nintedanib im Anwendungsgebiet IPF werden in Tabelle 4-1 zusammengefasst.

Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für die Nutzenbewertung des Arzneimittels Nintedanib

	Einschlusskriterien	Ausschlusskriterien	Aus-schluss-grund
Patientenpopulation	Erwachsene Patienten mit idiopathischer Lungenfibrose (IPF)	Patienten mit ausschließlich abweichenden Erkrankungen Kinder und Jugendliche <18 Jahre	A1
Intervention	Das zu bewertende Arzneimittel ist Nintedanib (Ofev®). In die Nutzenbewertung sind Studien einzuschließen, in denen Nintedanib gemäß Angaben in der Fachinformation gegeben wurde.	Studien ohne Behandlungsarm mit Nintedanib Studienarme mit anderen als den zugelassenen Therapie- und Dosierungsschemata für Nintedanib	A2
Vergleichstherapie	• BSC • Placebo^a • Pirfenidon	Abweichende Vergleichsarme bzw. Komparatoren	A3

	Einschlusskriterien	Ausschlusskriterien	Aus-schluss-grund
Endpunkte	Mindestens ein patientenrelevanter Endpunkt aus den Dimensionen • Mortalität • Morbidität • HRQoL/Gesundheitszustand • Nebenwirkungen	Kein relevanter Endpunkt aus den Dimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität oder Nebenwirkungen berichtet	A4
Studientypen	Randomisierte kontrollierte Studien (RCT)	Nicht-randomisierte Studien	A5
Studiendauer	Studien mit einer Behandlungsdauer von ≥ 24 Wochen	Studien mit einer Behandlungsdauer von < 24 Wochen	A6
Publikationstyp	Vollpublikationen, Studienberichte oder ausführliche Ergebnisberichte aus Studienregistern	Andere Publikationstypen Mehrfachpublikationen Studienregistereinträge oder Studienergebnisdatenbank-einträge ohne verfügbare Ergebnisse	A7
a: In einem Setting einer unheilbaren und potenziell tödlichen Krankheit wie der IPF, welche auch mit einer entsprechend stark ausgeprägten Symptomatik einhergeht, ist davon auszugehen, dass auch Patienten, die in einen Placebo-Arm randomisiert wurden, eine umfangreiche Therapie zur Linderung der Symptomatik zur Verfügung steht. Die Umsetzung einer BSC entsprechenden Therapie auch innerhalb von Placebo-Armen ist daher anzunehmen und wird für entsprechende Studien geprüft. BSC: Best Supportive Care; HRQoL: Health-Related Quality of Life (Gesundheitsbezogene Lebensqualität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)			

In die Bewertung eingeschlossene Studien

Das Ausmaß des Zusatznutzens für Nintedanib in der Indikation IPF wird im vorliegenden Dossier auf der Grundlage der randomisierten kontrollierten Phase III-Studien INPULSIS 1 und INPULSIS 2, welche bereits zur Erstbewertung herangezogenen wurden, sowie auf Basis einer weiteren randomisierten kontrollierten doppelblinden Phase IIb-Studie 1199.187 (kontrollierte Phase) bestimmt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2017).

In den beiden Zwillingsstudien INPULSIS 1 und 2 mit identischem Studiendesign betrug die Behandlungsphase 52 Wochen. Für die Studie 1199.187 war ursprünglich ebenfalls eine kontrollierte doppelblinde Behandlungsphase von 52 Wochen vorgesehen. Aufgrund ethischer Bedenken in Bezug auf das Weiterführen einer Placebo-Kontrolle im Anwendungsgebiet einer potenziell tödlich verlaufenden Erkrankung mit schwerer Symptomatik auch nach Marktzulassung, wurde die kontrollierte doppelblinde Behandlungsphase der Studie im Rahmen des globalen Amendments 1 auf 24 Wochen verkürzt. Daher wurde im Anschluss an die kontrollierte doppelblinde Phase eine einarmige offene Phase zur Weiterbehandlung der

Patienten aus beiden Studienarmen mit Nintedanib eingeführt, wobei eine Gesamtbehandlungsdauer von bis zu 78 Wochen festgelegt wurde (kontrollierte Phase *plus* einarmige Phase). Da diese Verkürzung mit einer Reduktion der Stichprobenzahl einherging, wurden alle Studienendpunkte explorativ ausgewertet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Der medizinische Nutzen von Nintedanib im Anwendungsgebiet IPF wurde mit der Zulassung durch die Europäische Arzneimittel-Agentur (EMA) bestätigt. Nachfolgend werden die Ergebnisse zum Ausmaß des medizinischen Zusatznutzens von Nintedanib bei erwachsenen IPF-Patienten gegenüber der vom G-BA benannten zVT BSC beschrieben.

Mit den pivotalen Zulassungsstudien INPULSIS 1 und INPULSIS 2 sowie der Phase IIIB-Studie 1199.187 liegen drei doppelblinde randomisierte kontrollierte Studien (RCT) vor, sodass die zur Bewertung des Zusatznutzens herangezogene Evidenz, die Meta-Analyse dieser Studien, der höchsten Evidenzstufe entspricht (Evidenzstufe Ia).

Die Phase III-Studien INPULSIS 1 und INPULSIS 2 sind analog aufgebaut und daher für eine gepoolte Auswertung auf Patientenebene geeignet. Auch wenn die Studiendauer unterschiedlich ist, ist Studie 1199.187 den INPULSIS Studien in Bezug auf die verwendeten Einschlusskriterien, Design-Aspekte sowie die Operationalisierung von Endpunkten hinreichend ähnlich, sodass eine meta-analytische Zusammenfassung aller drei Studien sinnvoll ist. Zudem wiesen auch die Ergebnisse der statistischen Tests auf Vorliegen von Heterogenität für alle Endpunkte auf keine heterogene Datenlage hin, sodass die meta-analytische Zusammenfassung sowohl aus medizinischer als auch aus methodischer und statistischer Sicht im Hinblick auf die Fragestellung zielführend ist.

Da sich die Studie 1199.187 hinsichtlich der Studiendauer von den INPULSIS Studien unterscheidet (Kürzung der kontrollierten doppelblinden Beobachtungsphase auf einen Zeitraum von 24 Wochen mit anschließender Weiterbehandlung der Patienten aus beiden Studienarmen mit Nintedanib gemäß Amendment 1), wurde bei gepoolten Analysen wie im Folgenden beschrieben vorgegangen:

- Für Endpunkte, die durch ihre Operationalisierung - insbesondere durch die Verwendung des statistischen Modells zur Auswertung - unterschiedliche Beobachtungsdauern einbeziehen und berücksichtigen, wurden alle drei Studien meta-analytisch zusammengefasst. Hierunter fallen z. B. Endpunkte, die über die Zeit bis zum ersten Ereignis oder annualisierte Veränderungen operationalisiert wurden, sowie alle Sicherheitsendpunkte. Hier wurde für jede Einzelstudie die gesamte kontrollierte doppelblinde Beobachtungsphase berücksichtigt.
- Für Endpunkte, die sich auf feste Auswertungszeitpunkte beziehen und als Veränderung zu Woche 52 bzw. zu Woche 24 operationalisiert wurden, wurde für den Zeitpunkt zu Woche 52 auf Basis der beiden INPULSIS Studien (INPULSIS [gepoolt]) und für den Zeitpunkt zu Woche 24 über alle drei Studien gepoolt (INPULSIS [gepoolt] + Studie 1199.187).

Das Studiendesign der INPULSIS Studien und der Studie 1199.187 sowie die für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte mit den jeweils herangezogenen Studienpools ist in Abbildung 1 dargestellt.

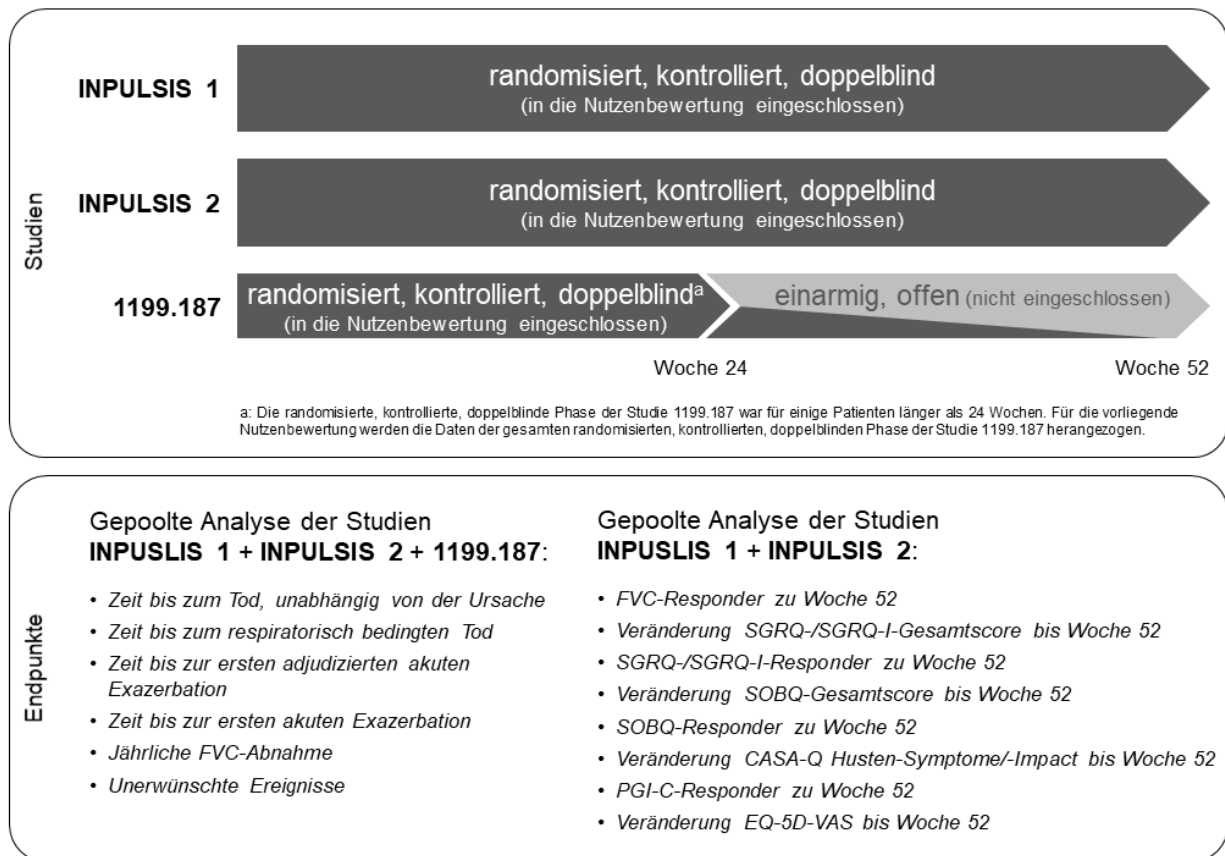


Abbildung 1: Studiendesign und für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte der INPULSIS Studien und Studie 1199.187

Auf Basis der gepoolten Analyse wird eine wesentlich höhere Ergebnissicherheit im Vergleich zu den einzelnen Studien erreicht. Deshalb wird in der Ergebnisdarstellung und bei der Bewertung des Ausmaßes des Zusatznutzens auf Ergebnisse der gepoolten Auswertung fokussiert. Im Setting einer chronischen Erkrankung sind Analysen, die einen möglichst langen Beobachtungszeitraum abdecken, zu bevorzugen. Daher werden zur Ableitung des Zusatznutzens primär die Ergebnisse der gepoolten Analysen mit der längstmöglichen Studiendauer herangezogen.

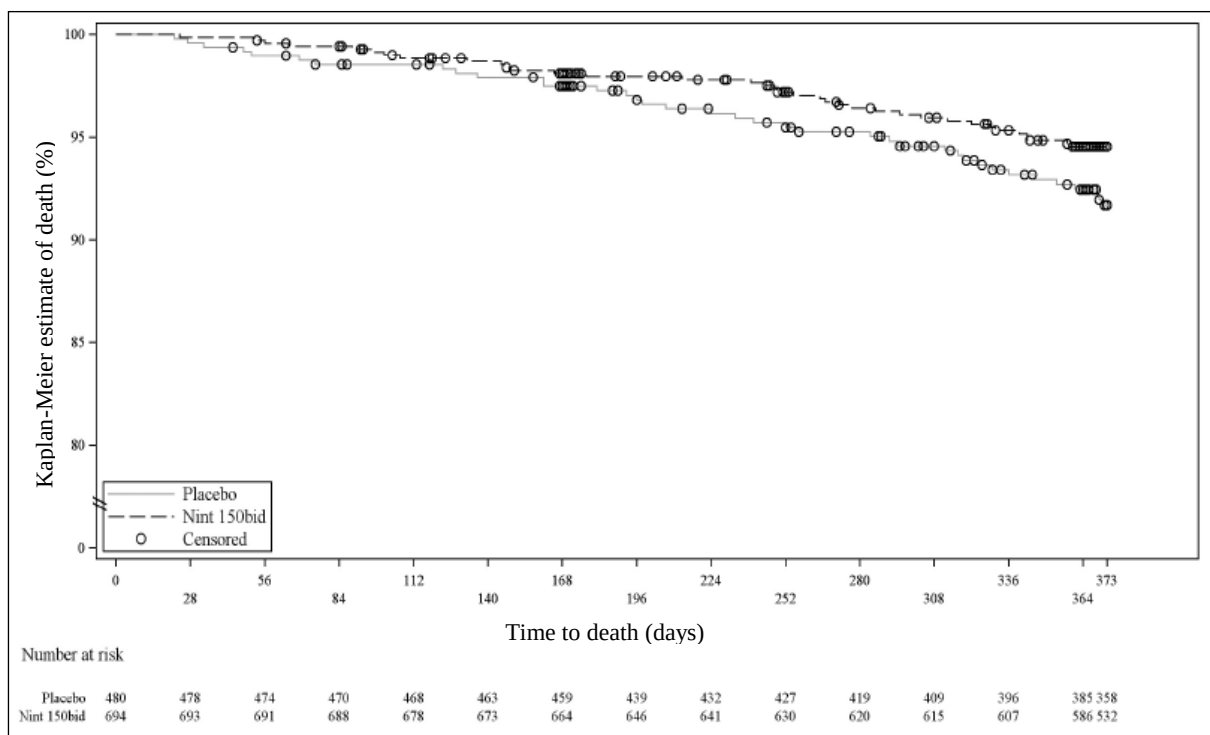
Mortalität

Die Mortalität wurde in diesem Dossier durch die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* und *Zeit bis zum respiratorisch bedingten Tod* sowie durch die *jährliche Abnahme der forcierten Vitalkapazität (FVC)* in ml und durch den Anteil *FVC-Responder* (Abnahme weniger als 10% und 5% des Sollwertes) operationalisiert und hinsichtlich des medizinischen Zusatznutzens bewertet.

Die Meta-Analyse über die drei relevanten Studien zeigt eine höhere Überlebenswahrscheinlichkeit unter Nintedanib als im Vergleichsarm. Zudem waren die *jährliche FVC-Abnahme (in ml)* geringer und der Anteil der FVC-Responder zu Woche 52 unter Nintedanib höher als im Vergleichsarm; der Unterschied ist jeweils statistisch signifikant. Als valides Surrogat für die Mortalität bedeutet die geringere Abnahme der FVC eine Reduktion der Sterbewahrscheinlichkeit bei IPF-Patienten unter der Therapie mit Nintedanib im Vergleich zur zVT.

Zeit bis zum Tod

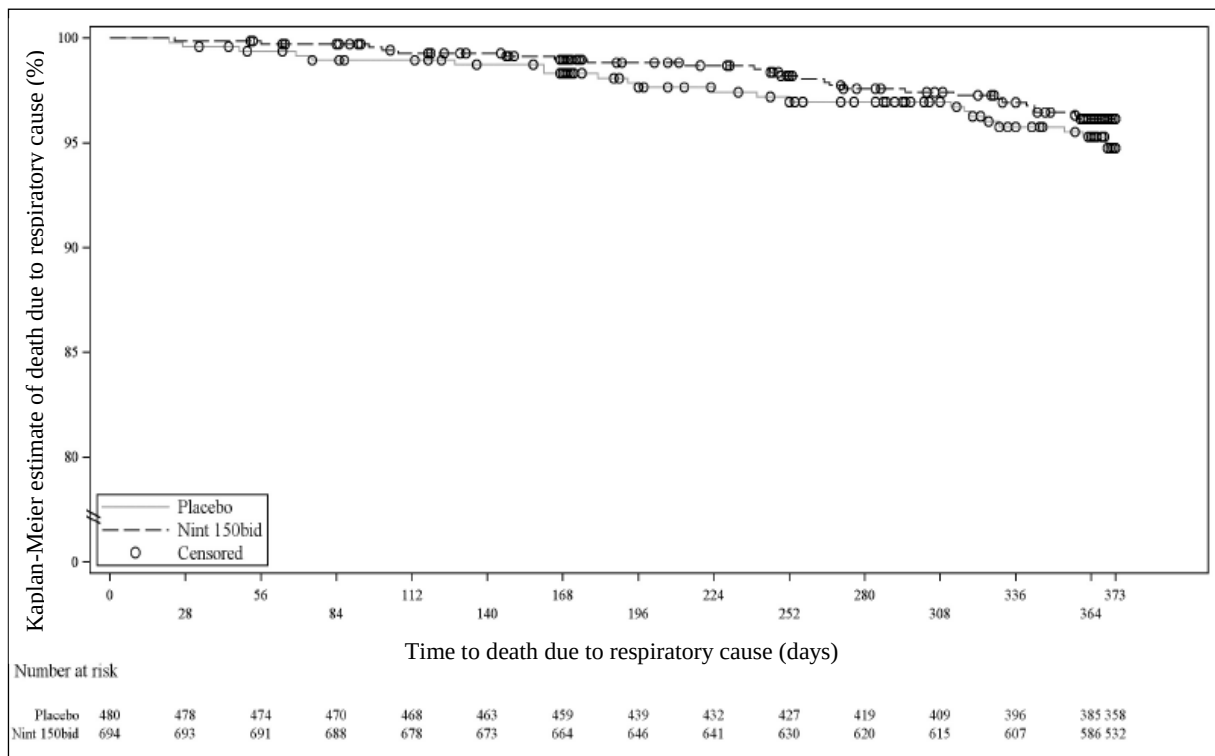
Die *Zeit bis zum Tod, unabhängig von der Ursache* wird als Überlebenswahrscheinlichkeit anhand des Kaplan-Meier-Schätzers geschätzt. Basierend auf einer Cox-Regression war das Sterberisiko über 52 Wochen in den gepoolten INPULSIS Studien und der Studie 1199.187 unter Nintedanib geringer als im Vergleichsarm (Hazard Ratio [HR] [95%-Konfidenzintervall (KI)]: 0,65 [0,41; 1,03]; $p=0,0775$) (Abbildung 2). Die Kaplan-Meier-Kurven zeigen, dass mit zunehmender Beobachtungszeit der Unterschied im Anteil der verstorbenen Studienteilnehmer zwischen den Behandlungsgruppen Nintedanib und zVT größer wird.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 2: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zum Tod, unabhängig von der Ursache* für die gepoolte Analyse der INPULSIS Studien und der Studie 1199.187 (kontrollierte Phase)

Ein numerisch vergleichbarer Trend zugunsten einer Behandlung mit Nintedanib wurde auch für die *Zeit bis zum respiratorisch bedingten Tod* beobachtet (HR [95%-KI]: 0,71 [0,40; 1,26]; $p=0,2960$) (Abbildung 3).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 3: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zum respiratorisch bedingten Tod* für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

Forcierte Vitalkapazität (FVC)

Die *jährliche FVC-Abnahme (ml)* war in den gepoolten INPULSIS Studien und der Studie 1199.187 unter Nintedanib mit 108,93 ml klinisch relevant und statistisch signifikant geringer als im Vergleichsarm (222,97 ml). Der Unterschied zwischen den beiden Behandlungsgruppen hat eine Mittelwertdifferenz (MWD) von 112,42 ml ([95%-KI]: [79,06; 145,77]; $p < 0,0001$). Die Abnahme der FVC im Behandlungsarm mit Nintedanib fiel somit durchschnittlich um etwa 50% geringer aus als im Vergleichsarm.

Der Anteil der *FVC-Responder zu Woche 52* in den INPULSIS Studien war definiert als Anteil der Studienteilnehmer mit einem absoluten Abfall der FVC (% des Sollwertes) nicht größer als 10% bzw. 5% bis Woche 52.

Die Analyse der *FVC-Responder zu Woche 52* der gepoolten INPULSIS Studien bezogen auf $<10\%$ des Sollwertes zeigt durch eine logistische Regression mit einem Odds Ratio (OR) von 1,577 einen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 1,577 [1,21; 2,05]; $p=0,0007$). Bei einem Schwellenwert von 5% zeigt sich ebenfalls ein statistisch signifikanter Unterschied: Auf Nintedanib sprachen zu Woche 52 gemäß dieser Analyse mehr Studienteilnehmer an als auf die Vergleichstherapie (OR [95%-KI]: 1,835 [1,43; 2,36]; $p < 0,0001$).

Im vorliegenden Dossier wird in Abschnitt 4.5.4 zusätzlich eine auf einem Metaregressionsmodell basierende Ableitung des Therapieeffekts von Nintedanib im Endpunkt FVC auf das Mortalitätsrisiko vorgenommen. Demnach entspricht eine Reduktion der *jährlichen FVC-Abnahme* von 112,42 ml, wie sie über die gepoolten INPULSIS Studien und die Studie 1199.187 beobachtet wurde, einer Reduktion des Mortalitätsrisikos um über 20%.

Morbidität

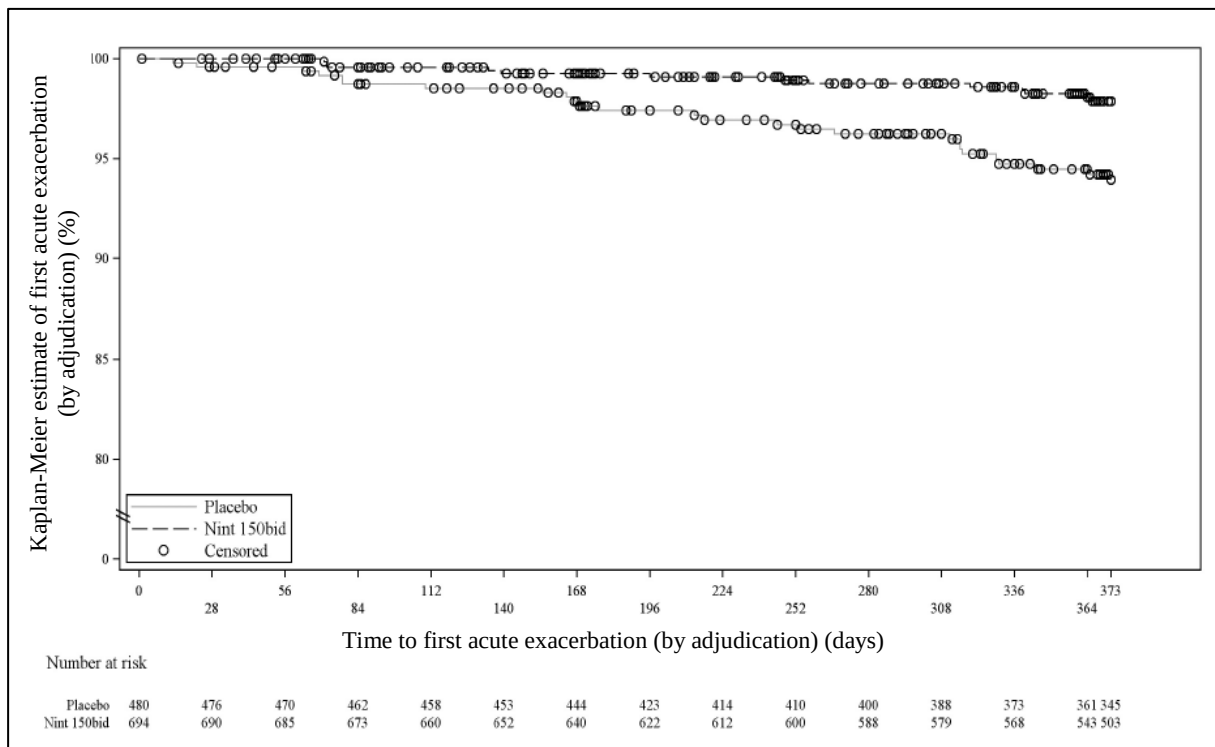
Die Morbidität wurde in diesem Dossier durch die Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* (bewertet durch ein Unabhängiges Bewertungskomitee, IAC), *Zeit bis zur ersten akuten Exazerbation* (bewertet durch den Prüfarzt), *Veränderung des Quantitativen Lungenfibrosescores (QLF) anhand Hochauflösender Computer-Tomografie (HRCT) zu Woche 24 (HRCT-QLF)*, *Veränderung der Gehstrecke im 6-Minuten-Gehtest zu Woche 24 (6 MWT)* sowie *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* operationalisiert und hinsichtlich des Ausmaßes des medizinischen Zusatznutzens bewertet. Endpunkte, welche ausschließlich in der Studie 1199.187 erhoben wurden (*HRCT-QLF*, *6 MWT* und *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung*), werden im vorliegenden Dossier und wie im Studienprotokoll (CTP) vorgesehen rein deskriptiv ausgewertet. Dies ist eine Konsequenz des globalen Amendments 1, im Zuge dessen eine Verkürzung der kontrollierten Phase von 52 Wochen auf 24 Wochen (mit anschließender Weiterbehandlung der Patienten aus beiden Studienarmen mit Nintedanib) sowie eine Reduktion der Patientenzahlen von 275 auf 150 vorgenommen wurde und die Studie in ein exploratives Design überführt wurde.

Die Ergebnisse aller Endpunkte zur Morbidität zeigen einen Vorteil bzw. einen numerischen Trend zugunsten der Nintedanib-Therapie im Vergleich zur zVT. Für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* war der Unterschied statistisch signifikant. Ausschließlich deskriptiv ausgewertete Endpunkte aus der Dimension Morbidität werden in Abschnitt 4.3.1.3.2 dargestellt.

Exazerbation

Adjudizierte akute Exazerbationen traten in den drei gepoolten Studien INPULSIS 1, INPULSIS 2 und der Studie 1199.187 unter Nintedanib seltener auf als unter der Vergleichstherapie (1,9% vs. 5,4%). Die Wahrscheinlichkeit, keine Exazerbation zu erleiden (ausgedrückt durch den Kaplan-Meier-Schätzer), beträgt nach 52 Wochen in der Nintedanib-Gruppe 0,979 und im Vergleichsarm 0,939. Auch das HR [95%-KI] von 0,32 für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* zeigt einen Vorteil zugunsten von Nintedanib (HR [95%-KI]: 0,32 [0,17; 0,63]; p=0,0010). Der Unterschied ist statistisch signifikant. Dies entspricht einer statistisch signifikanten Reduktion des Risikos einer ersten adjudizierten akuten Exazerbation um 68%.

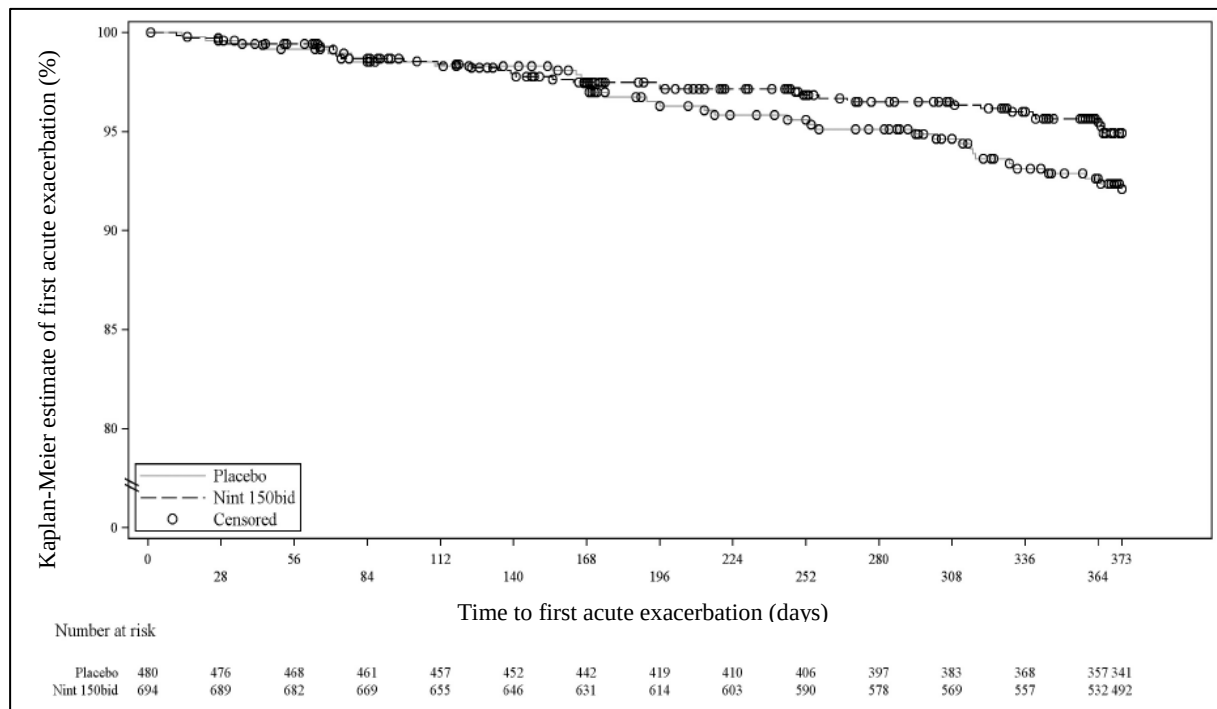
Die Kaplan-Meier-Kurven zeigen deutlich, dass mit zunehmender Beobachtungszeit die Kurven der *Zeit bis zur ersten adjudizierten akuten Exazerbation* zwischen den Behandlungsgruppen Nintedanib und der zVT kontinuierlich auseinandergehen (Abbildung 4).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 4: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zur ersten adjudizierten akuten Exazerbation* für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

Ein numerischer Trend zugunsten von Nintedanib wurde auch für die *Zeit bis zur ersten akuten Exazerbation* in den gepoolten Daten der drei Studien (HR [95%-KI]: 0,63 [0,39; 1,02]; $p=0,0695$) beobachtet (Abbildung 5). Betrachtet man nur Exazerbationen, die als schwerwiegendes unerwünschtes Ereignis (SUE) eingestuft wurden, zeigt sich sowohl für adjudizierte Exazerbationen als auch für Exazerbationsereignisse, die vom Prüfarzt als solche dokumentiert wurden, ein Vorteil zugunsten der Therapie mit Nintedanib (HR [95%-KI]: 0,30 [0,15; 0,62]; $p=0,0015$ (p-Wert aus Wald-Test im adjustierten Cox-Modell: 0,0012) bzw. 0,55 [0,32; 0,96]; $p=0,0545$ (p-Wert aus Wald-Test im adjustierten Cox-Modell: 0,0337)) (Tabelle 4-28).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 5: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zur ersten akuten Exazerbation* für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

Gesundheitsbezogene Lebensqualität

Die krankheitsspezifische Lebensqualität wurde in diesem Dossier durch die Endpunkte *Veränderung des St. George's Respiratory Questionnaire (SGRQ)*, *Veränderung des IPF-spezifischen SGRQ (SGRQ-I)* und –Responderanalysen sowie *Veränderung des Shortness of Breath Questionnaire (SOBQ)* und der *Veränderung des Cough and Sputum Assessment Questionnaire (CASA-Q)* bewertet. Die allgemeine gesundheitsbezogene Lebensqualität bzw. der allgemeine Gesundheitszustand wurde durch die Endpunkte *Patient's Global Impression Change (PGI-C)-Responder* und *Veränderung der European Quality of Life-5 Dimensions Analogskala (EQ-5D-VAS)* dargestellt.

Die Ergebnisse zur Veränderung der Lebensqualität und des Gesundheitszustands basierend auf den gepoolten Daten der INPULSIS Studien zeigen zu Woche 52 einen Vorteil oder einen numerischen Trend zugunsten der Nintedanib-Therapie in Form einer geringeren Verschlechterung der gesundheitsbezogenen Lebensqualität im Vergleich zur zVT. Für die Endpunkte *Veränderung SGRQ-I-Gesamtscore*, *PGI-C-Responder* und *Veränderung der EQ-5D-VAS* zu Woche 52 sind die Unterschiede statistisch signifikant.

Endpunkte zum SGRQ und SGRQ-I

Die gesundheitsbezogene Lebensqualität, ausgedrückt durch die Differenz des Wertes zu Woche 52 minus des Wertes zum Ausgangswert, führt in den gepoolten INPULSIS Studien zu einer Zunahme des *SGRQ-Gesamtscores* in beiden Studienarmen, was einer Verschlechterung

der gesundheitsbezogenen Lebensqualität entspricht (Wertebereich: 0-100; ein hoher Wert entspricht einer hohen Beeinträchtigung). Dabei fällt die Verschlechterung der gesundheitsbezogenen Lebensqualität unter Nintedanib geringer aus als bei Studienteilnehmern im Vergleichsarm (3,53 Punkte vs. 4,96 Punkte). Die *Veränderung des SGRQ-Gesamtscores* zu Woche 52 zeigt somit einen numerischen Trend zugunsten von Nintedanib gegenüber der zVT (MWD [95%-KI]: -1,43 Punkte [-3,09; 0,23]; p=0,0923).

Die dichotomisierte Analyse der *SGRQ-(≤ -4 Punkte)-Responder* (Änderung des SGRQ um ≤ -4 Punkte zu Woche 52), welche den Anteil der Patienten mit einer klinisch relevanten Verbesserung betrachtet, zeigt für die gepoolten Auswertungen, dass unter Nintedanib in etwa gleich viele Studienteilnehmer den Schwellenwert von ≤ -4 Punkten im *SGRQ-Gesamtscore* erreichten wie unter der Vergleichstherapie: Für die gepoolte Analyse der INPULSIS Studien zu Woche 52 ergaben sich 22,88% gegenüber 20,33% Responder (OR [95%-KI]: 1,186 [0,88; 1,61]; p=0,2715).

Die Verschlechterung der gesundheitsbezogenen Lebensqualität vom Ausgangswert bis Woche 52, ausgedrückt durch eine Zunahme des SGRQ-I-Gesamtscores, fällt in der gepoolten Auswertung der INPULSIS Studien unter Nintedanib geringer aus als bei den Studienteilnehmern unter der Vergleichstherapie (3,48 Punkte vs. 5,47 Punkte). Dies entspricht einem statistisch signifikanten Unterschied zugunsten der Studienteilnehmer unter der Behandlung mit Nintedanib im Vergleich zur zVT (MWD [95%-KI]: -1,99 Punkte [-3,71; -0,27]; p=0,0230).

Die Analyse der *SGRQ-I-(≤ -4 Punkte)-Responder* für die gepoolte Auswertung der INPULSIS Studien zu Woche 52 zeigt, dass unter Nintedanib mehr Studienteilnehmer den Schwellenwert von ≤ -4 Punkten im SGRQ-I-Gesamtscore erreichten als Studienteilnehmer unter der Vergleichstherapie (23,51% vs. 19,62% Responder; OR [95%-KI]: 1,284 [0,95; 1,74]; p=0,1091).

Weitere Endpunkte zur krankheitsspezifischen Lebensqualität

Die Veränderung des SOBQ-Scores, der die subjektiv bewertete Dyspnoe der Studienteilnehmer misst (Wertebereich: 0-120; ein hoher Wert entspricht einer hohen Beeinträchtigung), zeigt zu Woche 52 eine geringere Verschlechterung unter Nintedanib im Vergleich zur zVT in der gepoolten Auswertung der INPULSIS Studien (6,71 Punkte vs. 8,40 Punkte; MWD [95%-KI]: -1,69 [-4,09; 0,70]; p=0,1648).

Auch die Ergebnisse zum CASA-Q-Score (Wertebereich: 0-100; ein hoher Wert entspricht einer geringeren Beeinträchtigung) in der gepoolten Auswertung der INPULSIS Studien zeigen für die Veränderung des CASA-Q Husten-Impact bis Woche 52 eine geringere Verschlechterung der krankheitsspezifischen Lebensqualität unter der Behandlung mit Nintedanib im Vergleich zur zVT (Impact: MWD [95%-KI]: 1,73 Punkte [-0,46; 3,92]; p=0,1212; Symptome: MWD [95%-KI]: 0,95 Punkte [-1,49; 3,38]; p=0,4450).

Endpunkte zur allgemeinen Gesundheitsbezogenen Lebensqualität und zum allgemeinen Gesundheitszustand

Der Anteil *PGI-C-Responder*, definiert als Studienteilnehmer, die ihre gesundheitsbezogene Lebensqualität in Woche 52 im Vergleich zum Ausgangswert als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „ohne Veränderung“ einschätzen, ist in der gepoolten Auswertung der INPULSIS Studien unter Nintedanib höher als unter der zVT (61,29% vs. 54,37% Responder). Der Unterschied ist statistisch signifikant (OR [95%-KI]: 1,328 [1,04; 1,70]; $p=0,0254$).

Auch die in den INPULSIS Studien erhobene *Veränderung der EQ-5D-VAS* (Wertebereich: 0-100) vom Ausgangswert bis Woche 52 zeigt in der gepoolten Auswertung eine geringere Verschlechterung des subjektiv empfundenen allgemeinen Gesundheitszustandes unter Nintedanib im Vergleich zur zVT (-2,70 Punkte vs. -6,51 Punkte). Der beobachtete Unterschied ist statistisch signifikant, jedoch von unklarer klinischer Relevanz (adjustierte MWD [95%-KI]: 3,81 [1,78; 5,85]; $p=0,0002$; Hedges' g [95%-KI]: 0,252 [0,12; 0,39]).

Unerwünschte Ereignisse

Im Rahmen dieser Nutzenbewertung wurden *unerwünschte Ereignisse (UE)*, *SUE* sowie *UE, die zum Therapieabbruch führten* berücksichtigt. Zusätzlich wurden *UE von besonderem Interesse* sowie *UE, die in >10% der Studienteilnehmer eines Behandlungsarms auftraten*, betrachtet. Die Darstellung der UE erfolgt auf Grundlage der gepoolten INPULSIS Studien und der Studie 1199.187. In Modul 5 werden die Daten zu den UE auf Einzelstudienbene dargestellt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017; Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

- Alle *UE*: Relatives Risiko (RR) [95%-KI]: 1,07 [1,03; 1,10]; $p<0,001$
- Alle *SUE*: RR [95%-KI]: 1,01 [0,84; 1,21]; $p=0,941$
- *UE, die zum Therapieabbruch führten*: RR [95%-KI]: 1,54 [1,15; 2,04]; $p=0,003$
- *UE von besonderem Interesse: gastrointestinale Störungen, Leberwerterhöhungen, kardiale UE, respiratorische UE und Blutungsereignisse*

Die über die INPULSIS Studien und Studie 1199.187 gepoolten *UE von besonderem Interesse* sind in Tabelle 4-2 zusammengefasst.

Tabelle 4-2: Ergebnisse für *UE von besonderem Interesse* (INPULSIS Studien und Studie 1199.187 gepoolt)

Gastrointestinale Störungen														
Nintedanib vs. Placebo	Diarrhö		Übelkeit		Erbrechen		Appetitlosigkeit		Gewichtsverlust					
RR [95 %-KI]; p-Wert	3,09 [2,56; 3,72]; <0,001		3,02 [2,18; 4,18]; <0,001		4,20 [2,40; 7,35]; <0,001		2,02 [1,34; 3,05]; <0,001		3,27 [1,91; 5,61]; <0,001					
Unerwünschte Leberenzym- und Bilirubinveränderungen														
Nintedanib vs. Placebo	Leberenzymerrhöhung (AST, ALT)					Bilirubinerhöhung		AP-Erhöhung						
	≥3x ULN ^a		≥5x ULN ^a		≥8x ULN	≥1,5x ULN ^a		≥2x ULN		≥1,5x ULN ^a		≥2x ULN		
RR [95 %-KI]; p-Wert	5,99 [2,10; 17,07]; <0,001		9,30 [1,23; 70,53]; 0,008		3,32 [0,39; 28,36]; 0,244		3,60 [1,06; 12,17]; 0,027		1,00 [0,17; 5,94]; 0,997		5,44 [2,13; 13,87]; <0,001		11,29 [1,51; 84,64]; 0,003	
Unerwünschte kardiale Ereignisse														
Nintedanib vs. Placebo	Unerwünschte kardiale Ereignisse (SOC)		Schwerwiegende unerwünschte kardiale Ereignisse (SOC)			Tödliche kardiale Ereignisse		Arterielle thromboembolische Ereignisse		Myokardinfarkt / akuter Myokardinfarkt				
RR [95 %-KI]; p-Wert	0,98 [0,69; 1,40]; 0,916		0,92 [0,55; 1,55]; 0,763			0,33 [0,08; 1,32]; 0,100		3,82 [1,13; 12,85]; 0,020		3,65 [0,81; 16,36]; 0,070				
Unerwünschte respiratorische Ereignisse														
Nintedanib vs. Placebo	Progression der IPF		Husten			Dyspnoe		Bronchitis		Nasopharyngitis				
RR [95 %-KI]; p-Wert	0,69 [0,50; 0,95]; 0,023		0,94 [0,70; 1,27]; 0,683			0,74 [0,52; 1,05]; 0,093		0,97 [0,68; 1,38]; 0,866		0,84 [0,63; 1,12]; 0,249				
Unerwünschte Blutungsereignisse														
Nintedanib vs. Placebo	Blutungsereignisse													
RR [95 %-KI]; p-Wert	1,31 [0,89; 1,94]; 0,171													
a: Kategorien sind kumulativ, d. h. Studienteilnehmer, die in die höheren Kategorien fallen, sind bei den niedrigeren Kategorien mit angegeben. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode.														
ALT: Alanin-Aminotransferase; AP: Alkalische Phosphatase; AST: Aspartat-Aminotransferase; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; RR: Risk Ratio (Relatives Risiko); SOC: System Organ Class (Systemorganklasse); UE: Unerwünschtes Ereignis; ULN: Upper Limits of Normal (Oberer Normgrenzwert)														

Gastrointestinale Störungen und *Unerwünschte Leberenzym- und Bilirubinveränderungen*, wurden in der gepoolten Analyse der INPULSIS Studien und der Studie 1199.187 häufiger unter Nintedanib als im Vergleichsarm dokumentiert. *Unerwünschte kardiale Ereignisse* wurden überwiegend vergleichbar häufig unter Nintedanib und der zVT beobachtet, einzig *Arterielle thromboembolische Ereignisse* traten in der Nintedanib-Gruppe statistisch signifikant

häufiger auf. Aus der Gruppe der *Unerwünschten respiratorischen Ereignisse* ergab sich nur für das Auftreten von *Progression der IPF* ein statistisch signifikanter Unterschied. Dieser fiel zugunsten von Nintedanib aus. Zu *Unerwünschten Blutungsereignissen* fand sich zwischen den Gruppen kein Unterschied.

Schlussfolgerungen zum Ausmaß des therapeutisch bedeutsamen Zusatznutzens

Nintedanib weist gegenüber der zVT BSC in den Nutzendimensionen Mortalität, Morbidität und gesundheitsbezogene Lebensqualität Vorteile auf.

Die Vorteile zugunsten einer Nintedanib-Behandlung hinsichtlich Mortalität begründen sich in der statistisch signifikanten Reduktion der *jährlichen FVC-Abnahme* als valides Surrogat für eine Reduktion des Mortalitätsrisikos, dem numerischen Vorteil in der *Zeit bis zum Tod, unabhängig von der Ursache* sowie dem statistisch signifikanten Vorteil im Mortalitäts-assoziierten Endpunkt *adjudizierte akute Exazerbationen*. In der Gesamtbetrachtung der Ergebnisse zur Nutzendimension Mortalität ist für Nintedanib ein Beleg für einen beträchtlichen Zusatznutzen gegenüber der zVT abzuleiten.

Akute Exazerbationen sind akut lebensbedrohliche Ereignisse, die maßgeblichen Einfluss auf die Progression der IPF haben, zu einer weiteren rapiden Verschlechterung des Gesundheitszustands der Betroffenen führen sowie mit einer hohen Sterblichkeit assoziiert sind. Die Behandlung mit Nintedanib führte im Vergleich zur zVT zu einer statistisch signifikanten Reduktion der *Zeit bis zur ersten adjudizierten akuten Exazerbation* entsprechend einer Risikoreduktion um 68%. In der Nutzendimension Morbidität ist damit der gezeigte therapierelevante Vorteil unter Nintedanib gegenüber der zVT als Beleg für einen beträchtlichen Zusatznutzen einzustufen.

Die positiven Effekte in der Nutzendimension gesundheitsbezogene Lebensqualität, welche sich in einem signifikanten Vorteil in der *Veränderung des SGRQ-I-Gesamtscores* und der *Veränderung der EQ-5D-VAS* sowie in einem signifikant höheren Anteil der *PGI-C-Responder* begründen, sind als Hinweis für einen geringen Zusatznutzen für Nintedanib im Vergleich zur zVT einzustufen.

Bezüglich der *UE* liegt für die Endpunktkategorien *UE, die zum Therapieabbruch führten* sowie für *gastrointestinale Ereignisse, Leberenzymerrhöhung* und *Bilirubinerhöhung* sowie *arterielle thromboembolische Ereignisse* jeweils ein Nachteil vor.

In der Gesamtschau der dargestellten Ergebnisse und unter Berücksichtigung der Aussagekraft der vorgebrachten Nachweise zeigt sich, dass Nintedanib für Patienten mit IPF eine gegenüber der zVT nicht erreichte, bedeutsame Verbesserung des therapierelevanten Nutzens im Sinne einer Verlängerung des Überlebens, einer Verbesserung schwerwiegender Symptome sowie einer geringeren Verschlechterung der gesundheitsbezogenen Lebensqualität bedeutet. Unter Berücksichtigung des Nachteils hinsichtlich unerwünschter Ereignisse geht in der Gesamtbilanz aus den dargestellten Daten hervor, dass der therapierelevante Nutzen überwiegt.

Mit drei meta-analytisch zusammengefassten RCT im Anwendungsgebiet einer seltenen Erkrankung liegen für die vorliegende Nutzenbewertung insgesamt Ergebnisse der höchsten Evidenzstufe vor. Insbesondere die gezeigte Risikoreduktion für Exazerbationen sowie die vorliegenden Vorteile in den Mortalitätsendpunkten verdeutlichen, dass eine Behandlung mit Nintedanib für IPF-Patienten eine Verbesserung schwerwiegender Symptomatik bedeutet und weisen auf eine Verlängerung des Überlebens hin.

Damit ergibt sich insgesamt ein **Hinweis für einen beträchtlichen Zusatznutzen** für Nintedanib gegenüber der zVT für Patienten mit IPF (Tabelle 4-3).

Tabelle 4-3: Ergebnisse zum medizinischen Zusatznutzen von Nintedanib

Endpunkt	Nintedanib vs. BSC Effektschätzer [95%-KI]; p-Wert	Ausmaß des Zusatznutzens
Mortalität		
Zeit bis zum Tod, unabhängig von der Ursache ^a	HR: 0,65 [0,41; 1,03]; 0,0775	beträchtlich Wahrscheinlichkeit: Beleg
Zeit bis zum respiratorisch bedingten Tod ^a	HR: 0,71 [0,40; 1,26]; 0,2960	
FVC (als valider Surrogatendpunkt für Mortalität)		
Jährliche FVC-Abnahme (in ml) ^a	MWD: 112,42 [79,06; 145,77]; <0,0001	
FVC-Responder (10% des Sollwertes) ^{b,c}	OR: 1,577 [1,21; 2,05]; 0,0007	
FVC-Responder (5% des Sollwertes) ^{b,c}	OR: 1,835 [1,43; 2,36]; <0,0001	
Zeit bis zur ersten (adjudizierten) akuten Exazerbation (als Mortalitäts-assoziiierter Endpunkt) ⇒ Ergebnisse siehe Morbidität		
Morbidität		
Zeit bis zur ersten adjudizierten akuten Exazerbation ^{a,d}	HR: 0,32 [0,17; 0,63]; 0,0010	beträchtlich Wahrscheinlichkeit: Beleg
Zeit bis zur ersten akuten Exazerbation ^{a,d}	HR: 0,63 [0,39; 1,02]; 0,0695	
Gesundheitsbezogene Lebensqualität		
SGRQ	Veränderung Gesamtscore ^b : MWD: -1,43 [-3,09; 0,23]; 0,0923	gering Wahrscheinlichkeit: Hinweis
	Responder (≤-4 Punkte) ^{b,e} : OR: 1,186 [0,88; 1,61]; 0,2715	
SGRQ-I	Veränderung Gesamtscore ^b : MWD: -1,99 [-3,71; -0,27]; 0,0230	
	Responder (≤-4 Punkte) ^{b,e} : OR: 1,284 [0,95; 1,74]; 0,1091	
SOBQ	Veränderung Gesamtscore ^b : MWD: -1,69 [-4,09; 0,70]; 0,1648	
CASA-Q	Veränderung Husten-Symptome ^b : MWD: 0,95 [-1,49; 3,38]; 0,4450	
	Veränderung Husten-Impact ^b : MWD: 1,73 [-0,46; 3,92]; 0,1212	

Endpunkt	Nintedanib vs. BSC Effektschätzer [95%-KI]; p-Wert	Ausmaß des Zusatznutzens
PGI-C	Responder ^{b,f} : OR: 1,328 [1,04; 1,70]; 0,0254	
EQ-5D-VAS	Veränderung VAS-Score ^b : MWD: 3,81 [1,78; 5,85]; 0,0002 Hedges‘ g: 0,252 [0,12; 0,39]	
Unerwünschte Ereignisse		
Gesamtzahl UE ^a	RR: 1,07 [1,03; 1,10]; <0,001	Geringerer Nutzen
SUE ^a	RR: 1,01 [0,84; 1,21]; 0,941	Zusatznutzen nicht belegt
UE, die zum Therapieabbruch führten ^a	RR: 1,54 [1,15; 2,04]; 0,003	Geringerer Nutzen
UE von besonderem Interesse ^a		
Gastrointestinale Ereignisse	Nachteil	
Leberenzym-/Bilirubinveränderungen	Nachteil	
Arterielle thromboembolische Ereignisse	Nachteil	
Progression der IPF	Vorteil	
Blutungsereignisse	weder Vor- noch Nachteil	
a: Gepoolte Analyse INPULSIS Studien und Studie 1199.187 zu Woche 52 b: Gepoolte Analyse INPULSIS Studien zu Woche 52 c: FVC-Responder sind definiert als Patienten mit einer FVC-Abnahme von ≤10% bzw. 5% des Sollwertes. d: Zeit bis zur ersten adjudizierten akuten Exazerbation: nachträglich durch ein unabhängiges Bewertungskomitee bestätigt; Zeit bis zur ersten akuten Exazerbation: durch den Prüfarzt bewertet e: Für SGRQ-/SGRQ-I sind Responder definiert als Patienten mit einer Verbesserung im Gesamtscore von ≤-4 Punkten. f: Für PGI-C sind Responder definiert als Patienten, die ihre Lebensqualität im Vergleich zum Studienbeginn als unverändert oder besser einschätzten. CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HR: Hazard Ratio; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; MWD: Mittelwertdifferenz; OR: Odds Ratio; PGI-C: Patient´s Global Impression of Change; RR: Risk Ratio (Relatives Risiko); SGRQ: St. George´s Respiratory Questionnaire; SGRQ-I: St. George´s Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visual Analog Scale		

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.

Ausgewiesen als Arzneimittel zur Behandlung seltener Leiden, erhielt Nintedanib als solches am 15.01.2015 von der EC die Zulassung zur Behandlung der IPF (European Commission (EC), 2015). Das Anwendungsgebiet der Zulassung gemäß Fachinformation lautet: Ofev[®] wird angewendet bei Erwachsenen zur Behandlung der idiopathischen Lungenfibrose (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Der medizinische Nutzen und medizinische Zusatznutzen für Arzneimittel zur Behandlung seltener Leiden gilt gemäß § 35a Abs. 1 Satz 10 SGB V bereits durch die Zulassung als belegt. Die Einstufung des Ausmaßes des Zusatznutzens in der Erstbewertung erfolgte auf Basis der zulassungsbegründenden multizentrischen, randomisierten, kontrollierten klinischen Studien INPULSIS 1 und INPULSIS 2 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b). Im Beschluss zum Nutzenbewertungsverfahren vom 03.09.2015 wurde für Nintedanib (Ofev[®]) durch den G-BA ein geringer Zusatznutzen vergeben (Gemeinsamer Bundesausschuss, 2015a).

Nach § 35a Abs. 1 Satz 10 SGB V hat der pharmazeutische Unternehmer bei Überschreiten eines Umsatzes von 50 Mio. Euro des Arzneimittels mit der gesetzlichen Krankenversicherung (GKV) zu Apothekenverkaufspreisen einschließlich Umsatzsteuer innerhalb der letzten 12 Kalendermonate erneut ein Dossier zur Nutzenbewertung beim G-BA einzureichen. Eine vorsorgliche Aufforderung an BI zur Einreichung eines Dossiers durch den G-BA erging am 11.01.2019 (Gemeinsamer Bundesausschuss, 2019).

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Zusatznutzens im Sinne des § 35a SGB V von Nintedanib (Ofev®) bei der Behandlung von erwachsenen Patienten mit IPF gegenüber der vom G-BA festgelegten zVT. Aus den vom G-BA im Beratungsgespräch benannten Therapiealternativen, BSC oder Pirfenidon, wählt BI BSC. Hierzu liegt neben den bereits zur Erstbewertung berücksichtigten Studien INPULSIS 1 und INPULSIS 2 eine weitere randomisierte kontrollierte Phase IIIB-Studie (Studie 1199.187) vor (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017). Wie in Abschnitt 4.2.5.2 beschrieben, wird nur die kontrollierte Phase dieser Studie als nutzenbewertungsrelevant betrachtet, da während der auf diese folgenden offenen Phase die Behandlung der Patienten nur mit Nintedanib als Studienmedikation erfolgte. Entsprechend werden nur Ergebnisse der kontrollierten Phase im Dossier dargestellt.

Die Bewertung erfolgt auf der Basis randomisierter kontrollierter Studien anhand patientenrelevanter Endpunkte aus den Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen.

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.

Zur Beantwortung der Fragestellung gemäß Abschnitt 4.2.1 wurden solche Studien berücksichtigt, die anhand der im Folgenden aufgeführten Ein- und Ausschlusskriterien als relevant bewertet wurden.

Population

Die Einschlusskriterien zur Zielpopulation ergaben sich aus der Fachinformation zu Nintedanib im vorliegenden Anwendungsgebiet IPF (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b). Berücksichtigt wurden demnach Studien, die erwachsene Patienten mit IPF einschließen.

Intervention

Die zu bewertende Intervention ist Nintedanib im Anwendungsgebiet IPF. In die Nutzenbewertung wurden demnach Studien eingeschlossen, in welchen Nintedanib gemäß der Zulassung verabreicht wurde (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b). Entsprechend den Angaben der Fachinformation wird eine zweimal tägliche Einnahme von 150 mg Nintedanib im Abstand von ca. 12 Stunden empfohlen. Für Patienten, die eine zweimal tägliche Einnahme von 150 mg Nintedanib nicht tolerieren, wird eine Dosis von zweimal täglich 100 mg empfohlen. Nach dem Abklingen der Nebenwirkungen bzw. der Normalisierung der zur Dosisreduktion führenden Faktoren, kann gemäß Fachinformation eine Dosis-Re-Eskalation auf 150 mg erfolgen.

Vergleichstherapie

In der Niederschrift zum Beratungsgespräch zur Nutzenbewertung der Progressiven Fibrosierenden Interstitiellen Lungenerkrankungen (PF-ILD), unter welche auch das Anwendungsgebiet IPF fällt (Beratungsanforderung 2016-B-049), sowie im Schreiben zur vorsorglichen Aufforderung zur Dossiereinreichung vom 21.06.2018 bestätigte der G-BA die zVT BSC *oder* Pirfenidon, welche auch bereits im Zuge der Erstbewertung von Nintedanib im Anwendungsgebiet IPF festgelegt worden war ((Gemeinsamer Bundesausschuss, 2016b; Gemeinsamer Bundesausschuss, 2018a), Modul 3 Abschnitt 3.1). Der G-BA führt aus, dass unter BSC eine Therapie verstanden wird, welche eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und der Verbesserung der gesundheitsbezogenen Lebensqualität gewährleistet.

Die gewählte Suchstrategie folgte dem Hinweis des G-BA, im Zuge der Studienregister- und bibliografischen Literaturrecherche nicht auf die ausgewählte zVT einzuschränken und berücksichtigt Studien sowohl von Nintedanib gegenüber BSC als auch gegenüber Pirfenidon.

Da innerhalb des Settings einer unheilbaren und potenziell tödlichen Krankheit wie der IPF, welche auch mit einer entsprechend stark ausgeprägten Symptomatik einhergeht, davon auszugehen ist, dass auch Patienten, die in einen Placebo-Arm randomisiert wurden, eine umfangreiche Therapie zur Linderung ihrer Symptomatik zur Verfügung steht, wurden auch Studien gegenüber Placebo berücksichtigt und auf die Umsetzung einer BSC entsprechenden Therapie geprüft.

Prinzipiell konnten für die Nutzenbewertung von Nintedanib gegenüber BSC Studien berücksichtigt werden, die Nintedanib in der Monotherapie oder in Kombination mit BSC mit einer Behandlung mit BSC allein verglichen hatten. Im vorliegenden Dossier erfolgte die Bewertung durch den in den eingeschlossenen Studien 1199.32 (INPULSIS 1), 1199.34 (INPULSIS 2) und 1199.187 durchgeführten Vergleich von Nintedanib in Kombination mit BSC (Nintedanib/BSC) mit einer Behandlung mit BSC allein (Placebo/BSC). Bezüglich der Studie 1199.187 wurden hierbei nur Daten zur Nutzenbewertung herangezogen, welche innerhalb der kontrollierten Phase erhoben wurden (siehe Abschnitt 4.2.5.2). Die Anforderungen des G-BA an die im Vergleichsarm zur adäquaten Umsetzung von BSC zur Verfügung stehenden Therapien, waren in jeder der drei Studien erfüllt (siehe Abschnitt 4.3.1.2.1) (Gemeinsamer Bundesausschuss, 2016b).

Endpunkte

Gemäß § 2 Absatz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) besteht der Nutzen eines Arzneimittels aus einem patientenrelevanten therapeutischen Effekt, wenn eine Verbesserung des Gesundheitszustandes, eine Verkürzung der Krankheitsdauer, eine Verlängerung des Überlebens, eine Verringerung von Nebenwirkungen oder eine Verbesserung der gesundheitsbezogenen Lebensqualität erreicht wird.

Für die vorliegende Nutzenbewertung wurden Studien berücksichtigt, die Ergebnisse zu mindestens einem patientenrelevanten Endpunkt berichten. Diese umfassen z. B.:

- Mortalität
 - Gesamtüberleben
 - Abnahme der FVC (als valides Surrogat für die Mortalität)
 - Andere Endpunkte der Dimension Mortalität
- Morbidität/Symptomatik
 - Akute Exazerbationen
 - Veränderung des HRCT-QLF Scores
 - Veränderung der Gehstrecke im 6 MWT
 - Respiratorisch bedingte Hospitalisierung
 - Andere Endpunkte der Dimension Morbidität/Symptomatik
- Endpunkte der Dimension gesundheitsbezogene Lebensqualität (z. B. SGRQ, SOBQ, CASA-Q, PGI-C) bzw. des allgemeinen Gesundheitszustandes (EQ-5D)
- Nebenwirkungen

Die Patientenrelevanz der zur Ableitung des Zusatznutzens berücksichtigten Endpunkte sowie die Validität der Messinstrumente zur Erhebung der gesundheitsbezogenen Lebensqualität werden in Abschnitt 4.2.5.2 dargelegt.

Studientyp

RCT sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt werden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des medizinischen Nutzens und Zusatznutzens einer Intervention. Somit werden für diese Nutzenbewertung nur RCT berücksichtigt.

Studiendauer

Die IPF ist eine chronisch progrediente Erkrankung, welche einer entsprechend langfristig angelegten Therapie bedarf. Um Aussagen zu langfristigen Effekten auf nutzenbewertungsrelevante Endpunkte treffen zu können, ist daher eine gewisse Mindeststudiendauer vorauszusetzen.

Seitens der EMA und der U. S. Food and Drug Administration (FDA) liegen derzeit keine Empfehlungen vor, anhand derer sich eine Mindeststudiendauer im Anwendungsgebiet IPF ableiten ließe. Entsprechende Empfehlungen existieren jedoch für die hinsichtlich Symptomatik, Chronizität und Progredienz vergleichbare Chronisch obstruktive Lungenerkrankung (COPD). BI sieht daher die Empfehlungen zur Mindeststudiendauer bei COPD auch als relevant im Anwendungsgebiet IPF an.

Als Mindeststudiendauer wird durch die EMA 12 bis 24 Wochen angegeben. Zur Beurteilung eines Therapieeffekts auf das Auftreten von COPD-Exazerbationen gibt die EMA eine Dauer von mindestens einem Jahr an (European Medicines Agency (EMA), 2012).

In zurückliegenden Nutzenbewertungsverfahren im Anwendungsgebiet COPD wurde seitens des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und des G-BA eine Mindeststudiendauer von 24 Wochen angegeben (IQWiG, 2012; IQWiG, 2014; IQWiG, 2015b; IQWiG, 2015a; Gemeinsamer Bundesausschuss, 2016a; IQWiG, 2016b; IQWiG, 2016a). Entsprechend wurden im vorliegenden Dossier RCT mit einer Mindeststudiendauer von 24 Wochen eingeschlossen.

Publikationstyp

Für die Nutzenbewertung wurden Studienberichte und Vollpublikationen berücksichtigt, die dem Consolidated Standards of Reporting Trials (CONSORT)-Standard genügen. Darüber hinaus konnten auch Studien berücksichtigt werden, für die ein ausführlicher Ergebnisbericht aus einem Studienregister vorlag. Andere Publikationstypen wie z. B. Konferenzabstracts oder Reviews werden als nicht ausreichend zur Bewertung einer Studie angesehen.

Tabelle 4-4 fasst die Ein- und Ausschlusskriterien zusammen.

Tabelle 4-4: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien für die Nutzenbewertung des Arzneimittels Nintedanib

	Einschlusskriterien	Ausschlusskriterien	Aus- schluss- grund
Patientenpopulation	Erwachsene Patienten mit idiopathischer Lungenfibrose (IPF) <u>Begründung:</u> Die im Einschlusskriterium definierte Population entspricht der Patientenpopulation, für die Nintedanib zugelassen ist. Eine weitere Differenzierung der Population ist nicht notwendig, da die Festlegung der zVT durch den G-BA für das gesamte Anwendungsgebiet zutrifft.	Patienten mit ausschließlich abweichenden Erkrankungen Kinder und Jugendliche <18 Jahre	A1
Intervention	Das zu bewertende Arzneimittel ist Nintedanib (Ofev®). In die Nutzenbewertung sind Studien einzuschließen, in denen Nintedanib gemäß Angaben in der Fachinformation gegeben wurde. Diese sieht eine zweimal tägliche Einnahme von 150 mg Nintedanib vor. Die beiden Einnahmen von Nintedanib sollen im Abstand von ungefähr 12 Stunden erfolgen. Für Patienten, die eine zweimal tägliche Einnahme von 150 mg Nintedanib nicht tolerieren, wird eine zweimal tägliche Einnahme von 100 mg Nintedanib empfohlen. Nach dem Abklingen von Nebenwirkungen bzw. der Normalisierung der zur Dosisreduktion führenden Ursachen, kann eine Dosis-Re-Eskalation auf 150 mg erfolgen. Nintedanib kann zusammen mit weiteren unterstützenden medikamentösen oder nicht-medikamentösen Maßnahmen zur Linderung der Symptomatik eingesetzt werden. <u>Begründung:</u> Intervention gemäß Zulassungsstatus von Nintedanib	Studien ohne Behandlungsarm mit Nintedanib Studienarme mit anderen als den zugelassenen Therapie- und Dosierungsschemata für Nintedanib	A2

	Einschlusskriterien	Ausschlusskriterien	Aus- schluss- grund
Vergleichstherapie	• BSC (bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und der Verbesserung der gesundheitsbezogenen Lebensqualität) • Placebo^a • Pirfenidon <u>Begründung:</u> BI folgt den Empfehlungen des G-BA bezüglich der zVT (siehe Modul 3 Abschnitt 3.1, (Gemeinsamer Bundesausschuss, 2016b)) und vergleicht Nintedanib gegen BSC, da drei vergleichende Studien vorliegen (INPULSIS 1, INPULSIS 2 und Studie 1199.187), in welchen die Vorgaben des G-BA zur Ausgestaltung von BSC umgesetzt wurden. Es wird in diesem Dossier die beste verfügbare Evidenz dargestellt und der medizinische Zusatznutzen gegenüber BSC abgeleitet.	Abweichende Vergleichsarme bzw. Komparatoren	A3
Endpunkte	Mindestens einer der folgenden patientenrelevanten Endpunkte wird berichtet: • Mortalität - Gesamtüberleben - Abnahme der FVC - Andere patientenrelevante Endpunkte der Dimension Mortalität • Morbidität - Akute Exazerbationen - HRCT-QLF - Veränderung der Gehstrecke im 6 MWT - Respiratorisch bedingte Hospitalisierung - Andere patientenrelevante Endpunkte der Dimension Morbidität • Endpunkte der Dimension HRQoL/Gesundheitszustand (z. B. SGRQ, SOBQ, CASA-Q, PGI-C, EQ-5D) • Nebenwirkungen	Kein relevanter Endpunkt aus den Dimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität oder Nebenwirkungen berichtet	A4

	Einschlusskriterien	Ausschlusskriterien	Ausschlussgrund
	<u>Begründung:</u> Zur Darstellung des Zusatznutzens soll nach § 35 Abs. 1b Satz 4 SGB V, § 5 Abs. 2 Satz 3 AM-NutzenV und 5. Kapitel, 1. Abschnitt 2. Titel § 5 Abs. 2 Satz 3 Verfahrensordnung (VerfO) des G-BA (Bundesministerium der Justiz und für Verbraucherschutz, 1988; Bundesministerium der Justiz und für Verbraucherschutz, 2010; Gemeinsamer Bundesausschuss, 2018b) auf patientenrelevante Endpunkte insbesondere im Hinblick auf • Mortalität • Morbidität • Gesundheitsbezogene Lebensqualität Bezug genommen werden.		
Studientypen	Randomisierte kontrollierte Studien (RCT) <u>Begründung:</u> Bei der Prüfung von Interventionseffekten wird „der höchste Evidenzgrad RCT und systematischen Übersichten von RCT zugeordnet“ (IQWiG, 2017)	Nicht-randomisierte Studien (z. B. einarmige Studien, Beobachtungsstudien etc.)	A5
Studiendauer	Studien mit einer Behandlungsdauer von ≥ 24 Wochen <u>Begründung:</u> Der G-BA und das IQWiG bewerteten eine kürzere Studiendauer in zurückliegenden Verfahren im Anwendungsgebiet COPD als nicht adäquat. Aufgrund der vergleichbaren Symptomatik, Chronizität und Progredienz der COPD wird dieses Kriterium auf das Anwendungsgebiet IPF übertragen.	Studien mit einer Behandlungsdauer von < 24 Wochen	A6
Publikationstyp	Vollpublikationen, Studienberichte oder ausführliche Ergebnisberichte aus Studienregistern <u>Begründung:</u> Zur Bewertung des Verzerrungspotenzials einer Studie und damit der Festlegung der Ergebnissicherheit sind umfassende Informationen über die klinische Studie erforderlich. Dies ist nur durch einen Studienbericht oder eine Vollpublikation gewährleistet. Auch ein ausführlicher Ergebnisbericht in Studienregistern kann eine Bewertung der Studie ermöglichen und wird im Einzelfall geprüft.	Andere Publikationstypen (z. B. Reviews, Comments, Editorials, Notes, Letters, Conference Abstracts) Mehrfachpublikationen Studienregistereinträge oder Studienergebnisdatenbank-einträge ohne verfügbare Ergebnisse	A7

	Einschlusskriterien	Ausschlusskriterien	Ausschlussgrund
a: In einem Setting einer unheilbaren und potenziell tödlichen Krankheit wie der IPF, welche auch mit einer entsprechend stark ausgeprägten Symptomatik einhergeht, ist davon auszugehen, dass auch Patienten, die in einen Placebo-Arm randomisiert wurden, eine umfangreiche Therapie zur Linderung der Symptomatik zur Verfügung steht. Die Umsetzung einer dem BSC entsprechenden Therapie auch innerhalb von Placebo-Armen ist daher anzunehmen und wird für entsprechende Studien geprüft. 6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); AM-NutzenV: Arzneimittel-Nutzenbewertungsverordnung; BI: Boehringer Ingelheim; BSC: Best Supportive Care; CASA-Q: Cough and Sputum Assessment Questionnaire; COPD: Chronic Obstructive Pulmonary Disease (Chronisch obstruktive Lungenerkrankung); EQ-5D: European Quality of Life-5 Dimensions; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); G-BA: Gemeinsamer Bundesausschuss; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); HRQoL: Health-Related Quality of Life (Gesundheitsbezogene Lebensqualität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); IQWiG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; PGI-C: Patient's Global Impression of Change; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGB: Sozialgesetzbuch; SGRQ: St. George's Respiratory Questionnaire; SOBQ: Shortness of Breath Questionnaire; VerFO: Verfahrensordnung; zVT: zweckmäßige Vergleichstherapie			

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.

Ziel der systematischen bibliografischen Literaturrecherche ist es, RCT mit dem zu bewertenden Arzneimittel Nintedanib (Ofev®) im Anwendungsgebiet IPF zu identifizieren, die einen Vergleich von Nintedanib zu einer Vergleichstherapie in Übereinstimmung mit den Festlegungen des G-BA zur zVT erlauben.

Die Suche erfolgte in folgenden Datenbanken:

- MEDLINE
- EMBASE
- Cochrane Central Register of Controlled Trials

Der Informationsstand bezieht sich auf den 17.01.2019. Für jede einzelne Datenbank wurde eine sequenzielle und individuell adaptierte Suchstrategie entwickelt und separat angewandt. Für die Suche wurde die Plattform Ovid verwendet. Der Suchzeitraum wurde nicht eingeschränkt. Die individuellen Suchstrategien sind in Anhang 4-A dokumentiert.

Im Zuge der Suche in der Cochrane-Datenbank können auch Einträge aus dem clinicaltrials.gov-Register gefunden werden. Entsprechende Einträge wurden auf Basis des Abstracts sowie des entsprechenden Studienregistereintrages nach relevanten Studien durchsucht und anhand der in diesem Abschnitt beschriebenen Kriterien bewertet. Wurden im Zuge der Bewertung relevante Studien identifiziert, wurden die entsprechenden Einträge mit den Ergebnissen der Suche im clinicaltrials.gov-Register abgeglichen. Da die Einträge des

Cochrane-Registers selbst keine verwendbare Vollpublikation darstellen, wurden entsprechende Einträge im Zuge der bibliografischen Recherche ausgeschlossen.

Die Studien wurden nach den in Tabelle 4-4 prädefinierten Ein- und Ausschlusskriterien selektiert. Die Selektion nach den einzelnen relevanten Endpunkten fand erst auf Volltextbasis statt.

Die Ergebnisse der systematischen bibliografischen Literaturrecherche sind in Abschnitt 4.3.1.1.2 beschrieben.

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.

Zur Identifikation relevanter Studien zum zu bewertenden Arzneimittel Nintedanib wurden entsprechend den oben beschriebenen Vorgaben folgende Register durchsucht:

- ClinicalTrials.gov
- International Clinical Trials Registry Platform (ICTRP)
- European Union Clinical Trials Register (EU-CTR)
- PharmNet.Bund

Der Informationsstand bezieht sich auf den 17.01.2019. Die individuellen Suchstrategien sind in Anhang 4-B dokumentiert. Die Ergebnisse der Registersuche werden in Abschnitt 4.3.1.1.3 beschrieben.

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.

Die durch die bibliografische Literaturrecherche (Abschnitt 4.2.3.2) und die Studienregisterabfrage (Abschnitt 4.2.3.3) identifizierten Studien wurden anhand der in Abschnitt 4.2.2 beschriebenen Kriterien durch zwei Personen unabhängig voneinander auf Relevanz geprüft und selektiert. Abweichende Einschätzungen zur Relevanz einzelner Publikationen oder Studienregistereinträge wurden im Konsens durch Diskussion zwischen den selektierenden Personen gelöst.

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.

Die Bewertung der eingeschlossenen Studien hinsichtlich ihrer Aussagekraft wurde gemäß der Vorgaben der Dossievorlage anhand der zur Verfügung stehenden Informationen aus den Studienberichten vorgenommen. Sie erfolgte zunächst endpunktübergreifend auf Studienebene, anschließend für jeden patientenrelevanten Endpunkt und gegebenenfalls für jede Operationalisierung separat.

Es wurden die folgenden endpunktübergreifenden (A) und endpunktspezifischen (B) Aspekte systematisch extrahiert und bewertet (Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung der Patienten sowie des behandelnden Arztes
- Ergebnissteuerte Berichterstattung
- Sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-Treat (ITT)-Prinzips
- Ergebnissteuerte Berichterstattung
- Sonstige Aspekte

Gemäß der Verfahrensordnung (VerfO) des G-BA wurde das Verzerrungspotenzial auf Studienebene für RCT zusammenfassend als „niedrig“ oder „hoch“ eingestuft.

Bei niedrigem Verzerrungspotenzial kann mit großer Wahrscheinlichkeit davon ausgegangen werden, dass die Ergebnisse der Studie nicht relevant verzerrt sind. Bei hohem Verzerrungspotenzial kann eine relevante Verzerrung nicht ausgeschlossen werden. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden (Gemeinsamer Bundesausschuss, 2018b).

Zur Bewertung eines Endpunkts wurde für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend als „niedrig“ oder „hoch“ eingestuft. Falls die Einstufung als „hoch“ erfolgte, wurde auch das Verzerrungspotenzial für den Endpunkt als „hoch“ bewertet.

Die Einstufung „hoch“ für das Verzerrungspotenzial eines Endpunktergebnisses führte nicht zum Ausschluss der Daten, sondern diente zur Einschätzung der Aussagekraft der Ergebnisse zum betreffenden Endpunkt. Unter „sonstige Aspekte“ wurden z. B. relevante Dateninkonsistenzen innerhalb der Studienunterlagen oder zwischen den Studienunterlagen oder die Anwendung inadäquater statistischer Verfahren bewertet. Die Vorgehensweise bei der Bewertung des Verzerrungspotenzials für jede Studie und jeden Endpunkt (Operationalisierung) richtete sich nach den Hinweisen der Bewertungsbögen (Anhang 4-F).

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.

Die Darstellung von Design und Methodik erfolgte bei den eingeschlossenen RCT entsprechend der Anforderungen des CONSORT-Statements (Items 2b bis 14 sowie CONSORT-Flussdiagramm). Die Anwendung des CONSORT-Statements stützte sich neben der Checkliste zusätzlich auf weitere Ausführungen zu den einzelnen Items. Danach sollten das Ziel der Studie und die dazugehörigen Hypothesen, das Studiendesign, die Studienpopulation, die Interventionen und die betrachteten Endpunkte mittels hinreichender Kriterien klar definiert sein (Item 2b-6). Zur Bewertung der Aussagekraft der Studien waren die Details zur

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.

Bestimmung der Fallzahl, die Erzeugung der Randomisierungssequenz, die Zuteilung der Studienteilnehmer zu den Behandlungsarmen, die Verblindung und zuletzt die Verwendung adäquater statistischer Analysemethoden darzustellen (Item 7-12). Darüber hinaus wurden Angaben zur Zahl der eingeschlossenen, behandelten und analysierten Studienteilnehmer gemacht. Studienabbrüche oder Ausschlüsse von Studienteilnehmern nach der Randomisierung wurden klar und anschaulich beschrieben. Zudem wurde der Zeitraum, in dem die Rekrutierung und die Studie selbst stattfanden, angegeben, um diese später möglicherweise in einen historischen Zusammenhang stellen zu können (Item 13-14) (Moher et al., 2010; Schulz et al., 2010).

Als Informationsquellen für die Beschreibung des Designs und der Methodik der in die Nutzenbewertung eingeschlossenen Studien wurden Studienbericht, Protokoll und Statistischer Analyseplan (SAP) herangezogen.

Die Methodik der eingeschlossenen Studien, die Interventionen sowie die Charakteristika der Studienpopulationen wurden in Tabelle 4-12, Tabelle 4-13 und Tabelle 4-14 dargestellt.

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 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.

Patientencharakteristika

Die Charakterisierung der Studienpopulation für jede der eingeschlossenen Studien erfolgte separat für die einzelnen Behandlungsarme (Tabelle 4-5). Für die Beschreibung der Studienpopulation in Bezug auf Krankheitsschwere und die Verteilung prognostischer Faktoren wurden die folgenden demografischen und krankheitsspezifischen Faktoren berücksichtigt:

Tabelle 4-5: Faktoren zur Charakterisierung der Studienpopulation der RCT mit dem zu bewertenden Arzneimittel

Patientencharakteristika	INPULSIS 1	INPULSIS 2	Studie 1199.187
Demografie			
Alter (Jahre)	ja	ja	ja
Geschlecht (männlich, weiblich)	ja	ja	ja
Gewicht (kg)	ja	ja	ja
BMI (kg/m ²)	ja	ja	ja
Ethnie (kaukasisch, asiatisch)	ja	ja	ja
Raucherstatus (Nichtraucher, Ex-Raucher, Raucher)	ja	ja	ja
Zeit seit IPF-Diagnose (Jahre)	ja	ja	ja
Krankheitsspezifische Charakteristika			
Radiologische Beurteilung der IPF (vermutlich UIP, mögliche UIP, keine UIP)	ja	ja	nein
Lungenbiopsie (ja, nein)	ja	ja	nein
FVC (ml, % des Sollwertes)	ja	ja	ja
FEV ₁ /FVC-Verhältnis (%) ^a	ja	ja	nein
DL _{CO} (mmol/min/kPa)	ja	ja	ja
SpO ₂ (%)	ja	ja	ja
SGRQ-Gesamtscore (Punkte)	ja	ja	ja
Systemische Kortikosteroid-Therapie (ja, nein)	ja	ja	nein
a: Nur INPULSIS Studien und für den Gesamtstudienpool aller drei Studien BMI: Body-Mass-Index; DL _{CO} : Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); FEV ₁ : Forced Expiratory Volume in one Second (Forcierte expiratorische Einsekunden-Kapazität); FVC: Forced Vital Capacity (Forcierte Vitalkapazität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SpO ₂ : Sauerstoffsättigung; UIP: Usual Interstitial Pneumonia			

Ergänzt wurde die Charakterisierung der Studienpopulation durch eine Gegenüberstellung der Anzahl der Studienabbrecher in den Behandlungsarmen und eine Gegenüberstellung der medianen Behandlungs- und Beobachtungsdauer.

Die Patientenflüsse der berücksichtigten Studien wurden in Flow-Charts dargestellt (Abbildung 39, Abbildung 40, Abbildung 41), die die Verteilung der Studienteilnehmer auf die Behandlungsgruppen sowie deren Gründe für einen ggf. vorzeitigen Studienabbruch enthalten.

Studienpools

Zusätzlich zur Beschreibung der Studienpopulation jeder Einzelstudie wurden auch Patientencharakteristika für Studienpools der eingeschlossenen Studien dargestellt; zum einen über alle Studien hinweg (INPULSIS Studien und Studie 1199.187) sowie gesondert für die beiden INPULSIS Studien (INPULSIS gepoolt). Die Überprüfung der Ähnlichkeitsannahme – eine Voraussetzung für die meta-analytische Zusammenfassung der drei Studien – wurde deskriptiv anhand des Vergleichs der Charakteristika der Studie 1199.187 mit den Charakteristika der gepoolten Population der INPULSIS Studien durchgeführt. Von der Ähnlichkeit der INPULSIS Studien ist aufgrund des identischen CTP und der identischen Einschlusskriterien auszugehen.

Endpunkte

Die in Tabelle 4-6 genannten Endpunkte wurden für die vorliegende Nutzenbewertung herangezogen.

Tabelle 4-6: Endpunkte der RCT mit dem zu bewertenden Arzneimittel für die vorliegende Nutzenbewertung

Endpunkt	INPULSIS 1	INPULSIS 2	Studie 1199.187
Mortalität			
Zeit bis zum Tod, unabhängig von der Ursache	ja	ja	ja
Zeit bis zum respiratorisch bedingten Tod	ja	ja	ja
Jährliche FVC-Abnahme (ml)	ja	ja	ja
FVC-Responder zu Woche 52 (10% des Sollwertes)	ja	ja	nein
FVC-Responder zu Woche 52 (5% des Sollwertes)	ja	ja	nein
FVC-Responder zu Woche 24 (10% des Sollwertes)	ja	ja	ja
FVC-Responder zu Woche 24 (5% des Sollwertes)	ja	ja	ja
Morbidität			
Zeit bis zur ersten adjudizierten akuten Exazerbation	ja	ja	ja
Zeit bis zur ersten akuten Exazerbation	ja	ja	ja
Veränderung HRCT-QLF Score (bis Woche 24)	nein	nein	ja
Veränderung der Gehstrecke im 6 MWT (bis Woche 24)	nein	nein	ja
Zeit bis zur ersten respiratorisch bedingten Hospitalisierung	nein	nein	ja

Endpunkt	INPULSIS 1	INPULSIS 2	Studie 1199.187
Gesundheitsbezogene Lebensqualität/allgemeiner Gesundheitszustand			
Veränderung SGRQ-Gesamtscore ^a (bis Woche 24 bzw. Woche 52)	ja	ja	ja
SGRQ-(≤-4 Punkte)-Responder ^a (zu Woche 52)	ja	ja	nein
SGRQ-(≤-4 Punkte)-Responder ^a (zu Woche 24)	ja	ja	ja
Veränderung SGRQ-I-Gesamtscore ^a (bis Woche 24 bzw. Woche 52)	ja	ja	ja
SGRQ-I-(≤-4 Punkte)-Responder ^a (zu Woche 52)	ja	ja	nein
SGRQ-I-(≤-4 Punkte)-Responder ^a (zu Woche 24)	ja	ja	ja
Veränderung SOBQ-Gesamtscore ^a (bis Woche 24 bzw. Woche 52)	ja	ja	ja
Veränderung CASA-Q Husten-Symptome ^a bis Woche 52	ja	ja	nein
Veränderung CASA-Q Husten-Impact ^a bis Woche 52	ja	ja	nein
PGI-C-Responder ^b zu Woche 52	ja	ja	nein
Veränderung EQ-5D-VAS ^b bis Woche 52	ja	ja	nein
Unerwünschte Ereignisse			
UE	ja	ja	ja
SUE	ja	ja	ja
UE, die zum Therapieabbruch führten	ja	ja	ja
UE von besonderem Interesse	ja	ja	ja
a: Krankheitsspezifische Messinstrumente b: Generische Messinstrumente 6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); PGI-C: Patient's Global Impression of Change; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visual Analog Scale			

Die Darstellung der Ergebnisse für die einzelnen Endpunkte einschließlich der Beschreibung der Operationalisierung und des Verzerrungspotenzials erfolgt in Abschnitt 4.3.1.3.

Im Folgenden werden die Endpunkte bezüglich ihrer Patientenrelevanz bzw. Validität bewertet:

Mortalität

Die IPF ist eine schwerwiegende und in der Regel zum Tod führende Erkrankung (Behr et al., 2013). Somit ist die Verlängerung der Gesamtüberlebenszeit ein wichtiges Therapieziel und daher patientenrelevant. Auch seitens des IQWiG und des G-BA ist die Verlängerung des Überlebens als patientenrelevante Nutzendimension anerkannt. Aufgrund der eindeutigen und klaren Bestimmbarkeit ist der Überlebensstatus auch im Fall des vorzeitigen Studienabbruchs eines Patienten in der Regel zu späteren Zeitpunkten noch zweifelsfrei feststellbar.

Effekte der Behandlung mit Nintedanib auf die Mortalität wurden unmittelbar durch die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* und *Zeit bis zum respiratorisch bedingten Tod*, sowie indirekt durch die validen Surrogatendpunkte *jährliche FVC-Abnahme* (ml) und *Anteil FVC-Responder* (Schwellenwerte 10% des Sollwertes und 5% des Sollwertes) erfasst und bewertet. Zur Validität dieser Surrogatendpunkte siehe Abschnitt 4.5.4.

Zeit bis zum Tod

Zeit bis zum Tod, unabhängig von der Ursache

Die *Zeit bis zum Tod, unabhängig von der Ursache* wurde als die Zeit von Randomisierung bis zum Todeszeitpunkt eines Studienteilnehmers definiert. Überlebende Studienteilnehmer wurden in den INPULSIS Studien zum Zeitpunkt des letzten Patientenkontaktes (höchstens 372 Tage) zensiert. Die Zensierung in der Studie 1199.187 erfolgte zum Zeitpunkt des letzten Patientenkontaktes bzw. am End-Datum der kontrollierten Phase. Für diesen Endpunkt wurden tabellarisch die Anzahl der Studienteilnehmer unter Risiko, sowie die Anzahl und der Anteil Studienteilnehmer mit dem betreffenden Ereignis angegeben. Aufgrund der niedrigen Ereignisrate konnte kein Median angegeben werden. Stattdessen wurde die Wahrscheinlichkeit, das betreffende Ereignis nicht zu erleiden, anhand des Kaplan-Meier-Schätzers pro Behandlungsgruppe angegeben. Zusätzlich zur tabellarischen Darstellung wurden Kaplan-Meier-Kurven erstellt und präsentiert. Der Behandlungseffekt wurde als HR mit 95%-KI und p-Wert basierend auf Cox-Regression berichtet und der p-Wert mittels Log-rank-Test berechnet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten.

Zeit bis zum respiratorisch bedingten Tod

Die *Zeit bis zum respiratorisch bedingten Tod* wurde als die Zeit von Randomisierung bis zum Zeitpunkt des Todes aufgrund einer respiratorisch bedingten Ursache eines Studienteilnehmers definiert. Studienteilnehmer, die aufgrund einer anderen Ursache verstarben, wurden zum Zeitpunkt ihres Todes zensiert. Überlebende Studienteilnehmer in den INPULSIS Studien wurden zum Zeitpunkt des letzten Patientenkontaktes (höchstens 372 Tage) zensiert. In Studie 1199.187 fand die Zensierung für den Therapievergleich mittels Cox-Regression zum Zeitpunkt des letzten Patientenkontaktes bzw. am End-Datum der kontrollierten Phase statt. Für die Kaplan-Meier-Schätzung wurden alle Daten 176 Tage nach der Einnahme der ersten Studienmedikation zensiert. Auswertung und Darstellung der Ergebnisse zu diesem Endpunkt erfolgten analog zu denjenigen für *Zeit bis zum Tod, unabhängig von der Ursache*.

Alle Todesfälle wurden hinsichtlich ihrer Ursache durch ein verblindetes IAC bewertet, dem drei ausgewiesene Experten aus dem Bereich der Pneumologie mit besonderen Fachkenntnissen in der Behandlung der IPF angehörten. Die Bewertung durch die Mitglieder des IAC fand unabhängig voneinander auf Basis der vorliegenden Informationen statt. Abweichende Einschätzungen wurden im Zuge eines Panel Meetings konsentiert oder per Mehrheitsbeschluss entschieden. Konnte eine Todesursache nicht eindeutig bestimmt werden, wurde ein nicht respiratorisch bedingter Tod angenommen. Die Entscheidung durch das IAC erfolgte auf Grundlage klinisch relevanter Daten wie z. B.:

- Narrative der SUE mit der Zusammenfassung des klinischen Verlaufs, den Symptomen, dem Totenschein, dem Autopsiebericht sowie der Einschätzung des Prüfarztes zur Todesursache
- Patientenprofil mit der Krankengeschichte, der Begleitmedikation und den berichteten UE
- Zur Behandlung der Exazerbation der IPF verabreichte Behandlung

In der Studie 1199.187 wurden Daten zu UE, wie z. B. gastrointestinale Erkrankungen, vor Übermittlung an das IAC zensiert. Eine detaillierte Beschreibung des Adjudizierungsprozesses in den verschiedenen Studien sowie der dem IAC dazu vorgelegten Unterlagen findet sich in Modul 5 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012c; Boehringer Ingelheim Pharma GmbH & Co. KG, 2016a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2016b).

Forcierte Vitalkapazität (FVC)

Die FVC ist das Lungenvolumen, das nach maximaler Einatmung forciert ausgeatmet werden kann. Bei der IPF kommt es aufgrund der zunehmenden Vernarbung der Lunge zu einer mit der Zeit fortschreitenden Abnahme der FVC, sodass die Belüftung der Lunge zunehmend eingeschränkt wird bzw. eine erhebliche Atemanstrengung des Patienten erfordert. Da die Vernarbungsprozesse bei der IPF gegenwärtig weder rückgängig gemacht noch gestoppt werden können, zielt die Therapie der IPF darauf ab, den Abfall der FVC und damit die Krankheitsprogression zu verlangsamen und folglich das Mortalitätsrisiko zu verringern. Eine detaillierte Beschreibung der FVC als Surrogatendpunkt für die Mortalität erfolgt in Abschnitt 4.5.4.

Im Nutzenbewertungsverfahren zu Nintedanib im Anwendungsgebiet IPF als Arzneimittel zur Behandlung eines seltenen Leidens hatte der G-BA den Surrogatcharakter der FVC in Bezug auf Mortalität nicht anerkannt (Gemeinsamer Bundesausschuss, 2015c). Seit diesem zurückliegenden Nutzenbewertungsverfahren wurde nun weitere Evidenz verfügbar, welche allesamt den Surrogatcharakter der FVC auf Mortalität stützt. BI sieht daher die Voraussetzungen zur Berücksichtigung als Surrogatendpunkt in der vorliegenden Nutzenbewertung als gegeben an. In diesem Zusammenhang verweist BI auf die seit der Erstbewertung publizierte Evidenz, welche in Abschnitt 4.5.4 dargestellt ist. Anhand der vorliegenden Daten zur FVC wird in diesem Abschnitt zusätzlich eine auf einem

Metaregressionsmodell basierende Ableitung des Therapieeffekts von Nintedanib auf das Mortalitätsrisiko dargestellt.

Neben dem Surrogatcharakter für die Mortalität wirkt sich eine Abnahme der FVC auch unmittelbar auf den Patienten aus. Dies zeigt sich in der deutlich höheren gesundheitsbezogenen Lebensqualität von Patienten, welche keine Abnahme der FVC erfahren. So verbesserte sich die Lebensqualität gemessen anhand des SGRQ in einer Post-hoc-Analyse auf Basis der INPULSIS Studien unter Patienten ohne Abnahme der FVC innerhalb eines Jahres um 1,9 Punkte (Nintedanib- und Vergleichsarm). Dagegen verschlechterte sich der SGRQ-Wert unter Patienten, die innerhalb eines Jahres eine Abnahme der FVC erfuhren, um 4,9 und 5,2 Punkte (Nintedanib-Arm bzw. Vergleichsarm) (Flaherty et al., 2018). Dies verdeutlicht den starken Zusammenhang zwischen dem Erhalt des Lungenvolumens und der direkt damit in Verbindung stehenden Lebensqualität der Patienten und verdeutlicht somit die Patientenrelevanz FVC-basierter Endpunkte neben ihrer Funktion als Surrogat für Mortalität.

Jährliche FVC-Abnahme (ml)

Die Bestimmung der FVC erfolgte in den INPULSIS Studien bei Studienbeginn sowie zu Woche 2, 4, 6, 12, 24, 36 und 52 mittels Spirometrie unter Einhaltung der Kriterien der American Thoracic Society/European Respiratory Society/Japanese Respiratory Society/ Latin American Thoracic Association (ATS/ERS/JRS/ALAT) (Miller et al., 2005). In der Studie 1199.187 erfolgten bei allen Patienten entsprechende Erhebungen vor bzw. bei Studienbeginn sowie zu Woche 12 und 24. Zudem erfolgten zu Woche 36 und 52 weitere Erhebungen, wenn dieser Zeitrahmen in die verblindete Periode fiel. Die *jährliche FVC-Abnahme* (ml) wurde mittels Random-Coefficient-Modell unter Berücksichtigung aller vorhandenen Messwerte berechnet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Insbesondere berücksichtigt dieses Modell auch Daten von Studienteilnehmern mit fehlenden Messungen. Die Anzahl der analysierten Studienteilnehmer sowie die beobachteten Mittelwerte (MW) mit Standardabweichung (SD) zu Studienbeginn und Studienende (INPULSIS Studien) bzw. zum Ende der kontrollierten Phase (Studie 1199.187) wurden tabellarisch dargestellt. Weiterhin zeigt die Tabelle die auf dem Random-Coefficient-Modell basierende Veränderung der FVC von Studienbeginn bis Ende des Auswertungszeitraumes als adjustierte MWD mit Standardfehler (SE) und den daraus resultierenden Behandlungseffekt ebenfalls als adjustierte MWD mit SE zwischen den Behandlungsgruppen mit 95%-KI und p-Wert. Von der Endpunktoperationalisierung der INPULSIS Studien abweichend war gemäß CTP und SAP der Studie 1199.187 eine Mixed Effect Model for Repeated Measures (MMRM)-Analyse für die absoluten und relativen Änderungen der FVC vorgesehen. Im statistischen Modell waren Behandlung und Visite als feste Effekte, der Ausgangswert als Kovariate sowie die Interaktionsterme Behandlung×Visite und Ausgangswert×Visite vorgesehen. Als bekannte Prädiktoren waren zudem Geschlecht, Alter und Größe im Modell eingeschlossen. Der Patienteneffekt wurde als zufällig und die Kovarianzmatrix als unstrukturiert gewählt, um intraindividuelle Variationen abzubilden. Abweichend von der Präspezifizierung wurde für die vorliegende Nutzenbewertung für Studie 1199.187 dieselbe Operationalisierung und

Analysemethode verwendet wie für die INPULSIS Studien, um die Ergebnisse meta-analytisch zusammenfassen zu können.

FVC-Responder (Schwellenwerte 10% des Sollwertes und 5% des Sollwertes)

Als *FVC-Responder* (Schwellenwerte 10% des Sollwertes und 5% des Sollwertes) war ein Patient definiert, dessen FVC in Woche 24 bzw. Woche 52 gegenüber dem individuellen Messwert zu Studienbeginn um nicht mehr als 10% bzw. 5% (bezogen auf den Sollwert) abfiel. Falls zu Woche 24 bzw. Woche 52 kein Messwert vorlag, wurde der entsprechende Studienteilnehmer konservativ als Non-Responder bewertet (siehe Abschnitt 4.5.4 bezüglich der Validität der gewählten Schwellenwerte von 10% und 5% des Sollwertes für den Surrogatendpunkt FVC).

Der Endpunkt *FVC-Responder* (Schwellenwerte 10% des Sollwertes und 5% des Sollwertes) wurde mittels logistischer Regression ausgewertet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Die tabellarische Ergebnisdarstellung zeigt die Anzahl der Studienteilnehmer, die in die Analyse eingegangen sind, sowie die Anzahl und den Anteil der *FVC-Responder*. Der Behandlungseffekt wird durch das OR mit 95%-KI und p-Wert dargestellt.

Morbidität

Die IPF geht mit einer erheblichen Beeinträchtigung des Gesundheitszustandes einher. Diese Beeinträchtigung hat zwei wesentliche Komponenten: Einerseits die alltäglich vom Patienten verspürten Symptome, andererseits das Auftreten unmittelbar lebensbedrohlicher krisenhafter Verschlechterungen, sogenannte *akute Exazerbationen*, welche mit einer irreversiblen und mitunter in kurzer Zeit tödlich verlaufenden weiteren Schädigung der Lungenfunktion einhergehen. Effekte der Behandlung mit Nintedanib auf den Gesundheitszustand wurden in den eingeschlossenen Studien mithilfe der Erfassung der patientenrelevanten Ereignisse *adjudizierte akute Exazerbationen* und *akute Exazerbationen* beurteilt. Darüber hinaus wurde die Beeinträchtigung durch Symptome mithilfe von Fragebogeninstrumenten erfasst, die die Auswirkungen der Symptomintensität auf die gesundheitsbezogene Lebensqualität abbilden. Daher werden letztere nicht in der Nutzenkategorie Morbidität, sondern unter Verbesserung der gesundheitsbezogenen Lebensqualität betrachtet.

Akute Exazerbationen

Akute Exazerbationen sind akut lebensbedrohliche Ereignisse, welche bei Patienten mit IPF mit einer Inzidenz von 5-15% pro Jahr auftreten (Azuma et al., 2005; Collard et al., 2007b; King et al., 2008; Song et al., 2011). Sie können in jedem Erkrankungsstadium, jedoch vermehrt bei höherem Schweregrad der IPF, ohne Vorwarnung auftreten. Leitsymptom der *akuten Exazerbation* ist das Neuauftreten oder die Zunahme einer Dyspnoe. Weitere Symptome können Verschlimmerung von Husten, Auftreten von Fieber und ggf. vermehrtes Sputum sein (Parambil et al., 2005; Kim et al., 2006; Johansson und Collard, 2013). Ätiologisch werden sie als akute, intrinsische Akzeleration des IPF-Krankheitsfortschreitens gewertet (Johansson und Collard, 2013).

Klinisch führen *akute Exazerbationen* unmittelbar zu einer rapiden Verschlechterung des Gesundheitszustandes der Patienten und gehen in den meisten Fällen mit einer stationären Behandlung einher (Brown et al., 2015). Überdies kommt es aufgrund des Alveolarschadens und der Zunahme der Fibrosierung zu einer Verschlechterung der Lungenfunktionsparameter (Johannson und Collard, 2013). *Akute Exazerbationen* sind mit einer erhöhten Mortalitätsrate assoziiert: In einer retrospektiven Studie lag das mediane Überleben von Patienten mit einer *akuten Exazerbation* bei 2,2 Monaten und die Letalität bei stationärer Behandlung bei 50% (Song et al., 2011). Innerhalb eines Jahres nach einer Exazerbation liegt die Sterblichkeit bei 80-90% der betroffenen Patienten (Collard et al., 2007b). In einer weiteren Analyse basierend auf sechs Studien zur klinischen Entwicklung von Nintedanib und Pirfenidon ergab sich für Patienten, die im Studienverlauf mindestens eine Exazerbation erlitten hatten, ein mehr als 10-fach höheres Sterberisiko (Paterniti et al., 2017). Schwere und Häufigkeit *akuter Exazerbationen* stellen somit ein wichtiges prognostisches Kriterium dar. Daher sind *akute Exazerbationen* patientenrelevante Ereignisse und deren Vermeidung ist ein wichtiges Therapieziel.

Eine *akute Exazerbation* war in den berücksichtigten Studien definiert als eine erhebliche akute Verschlechterung der respiratorischen klinischen Situation des Patienten, die auf keine andere Ursache als die zugrunde liegende IPF zurückgeführt werden konnte. Gemäß CTP mussten zur Einstufung als *akute Exazerbation* alle im Folgenden aufgeführten Kriterien erfüllt sein:

- Unerklärtes Neuauftreten oder Zunahme von Dyspnoe innerhalb der letzten 30 Tage
- Neue pulmonale Infiltrate im Thorax-Röntgenbild und/oder neue parenchymale Anomalien ohne Pneumothorax oder Pleuraerguss (neue Milchglasverschattungen) in der HRCT
- Ausschluss von Infektionen
- Ausschluss anderer möglicher Erklärungen, wie beispielsweise Linksherzversagen, Lungenembolie oder akute Lungenschädigung bekannter Ursache

Adjudizierte akute Exazerbationen

Bei *akuten Exazerbationen* handelt es sich um in klinischen Studien relativ seltene Ereignisse. Zudem stellt ihre Abgrenzung zu akuten Verschlechterungen des Gesundheitszustandes aufgrund anderer medizinischer Ursachen eine besondere Herausforderung dar. Aus diesem Grund ist die alleinige Bewertung durch den Prüfarzt, ob ein Verdachtsereignis tatsächlich eine *akute Exazerbation* ist oder nicht, mit einer gewissen Unsicherheit behaftet. Daher wurden die durch die Prüfarzte erhobenen Verdachtsereignisse adjudiziert: Die Daten für jede vom Prüfarzt vermutete *akute Exazerbation* wurden durch ein verblindetes IAC überprüft, dem drei ausgewiesene Experten aus dem Bereich der Pneumologie mit besonderen Fachkenntnissen in der Behandlung der IPF angehörten. Jedes aufgetretene Ereignis wurde einer der drei Kategorien *bestätigte Exazerbation*, *vermutete Exazerbation* oder *keine Exazerbation* zugeordnet. Die Bewertung durch die Mitglieder des IAC fand unabhängig voneinander auf Basis der vorliegenden Informationen statt:

- Narrative der SUE mit der Zusammenfassung des klinischen Verlaufs und den Symptomen
- Einen spezifischen Fragebogen zur Adjudizierung akuter Exazerbationen (beinhaltend Laborparameter, Lungenfunktion und Bildgebung)
- Patientenprofil mit der Krankengeschichte, der Begleitmedikation und den berichteten UE
- Zur Behandlung der Exazerbation der IPF verabreichte Behandlung

Abweichende Einschätzungen wurden im Zuge eines Panel Meetings konsentiert oder per Mehrheitsbeschluss entschieden. Daten zu UE, wie z. B. gastrointestinale Erkrankungen, wurden in der Studie 1199.187 vor Übermittlung an das IAC zensiert. *Bestätigte* oder *vermutete akute Exazerbationen* wurden dem Endpunkt *adjudizierte akute Exazerbationen* als Ereignis zugeordnet. Eine detaillierte Beschreibung des Adjudizierungsprozesses in den verschiedenen Studien sowie der dem IAC dazu vorgelegten Unterlagen findet sich in Modul 5 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012c; Boehringer Ingelheim Pharma GmbH & Co. KG, 2016a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2016b).

Aufgrund der höheren Ergebnissicherheit durch die der Einschätzung des Prüfarztes nachfolgenden Adjudizierung durch das IAC, ist hinsichtlich der Bewertung, ob eine plötzlich auftretende Verschlechterung des Gesundheitszustandes tatsächlich einer Exazerbation der IPF entspricht, daher von einer höheren Aussagekraft des Endpunkts *adjudizierte akute Exazerbationen* gegenüber dem Endpunkt *akute Exazerbationen* auszugehen.

Im ersten Nutzenbewertungsverfahren zu Nintedanib im Anwendungsgebiet IPF hatte der G-BA die Patientenrelevanz der Endpunkte *akute Exazerbationen* und *adjudizierte akute Exazerbationen* mit dem Verweis daraufhin infrage gestellt, dass nicht beurteilt werden könne, ob die Symptomintensität bei Auftreten von Exazerbationen ein Ausmaß erreichen würde, welches als patientenrelevant eingestuft werden kann (Gemeinsamer Bundesausschuss, 2015c).

Das Auftreten *akuter Exazerbationen* wurde in den berücksichtigten Studien im Rahmen der Erfassung der unerwünschten Ereignisse dokumentiert. Wie schon im Zuge des Stellungnahmeverfahrens verweist BI in diesem Zusammenhang darauf, dass 76,2% der während der INPULSIS Studien bzw. 100% der während der Studie 1199.187 aufgetretenen ersten Exazerbationen durch den Prüfarzt als SUE bewertet wurden (Tabelle 4-7). Betrachtet man ausschließlich die durch das IAC als *vermutete Exazerbation* oder *bestätigte Exazerbation* bewerteten Ereignisse, so erhöht sich der Anteil der als SUE dokumentierten Exazerbationen in den INPULSIS Studien auf 84,4%. Bei den in den berücksichtigten Studien dokumentierten *akuten Exazerbationen* handelt es sich demnach zum überwiegenden Teil um als SUE bewertete Ereignisse. Diese sind gemäß der AM-NutzenV bei der Bestimmung des Zusatznutzens grundsätzlich als patientenrelevant zu bewerten (Bundesministerium der Justiz und für Verbraucherschutz, 2010).

Tabelle 4-7: Akute Exazerbationen gemäß Adjudizierungsergebnis, aufgeschlüsselt nach SUE und nicht SUE in den RCT mit dem zu bewertenden Arzneimittel

	Bestätigte/vermutete Exazerbation gemäß IAC	Keine Exazerbation gemäß IAC	Summe
INPULSIS 1 + 2			
SUE	28	20	48
nicht SUE	5	10	15
Studie 1199.187^a			
SUE	3	0	3
nicht SUE	0	0	0
a: Aufgetreten während der kontrollierten Phase IAC: Independent Adjudication Committee (Unabhängiges Bewertungskomitee); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SUE: Schwerwiegendes unerwünschtes Ereignis			

Die Patientenrelevanz der in den berücksichtigten Studien dokumentierten *akuten Exazerbationen* zeigt sich zudem durch einen Vergleich der jeweiligen Mortalitätsraten: Unter den 63 Patienten, die in den INPULSIS Studien eine Exazerbation erlitten, verstarben während der Nachverfolgung 31 Patienten, entsprechend einer Mortalität von 49,2%. Demgegenüber kann für Patienten ohne Exazerbation eine Mortalität von 5,3% ermittelt werden. Somit ging die Tatsache, mindestens eine Exazerbation erlitten zu haben, im Vergleich zu Patienten ohne Exazerbation mit einer fast 10-fach höheren Mortalität einher.

Die Operationalisierung *akute Exazerbationen* und *adjudizierte akute Exazerbationen* erfolgte als *Zeit bis zur ersten akuten Exazerbation* und *Zeit bis zur ersten adjudizierten akuten Exazerbation*.

Zeit bis zur ersten akuten Exazerbation

Der Endpunkt *Zeit bis zur ersten akuten Exazerbation* wurde als die Zeit von Randomisierung bis zum Auftreten einer für den betroffenen Studienteilnehmer ersten vom Prüfarzt berichteten *akuten Exazerbation* definiert. Studienteilnehmer ohne Exazerbationsereignis wurden in den INPULSIS Studien zum Zeitpunkt des letzten Patientenkontaktes zensiert. In Studie 1199.187 fand die Zensierung für den Therapievergleich mittels Cox-Regression zum Zeitpunkt des letzten Patientenkontaktes bzw. am End-Datum der kontrollierten Phase statt. Für die Kaplan-Meier-Schätzung wurden alle Daten 176 Tage nach der Einnahme der ersten Studienmedikation zensiert. Die Anzahl der Studienteilnehmer unter Risiko, sowie die Anzahl und der Anteil Studienteilnehmer mit dem betreffenden Ereignis wurden tabellarisch dargestellt. Aufgrund der niedrigen Ereignisrate konnte kein Median angegeben werden. Stattdessen wurde die Wahrscheinlichkeit, das betreffende Ereignis nicht zu erleiden, anhand des Kaplan-Meier-Schätzers pro Behandlungsgruppe zu Tag 372 bzw. Tag 176 angegeben. Der Behandlungseffekt wurde als HR mit 95%-KI basierend auf Cox-Regression und der p-Wert berechnet mittels Log-rank-Test angegeben. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Die Erfassung von

Exazerbationsereignissen erfolgte in den INPULSIS Studien und in Studie 1199.187 im Rahmen der Erhebung von UE.

Zeit bis zur ersten adjudizierten akuten Exazerbation

Der Endpunkt *Zeit bis zur ersten adjudizierten akuten Exazerbation* wurde analog zum Endpunkt *Zeit bis zur ersten akuten Exazerbation* operationalisiert. Darüber hinaus wurden Kaplan-Meier-Kurven erstellt und präsentiert.

6-Minuten-Gehtest (6 MWT)

Der 6 MWT ist ein funktioneller Test zur Beurteilung der Belastungseinschränkung und des Krankheitsverlaufs bei kardiopulmonalen Erkrankungen. Der 6 MWT wurde ursprünglich für Patienten mit COPD entwickelt, wurde nachfolgend aber auch im Anwendungsgebiet IPF validiert und ist dort mittlerweile in der klinischen Praxis weit verbreitet (du Bois et al., 2011c; Behr et al., 2013). Im Zuge des Tests wird die Gehstrecke ermittelt, welche ein Patient innerhalb von sechs Minuten auf ebenem Terrain zurücklegen kann. Während Tempowechsel und Pausen erlaubt sind, werden die Patienten angehalten innerhalb der sechs Minuten die für sie maximale Strecke zurückzulegen (Matthys und Seeger, 2008).

Die anhand des 6 MWT ermittelte Belastbarkeit des kardiopulmonalen Systems ist für Patienten insofern relevant, als dass diese Aufschluss darüber gibt, wie gut viele Aktivitäten des täglichen Lebens, wie z. B. das Überqueren eines beampelten Fußgängerüberwegs, ausgeführt werden können. Zudem deutet der Abfall des 6 MWT-Wertes in einer Reihe von Erhebungen im Anwendungsgebiet IPF auf ein erhöhtes Sterberisiko hin (Lettieri et al., 2006; Caminati et al., 2009; du Bois et al., 2011c; Nathan et al., 2015).

Zum Minimalen Schwellenwert für eine klinisch relevante Verbesserung (MCID) des 6 MWT, also der Gehstrecke, deren Erhalt vom Patienten als spürbarer und klinisch relevanter Unterschied gewertet wird, gibt es unterschiedliche Erhebungen. Entsprechende Studien im Anwendungsgebiet IPF ergaben Gehstrecken von 21,7-37 m (Nathan et al., 2015), 24-45 m (du Bois et al., 2011c) oder 28 m (Swigris et al., 2010b).

Der 6 MWT wurde in der Studie 1199.187 bis zu 4 Wochen vor Ausgangswert, zu Woche 0 (Ausgangswert) sowie zu Woche 12 und 24 erhoben. Die im Dossier dargestellte Operationalisierung erfolgte als Unterschied in der mittleren Gehstrecke zum Ausgangswert und Woche 24.

Veränderung des HRCT-QLF Scores

Die HRCT ist ein elementarer Bestandteil in der Diagnosestellung der IPF (Behr, 2013). Zudem wurden HRCT-basierte Algorithmen wie der HRCT-QLF Score entwickelt, anhand derer eine Quantifizierung des Fibrosierungsgrades bei fibrotisierenden Lungenerkrankungen vorgenommen werden kann (Kim et al., 2010).

Der HRCT-QLF Score wurde ursprünglich für Patienten mit systemischer Sklerodermie mit Lungenbeteiligung entwickelt und ist ein standardisiertes computergestütztes Verfahren zur automatischen Klassifizierung und Quantifizierung des Fortschritts der Fibrosierung der Lunge

(Kim et al., 2010). Der Fibrosierungsgrad wird dabei anhand von HRCT-Aufnahmen über die gesamte Ausdehnung der Lunge in einem Score abgebildet und ermöglicht im Setting klinischer Studien so eine objektive Einschätzung des Krankheitsprogresses. Niedrigere HRCT-QLF Score-Werte entsprechen dabei einer geringer ausgeprägten Lungenfibrose. Die Veränderung des Scores im Krankheitsverlauf korreliert dabei mit Veränderungen der FVC (Kim et al., 2011).

Ein entsprechender Zusammenhang zwischen dem HRCT-QLF Score und der Abnahme der FVC konnte ebenfalls im Anwendungsgebiet IPF nachgewiesen werden: Nach einer Studie von Goldin et al. hat die Veränderung des HRCT-QLF Scores zu Woche 24 prädiktive Aussagekraft über die weitere Abnahme der FVC. Demnach ging eine Zunahme des HRCT-QLF Scores von mehr als 4% mit einer Abnahme der FVC von 7,7% innerhalb eines Jahres einher, während für Patienten mit geringerer Zunahme des HRCT-QLF Scores eine durchschnittliche Abnahme der FVC von 3,5% beobachtet wurde (Goldin et al., 2017).

Als Endpunkt, welcher Rückschlüsse auf die Verschlechterung des Krankheitsverlaufs der IPF zulässt, ist der HRCT-QLF Score als patientenrelevant anzusehen.

Die *Veränderung des HRCT-QLF Scores* vom Ausgangswert bis Woche 24 wurde als primärer Endpunkt in der Studie 1199.187 erhoben. Der Endpunkt war definiert als Verhältnis des HRCT-QLF Scores von Woche 24 zum Ausgangswert.

Respiratorisch bedingte Hospitalisierungen

Die IPF ist eine progredient verlaufende Erkrankung, im Zuge derer durch auftretende Verschlechterungen des Gesundheitszustandes Hospitalisierungen notwendig werden können. Im vorliegenden Dossier werden Hospitalisierungen, welche aufgrund respiratorischer Probleme auftraten, als Wirksamkeitsendpunkt dargestellt.

Respiratorisch bedingte Hospitalisierungen sind mit einer sehr schlechten Prognose für den Patienten assoziiert, was letztendlich zu einer hohen Krankenhausmortalität unter IPF-Patienten führt (Brown et al., 2015; Cottin et al., 2017). In einer Studie von Cottin et al. erhöhte sich die Krankenhausmortalität von IPF-Patienten von 12,8% während der ersten Hospitalisierung auf 25,9% in Jahr 1 und stieg schließlich bis auf 43,0% in Jahr 5 an (Cottin et al., 2017). Das Transplantations-freie Überleben nach einer Hospitalisierung war deutlich geringer für Patienten, die respiratorisch bedingt hospitalisiert wurden (2,8 Monate) im Vergleich zu Patienten, die nicht respiratorisch bedingt hospitalisiert wurden (27,7 Monate) (Brown et al., 2015).

Die limitierten therapeutischen Optionen, die eingeschränkte Lebensdauer der Patienten nach Ende der Hospitalisierung sowie die der Hospitalisierung vorausgehende Symptomatik, welche für den Patienten deutlich spürbar und äußerst belastend sein können, verdeutlichen die Patientenrelevanz dieses Endpunkts.

Der Endpunkt *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* wurde in der Studie 1199.187 erhoben. Die Einschätzung, ob eine Hospitalisierung aufgrund respiratorischer Probleme auftrat, wurde durch den Prüfarzt vorgenommen.

Gesundheitsbezogene Lebensqualität

Die IPF ist eine schwere chronische Erkrankung, die mit zunehmender Dauer mit einer erheblichen Einschränkung der Gesundheitsbezogenen Lebensqualität (HRQoL) einhergeht. Deren Verbesserung oder Erhalt stellt ein wichtiges Therapieziel dar. Gemäß der AM-NutzenV gehört die Verbesserung der HRQoL zu den zentralen betrachteten Nutzendimensionen. Die HRQoL der Studienteilnehmer wird in klinischen Studien anhand von validierten Fragebogeninstrumenten sowohl spezifisch für ein bestimmtes Krankheitsbild oder Organsystem als auch generisch erfasst.

In den berücksichtigten Studien wurde die HRQoL mittels der folgenden spezifischen Fragebogeninstrumente beurteilt.

Krankheitsspezifische Messinstrumente:

- *SGRQ*
- *SGRQ-I*
- *SOBQ*
- *CASA-Q* (nur INPULSIS Studien)

Generische Messinstrumente:

- *PGI-C* (nur INPULSIS Studien)
- *EQ-5D-VAS* (nur INPULSIS Studien)

Bei der Beurteilung der krankheitsspezifischen und der generischen gesundheitsbezogenen Lebensqualität mittels der oben aufgeführten Fragebögen handelt es sich um weltweit eingesetzte und validierte Instrumente zur Bewertung der gesundheitsbezogenen Lebensqualität bzw. des allgemeinen Gesundheitszustandes (Aronson, 1993; Osoba et al., 1998; Velikova et al., 2010).

St. George's Respiratory Questionnaire (SGRQ)

Der *SGRQ* ist ein mehrdimensionales Erhebungsinstrument zur Dokumentation der krankheitsspezifischen Lebensqualität speziell für Patienten mit einer respiratorischen Erkrankung. Dieser validierte Fragebogen wurde von Jones et al. ursprünglich für die COPD entwickelt (Jones et al., 1992). Betrachtet wird die gesundheitsbezogene Lebensqualität bezogen auf drei Domänen: *Symptome* (Leiden aufgrund respiratorischer Symptome), *Aktivität* (Beeinträchtigung der Mobilität und körperlichen Aktivität) und *Impact* (Auswirkungen psychosozialer Einflüsse der Erkrankung auf das tägliche Leben und das Wohlbefinden).

Zwar existiert mit dem SGRQ-I eine IPF-spezifische Form des SGRQ, allerdings ist dieser auch in seiner ursprünglichen Form ein im Anwendungsgebiet IPF häufig verwendetes Messinstrument (Swigris et al., 2014; Furukawa et al., 2017). Es zeigt sich eine moderate bis starke Korrelation des SGRQ mit einer Vielzahl anderer Scores zur gesundheitsbezogenen Lebensqualität oder Symptomatik (Furukawa et al., 2017; Swigris et al., 2017). Seit Kurzem liegt zudem eine Validierungsstudie zum SGRQ im Anwendungsgebiet IPF vor, im Zuge derer eine moderate bis starke Korrelation zwischen dem SGRQ mit einer Reihe weiterer im Anwendungsgebiet IPF verbreiteter und anerkannter Endpunkte zur gesundheitsbezogenen Lebensqualität wie CASA-Q, SOBQ und der VAS des EQ-5D nachgewiesen wurde (Swigris et al., 2018). Im Rahmen der frühen Nutzenbewertung von Pirfenidon hat das IQWiG den SGRQ als validiertes Instrument zur Bewertung der HRQoL von Patienten mit IPF anerkannt (IQWiG, 2011). Auch der G-BA hat im Zuge des ersten Nutzenbewertungsverfahrens von Nintedanib im Anwendungsgebiet IPF den SGRQ unterstützend zur Ableitung des Zusatznutzens herangezogen (Gemeinsamer Bundesausschuss, 2015b).

Die Patienten beantworteten den Fragebogen selbstständig. Aus den Antworten wurde ein Grad der Beeinträchtigung der krankheitsspezifischen Lebensqualität in einem Bereich von 0 (keine Beeinträchtigung/gute gesundheitsbezogene Lebensqualität) bis 100 (maximale Beeinträchtigung/schlechte gesundheitsbezogene Lebensqualität) ermittelt. Eine Verminderung des SGRQ-Gesamtscores um vier Punkte im Vergleich zum Ausgangswert ist allgemein als MCID der subjektiv empfundenen gesundheitsbezogenen Lebensqualität akzeptiert (Swigris et al., 2010a; Jones et al., 2011; Jones, 2014). Spezifisch im Anwendungsgebiet IPF wurde für Responderanalysen eine Änderung des SGRQ-Gesamtscores von vier bis fünf Punkten vorgeschlagen (Swigris et al., 2018).

Veränderung des SGRQ-Gesamtscores

Die Bestimmung des SGRQ erfolgte in den INPULSIS Studien bei Studienbeginn sowie zu Woche 4, 6, 12, 24 und 52. In der Studie 1199.187 wurde der SGRQ zu Studienbeginn sowie zu Woche 12 und 24 erhoben. Einige Patienten in der Studie 1199.187 wiesen eine kontrollierte Phase über 52 Wochen auf. Für diese Patienten fand zudem eine Bestimmung des SGRQ zu Woche 52 statt.

Die *Veränderung des SGRQ-Gesamtscores* wurde mittels eines MMRM unter Berücksichtigung aller vorhandenen Messwerte berechnet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Insbesondere berücksichtigt dieses Modell auch Daten von Studienteilnehmern mit fehlenden Messungen. Die Anzahl der analysierten Studienteilnehmer sowie die beobachteten MW mit SD zu Studienbeginn und –ende wurden tabellarisch dargestellt. Weiterhin zeigt die Tabelle die auf dem MMRM basierende *Veränderung des SGRQ-Gesamtscores* über den Auswertungszeitraum pro Behandlungsgruppe als adjustierte MWD mit SE und den daraus resultierenden Behandlungseffekt ebenfalls als adjustierte MWD mit SE zwischen den Behandlungsgruppen mit 95%-KI und p-Wert.

SGRQ-($\leq$ -4 Punkte)-Responder

Als *SGRQ-($\leq$ -4 Punkte)-Responder* wurde in den INPULSIS Studien a priori ein Patient definiert, dessen *SGRQ-Gesamtscore* in Woche 52 gegenüber dem individuellen Messwert zu Studienbeginn um mindestens 4 Punkte abfiel. Falls zu Woche 52 kein Messwert vorlag, wurde der entsprechende Patient konservativ als Non-Responder bewertet. Der Endpunkt Anteil *SGRQ-($\leq$ -4 Punkte)-Responder* wurde mittels logistischer Regression ausgewertet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Die tabellarische Ergebnisdarstellung zeigt die Anzahl der Studienteilnehmer, die in die Analyse eingegangen sind, sowie die Anzahl und den Anteil der *SGRQ-($\leq$ -4 Punkte)-Responder*. Der Behandlungseffekt wird durch das OR mit 95%-KI und p-Wert dargestellt. Eine entsprechende Analyse war in Studie 1199.187 nicht präspezifiziert, wurde jedoch für die vorliegende Nutzenbewertung post hoc zu Woche 24 analog zu den INPULSIS Studien durchgeführt.

IPF-spezifischer St. George's Respiratory Questionnaire (SGRQ-I)

Mit dem *SGRQ-I* wurde eine IPF-spezifische Form des *SGRQ* entwickelt, die auf Basis des erhobenen *SGRQ* durch einen präspezifizierten Algorithmus ermittelt werden kann (Yorke et al., 2010). Das Ziel dieser Weiterentwicklung war laut Autoren die Generierung einer IPF-spezifischen *SGRQ*-Fragebogenversion, die spezifisch die Einschränkungen der HRQoL durch die IPF abbildet. Dieser für IPF-Patienten angepasste Fragebogen wurde hinsichtlich seiner Reliabilität im Vergleich zum *SGRQ* und im Hinblick auf seine Validität mit Messungen der Atemphysiologie, der funktionalen Kapazität, der Dyspnoe und des Gesundheitsstatus evaluiert (Swigris et al., 2010a; Yorke et al., 2010). Der *SGRQ-I* besteht ebenfalls aus den drei Domänen *Symptome*, *Aktivität* und *Impact*. Aus den Antworten wird – analog dem *SGRQ* – ein Grad der Beeinträchtigung der krankheitsspezifischen Lebensqualität in einem Bereich von 0 (keine Beeinträchtigung/gute gesundheitsbezogene Lebensqualität) bis 100 (maximale Beeinträchtigung/schlechte gesundheitsbezogene Lebensqualität) ermittelt. Für die Indikation IPF gibt es derzeit noch keine allgemein akzeptierte MCID. Um eine Responderanalyse durchführen zu können, wurde in Anlehnung an die MCID für einen klinisch relevanten Effekt bzgl. des *SGRQ-Gesamtscores* auch für den *SGRQ-I* ein Schwellenwert von -4 Punkten als MCID festgelegt.

Veränderung des SGRQ-I-Gesamtscores

Die Bestimmung des *SGRQ-I* erfolgte, wie im CTP präspezifiziert, aus den vorliegenden *SGRQ*-Daten. Die Auswertung und Darstellung der *Veränderung des SGRQ-I-Gesamtscores* erfolgte analog zu derjenigen für die *Veränderung des SGRQ-Gesamtscores*. Eine entsprechende Analyse war in Studie 1199.187 nicht präspezifiziert, wurde jedoch für die vorliegende Nutzenbewertung post hoc analog zu den INPULSIS Studien durchgeführt.

SGRQ-I-($\leq$ -4 Punkte)-Responder

Die Bestimmung der *SGRQ-I-($\leq$ -4 Punkte)-Responder* erfolgte analog zur Bestimmung der *SGRQ-($\leq$ -4 Punkte)-Responder*. Dieser Endpunkt war post hoc definiert.

Shortness of Breath Questionnaire (SOBQ)

Beim SOBQ handelt es sich um ein validiertes Fragebogeninstrument, das die Selbsteinschätzung der Beeinträchtigung der Aktivitäten des täglichen Lebens durch Dyspnoe widerspiegelt (Archibald und Guidotti, 1987; Eakin et al., 1998). Zudem liegt Evidenz vor, die die Validität dieses ursprünglich für COPD-Patienten entwickelten Messinstruments spezifisch für das Anwendungsgebiet IPF stützt (Swigris et al., 2012; Gries et al., 2013). Die Beurteilung der Schwere der Atemnot während täglicher Aktivitäten erfolgt auf einer 120-Punkte-Skala, wobei ein großer Gesamtscore eine schwere Beeinträchtigung durch Atemnot reflektiert. Als MCID wurde eine Veränderung von fünf Punkten angegeben (Kupferberg et al., 2005).

Veränderung des SOBQ-Gesamtscores

Die Bestimmung des SOBQ-Gesamtscores erfolgte in den INPULSIS Studien bei Studienbeginn sowie zu Woche 4, 6, 12, 24 und 52. In der Studie 1199.187 wurde der SOBQ-Gesamtscore zu Studienbeginn sowie zu Woche 12 und 24 erhoben. Für Patienten in der Studie 1199.187 mit einer kontrollierten Phase über 52 Wochen fand zudem eine Erhebung des Endpunkts zu diesem Zeitpunkt statt. Die Veränderung des SOBQ-Gesamtscores wurde analog zum Vorgehen für die Veränderung des SGRQ-Gesamtscores berechnet und dargestellt.

Cough and Sputum Assessment Questionnaire (CASA-Q)

Der CASA-Q ist in der Indikation COPD und chronische Bronchitis ein validiertes Fragebogeninstrument und dient der Evaluierung der Intensität der klinischen Symptome Husten und Sputum (*Husten-Symptome*, *Sputum-Symptome*), sowie deren Auswirkungen auf die HRQoL (*Husten-Impact*, *Sputum-Impact*) (Crawford et al., 2008). Die Validität dieses ursprünglich für COPD-Patienten entwickelten Messinstruments wurde für IPF-Patienten bestätigt (Gries et al., 2013). Die Beurteilung erfolgt auf einer 100-Punkte-Skala, wobei ein geringerer Wert eine stärkere Intensität bzw. Beeinträchtigung durch die genannten Symptome anzeigt.

Veränderung des CASA-Q (Husten-Symptome bzw. Husten-Impact)

In den INPULSIS Studien wurden ausschließlich die beiden Husten-Domänen (*Husten-Symptome* und *Husten-Impact*) des CASA-Q erhoben, da Sputum als Symptom bei der IPF – im Gegensatz zur COPD – eine untergeordnete Rolle spielt. Die Bestimmung des CASA-Q erfolgte bei Studienbeginn sowie zu Woche 4, 6, 12, 24 und 52. Die Veränderung des CASA-Q (*Husten-Symptome* bzw. *Husten-Impact*) wurde analog zum Vorgehen für die Veränderung des SGRQ-Gesamtscores berechnet und dargestellt.

Patient's Global Impression of Change (PGI-C)

Seit den Anfängen der Heilkunst ist die Frage an den Patienten nach dem aktuellen Befinden im Vergleich zum Zustand vor Beginn einer Behandlung ein zentrales Element bei der Wirksamkeitsbewertung durch den Arzt. Der *PGI-C* stellt den Versuch dar, die Antwort auf diese ureigene ärztliche Frage („Wie geht es Ihnen im Vergleich zu früher?“) im Kontext einer klinischen Studie in standardisierter Weise zu dokumentieren. Mit Hilfe des *PGI-C* gibt der Patient seinen subjektiven Eindruck zur Verbesserung oder Verschlechterung seines Gesundheitszustandes im Vergleich zum Referenzzeitpunkt bei Studienbeginn an. Die Änderungen werden auf einer 7-Punkte-Likert-Skala von eins (sehr viel besser) über vier (unverändert) bis sieben (sehr viel schlechter) dokumentiert. Bei einer chronisch progredienten Erkrankung wie der IPF ist ohne eine adäquate Therapie von einer stetigen Verschlechterung über die Zeit auszugehen. Das Ausbleiben einer solchen subjektiven Verschlechterung stellt einen Therapieerfolg und somit einen patientenrelevanten Nutzen dar.

PGI-C-Responder

Die *PGI-C* wurde ausschließlich in den INPULSIS Studien erhoben und wurde zu Woche 12, 24 und 52 bestimmt. Als *PGI-C-Responder* wurde a priori ein Patient definiert, der bei der Erhebung zu Woche 52 verglichen mit seinem subjektiv erinnerten Befinden bei Studienbeginn auf der 7-Punkte-Likert-Skala keine Kategorie der Verschlechterung seines Zustandes angegeben hatte, also seinen Gesundheitszustand gegenüber Studienbeginn als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „unverändert“ bewertete. Falls zu Woche 52 kein Messwert vorlag, wurde der entsprechende Patient konservativ als Non-Responder bewertet. Der Endpunkt Anteil *PGI-C-Responder* wurde mittels logistischer Regression ausgewertet. Die entsprechenden Adjustierungsparameter finden sich in den Ergebnistabellen des Abschnitts 4.3.1 als Fußnoten. Die tabellarische Ergebnisdarstellung erfolgte in Analogie zur Darstellung für den Endpunkt Anteil *SGRQ*-(≤ 4 Punkte)-Responder.

European Quality of Life-5 Dimensions (EQ-5D)

Der *EQ-5D* ist ein unter anderem für respiratorische Erkrankungen validiertes Fragebogeninstrument zur generischen Erfassung des allgemeinen Gesundheitszustandes zu den fünf Dimensionen *Mobilität, Fähigkeit sich selbst zu versorgen, alltägliche Aktivitäten, Schmerz/körperliche Beschwerden* und *Ängste/Niedergeschlagenheit* (Brazier et al., 1993; Rabin und de Charro, 2001; de Boer et al., 2004; Pickard et al., 2008). Durch eine weitere Domäne, die *EQ-5D-VAS*, bewertet der Patient zudem gesamthaft seinen Gesundheitsstatus auf einer 100-Punkte-Skala, wobei niedrigere Werte einen schlechteren Gesundheitszustand anzeigen. Im Kontext medizinischer Nutzenbewertungen wird der *EQ-5D* (beispielsweise durch das britische National Institute for Health and Care Excellence [NICE]) breit eingesetzt.

Veränderung der EQ-5D-VAS

In den INPULSIS Studien wurde der *EQ-5D* bei Studienbeginn sowie zu Woche 12, 24 und 52 bestimmt. Die *Veränderung der EQ-5D-VAS* wurde analog zum Vorgehen für die *Veränderung des SGRQ-Gesamtscores* berechnet und dargestellt.

Unerwünschte Ereignisse

Die Verringerung von UE stellt nach der AM-NutzenV einen patientenrelevanten Endpunkt dar und ist geeignet, im Rahmen einer Nutzenbewertung für die Ableitung eines Zusatznutzens herangezogen zu werden (Bundesministerium der Justiz und für Verbraucherschutz, 2010). Auch das IQWiG erkennt die Verringerung von Nebenwirkungen als patientenrelevanten Endpunkt an (IQWiG, 2017).

In den berücksichtigten Studien wurden UE durch die folgenden Endpunkte erfasst:

- *UE*
- *SUE*
- *UE, die zum Therapieabbruch führten*
- *UE von besonderem Interesse*

Als *UE von besonderem Interesse* wurden in der vorliegenden Nutzenbewertung *gastrointestinale Störungen (Diarrhö, Übelkeit, Erbrechen, Appetitlosigkeit, Gewichtsverlust, gastrointestinale Perforation), Leberenzym- und Bilirubinerhöhungen, kardiale Ereignisse, respiratorische Ereignisse (Progression der IPF, Husten, Dyspnoe, Bronchitis, Nasopharyngitis), Blutungsereignisse* sowie alle weiteren Störungen, die in mindestens einem Studienarm mit einer Häufigkeit von $\geq 10\%$ der Studienteilnehmer auftraten, betrachtet (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a). Alle *UE* wurden in den berücksichtigten Studien routinemäßig bei jeder Studienvisite ab der Randomisierung bis 28 Tage nach der letzten Einnahme der Studienmedikation erfasst. Die Anzahl der Studienteilnehmer, die Anzahl und der Anteil der Studienteilnehmer mit einem UE der betreffenden Kategorie wurden tabellarisch dargestellt. Der Gruppenunterschied wurde als RR, OR und Absolute Risikodifferenz (RD), jeweils mit 95%-KI und p-Wert mittels stratifiziertem Cochran-Mantel-Haenszel Test angegeben. In den INPULSIS Studien sowie in Studie 1199.187 waren *Leberenzym erhöhungen*, die Hy's law-Kriterien erfüllen sowie Ereignisse, die im Zusammenhang mit einer *gastrointestinalen Perforation* stehen, als UE von besonderem Interesse präspezifiziert (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2012a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013).

Darstellung und Auswertung der Endpunkte

Analysesets

In den eingeschlossenen Studien wurden Analysen bezüglich der Mortalität, Morbidität, gesundheitsbezogenen Lebensqualität und UE gemäß SAP auf Basis des Treated Set ausgewertet. Das Treated Set bestand aus den Studienteilnehmern, für die die Einnahme von mindestens einer Dosis der Studienmedikation dokumentiert wurde. Dem gegenüber schloss das Randomized Set alle randomisierten Studienteilnehmer ein, unabhängig davon, ob mindestens eine dokumentierte Dosis der Studienmedikation verabreicht wurde. Die Anzahl der Studienteilnehmer im Treated Set unterschied sich lediglich in zwei von insgesamt 515 randomisierten Studienteilnehmern der INPULSIS 1 Studie bzw. drei von insgesamt 551 randomisierten Studienteilnehmern der INPULSIS 2 Studie, die keine Dosis der Studienmedikation erhalten hatten, von dem Randomized Set (entsprechend 0,39% bzw. 0,54% des Randomized Set); in der Studie 1199.187 wurden alle randomisierten Patienten behandelt.

Studienphase

Für die Studie 1199.187 wurden, abweichend von der Präspezifizierung und der Darstellung im Studienbericht, für die vorliegende Nutzenbewertung die Daten der gesamten kontrollierten Phase herangezogen, um den Anforderungen der Nutzenbewertung entsprechend die längste unverzerrt bewertbare Beobachtungsdauer zugrunde zu legen.

Im Rahmen eines globalen Amendments (nach erfolgter Zulassung) war die ursprünglich geplante kontrollierte Behandlungsphase der Studie von 52 Wochen auf 24 Wochen verkürzt⁴ worden. Im Studienbericht waren der Präspezifizierung durch den SAP folgend für Wirksamkeitsendpunkte nur Daten bis zu Woche 24 berücksichtigt und auf einer *observed cases* (OC)-Basis ausgewertet worden.

Die Ergebnisse der Wirksamkeits-Endpunkte wurden für die Studien INPULSIS 1, INPULSIS 2, die Studie 1199.187 sowie für die gepoolten INPULSIS Studien mit und ohne Hinzunahme der Studie 1199.187 ausgewertet und dargestellt. Für die Bewertung der UE, die auf dem Zählen von Einzelereignissen in der Gesamtpopulation beruht, wurde auf eine Darstellung auf Einzelstudien-Ebene verzichtet und nur die gepoolte Patientenpopulation betrachtet. Vor dem Hintergrund des identischen Studiendesigns zwischen den INPULSIS Studien und des vergleichbaren Designs der Studie 1199.187, sowie ähnlicher Inzidenzen der erhobenen UE in den Einzelstudien, hat die gepoolte Darstellung der UE eine höhere Trennschärfe und ist damit aussagekräftiger für die Grundgesamtheit. Eine Darstellung auf Einzelstudien-Ebene findet sich in Modul 5 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2017; Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

⁴ Es ist zu beachten, dass bei einem Teil der Patienten die Dauer der kontrollierten Phase nach oben abweicht (in Abhängigkeit vom Zeitpunkt des Studieneinschlusses relativ zum Zeitpunkt des globalen Amendments).

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 Effekten⁵ 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 verwendet⁶ werden.

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 Heterogenität⁷ 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.

⁵ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

⁶ 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.

⁷ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Für die Nutzenbewertung für Nintedanib im Anwendungsgebiet IPF liegen drei RCT vor. Die Studien INPULSIS 1 und INPULSIS 2 sind sogenannte Zwillingstudien (Replikationsstudien) und folgen einem identischen Protokoll. Studie 1199.187 ist in Bezug auf die verwendeten Einschlusskriterien, Design-Aspekte sowie die Operationalisierung von Endpunkten hinreichend ähnlich, sodass eine meta-analytische Zusammenfassung der drei Studien sinnvoll ist, unterscheidet sich allerdings in der Dauer der kontrollierten Beobachtung.

Für die vorliegende Nutzenbewertung wurde eine Meta-Analyse auf Basis Individueller Patientendaten (IPD) durchgeführt. Wie im IQWiG-Methodenpapier 5.0 beschrieben (IQWiG, 2017), erfolgte die Auswertung der Daten dabei auf Patientenebene im Rahmen eines allgemeinen statistischen Modells, in das die Studie als Effekt und nicht als Beobachtungseinheit einging. Entsprechend der Empfehlung zur Methodik der Evidenz-Synthese in Situationen, in denen nur wenige Studien vorliegen (Bender et al., 2018), wurde für die drei Studien ein gemeinsamer Effekt angenommen und für die Meta-Analysen zu den einzelnen Endpunkten jeweils ein Modell mit festen Effekten (FEM) verwendet. Modelle mit zufälligen Effekten kamen nicht zur Anwendung.

Zur Einschätzung einer möglichen Heterogenität zwischen den Studienergebnissen wurden im Rahmen der IPD Meta-Analyse Interaktionstests durchgeführt (Interaktionsterm Behandlung×Studie). Die p-Werte der entsprechenden Tests wurden im Dossier berichtet. War die statistische Heterogenität nicht bedeutsam, d. h. $p > 0,05$ (Heterogenitätstest auf Basis der IPD), so wurde ein zusammenfassender Schätzer berechnet. Im Fall relevanter Heterogenität ($p < 0,05$) wurden zusammenfassende Effektschätzer für die Zwillingstudien INPULSIS berechnet und dem Effektschätzer aus der Studie 1199.187 qualitativ gegenübergestellt; mögliche Ursachen der Heterogenität wurden ggf. mittels Sensitivitäts- und Subgruppenanalysen untersucht. Die Interpretierbarkeit der Schätzer wurde in diesem Fall kritisch diskutiert.

Zur Visualisierung der Heterogenität und der Ergebnisse der Einzelstudien, sowie um den Vorgaben der Formatvorlage gerecht zu werden, wurden die Resultate der Einzelstudien mit Hilfe eines Forest-Plots dargestellt. Als zusammenfassender Schätzer wurde jeweils das Resultat der IPD Meta-Analyse dargestellt, dieses ist maßgeblich für die Bewertung des Zusatznutzens von Nintedanib.

Auswertungszeitpunkte und Studienpools

Studie 1199.187 weist mit sechs Monaten eine kürzere doppelblinde und placebokontrollierte Beobachtungsdauer auf als die INPULSIS Studien (zwölf Monate). Für alle drei Studien lagen Ergebnisse nach 24 Wochen, für die INPULSIS Studien auch nach Woche 52 vor. Dementsprechend erfolgte eine Meta-Analyse für Endpunkte, die sich auf feste Auswertungszeitpunkte beziehen und als Veränderung zu Woche 52 bzw. zu Woche 24 operationalisiert wurden, für den 24 Wochen Zeitpunkt über alle drei Studien (INPULSIS und Studie 1199.187 [gepoolt]) und für den 52 Wochen Zeitpunkt auf Basis der beiden INPULSIS Studien (INPULSIS [gepoolt]). Darüber hinaus wurden für Endpunkte, die durch ihre Operationalisierung unterschiedliche Beobachtungsdauern einbeziehen und berücksichtigen, wie z. B. Endpunkte, die über die Zeit bis zum ersten Ereignis operationalisiert oder

mithilfe analysierter Ereignisraten analysiert wurden, sowie für alle Sicherheitsendpunkte, alle drei Studien meta-analytisch zusammengefasst ungeachtet der unterschiedlichen Beobachtungsdauern der Einzelstudien. Für Studie 1199.187 wurde hierfür die gesamte kontrollierte Beobachtungsphase herangezogen. Für die Auswertungen zu Woche 24 wurden nur Daten bis Woche 24 berücksichtigt.

Da es sich bei der IPF um eine chronische Erkrankung handelt, sind die Ergebnisse der Analysen zum späteren Zeitpunkt (52 Wochen) bzw. der Analysen, die die Informationen aus dem gesamten Zeitraum der kontrollierten Beobachtungsphase berücksichtigen, für die Nutzenbewertung maßgeblich.

Das Studiendesign der INPULSIS Studien und der Studie 1199.187, sowie die für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte mit den jeweils herangezogenen Studienpools ist in Abbildung 6 dargestellt.

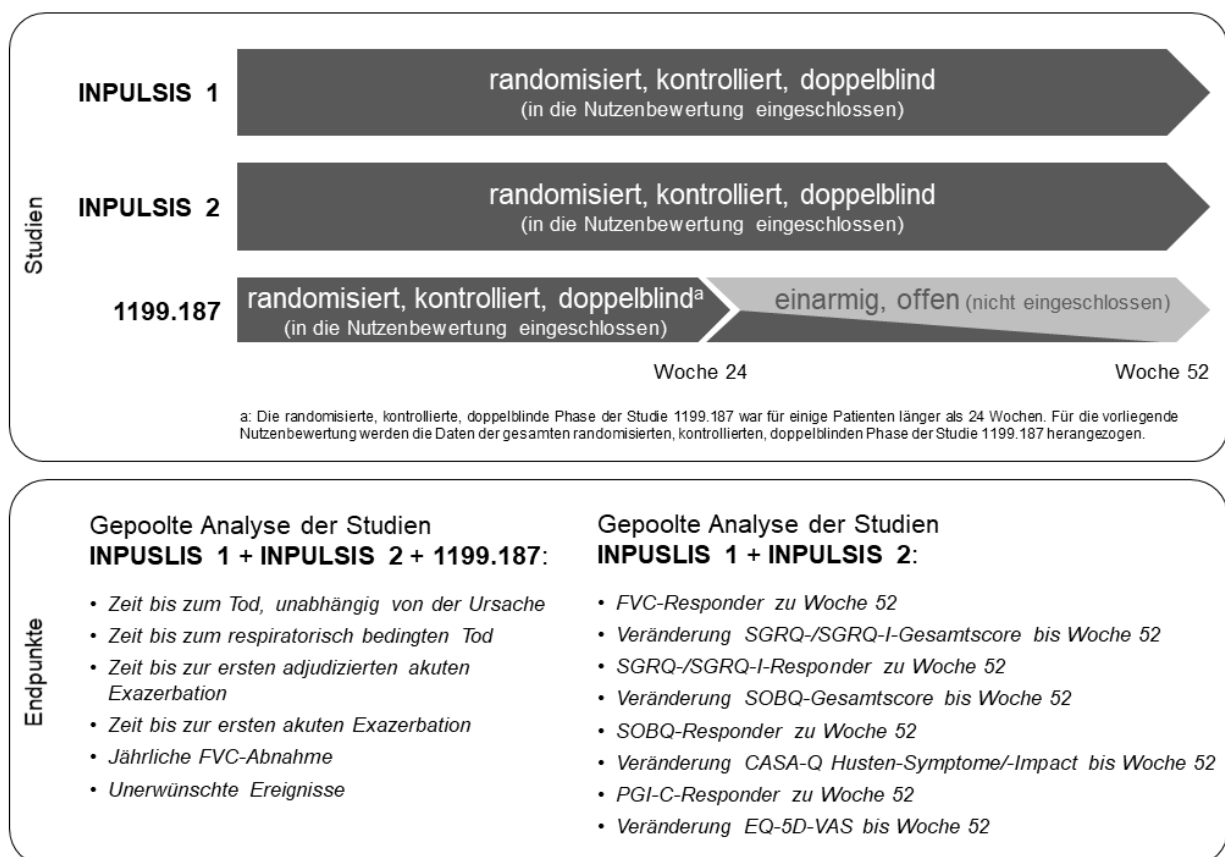


Abbildung 6: Studiendesign und für die Nutzenbewertung herangezogene Auswertungsmethodik der Endpunkte der INPULSIS Studien und Studie 1199.187

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.

In den Einzelstudien wurden sowohl für den primären Endpunkt als auch für weitere patientenrelevante Endpunkte Sensitivitätsanalysen durchgeführt. Diese bezogen sich beispielsweise auf das verwendete Analyseset, den *Estimand* (Beschränkung der Messwerte auf Werte unter Behandlung), das verwendete statistische Modell oder Ähnliches. Alle Sensitivitätsanalysen konnten die Robustheit der berichteten Ergebnisse demonstrieren. Eine Übersicht der Sensitivitätsanalysen, die sich auf die in diesem Dossier dargestellten Endpunkte beziehen, befindet sich in Modul 5 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018c).

Bewertung der klinischen Relevanz

Für stetige Endpunkte – und hier insbesondere skalenbasierte Endpunkte – wurden zur Bewertung der klinischen Relevanz von beobachteten Unterschieden Responderanalysen präsentiert, vorausgesetzt ein validiertes bzw. etabliertes Responsekriterium war vorhanden. War dies nicht der Fall, so wurde die klinische Relevanz eines beobachteten Unterschiedes anhand der standardisierten MWD (Hedges' *g*) bewertet. Gemäß des IQWiG-Methodenpapiers wurde von einem klinisch relevanten Effekt ausgegangen, wenn der Intervallschätzer für das Hedges' *g* vollständig außerhalb des klinisch irrelevanten Bereiches lag; der Irrelevanzbereich ist der Korridor von -0,2 bis 0,2 (IQWiG, 2017).

Analyse unerwünschter Ereignisse unter Ausschluss von Exazerbationen und Progression der Grunderkrankung

Zusätzlich zu der im SAP vorgesehenen Analyse aller erhobenen UE wurde für die vorliegende Nutzenbewertung eine Analyse der UE durchgeführt, die alle Ereignisse unberücksichtigt ließ, die im Case Report Form (CRF) im Rahmen der Erhebung der UE als Exazerbationen der Grunderkrankung erfasst wurden. Für die Ableitung des Zusatznutzens maßgeblich ist die Analyse, die alle erfassten UE einschloss.

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 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.

Die Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen wurden anhand von Interaktionstests bewertet.

Interaktionen wurden, basierend auf den kombinierten IPD, durch den Einschluss des subgruppenbildenden Faktors und des entsprechenden Interaktionsterms (Behandlung×subgruppenbildender Faktor) im statistischen Modell untersucht. Im Random-Coefficient-Regression-Modell wurden der Effekt Subgruppe und die Interaktionsterme Behandlung×Subgruppe, Zeit×Subgruppe und Behandlung×Subgruppe×Zeit im Modell aufgenommen. Im MMRM wurden der Effekt Subgruppe und der Interaktionsterm Behandlung×Subgruppe im Modell aufgenommen.

Ein p-Wert <0,05 wurde als Beleg für unterschiedliche Effekte bezeichnet. In diesem Fall wurden die Ergebnisse separat für die Subgruppen präsentiert und diskutiert. Die Ergebnisse von Subgruppenanalysen wurden stets hinsichtlich ihrer Validität (biologische Plausibilität, multiples Testen, Power, Präspezifizierung, qualitative versus quantitative Interaktion) diskutiert. Falls kein Beleg auf eine signifikante Interaktion zwischen Behandlung und dem subgruppenbildenden Faktor identifiziert wurde, wurde keine nach Subgruppen getrennte Analyse präsentiert.

In den INPULSIS Studien waren im CTP Subgruppenanalysen für folgende Merkmale vorgesehen: Geschlecht, Alter, japanische versus nicht-japanische Patienten, Ethnie, Körpergewicht, Körperoberfläche, Nierenfunktion basierend auf Kreatinin-Clearance und glomeruläre Filtrationsrate (GFR). Mit Ausnahme der Subgruppenbetrachtung für japanische versus nicht-japanische Patienten wurden alle Subgruppenanalysen nur für pharmakokinetische Endpunkte geplant und in den Studienberichten dargestellt.

In der Studie 1199.187 wurden gemäß CTP und SAP keine Subgruppen untersucht.

Für die Einreichung bei den Zulassungsbehörden wurden die folgenden Subgruppen und Trennpunkte für die gepoolte Analyse der INPULSIS Studien präspezifiziert:

- Geschlecht (männlich; weiblich)
- Alter (<65 Jahre; ≥65 Jahre)
- Ethnie (kaukasisch; asiatisch)
- Region (Europa; Amerika; Asien/Australien)
- FVC (% des Sollwertes) (≤70%; >70%)
- SGRQ-Gesamtscore (≤40 Punkte; >40 Punkte)
- Raucherstatus (Nichtraucher; Ex-Raucher/Raucher)
- Systemische Kortikosteroide zu Studienbeginn (ja; nein)
- Bronchodilatator zu Studienbeginn (ja; nein)
- Studie (INPULSIS 1; INPULSIS 2; Studie 1199.187)
- Niereninsuffizienz (ja; nein)
- Körpergewicht zu Studienbeginn (<65 kg; ≥65 kg)

Unabhängig davon, ob diese Faktoren für die Subgruppenbetrachtungen von Wirksamkeits- oder Sicherheitendpunkten festgelegt waren, wurden für das vorliegende Nutzendossier alle Subgruppenmerkmale für alle Endpunkte herangezogen.

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 möglich⁸. 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“⁹, „Multiple-Treatment-Meta-Analysen“¹⁰ oder auch „Netzwerk-Meta-Analysen“¹¹, 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 ist¹².

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 untersuchen¹³.

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.*

⁸ 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.

⁹ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹⁰ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹¹ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

¹² B. Schöttker, D. Lühmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹³ 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.

- *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.*
- *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.*

Es wurden im Rahmen der vorliegenden Nutzenbewertung keine indirekten Vergleiche durchgeführt.

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 4-8: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zu- lassungs- studie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen /abgebrochen /laufend)	Studiendauer	Therapiearme
TOMORROW (1199.30)	nein	ja	abgeschlossen	52 Wochen	Nintedanib 50 mg (1x täglich) 50 mg 100 mg 150 mg vs. Placebo je 2x täglich, per os
1199.31	nein	ja	abgeschlossen	14 Tage bzw. 28 Tage	Nintedanib 50 mg 100 mg 150 mg vs. Placebo je 2x täglich, per os
INPULSIS 1 (1199.32)	ja	ja	abgeschlossen	52 Wochen	Nintedanib 150 mg vs. Placebo je 2x täglich, per os
INPULSIS- ON (1199.33)	nein	ja	abgeschlossen	-	Nintedanib 100 mg 150 mg je 2x täglich, per os
INPULSIS 2 (1199.34)	ja	ja	abgeschlossen	52 Wochen	Nintedanib 150 mg vs. Placebo je 2x täglich, per os
1199.35	nein	ja	abgeschlossen	-	Nintedanib 50 mg (1x täglich) 50 mg 100 mg 150 mg je 2x täglich, per os; Studienteilnehmern wurde nach 3 Monaten angeboten auf die Dosierung 150 mg 2x täglich zu eskalieren
INSTAGE (1199.36)	nein	ja	abgeschlossen	24 Wochen	Nintedanib 150 mg 2x täglich plus Sildenafil 20 mg oder Placebo 3x täglich, per os
1199.40	nein	ja	abgeschlossen	-	Nintedanib 150 mg 2x täglich, per os (begleitend zu Pirfenidon)

Studie	Zu- lassungs- studie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen /abgebrochen /laufend)	Studiendauer	Therapiearme
1199.187	nein	ja	abgeschlossen	24-78 Wochen	Nintedanib 150 mg vs. Placebo je 2x täglich, per os
INJOURNEY (1199.222)	nein	ja	abgeschlossen	12 Wochen	Nintedanib 150 mg 2x täglich vs. Nintedanib 150 mg 2x täglich plus Pirfenidon bis zu 3x täglich, per os
INMARK (1199.227)	nein	ja	abgeschlossen	12-52 Wochen	Nintedanib 150 mg vs. Placebo je 2x täglich, per os in den ersten 12 Wochen, danach Nintedanib 150 mg 2x täglich, per os, für alle Patienten
1199.229	nein	ja	abgeschlossen	24 Tage	Nintedanib 150 mg bis zu 2x täglich plus Pirfenidon bis zu 3x täglich, per os
1199.305	nein	ja	laufend	12 Wochen	4 Zyklen Carboplatin plus nab-Paclitaxel (100 mg/m ²) alle 3 Wochen vs. 4 Zyklen Carboplatin plus nab-Paclitaxel (100 mg/m ²) kombiniert mit Nintedanib (150 mg bis zu 2x täglich) alle 3 Wochen gefolgt von Nintedanib (150 mg bis zu 2x täglich) als alleinige Medikation
1305.12	nein	ja	laufend	12 Wochen (Teil 1) + 4 Wochen (Teil 2)	Teil 1: BI 1015550 vs. Placebo Teil 2: Pirfenidon plus Nintedanib plus BI 1015550 vs. Pirfenidon plus Nintedanib plus Placebo
FITNESS (17082106)	nein	ja	laufend	52 Wochen	Nintedanib vs. Nintedanib plus pulmonale Rehabilitation

Studie	Zu- lassungs- studie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen /abgebrochen /laufend)	Studiendauer	Therapiearme
HOPE-IPF (1199.0253)	nein	nein	laufend	26 Wochen	Standard pulmonale Rehabilitation vs. pulmonale Rehabilitation mit 60% Sauerstoff
1199-0329	nein	ja	laufend	24 Monate	Nintedanib 150 mg 2x täglich vs. Placebo in IPF-Patienten mit Lungentransplantation
1199-0324	nein	ja	laufend	24 Wochen	Nintedanib 150 mg 2x täglich vs. Nintedanib 150 mg 2x täglich plus pulmonale Rehabilitation
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)					

Geben Sie an, welchen Stand die Information in Tabelle 4-8 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 Dokuments maßgeblichen Zeitpunkt liegen.

Der Informationsstand bezieht sich auf den 08.02.2019.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-8 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-9: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
1199.30 (TOMORROW)	Intervention abweichend von Vorgaben der Fachinformation aufgrund fehlender Option zur Re-Eskalation auf Startdosis nach zwischenzeitlicher Behandlungsunterbrechung oder Dosisreduktion
1199.31	Zu kurze Studiendauer (14-28 Tage)
1199.33 (IMPULSIS-ON)	Open-label Extension der 1199.32/1199.34; einarmige Studie
1199.35	Open-label Extension der 1199.30; nicht-randomisiert
1199.36 (INSTAGE)	Falscher Komparator
1199.40	Open-label Extension der 1199.31; einarmige Studie
1199.222 (INJOURNEY)	Falscher Komparator
1199.227 (INMARK)	Zu kurze Studiendauer (12 Wochen, danach einarmig)
1199.229	Nicht-randomisiert
1199.305	Zu kurze Studiendauer (12 Wochen)
1305.12	Zu kurze Studiendauer (16 Wochen)
FITNESS (17082106)	Falscher Komparator
HOPE-IPF (1199.0253)	Falscher Komparator
1199-0329	Laufende Studie ohne Ergebnisse
1199-0324	Falscher Komparator
RCT: Randomized Controlled Trial (randomisierte kontrollierte 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.

Die bibliografische Literaturrecherche wurde am 17.01.2019 durchgeführt und ergab insgesamt 614 Treffer. Nach Entfernung von Duplikaten (170) verblieben 444 Treffer, die anhand von Titel und Abstract gemäß den in Abschnitt 4.2.2 beschriebenen Kriterien selektiert wurden. Davon wurden 405 Publikationen als nicht relevant bewertet, die verbliebenen 39 Publikationen wurden im Volltext gesichtet. Auf Volltext-Ebene wurden 21 Publikationen als nicht relevant ausgeschlossen, 18 Publikationen wurden als relevant bewertet. Diesen liegen zwei Studien zugrunde.

Im Zuge der Suche in der Cochrane-Datenbank wurden auch Einträge aus dem clinicaltrials.gov-Register gefunden. Entsprechende Einträge wurden auf Basis des Abstracts sowie des entsprechenden Studienregistereintrages nach relevanten Studien durchsucht und anhand der in Abschnitt 4.2.2 beschriebenen Kriterien bewertet. Dabei wurden auf Basis des Abstracts und des Studienregistereintrages die Studien NCT01335477 (INPULSIS 2), NCT01979952 (Studie 1199.187) als relevant identifiziert. Beide Studien wurden auch im Zuge der Suche im clinicaltrials.gov-Register gefunden. Als nicht relevant bewertet wurden die Einträge zu den Studien NCT03283007, NCT02412020, NCT01890265, NCT02606877, NCT02579603, NCT02788474, NCT02802345, NCT02999178, NCT03069989 und NCT03422068.

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

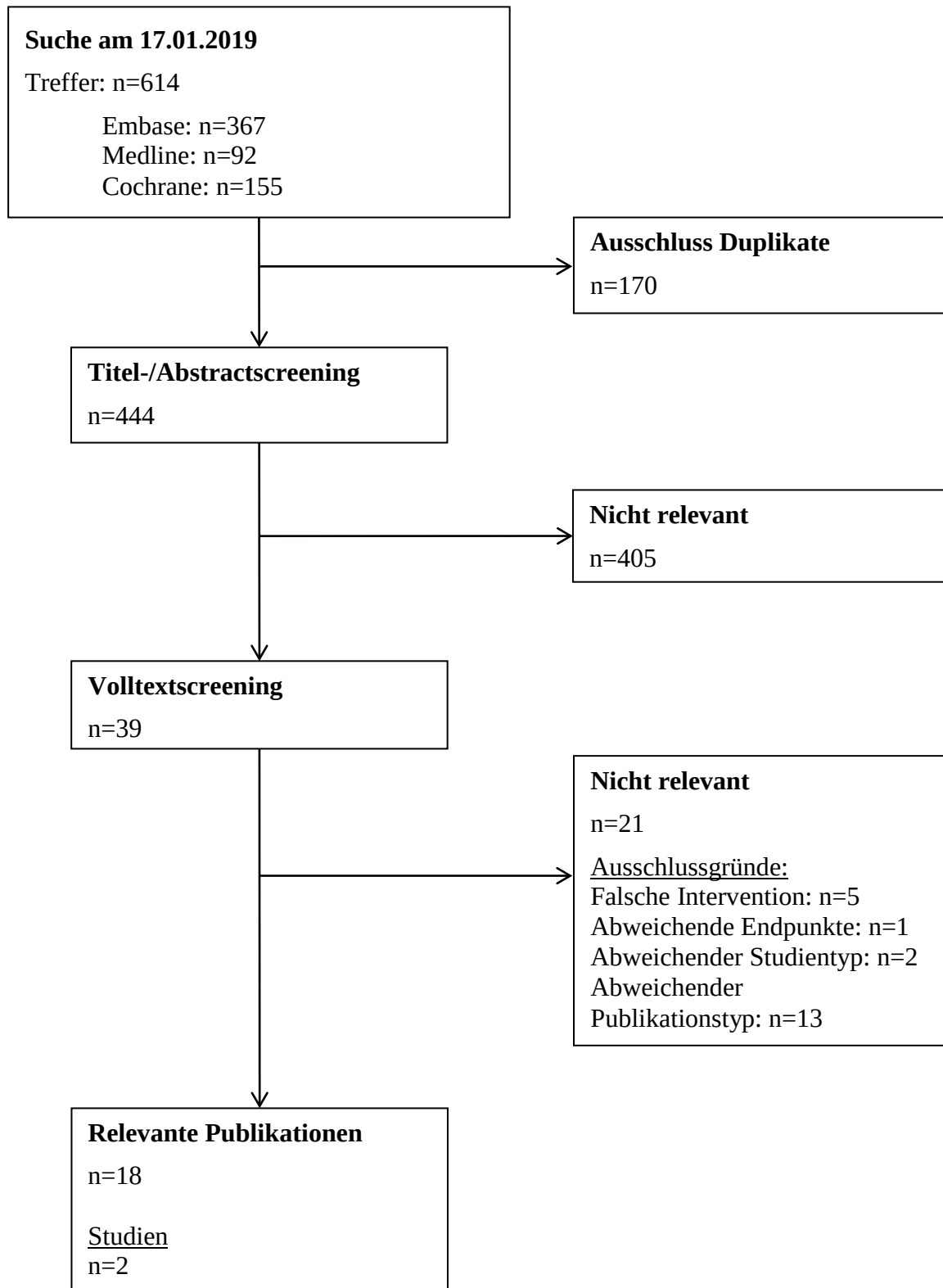


Abbildung 7: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Folgende Publikationen wurden im Zuge der bibliografischen Literaturrecherche als relevant identifiziert: (Richeldi et al., 2014; Corte et al., 2015; Goyard und Crestani, 2015; Keating, 2015; Costabel et al., 2016; Richeldi et al., 2016; Taniguchi et al., 2016; Wuyts et al., 2016; Azuma et al., 2017; Collard et al., 2017; Hajari Case und Johnson, 2017; Kolb et al., 2017; Raghu et al., 2017; Rinciog et al., 2017; Costabel et al., 2018; Flaherty et al., 2018; Kreuter et al., 2018; Brown et al., 2019).

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 4-8) 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 4-10: 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)
INPULSIS 1	ClinicalTrials.gov (NCT01335464) (ClinicalTrials.gov, 2016a) EU-CTR (2010-024251-87) (EU Clinical Trials Register, 2016b) WHO-ICTRP (EUCTR2010-024251-87- DE) (World Health Organization - International Clinical Trials Registry Platform, 2014)	ja	ja	abgeschlossen

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)
	WHO-ICTRP (CTRI/2011/12/002248) (World Health Organization - International Clinical Trials Registry Platform, 2018b) PharmNet.Bund (2010-024251-87) (PharmNet.bund, 2013b)			
INPULSIS 2	ClinicalTrials.gov (NCT01335477) (ClinicalTrials.gov, 2016b) EU-CTR (2010-024252-29) (EU Clinical Trials Register, 2016a) WHO-ICTRP (EUCTR2010-024252-29-FI) (World Health Organization - International Clinical Trials Registry Platform, 2013) WHO-ICTRP (CTRI/2011/12/002245) (World Health Organization - International Clinical Trials Registry Platform, 2018a) PharmNet.Bund (2010-024252-29) (PharmNet.bund, 2013a)	ja	ja	abgeschlossen
1199.187	ClinicalTrials.gov (NCT01979952) (ClinicalTrials.gov, 2018) WHO-ICTRP (NCT01979952) (World Health Organization - International Clinical Trials Registry Platform, 2018c)	ja	nein ^b	abgeschlossen

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)
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder –ergebnisse. b: Im Zuge der Suche in der Cochrane-Datenbank wurde ein Eintrag zur Studie 1199.187 im clinicaltrials.gov-Register gefunden. EU: Europäische Union; EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); WHO: World Health Organization (Weltgesundheitsorganisation)				

Geben Sie an, welchen Stand die Information in Tabelle 4-10 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 Dokuments maßgeblichen Zeitpunkt liegen.

Die Suche in den Studienregistern wurde am 17.01.2019 durchgeführt.

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 Zwischenzeilenüberschriften 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 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gespon-serte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studien-bericht (ja/nein [Zitat])	Register eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Placebokontrolliert						
Aktivkontrolliert, zVT(n)^d						
INPULSIS 1 (1199.32)	ja	ja	nein	ja (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a)	ja (PharmNet.bund, 2013b; World Health Organization - International Clinical Trials Registry Platform, 2014; ClinicalTrials.gov, 2016a; EU Clinical Trials Register, 2016b; World Health Organization - International Clinical Trials Registry Platform, 2018b)	ja (Richeldi et al., 2014; Corte et al., 2015; Goyard und Crestani, 2015; Keating, 2015; Costabel et al., 2016; Richeldi et al., 2016; Taniguchi et al., 2016; Wuyts et al., 2016; Azuma et al., 2017; Collard et al., 2017; Hajari Case und Johnson, 2017; Kolb et al., 2017; Raghu et al., 2017; Rinciog et al., 2017; Costabel et al., 2018; Flaherty et al., 2018; Kreuter et al., 2018; Brown et al., 2019)

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gespon-serte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studien-bericht (ja/nein [Zitat])	Register eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
INPULSIS 2 (1199.34)	ja	ja	nein	ja (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b)	ja (PharmNet.bund, 2013a; World Health Organization - International Clinical Trials Registry Platform, 2013; ClinicalTrials.gov, 2016b; EU Clinical Trials Register, 2016a; World Health Organization - International Clinical Trials Registry Platform, 2018a)	ja (Richeldi et al., 2014; Corte et al., 2015; Goyard und Crestani, 2015; Keating, 2015; Costabel et al., 2016; Richeldi et al., 2016; Taniguchi et al., 2016; Wuyts et al., 2016; Azuma et al., 2017; Collard et al., 2017; Hajari Case und Johnson, 2017; Kolb et al., 2017; Raghu et al., 2017; Rinciog et al., 2017; Costabel et al., 2018; Flaherty et al., 2018; Kreuter et al., 2018; Brown et al., 2019)
Studie 1199.187	nein	ja	nein	ja (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017)	ja (ClinicalTrials.gov, 2018; World Health Organization - International Clinical Trials Registry Platform, 2018c)	nein ^e

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gespon-sierte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Register eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 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. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder –ergebnisse. d: Patienten in den hier aufgeführten Studien erhielten im Vergleichsarm gemäß der Studienprotokolle ein Placebo. Wie in Abschnitt 4.3.1.2.1 beschrieben, stand diesen Patienten zudem eine Versorgung zur Verfügung, welche die Anforderungen des G-BA an die zVT BSC adäquat erfüllt. Die drei nutzenbewertungsrelevanten Studien werden im vorliegenden Dossier daher als Studien gegenüber einer Vergleichstherapie BSC angesehen. e: Im Zuge der Suche in der Cochrane-Datenbank wurde ein Eintrag zur Studie 1199.187 im clinicaltrials.gov-Register gefunden. BSC: Best Supportive Care; G-BA: Gemeinsamer Bundesausschuss; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); zVT: zweckmäßige Vergleichstherapie						

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 4-12: 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
INPULSIS 1 (1199.32)	zweiarmig, randomisiert (3:2), doppelblind, placebokontrolliert, multizentrisch, Phase III	Erwachsene (≥40 Jahre) mit IPF-Diagnose nach ATS/ERS/JRS/ ALAT- Richtlinien, DL _{CO} 30-79%, FVC ≥50%	<u>Nintedanib</u> 150 mg 2x täglich, per os: n=309 (bei Auftreten bestimmter UE ggf. Dosisreduktion auf 100 mg möglich) <u>Placebo</u> 2x täglich, per os: n=206	52 Wochen <i>Follow-up</i> : 28 Tage nach Beendigung der Therapie	98 Studienzentren in 13 Ländern (Australien, Belgien, China, Tschechische Republik, Frankreich, Deutschland, Indien, Irland, Israel, Italien, Japan, UK, USA) 5/2011 – 10/2013	<u>Primärer Endpunkt</u> : - Jährliche FVC-Abnahme (ml) <u>Key sekundäre Endpunkte</u> - Veränderung SGRQ Gesamtscore - Zeit bis zur ersten akuten Exazerbation (adjudiziert) <u>Sekundäre Endpunkte</u> : - Zeit bis zum Tod - Zeit bis zum respiratorisch bedingten Tod - Zeit bis zum Tod unter Behandlung - Zeit bis zum Tod oder Lungentransplantation ^a - Inzidenzrate einer akuten Exazerbation - absolute und relative Veränderung der FVC (ml) und FVC (%) - kategorielle Veränderung FVC (%) (Abfall >5%; Anstieg >5%; Änderung ≤5%) - kategorielle Veränderung FVC (%) (Abfall >10%; Anstieg >10%;

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<relevante Charakteristika, z. B. Schweregrad>		<ggf. Run-in, Behandlung, Nachbeobachtung>		Änderung $\leq 10\%$) - Anteil FVC-Responder (Schwellenwert 5% bzw. 10% des Sollwertes) - Veränderung SGRQ-Domänen vom Ausgangswert - Anteil SGRQ-Responder (≤ 4) - Veränderung SGRQ-I-Gesamtscore - Anteil SGRQ-I-Responder (≤ 4) - Veränderung SOBQ-Gesamtscore - Veränderung CASA-Q - Anteil PGI-C-Responder - Veränderung EQ-5D - Veränderung SpO₂ - Veränderung DL_{CO}
INPULSIS 2 (1199.34)	zweiarmig, randomisiert (3:2), doppelblind, placebokontrolliert, multizentrisch, Phase III	Erwachsene (≥ 40 Jahre) mit einer IPF- Diagnose nach ATS/ERS/JRS/ ALAT- Richtlinien, DL _{CO} 30-79%, FVC $\geq 50\%$	<u>Nintedanib</u> 150 mg 2x täglich, per os: n=331 (bei Auftreten bestimmter UE ggf. Dosisreduktion auf 100 mg möglich)	52 Wochen <i>Follow-up</i> : 28 Tage nach Beendigung der Therapie	107 Studienzentren in 17 Ländern (Kanada, Chile, China, Finnland, Frankreich, Deutschland, Griechenland, Indien, Japan, Korea, Mexiko, Niederlande, Portugal, Russland,	homogenes Studiendesign wie INPULSIS 1

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<relevante Charakteristika, z. B. Schweregrad>		<ggf. Run-in, Behandlung, Nachbeobachtung>		
			Placebo 2x täglich, per os: n=220		Spanien, Türkei, USA) 05/2011 – 10/2013	
Studie 1199.187 (1199.187)	Zweiarmig, randomisiert (1:1), doppelblind, placebokontrolliert, multizentrisch, Phase IIIb	Erwachsene (≥40 Jahre) mit einer IPF- Diagnose nach ATS/ERS/JRS ALAT- Richtlinien, DL _{CO} 30-79%, FVC ≥50%	<u>Nintedanib</u> 150 mg 2x täglich, per os: n=56 (bei Auftreten bestimmter UE ggf. Dosisreduktion auf 100 mg möglich) <u>Placebo</u> 2x täglich, per os: n=57	24-78 Wochen (≥24 Wochen kontrollierte Phase; danach offene Phase ^b) <i>Follow-up:</i> 28 Tage nach Beendigung der Therapie	26 Studienzentren in 3 Ländern (USA, Kanada, Türkei) 12/2013 – 10/2016	Nach Amendment 1 des Studienprotokolls, das die Verkürzung der kontrollierten Phase der Studie und die Reduktion der Stichprobenzahl bedeutete, wurden alle Studienendpunkte explorativ ausgewertet. <u>Primärer Endpunkt:</u> - Veränderung HRCT-QLF Score vom Ausgangswert bis Woche 24 <u>Sekundäre Endpunkte:</u> - Zeit bis zum Tod - Zeit bis zum respiratorisch bedingten Tod - Zeit bis zur ersten respiratorisch bedingten Hospitalisierung - Zeit bis zur ersten akuten Exazerbation - absolute und relative Veränderung der FVC (ml) und FVC (%) - kategorielle Veränderung FVC (%) (Abfall >5%; Anstieg >5%; Änderung ≤5%)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<relevante Charakteristika, z. B. Schweregrad>		<ggf. Run-in, Behandlung, Nachbeobachtung>		
						- kategorielle Veränderung FVC (%) (Abfall >10%; Anstieg >10%; Änderung ≤10%) - Veränderung SGRQ-Gesamtscore - Veränderung 6 MWT - Veränderung SOBQ-Gesamtscore
a: Zusammengesetzter Endpunkt, bei dem zwei verschiedene Ergebnissituationen für den Patienten vermischt werden. b: Während der offenen Phase erhielten die Patienten aus beiden Studienarmen Nintedanib 150 mg 2x täglich per os. 6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); ALAT: Latin American Thoracic Association; ATS: American Thoracic Society; CASA-Q: Cough and Sputum Assessment Questionnaire; DL_{CO}: Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); EQ-5D: European Quality of Life-5 Dimensions; ERS: European Respiratory Society; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); JRS: Japanese Respiratory Society; PGI-C: Patient's Global Impression of Change; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SpO₂: Sauerstoffsättigung; UE: Unerwünschtes Ereignis						

Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib	Placebo	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run- in-Phase etc.</i>
INPULSIS 1 (1199.32)	Nintedanib 150 mg 2x täglich per os	Placebo 2x täglich per os	Studienteilnehmer erhielten Nintedanib als Soft-Gelatine-Kapsel kontinuierlich in einer Dosis von 150 mg zweimal täglich. Placebo wurde ebenfalls als Soft-Gelatine-Kapsel zweimal täglich verordnet. Die Einnahme sollte mit einem Glas Wasser (ca. 250 ml) in einem Intervall von 12 Stunden, wenn möglich täglich zur gleichen Zeit, erfolgen (morgens zwischen 6.00 Uhr und 11.00 Uhr und abends zwischen 18.00 Uhr und 23.00 Uhr). Zur Vermeidung von Magenproblemen sollte die Einnahme nach dem Essen erfolgen. Beim Auftreten bestimmter UE konnte eine Dosisreduktion auf 100 mg 2x täglich oder eine Unterbrechung der Therapie erwogen werden. Bei therapieassoziierten UE sollte die Re-Initiierung nach ≤ 4 Wochen erfolgen (reduzierte Dosis von 100 mg 2x täglich empfohlen); bei nicht-therapieassoziierten UE sollte die Re-Initiierung nach ≤ 8 Wochen erfolgen (volle Dosis von 150 mg 2x täglich möglich). Eine Re-Eskalation auf 150 mg war bei Besserung der Beschwerden durch die UE innerhalb von 4 Wochen nach Initiierung der reduzierten Dosis zulässig. Im Falle einer erheblichen Toxizität oder wenn auch eine reduzierte Dosis nicht toleriert wurde, sollte die Therapie abgebrochen und nicht wieder aufgenommen werden.
INPULSIS 2 (1199.34)	Nintedanib 150 mg 2x täglich per os	Placebo 2x täglich per os	Homogenes Studiendesign wie INPULSIS 1
Studie 1199.187 (1199.187)	Nintedanib 150 mg 2x täglich per os	Placebo 2x täglich per os	Analog zu INPULSIS 1 und INPULSIS 2
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis			

Tabelle 4-14: Charakterisierung der Studienpopulation (Ausgangswert) – RCT mit dem zu bewertenden Arzneimittel

Studie	INPULSIS 1		INPULSIS 2		INPULSIS (gepoolt)		Studie 1199.187		INPULSIS + Studie 1199.187 (gepoolt)	
Gruppe	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo
N ^a	309	204	329	219	638	423	56	57	694	480
Demografie										
Alter (Jahre) MW (SD)	66,9 (8,4)	66,9 (8,2)	66,4 (7,9)	67,1 (7,5)	66,6 (8,1)	67,0 (7,9)	68,8 (7,6)	66,2 (9,4)	66,8 (8,1)	66,9 (8,1)
Geschlecht , N (%) - männlich	251 (81,2)	163 (79,9)	256 (77,8)	171 (78,1)	507 (79,5)	334 (79,0)	45 (80,4)	37 (64,9)	552 (79,5)	371 (77,3)
Gewicht (kg) MW (SD)	82,03 (16,80)	81,17 (16,30)	76,59 (15,93)	76,28 (16,46)	79,23 (16,57)	78,64 (16,54)	85,19 (13,74)	86,42 (17,86)	79,708 (16,43)	79,560 (16,87)
BMI (kg/m ²) MW (SD)	28,63 (4,52)	28,12 (4,63)	27,56 (4,55)	27,17 (4,49)	28,08 (4,57)	27,63 (4,58)	29,4 (3,8)	30,3 (5,4)	28,189 (4,524)	27,943 (4,761)
Ethnie , N (%) ^b - kaukasisch - asiatisch	198 (64,1) 66 (21,4)	135 (66,2) 41 (20,1)	162 (49,2) 128 (38,9)	113 (51,6) 86 (39,3)	360 (56,4) 194 (30,4)	248 (58,6) 128 (30,3)	54 (96,4) 2 (3,6)	54 (94,7) 3 (5,3)	414 (59,7) 196 (28,2)	302 (62,9) 131 (27,3)
Raucherstatus , N (%) - Nichtraucher - Ex-Raucher - Raucher	71 (23,0) 217 (70,2) 21 (6,8)	51 (25,0) 144 (70,6) 9 (4,4)	103 (31,1) 218 (66,3) 8 (2,4)	71 (32,4) 139 (63,5) 9 (4,1)	174 (27,3) 435 (68,2) 29 (4,5)	122 (28,8) 283 (66,9) 18 (4,3)	14 (25,0) 41 (73,2) 1 (1,8)	17 (29,8) 40 (70,2) 0	188 (27,1) 476 (68,6) 30 (4,3)	139 (29,0) 323 (67,3) 18 (3,8)
Zeit seit IPF-Diagnose (Jahre), MW (SD)	1,66 (1,37)	1,59 (1,35)	1,64 (1,34)	1,55 (1,27)	1,65 (1,36)	1,57 (1,31)	1,46 (1,39)	1,46 (1,36)	1,63 (1,36)	1,56 (1,31)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	INPULSIS 1		INPULSIS 2		INPULSIS (gepoolt)		Studie 1199.187		INPULSIS + Studie 1199.187 (gepoolt)	
Gruppe	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo
N ^a	309	204	329	219	638	423	56	57	694	480
Krankheitsspezifische Charakteristika										
Radiolog. Beurteilung der IPF, N (%) - UIP - vermutlich UIP - mögliche UIP - keine UIP	301 (97,4) n. e. 8 (2,6) 0	198 (97,1) n. e. 6 (2,9) 0	321 (97,6) n. e. 8 (2,4) 0	215 (98,2) n. e. 4 (1,8) 0	622 (97,5) n. e. 16 (2,5) 0	413 (97,6) n. e. 10 (2,4) 0	nicht berichtet	nicht berichtet	nicht berichtet	nicht berichtet
Lungenbiopsie N (%) ^c	60 (19,4)	33 (16,2)	84 (25,5)	52 (23,7)	144 (22,6)	85 (20,1)	nicht berichtet	nicht berichtet	nicht berichtet	nicht berichtet
FVC - ml, MW (SD) - % des Sollwertes, MW (SD)	2.756,8 (735,1) 79,47 (17,03)	2.844,5 (820,1) 80,53 (17,34)	2.672,8 (776,0) 79,99 (18,08)	2.619,0 (787,3) 78,09 (18,97)	2.713,5 (757,0) 79,74 (17,57)	2.727,7 (810,2) 79,27 (18,22)	2.996,8 (831,3) 78,0 (17,4)	2.921,3 (833,9) 78,1 (19,4)	2.736,3 (766,5) 79,59 (17,55)	2.750,7 (814,6) 79,12 (18,35)
FEV₁:FVC (%) MW (SD)	81,54 (5,36)	80,83 (6,13)	81,77 (6,25)	82,44 (5,72)	81,66 (5,83)	81,66 (5,97)	nicht berichtet	nicht berichtet	nicht berichtet	nicht berichtet
DL_{co} (mmol/min/kPa) MW (SD)	3,96 (1,20)	3,96 (1,11)	3,77 (1,23)	3,75 (1,32)	3,87 (1,22)	3,85 (1,23)	4,4 (1,2)	4,3 (1,4)	3,9 (1,23)	3,9 (1,26)
SpO₂ (%) MW (SD)	95,9 (2,0)	95,9 (1,9)	95,8 (2,6)	95,7 (2,1)	95,9 (2,3)	95,8 (2,0)	95,1 (2,4)	95,1 (2,5)	95,8 (2,3)	95,7 (2,1)
SGRQ-Gesamtscore (Punkte), MW (SD)	39,55 (17,63)	39,79 (18,48)	39,46 (20,47)	39,39 (18,65)	39,51 (19,15)	39,58 (18,54)	35,75 (17,49)	44,39 (18,49)	39,20 19,04	40,16 (18,58)
Bronchodilatator	61 (19,7)	34 (16,7)	68 (20,7)	38 (17,4)	131 (20,5)	72 (17,0)	13 (23,2)	16 (28,1)	144 (20,7)	88 (18,3)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	INPULSIS 1		INPULSIS 2		INPULSIS (gepoolt)		Studie 1199.187		INPULSIS + Studie 1199.187 (gepoolt)	
Gruppe	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo	Nintedanib	Placebo
N ^a	309	204	329	219	638	423	56	57	694	480
Systemische Kortikosteroid-Therapie N (%)	68 (22,0)	43 (21,1)	68 (20,7)	46 (21,0)	129 (20,2)	86 (20,3)	3 (5,4)	8 (14,0)	131 (18,9)	94 (19,6)
Therapieabbrecher	78 (25,2)	36 (17,6)	78 (23,7)	44 (20,1)	156 (24,5)	80 (18,9)	13 (23,2) ^d	14 (24,6) ^d	169 (24,4)	94 (19,6)
Studienabbrecher	49 (15,9)	30 (14,7)	57 (17,3)	40 (18,3)	106 (16,6)	70 (16,5)	4 (7,1) ^d	6 (10,5) ^d	110 (15,9)	76 (15,8)
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 10.1: 1, Tab. 15.1.4.1, Tab. 15.1.4.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.1.4.1, Tab. 15.1.4.2, Tab. 10.1: 1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c): Tab. 1.1.6, Tab. 1.1.7, Tab. 1.1.8, Tab. 1.1.10, Tab. 1.1.11, Tab. 1.1.12; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017): Tab. 15.1.1: 1, Tab. 15.1.4.1, Tab. 15.1.4.2, Tab. Table 15.1.4.3.1.3; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.69.15.7.1, Tab. 1.3.65.1.1, Tab. 1.3.65.1.2, Tab. 1.3.65.1.3, Tab. 1.3.65.1.4, Tab. 1.3.65.1.5; Tab. 1.3.69.1.8; Tab. 1.3.69.8.1; Tab. 1.3.70.1, Tab. 1.1.1 a: Anzahl Studienteilnehmer im Treated Set b: Das Merkmal Ethnie wurde in Frankreich aufgrund entsprechender gesetzlicher Vorschriften nicht erfasst. c: Zahlen wurden aus Daten der oben genannten Quellen berechnet. d: Bezogen auf die kontrollierte Phase BMI: Body-Mass-Index; DL_{CO}: Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); FEV₁: Forced Expiratory Volume in one Second (Forcierte expiratorische Einsekunden-Kapazität); FVC: Forced Vital Capacity (Forcierte Vitalkapazität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); MW: Mittelwert; n. e.: nicht erhoben; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SGRQ: St. George's Respiratory Questionnaire; SpO₂: Sauerstoffsättigung; UIP: Usual Interstitial Pneumonia										

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.

Studiendesign

INPULSIS 1 und INPULSIS 2

Bei den Studien INPULSIS 1 und INPULSIS 2 handelt es sich um zweiarmige, kontrollierte, placebokontrollierte, multizentrische, multinationale, randomisierte (3:2) Phase III-Studien mit homogenem Studiendesign (Tabelle 4-12). Einzuschließende Studienteilnehmer waren Erwachsene mit einer nach ATS/ERS/JRS/ALAT-Richtlinien gesicherten IPF-Diagnose. Das Mindestalter der Studienteilnehmer betrug 40 Jahre. Weitere Einschlusskriterien waren eine Kohlenmonoxid-Diffusionskapazität (DL_{CO}) von 30-79% des Sollwertes, eine FVC von $\geq 50\%$ des Sollwertes bei Studienbeginn, das Vorliegen einer HRCT-Aufnahme der Brust innerhalb von zwölf Monaten vor Zeitpunkt der Randomisierung sowie – falls vorhanden – Lungenbiopsien, die die Diagnose IPF stützen.

In beiden Studien betrug die Behandlungsphase 52 Wochen, in der die Studienteilnehmer im Verum-Arm Nintedanib in einer kontinuierlichen Dosierung von 150 mg zweimal täglich per os erhielten. Die Einnahme sollte dabei im Intervall von 12 Stunden und nach einer Mahlzeit erfolgen. Im Falle des Auftretens intolerabler UE wurde in Rücksprache mit dem Studienteilnehmer eine Dosisreduktion auf 100 mg oder eine Therapieunterbrechung erwogen. Bei therapieassoziierten UE sollte die Re-Initiierung der Therapie nach spätestens vier Wochen erfolgen, wobei eine reduzierte Dosis von 100 mg zweimal täglich empfohlen wurde; bei nicht-therapieassoziierten UE sollte die Re-Initiierung der Therapie nach spätestens acht Wochen erfolgen, wobei die volle Dosis von 150 mg zweimal täglich verwendet werden konnte. Eine Dosis-Re-Eskalation auf 150 mg war bei Besserung der Beschwerden durch die UE innerhalb von vier Wochen nach Initiierung der reduzierten Dosis zulässig. Bei schwerwiegender Toxizität oder bei Intoleranz trotz Dosisreduktion wurde die Therapie abgebrochen und nicht fortgesetzt. Im Kontrollarm erhielten die Studienteilnehmer ein äußerlich nicht unterscheidbares Placebopräparat im gleichen Gabe- und Zeitintervall. An die Behandlungsphase schloss sich eine 28-tägige Nachbeobachtungsphase an. Studienteilnehmer hatten des Weiteren die Möglichkeit an einer unverblindeten Verlängerungsstudie zur Untersuchung der Langzeit-Sicherheit (INPULSIS-ON, 1199.33) teilzunehmen.

Als primärer Endpunkt war in beiden Studien die *jährliche Abnahme der FVC* in ml nach 52 Wochen definiert. Als key sekundäre Endpunkte wurden die *Veränderung des SGRQ-Gesamtscores* sowie das Auftreten von *akuten Exazerbationen* bis Woche 52 bestimmt. Weitere sekundäre Endpunkte waren Analysen bezüglich der *Veränderung des FVC* vom Ausgangswert bis Woche 52 (absolute, relative und kategoriale Veränderungen) sowie Responderanalysen zur Abnahme der FVC (angegeben als Prozent des Sollwertes) mit den Cut-Off-Werten für die Abnahme von 5- bzw. 10 Prozentpunkten sowie Analysen zur gesundheitsbezogenen Lebensqualität bzw. des allgemeinen Gesundheitszustandes mittels verschiedener Fragebögen (*Veränderungen im SGRQ, SGRQ-I, SOBQ, CASA-Q* und der

EQ-5D-VAS sowie SGRQ- [Cut-Off-Wert: -4 Punkte] und PGI-C-Responder). Zusätzlich wurden Überlebenszeitanalysen (*Zeit bis zum Tod, unabhängig von der Ursache, Zeit bis zum respiratorisch bedingten Tod, Zeit bis zum Tod unter Behandlung, Zeit bis zum Tod oder Lungentransplantation, Zeit bis zum Tod oder Lungentransplantation oder Qualifizierung bis Lungentransplantation*) sowie Bestimmungen der Veränderung vom Ausgangswert bis Woche 52 für die Parameter Sauerstoffsättigung (SpO₂) und DL_{CO} durchgeführt.

Beide Studien fanden in einem Zeitraum von Mai 2011 bis Oktober 2013 statt. Die Studie INPULSIS 1 wurde in 98 Studienzentren in 13 Ländern durchgeführt; insgesamt wurden 515 Patienten randomisiert, von denen 513 Patienten im Rahmen der Studie behandelt wurden (309 im Nintedanib-Arm und 204 im Kontrollarm). Die Studie INPULSIS 2 wurde in 107 Studienzentren in 17 Ländern weltweit durchgeführt; insgesamt wurden 551 Patienten randomisiert, von denen 548 Patienten im Rahmen der Studie behandelt wurden (329 im Nintedanib-Arm und 219 im Kontrollarm).

Die Behandlungsgruppen (Nintedanib und Placebo) waren in beiden Studien hinsichtlich der Patientencharakteristika vergleichbar (Tabelle 4-14). Auch zwischen den Studien wurden keine relevanten Unterschiede hinsichtlich der Patientencharakteristika festgestellt. Das mittlere Alter in den verschiedenen Behandlungsarmen lag zwischen 66,4 Jahren und 67,1 Jahren und der Anteil an männlichen Studienteilnehmern lag zwischen 78% und 81%. Die Mehrheit der Teilnehmer in beiden Studien war kaukasisch, wobei die Studie INPULSIS 1 einen geringfügig höheren Anteil kaukasischer Patienten (64% im Nintedanib-Arm, 66% im Kontrollarm) aufwies als die Studie INPULSIS 2 (52% bzw. 56%). Die mittlere Zeit seit der IPF-Diagnose war in den Behandlungsgruppen der Studien INPULSIS 1 und INPULSIS 2 mit 1,55 Jahren bzw. 1,66 Jahren nahezu identisch.

Studie 1199.187

Die Studie 1199.187 ist eine zweiarmige, kontrollierte, placebokontrollierte, multizentrische, multinationale, randomisierte (1:1) Phase IIb-Studie. Auch hier waren Erwachsene mit einer nach ATS/ERS/JRS/ALAT-Richtlinien gesicherten IPF-Diagnose einzuschließen. Die Einschlusskriterien stimmen weitgehend mit denen der INPULSIS Studien überein.

Ähnlich wie in den INPULSIS Studien war in Studie 1199.187 ursprünglich eine kontrollierte Behandlungsphase von 52 Wochen vorgesehen, in der die Studienteilnehmer im Verum-Arm Nintedanib in einer kontinuierlichen Dosierung von 150 mg zweimal täglich per os und die Studienteilnehmer im Kontrollarm ein entsprechendes Placebopräparat erhalten sollten. Aufgrund ethischer Bedenken in Bezug auf das Weiterführen einer Placebo-Kontrolle im Anwendungsgebiet einer potenziell tödlich verlaufenden Erkrankung mit schwerer Symptomatik auch nach Marktzulassung, wurde die kontrollierte Behandlungsphase der Studie

jedoch im Rahmen eines globalen Amendments auf 24 Wochen verkürzt^{14,15}. Zudem wurde im Anschluss an die kontrollierte Phase eine offene Phase zur Weiterbehandlung der Patienten aus beiden Studienarmen mit Nintedanib eingeführt, wobei eine Gesamtbehandlungsdauer von bis zu 78 Wochen festgelegt wurde (kontrollierte Phase *plus* einarmige Phase), nach der eine Nachbeobachtungsdauer von 28 Tagen vorgesehen war. Gleichzeitig wurde die Studie in ein exploratives Design überführt und die Fallzahl von 275 auf 150 Patienten reduziert. Wie in den INPULSIS Studien war auch in Studie 1199.187 ein Behandlungsintervall von 12 Stunden sowie die Einnahme mit einer Mahlzeit vorgesehen. Eine vorübergehende Therapieunterbrechung oder Dosisreduktion war ebenso analog zu den INPULSIS Studien möglich. Bei schwerwiegender Toxizität oder bei Intoleranz trotz Dosisreduktion war die Therapie abubrechen.

Der primäre Endpunkt der Studie *Veränderung des HRCT-QLF Scores* vom Ausgangswert bis Woche 52 wurde im Zuge des globalen Amendments des Protokolls auf *Veränderung des HRCT-QLF Scores* vom Ausgangswert bis Woche 24 geändert. Sekundäre Endpunkte umfassen die absolute, relative und kategoriale *Veränderung der FVC*, des *SGRQ-Gesamtscores*, des *6 MWT* und des *SOBQ-Gesamtscores*. Zudem wurden verschiedene Überlebenszeitanalysen betrachtet (*Zeit bis zum Tod, unabhängig von der Ursache; Zeit bis zum respiratorisch bedingten Tod; Zeit bis zur ersten respiratorisch bedingten Hospitalisierung; Zeit bis zur ersten akuten Exazerbation*). Aufgrund der Überführung in ein exploratives Design sah das revidierte CTP in seinen Vorgaben zur statistischen Analyse von der Durchführung formaler Hypothesentests ab. Ergebnisse zu den einzelnen Endpunkten sollten deskriptiv dargestellt werden.

Die Studie 1199.187 wurde im Zeitraum von Dezember 2013 bis Oktober 2016 an 26 Studienzentren in den USA, Kanada und der Türkei durchgeführt. Insgesamt wurden 113 Patienten randomisiert und behandelt (56 im Nintedanib-Arm und 57 im Kontrollarm).

Die beiden Behandlungsgruppen der Studie waren hinsichtlich der Patientencharakteristika vergleichbar (Tabelle 4-14). Das mittlere Alter betrug 68,8 Jahre im Nintedanib-Arm und 66,2 Jahre im Kontrollarm. Der Anteil männlicher Studienteilnehmer lag bei 80% unter Nintedanib bzw. 65% unter zVT. Die Patienten waren fast ausschließlich kaukasisch (96% bzw. 95%). Die mittlere Zeit seit der IPF-Diagnose war mit 1,46 Jahren in beiden Studienarmen identisch. Der durchschnittliche SGRQ-Gesamtscore zu Studienbeginn war im Nintedanib-Arm mit 35,75 Punkten niedriger als im Kontrollarm (44,39 Punkte); ein niedrigerer Score steht für eine bessere subjektiv empfundene krankheitsspezifische Lebensqualität.

¹⁴ In Studienbericht und Studienprotokoll der Studie 1199.187 wird die Dauer der kontrollierten Phase mit sechs Monaten angegeben. Diese Angaben sind synonym mit der im Dossier angegebenen Dauer der kontrollierten Phase von 24 Wochen zu sehen.

¹⁵ Es ist zu beachten, dass bei einem Teil der Patienten die Dauer der kontrollierten Phase nach oben abweicht (in Abhängigkeit vom Zeitpunkt des Studieneinschlusses relativ zum Zeitpunkt des globalen Amendments).

In Hinblick auf die Patientencharakteristika liegen mit Ausnahme des Anteils an kaukasischen Patienten, der in der Studie 1199.187 höher war als in den beiden INPULSIS Studien, insgesamt keine nennenswerten Unterschiede zwischen den drei eingeschlossenen Studien vor.

Der Anteil der Studienabbrecher lag in den INPULSIS Studien mit ungefähr 16,5% höher als der entsprechende Anteil in der kontrollierten Phase der Studie 1199.187, wo er mit 7,1% im Nintedanib-Arm und 10,5% im Kontrollarm leicht unbalanciert war. Der Anteil der Therapieabbrecher lag in den INPULSIS Studien bei 24,5% bzw. 18,9% und in Studie 1199.187 bei 23,2% bzw. 24,6%.

Übertragbarkeit auf den deutschen Versorgungskontext

Die INPULSIS Studien wurden unter anderem in deutschen Studienzentren durchgeführt.

Das Durchschnittsalter der Studienteilnehmer der eingeschlossenen Studien von ca. 67 Jahren spiegelt die Prävalenzverteilung in Deutschland gut wider (vgl. Modul 3, Abschnitt 3.2.3). Wie in den eingeschlossenen Studien, so tritt die IPF auch in der deutschen Versorgungsrealität bei Männern häufiger auf als bei Frauen. Allerdings ist der Anteil Männer in den Studien noch höher als unter IPF-Patienten in der Gesamtbevölkerung. Die in den Studien verwendeten Kriterien zur Diagnostik und Behandlung der IPF entsprechen den Empfehlungen der ATS/ERS/JRS/ALAT-Richtlinien, die auch im deutschen Versorgungsalltag zur Anwendung kommen (Behr et al., 2013).

Für Nintedanib liegen zudem keine Hinweise auf biodynamische- oder kinetische-Unterschiede zwischen einzelnen Bevölkerungsgruppen vor, die sich auf die Studienergebnisse auswirken könnten.

Daher kann davon ausgegangen werden, dass die Studienergebnisse auf den deutschen Versorgungskontext übertragbar sind.

Eignung der berücksichtigten Studien für die Bewertung des Zusatznutzens von Nintedanib gegenüber der zweckmäßigen Vergleichstherapie

Im Beratungsgespräch vom 29.06.2016 (Beratungsanforderung 2016-B-049) zur Ausgestaltung der zVT bei PF-ILD, welche unter anderem das Anwendungsgebiet der IPF umfassen, sowie im Schreiben zur vorsorglichen Aufforderung zur Dossiereinreichung, wurden durch den G-BA die Vergleichstherapien Pirfenidon *oder* BSC benannt (Gemeinsamer Bundesausschuss, 2016b; Gemeinsamer Bundesausschuss, 2018a). Der G-BA bestätigt damit die zVT, welche bereits im Zuge der Erstbewertung von Nintedanib im Anwendungsgebiet IPF festgelegt wurde. Wie in Abschnitt 4.2.1 dargestellt, wird im vorliegenden Dossier der Zusatznutzen gegenüber der Vergleichstherapie BSC bestimmt.

Der G-BA führt aus, dass unter BSC eine Therapie zu verstehen ist, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und der Verbesserung der Lebensqualität gewährleistet. Die Empfehlungen des G-BA, welche Behandlungen im Rahmen von BSC zu erwägen sind, werden durch die aktuelle Leitlinie gestützt (Behr et al., 2017). Darüber hinaus sind laut G-BA von BSC auch Behandlungen

umfasst, welche nicht im Anwendungsgebiet IPF zugelassen sind, wenn diese zur Behandlung spezifischer Symptome eingesetzt werden. Der G-BA stellt in seiner Niederschrift fest, dass die Vergleichstherapie dem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen hat. Die Geschäftsstelle führt aus, dass von der Anwendung der im Anwendungsgebiet IPF zugelassenen Kortikosteroide (Methylprednisolon, Prednisolon und Prednison) abgeraten wird bzw. deren Einsatz nur unter gründlicher Abwägung von Risiko und Nutzen erfolgen soll. Bezüglich der empfohlenen nicht-medikamentösen Therapien nennt der G-BA die Lungentransplantation, die Langzeit-Sauerstofftherapie sowie die pulmonale Rehabilitation. Diese Angaben sind im Einklang mit der aktuell gültigen deutschen S2k-Leitlinie (Behr et al., 2017). Zur Lungentransplantation schränkt der G-BA ein, dass aufgrund des kleinen Patientenkreises, der für eine Lungentransplantation infrage kommt, sowie den Limitierungen in der Verfügbarkeit von Spenderorganen, bei Lungentransplantationen nicht von einer im Rahmen von BSC regelhaft zu erwartenden Therapieoption auszugehen ist.

Hinsichtlich der durch den G-BA in der Niederschrift hervorgehobenen nicht-medikamentösen Therapien (Gemeinsamer Bundesausschuss, 2016b), fanden gemäß der CTP keinerlei Einschränkungen statt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2012a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013). Den Patienten standen somit im Rahmen aller drei Studien sowohl die Langzeit-Sauerstofftherapie als auch die pulmonale Rehabilitation in allen Studienarmen prinzipiell zur Verfügung. So begannen 10,3% bzw. 11,4% der Patienten in den Placebo-Armen der INPULSIS 1 und 2 Studien sowie 14% der Patienten in der Studie 1199.187 eine Langzeit-Sauerstofftherapie (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2017). Den Patienten stand zudem die Möglichkeit einer Lungentransplantation offen.

Hinsichtlich der eingesetzten Begleittherapien in den INPULSIS Studien und der Studie 1199.187 waren durch die Studienprotokolle folgende Vorgaben zu beachten:

Nicht zugelassen war in den betrachteten Studien der Einsatz von Pirfenidon oder anderen experimentellen Therapien zur Behandlung der IPF.

Zudem waren die Fibrinolysetherapie, Antikoagulantherapien mit voller Dosis (z. B. Vitamin K-Antagonisten, Dabigatran, Heparin, Hirudin) oder hoch dosierte Inhibitoren der Thrombozytenaggregation nicht vorgesehen.

Außerdem wurde der Einsatz der im Folgenden genannten Therapien für den Zeitraum der ersten 24 Wochen eingeschränkt.

- Azathioprin, Cyclophosphamid und Cyclosporin A (keine Verordnung innerhalb von acht Wochen nach erster Visite)
- N-Acetylcystein (NAC) (keine Verordnung innerhalb von zwei Wochen nach erster Visite)

- Prednison in einer Dosis von mehr als 15 mg pro Tag bzw. mehr als 30 mg alle zwei Tage oder einer äquivalenten Dosis eines anderen oralen Kortikosteroids (keine Verordnung innerhalb von zwei Wochen nach erster Visite)

Für keine dieser Therapien wird durch die deutsche Leitlinie oder die Leitlinie der ATS/ERS/JRS/ALAT eine Empfehlung ausgesprochen (Raghu et al., 2015; Behr et al., 2017). Vom Einsatz einer Monotherapie mit NAC wird in der deutschen Leitlinie sogar explizit abgeraten (Behr et al., 2017).

Die Behandlung *akuter Exazerbationen* kann von der Therapie während stetig-progressiver Phasen abweichen. So können während einer Exazerbation auch höher dosierte Kortikosteroide zum Einsatz kommen (Behr et al., 2013). Entsprechend ließ das Studiendesign der INPULSIS Studien und der Studie 1199.187 während *akuter Exazerbationen* die Anwendung eines jedweden Arzneimittels mit Ausnahme von Pirfenidon nach Ermessen des Arztes auch innerhalb der ersten 24 Wochen zu.

Unter Berücksichtigung der gültigen aktuellen Therapieempfehlungen ist daher nicht davon auszugehen, dass die hier genannten Therapien, für welche gemäß den CTP-Einschränkungen in der Verfügbarkeit bestanden, relevante Bestandteile einer zVT BSC ausmachen.

Aufgrund der sehr geringen Einschränkungen im zur Verfügung stehenden Therapiespektrum in allen drei Studien, welche ausschließlich von den Leitlinien nicht zur Therapie der IPF empfohlene oder explizit nicht empfohlene medikamentöse Therapien betreffen, sieht BI die Anforderungen des G-BA an die Ausgestaltung von BSC in den INPULSIS Studien und der Studie 1199.187 als adäquat umgesetzt an. Im vorliegenden Dossier wird daher der Zusatznutzen von Nintedanib plus BSC gegenüber der zVT BSC bewertet.

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 4-15: 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			
INPULSIS 1	ja	ja	ja	ja	ja	ja	niedrig
INPULSIS 2	ja	ja	ja	ja	ja	ja	niedrig
Studie 1199.187	ja	ja	ja	ja	ja	ja	niedrig
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)							

Begründen Sie für jede Studie die abschließende Einschätzung.

Alle drei Studien wurden als doppelblinde, randomisierte, placebokontrollierte, multizentrische, multinationale Studien durchgeführt.

Die Randomisierung erfolgte jeweils in adäquater Weise mittels *Interactive Voice Response System* bzw. *Interactive Response Technology* (IRT) und die Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass weder der Patient noch die behandelnden Personen Rückschlüsse auf die Behandlung ziehen konnten.

Die Auswertungskriterien für die Endpunkte waren in den Prüfplänen a priori definiert. Notwendige Protokolländerungen nach Studienbeginn wurden vor der Verfügbarkeit von Ergebnissen und einem möglichen Zugriff auf die Studiendaten durchgeführt, sodass eine Verzerrung ausgeschlossen werden kann. Die Ergebnisse wurden den Prüfplänen entsprechend ausgewertet und transparent und vollständig berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

Der Patientenfluss der jeweiligen Studien sowie die Methodik und Ergebnisse der verwendeten Analysen sind transparent dargestellt.

Es finden sich keine weiteren das Verzerrungspotenzial beeinflussenden Punkte, sodass das Verzerrungspotenzial dieser Studien als *niedrig* eingestuft wird.

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 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Endpunkt	INPULSIS 1	INPULSIS 2	INPULSIS (gepoolt)	Studie 1199.187	INPULSIS + Studie 1199.187 (gepoolt)
Mortalität					
Zeit bis zum Tod, unabhängig von der Ursache	ja	ja	ja	ja ^a	ja
Zeit bis zum respiratorisch bedingten Tod	ja	ja	ja	ja ^a	ja
Jährliche FVC-Abnahme (ml)	ja	ja	ja	ja ^b	ja
FVC-Responder zu Woche 52 (10% des Sollwertes)	ja	ja	ja	nein	nein
FVC-Responder zu Woche 52 (5% des Sollwertes)	ja	ja	ja	nein	nein
FVC-Responder zu Woche 24 (10% des Sollwertes)	ja ^b	ja ^b	ja	ja ^b	ja
FVC-Responder zu Woche 24 (5% des Sollwertes)	ja ^b	ja ^b	ja	ja ^b	ja
Morbidität					
Zeit bis zur ersten adjudizierten akuten Exazerbation	ja	ja	ja	ja ^a	ja
Zeit bis zur ersten akuten Exazerbation	ja	ja	ja	ja ^a	ja
Veränderung HRCT-QLF Score (bis Woche 24)	nein	nein	nein	ja	nein
Veränderung der Gehstrecke im 6 MWT (bis Woche 24)	nein	nein	nein	ja	nein
Zeit bis zur ersten respiratorisch bedingten Hospitalisierung	nein	nein	nein	ja	nein

Endpunkt	INPULSIS 1	INPULSIS 2	INPULSIS (gepoolt)	Studie 1199.187	INPULSIS + Studie 1199.187 (gepoolt)
Gesundheitsbezogene Lebensqualität					
Veränderung SGRQ-Gesamtscore (bis Woche 52)	ja	ja	ja	nein	nein
Veränderung SGRQ-Gesamtscore (bis Woche 24)	ja	ja	ja	ja	ja
SGRQ-(≤-4 Punkte)-Responder (zu Woche 52)	ja	ja	ja	nein	nein
SGRQ-(≤-4 Punkte)-Responder (zu Woche 24)	ja	ja	ja	ja	ja
Veränderung SGRQ-I-Gesamtscore (bis Woche 52)	ja	ja	ja	nein	nein
Veränderung SGRQ-I-Gesamtscore (bis Woche 24)	ja	ja	ja	ja	ja
SGRQ-I-(≤-4 Punkte)-Responder (zu Woche 52)	ja ^b	ja ^b	ja	nein	nein
SGRQ-I-(≤-4 Punkte)-Responder (zu Woche 24)	ja ^b	ja ^b	ja	ja	ja
Veränderung SOBQ-Gesamtscore (bis Woche 52)	ja	ja	ja	nein	nein
Veränderung SOBQ-Gesamtscore (bis Woche 24)	ja	ja	ja	ja	ja
Veränderung CASA-Q Husten-Symptome bis Woche 52	ja	ja	ja	nein	nein
Veränderung CASA-Q Husten-Impact bis Woche 52	ja	ja	ja	nein	nein
PGI-C-Responder zu Woche 52	ja	ja	ja	nein	nein
Veränderung EQ-5D-VAS bis Woche 52	ja	ja	ja	nein	nein
Verringerung von unerwünschten Ereignissen					
UE	ja	ja	ja	ja	ja
SUE	ja	ja	ja	Ja	ja
UE, die zum Therapieabbruch führten	ja	ja	ja	ja	ja
UE von besonderem Interesse	ja	ja	ja	ja	ja

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Endpunkt	INPULSIS 1	INPULSIS 2	INPULSIS (gepoolt)	Studie 1199.187	INPULSIS + Studie 1199.187 (gepoolt)
a: Abweichend vom Studienbericht der Studie 1199.187 wird im vorliegenden Nutzendossier für die Analyse dieser Wirksamkeitsendpunkte die gesamte kontrollierte Phase herangezogen, ohne auf die ersten 24 Wochen einzuschränken. Dieses Vorgehen ist in den üblichen Anforderungen der Nutzenbewertung begründet, soweit sinnvoll, die längste unverzerrt bewertbare Beobachtungsdauer zugrunde zu legen. Eine Ausnahme stellen die Auswertungen zu Woche 24 dar, hier werden nur Daten bis Woche 24 berücksichtigt. b: Post hoc berechneter Endpunkt (Operationalisierung im Analyseplan der Einzelstudie nicht vorgesehen) 6 MWT: 6 Minute Walk Test (6 Minuten Gehtest); CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); PGI-C: Patient's Global Impression of Change; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visual Analog Scale					

4.3.1.3.1 Mortalität – 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 4.2.5.2). 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.

4.3.1.3.1.1 Endpunkte Zeit bis zum Tod

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

Tabelle 4-17: Operationalisierung von Zeit bis zum Tod

Studie	Operationalisierung
INPULSIS 1 ^a	<u>Zeit bis zum Tod, unabhängig von der Ursache</u> Definition: Zeit von Randomisierung bis Todeszeitpunkt Zensierung von Patienten ohne Ereignis: 372 Tage nach Randomisierung (52 Wochen + 7 Tage Toleranz) bzw. beim letzten Patientenkontakt, definiert als Datum des letzten im CRF dokumentierten Ereignisses Erhebung: durch Prüfarzt, Ermittlung des letzten Patientenkontaktes aus CRF-Dokumentation vom Ausgangswert <u>Zeit bis zum respiratorisch bedingten Tod</u> Definition: Zeit von Randomisierung bis zum Zeitpunkt des respiratorisch bedingten Todes (basierend auf der Bewertung der Todesursachen durch das IAC) Zensierung von Patienten ohne Ereignis: 372 Tage nach Randomisierung (52 Wochen + 7 Tage Toleranz) bzw. beim letzten Patientenkontakt, definiert als Datum des letzten im CRF dokumentierten Ereignisses Erhebung: durch Prüfarzt, adjudiziert durch IAC, Ermittlung des letzten Patientenkontaktes aus CRF-Dokumentation vom Ausgangswert
INPULSIS 2 ^a	Homogenes Studiendesign wie INPULSIS 1
Studie 1199.187 ^a	<u>Zeit bis zum Tod, unabhängig von der Ursache</u> Definition: Zeit von Randomisierung bis Todeszeitpunkt Zensierung von Patienten ohne Ereignis: Für Kaplan-Meier-Schätzung: an Tag 176; für die Cox-Regression am End-Datum der kontrollierten Phase Erhebung: durch Prüfarzt, Ermittlung des letzten Patientenkontaktes aus CRF-Dokumentation <u>Zeit bis zum respiratorisch bedingten Tod</u> Definition: Zeit von Randomisierung bis zum Zeitpunkt des respiratorisch bedingten Todes (basierend auf der Bewertung der Todesursachen durch das IAC) Zensierung von Patienten ohne Ereignis: Für Kaplan-Meier-Schätzung: an Tag 176; für die Cox-Regression am End-Datum der kontrollierten Phase Erhebung: durch Prüfarzt, adjudiziert durch IAC, Ermittlung des letzten Patientenkontaktes aus CRF-Dokumentation
a: Die Operationalisierung für die gepoolten Auswertungen entspricht der Operationalisierung in den Einzelstudien. CRF: Case Report Form; IAC: Independent Adjudication Committee (Unabhängiges Bewertungskomitee)	

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 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt *Zeit bis zum Tod* 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
INPULSIS 1 ^a	<u>Zeit bis zum Tod, unabhängig von der Ursache</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zum respiratorisch bedingten Tod</u>					
	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2 ^a	<u>Zeit bis zum Tod, unabhängig von der Ursache</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zum respiratorisch bedingten Tod</u>					
	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187 ^a	<u>Zeit bis zum Tod, unabhängig von der Ursache</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zum respiratorisch bedingten Tod</u>					
	niedrig	ja	ja	ja	ja	niedrig
a: Das Verzerrungspotenzial der gepoolten Auswertungen der INPULSIS Studien ergibt sich aus dem Verzerrungspotenzial der Einzelstudien.						
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für die Endpunkte *Zeit bis zum Tod* und *Zeit bis zum respiratorisch bedingten Tod* erfolgte verblindet. Die Feststellung eines Todesfalles ist zudem objektiv möglich. Die Beurteilung der Todesursache (z. B. respiratorische Ursache) erfolgte primär durch den Prüfarzt. Zusätzlich wurde die Todesursache durch ein IAC anhand der Patientenakten adjudiziert. Dadurch kann von einer objektiven Einschätzung der Todesursache ausgegangen werden.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und der Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen haben. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für die Endpunkte *Zeit bis zum Tod* und *Zeit bis zum respiratorisch bedingten Tod* waren a priori definiert. Die Ergebnisse aller drei Studien wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Abweichungen der vorliegenden Berichterstattung vom Studienbericht sind durch die Anforderungen der Nutzenbewertungen begründet.¹⁶

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP ist transparent. Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die Endpunkte *Zeit bis zum Tod* ein niedriges Verzerrungspotenzial.

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

Die Endpunkte zum direkten Erheben der Mortalität wurden erfasst als *Zeit bis zum Tod*, unabhängig von der Ursache und *Zeit bis zum respiratorisch bedingten Tod*. Der in den INPULSIS Studien und der Analyse der über die gepoolten INPULSIS Studien abgedeckte Zeitraum umfasst 52 Wochen. In die Analyse der über die drei Studien gepoolten Daten (INPULSIS Studien und Studie 1199.187) sowie in die Einzeldarstellung der Studie 1199.187

¹⁶ Abweichend vom Studienbericht wird für die Studie 1199.187 im vorliegenden Nutzendossier für die Analyse der Endpunkte *Zeit bis zum Tod* die gesamte kontrollierte Phase herangezogen (ohne Einschränkung auf die ersten 24 Wochen), um den üblichen Anforderungen der Nutzenbewertung entsprechend die längste unverzerrt bewertbare Beobachtungsdauer zugrunde zu legen.

flossen Daten der gesamten kontrollierten Phase der Studie 1199.187 ein. Die Ergebnisse der über die drei Studien gepoolten Analyse sind als die im Kontext der Nutzenbewertung relevanten zu betrachten.

Tabelle 4-19: Ergebnisse für *Zeit bis zum Tod* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Überlebens wahrschein- lichkeit	N ^a	n ^b (%)	Überlebens wahrschein- lichkeit	HR ^c	95%-KI ^c	P- Wert ^c
Zeit bis zum Tod, unabhängig von der Ursache									
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	36 (5,2)	0,945 ^e	480	37 (7,7)	0,917 ^e	0,65	[0,41; 1,03]	0,0775
INPULSIS (gepoolt) ^d	638	35 (5,5)	0,945 ^e	423	33 (7,8)	0,921 ^e	0,70	[0,43; 1,12]	0,1399
INPULSIS 1 ^d	309	13 (4,2)	0,958 ^e	204	13 (6,4)	0,935 ^e	0,63	[0,29; 1,36]	0,2880
INPULSIS 2 ^d	329	22 (6,7)	0,933 ^e	219	20 (9,1)	0,908 ^e	0,74	[0,40; 1,35]	0,2995
Studie 1199.187 ^{d, f}	56	1 (1,8)	1,000 ^g	57	4 (7,0)	0,945 ^g	0,15	[0,02; 1,39]	0,1944
Zeit bis zum respiratorisch bedingten Tod									
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	25 (3,6)	0,961 ^e	480	23 (4,8)	0,948 ^e	0,71	[0,40; 1,26]	0,2960
INPULSIS (gepoolt) ^d	638	24 (3,8)	0,962 ^e	423	21 (5,0)	0,949 ^e	0,74	[0,41; 1,34]	0,3435
INPULSIS 1 ^d	309	10 (3,2)	0,967 ^e	204	10 (4,9)	0,950 ^e	0,61	[0,25; 1,47]	0,3515
INPULSIS 2 ^d	329	14 (4,3)	0,956 ^e	219	11 (5,0)	0,949 ^e	0,86	[0,39; 1,90]	0,6654
Studie 1199.187 ^{d, f}	56	1 (1,8)	1,000 ^g	57	2 (3,5)	0,963 ^g	0,32	[0,03; 3,54]	0,6025
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.4.1, Tab. 15.2.2.2.4.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.4.1, Tab. 15.2.2.2.4.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c): Tab. 2.3.1.1.1, Tab. 2.3.1.2.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.49.2.1.2, Tab. 1.3.49.2.2.2, Tab. 86.18.1.1, Tab. 86.19.1.1, Tab. 1.1.28.1.2, Tab. 1.1.28.2.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017): Tab. 15.2.2.8, Tab. 15.2.2.6 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: HR [95%-KI] berechnet mittels Cox-Regression mit Termen für Behandlung, Geschlecht, Alter und Größe (bei den gepoolten Analysen wurde zusätzlich der Term Studie hinzugefügt); p-Wert berechnet mittels Log-rank-Test; für die gepoolte Analyse aller drei Studien stratifiziert für Studie d: Treated Set e: 12-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 372)									

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Überlebens wahrschein lichkeit	N ^a	n ^b (%)	Überlebens wahrschein lichkeit	HR ^c	95%-KI ^c	p- Wert ^c
f: Kontrollierte Phase g: 6-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 176) HR: Hazard Ratio; KI: Konfidenzintervall; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)									

Für die *Zeit bis zum Tod, unabhängig von der Ursache* und die *Zeit bis zum respiratorisch bedingten Tod* war die Überlebenswahrscheinlichkeit nach 52 Wochen (INPULSIS Studien) bzw. nach 24 Wochen (Studie 1199.187) in allen Studien unter der Behandlung mit Nintedanib höher als im Vergleichsarm (Tabelle 4-19). Es zeigte sich jedoch weder in den Einzelstudien noch in den gepoolten Analysen ein statistisch signifikanter Effekt, wobei die Anzahl an Ereignissen in allen Studien relativ gering war. Numerisch zeigt sich in den Studien ein Trend zugunsten der Nintedanib-Gruppe bezüglich der Überlebenswahrscheinlichkeit gemäß Kaplan-Meier-Schätzung.

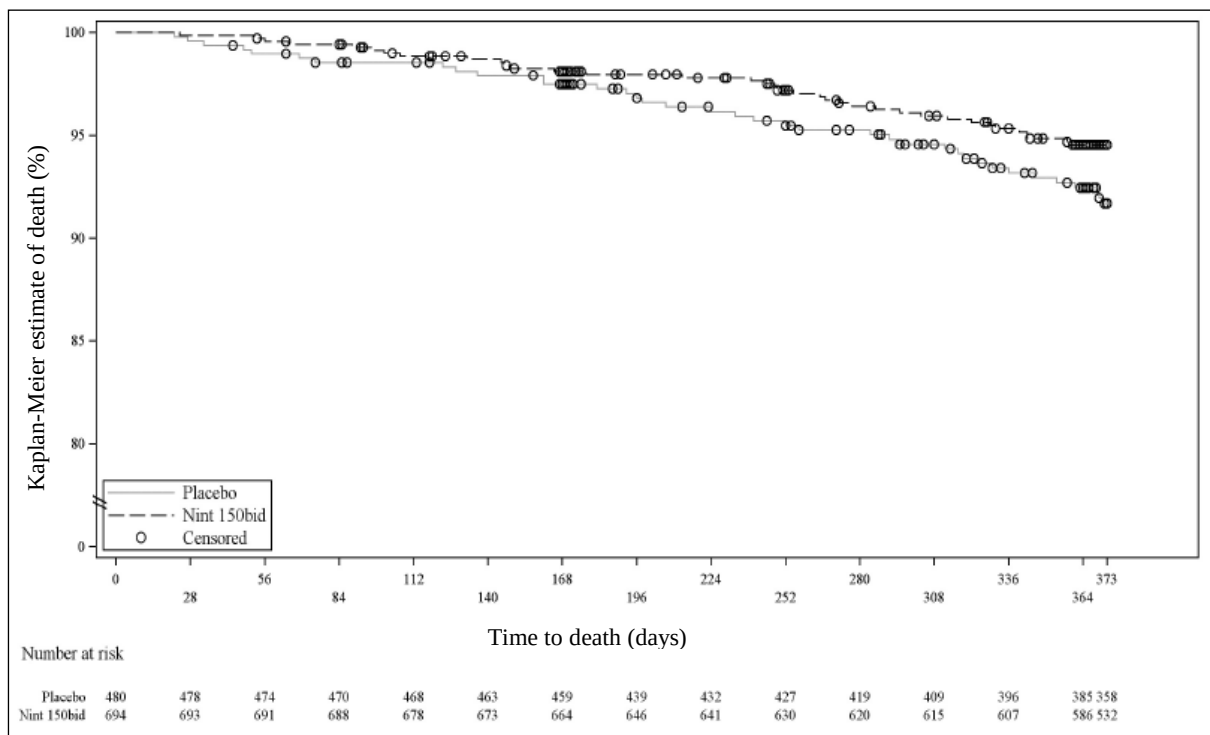
Der Anteil der im Beobachtungszeitraum verstorbenen Studienteilnehmer war unter der Behandlung mit Nintedanib geringer als im Vergleichsarm. Dies gilt sowohl für die über die drei Studien gepoolten Analyse (5,2% vs. 7,7%) als auch für die über die INPULSIS Studien gepoolten Analyse (5,5% vs. 7,8%). Das HR [95%-KI] für die *Zeit bis zum Tod, unabhängig von der Ursache* lag in der IPD Meta-Analyse der drei Studien bei 0,65 [0,41; 1,03]; p=0,0775 und in der IPD Meta-Analyse der gepoolten INPULSIS Studien bei 0,70 [0,43; 1,12]; p=0,1399.

In den Einzelstudien ergab sich ebenfalls ein geringerer Anteil verstorbener Studienteilnehmer unter Nintedanib (INPULSIS 1: 4,2% vs. 6,4%, INPULSIS 2: 6,7% vs. 9,1% und Studie 1199.187: 1,8% vs. 7,0%). Der entsprechende Effektschätzer lag in der INPULSIS 1 bei 0,63 [0,29; 1,36]; p=0,2880, in der INPULSIS 2 bei 0,74 [0,40; 1,35]; p=0,2995 und in der Studie 1199.187 bei 0,15 [0,02; 1,39]; p=0,1944.

Auch der Anteil der aus respiratorischen Gründen verstorbenen Studienteilnehmer war in der Nintedanib-Gruppe niedriger als in der Kontrollgruppe, sowohl in den gepoolten Daten aller drei Studien (3,6% vs. 4,8%) als auch in den gepoolten INPULSIS Studien (3,8% vs. 5,0%). Das HR [95%-KI] für die *Zeit bis zum respiratorisch bedingten Tod* lag in der IPD Meta-Analyse der drei Studien bei 0,71 [0,40; 1,26]; p=0,2960 und in der IPD Meta-Analyse der gepoolten INPULSIS Studien bei 0,74 [0,41; 1,34]; p=0,3435.

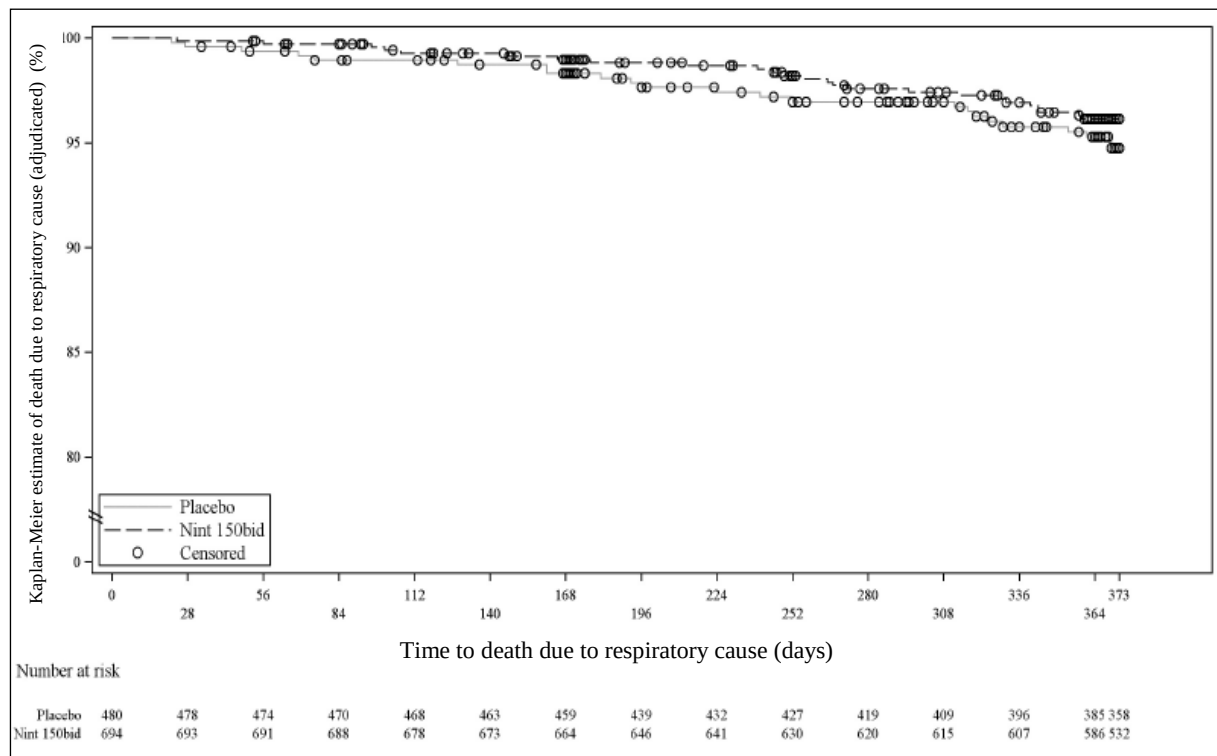
In den Einzelstudien ergab sich ebenfalls ein geringerer Anteil der aus respiratorischen Gründen verstorbenen Studienteilnehmer unter Nintedanib (INPULSIS 1: 3,2% vs. 4,9%, INPULSIS 2: 4,3% vs. 5,0% und Studie 1199.187: 1,8% vs. 3,5%). Der entsprechende Effektschätzer lag in der INPULSIS 1 bei 0,61 [0,25; 1,47]; p=0,3515, in der INPULSIS 2 bei 0,86 [0,39; 1,90]; p=0,6654 und in Studie 1199.187 bei 0,32 [0,03; 3,54]; p=0,6025.

Die folgenden Abbildungen zeigen die Kaplan-Meier-Kurven für die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* (Abbildung 8) und die *Zeit bis zum respiratorisch bedingten Tod* (Abbildung 9) für die gepoolten Daten der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187. Bei der Interpretation der Kaplan-Meier-Kurven für die gepoolten Daten aller drei Studien ist die von den INPULSIS Studien abweichende Nachbeobachtungszeit der Studie 1199.187 zu beachten. Die Daten der Patienten aus dieser Studie wurden für den Therapievergleich mittels Cox-Regression zum Zeitpunkt des letzten Patientenkontaktes bzw. am End-Datum der kontrollierten Phase zensiert. Für die Kaplan-Meier-Schätzung wurden alle Daten 176 Tage nach erstmaliger Einnahme der Studienmedikation zensiert.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 8: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zum Tod, unabhängig von der Ursache* für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 9: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zum respiratorisch bedingten Tod* für die gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

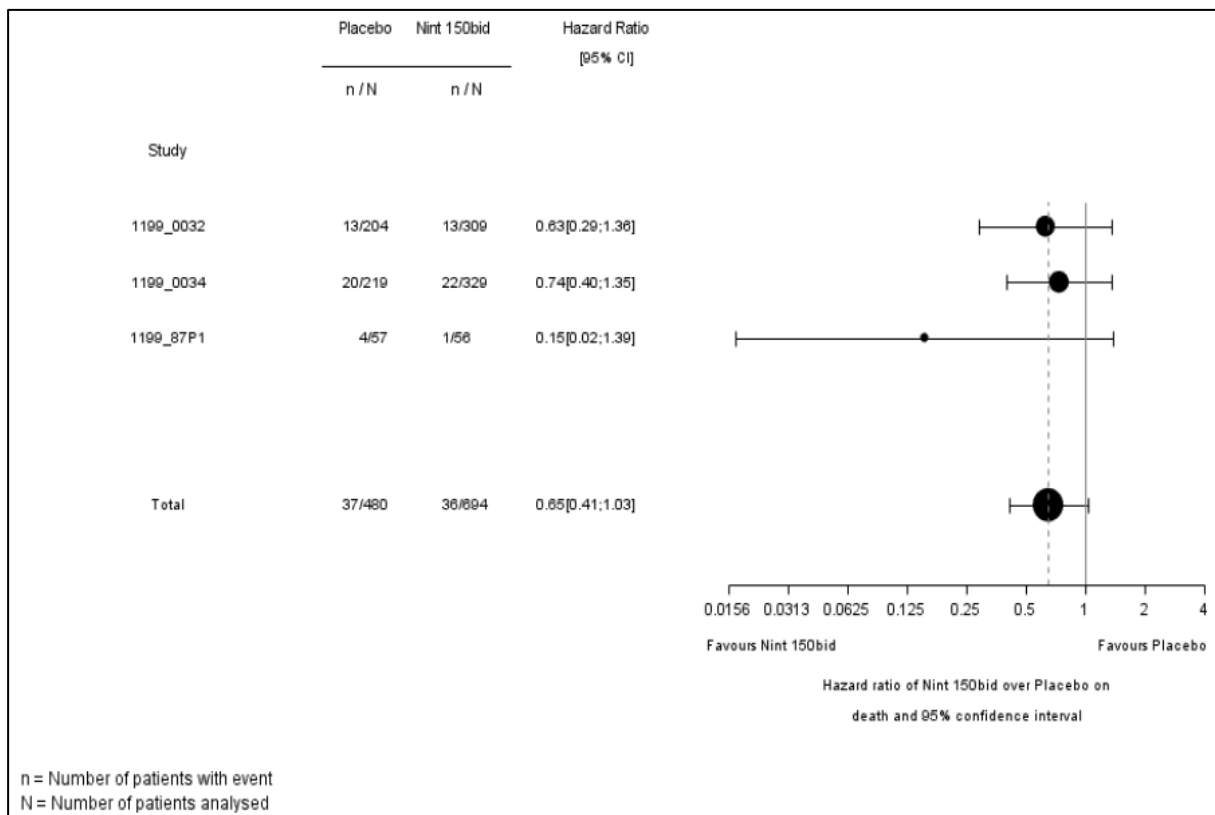
Die Ergebnisse für die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* und *Zeit bis zum respiratorisch bedingten Tod* zeigen in allen der Bewertung zugrunde liegenden Studien einen Trend zugunsten von Nintedanib (Abbildung 8, Abbildung 9). Dies deutet auf ein geringeres Mortalitätsrisiko für Nintedanib-behandelte Studienteilnehmer hin.

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.

Die Ergebnisse der IPD Meta-Analyse für die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* und *Zeit bis zum respiratorisch bedingten Tod* wurden bereits in Tabelle 4-19 dargestellt.

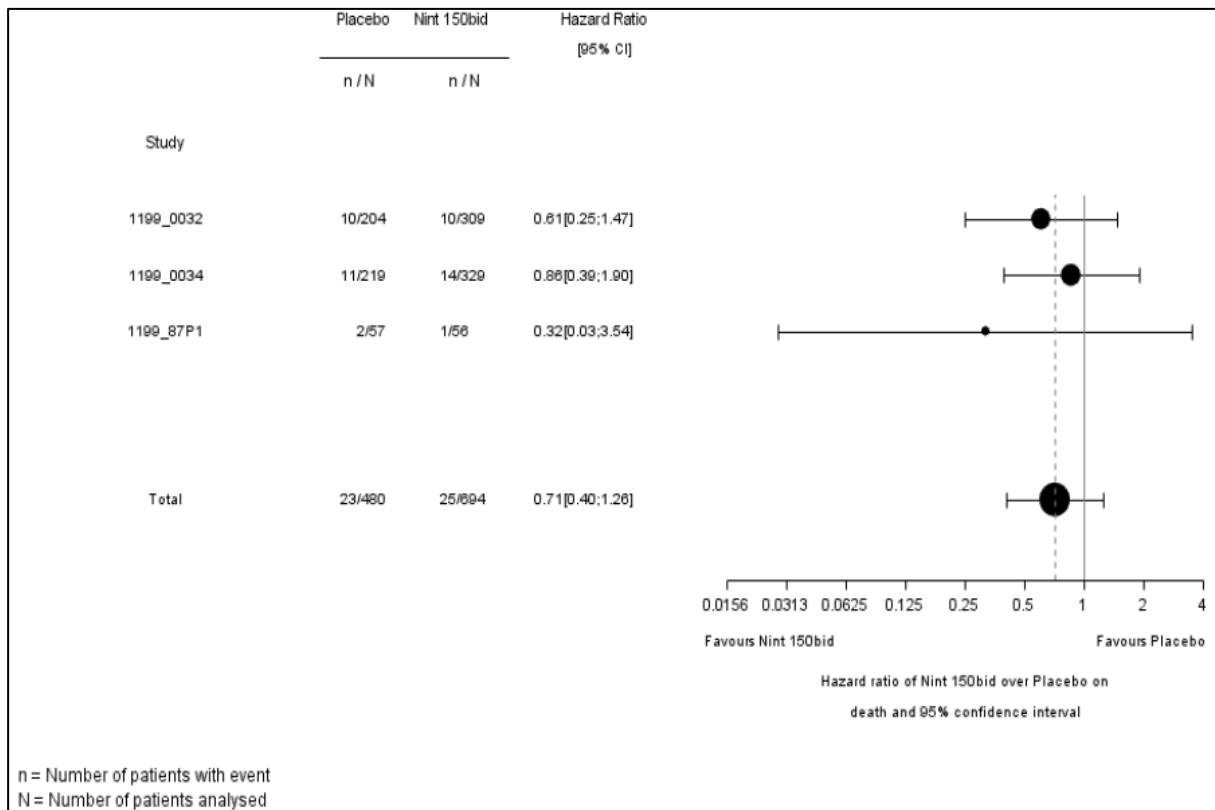
Die grafische Darstellung der Meta-Analysen mittels Forest-Plots befinden sich in Abbildung 10 und Abbildung 11. Sie zeigen die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der IPD Meta-Analyse. Die Heterogenitätstests (siehe auch Abschnitt 4.3.1.3.5.1) weisen auf eine homogene Datenlage hin (p-Wert der Interaktionstests [Merkmal Studie]: p=0,5795 für den Endpunkt *Zeit bis zum*

Tod, unabhängig von der Ursache bzw. $p=0,7910$ für den Endpunkt Zeit bis zum respiratorisch bedingten Tod).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 10: Forest-Plot zur Meta-Analyse für Endpunkt *Zeit bis zum Tod, unabhängig von der Ursache* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 11: Forest-Plot zur Meta-Analyse für Endpunkt *Zeit bis zum respiratorisch bedingten Tod* - RCT mit dem zu bewertenden Arzneimittel

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Der Endpunkt *Zeit bis zum Tod* ist eindeutig definiert und objektiv zu erheben. Die Bewertung der Todesursache für den Endpunkt *Zeit bis zum respiratorisch bedingten Tod* wurde durch ein verblindetes Adjudizierungskomitee aus drei ausgewiesenen Experten aus dem Bereich der Pneumologie mit besonderen Fachkenntnissen in der Behandlung der IPF vorgenommen. Für die Erhebung dieses Endpunkts wurde daher die größtmögliche Objektivität angestrebt. Für beide Endpunkte kann somit von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ausgegangen werden.

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.

4.3.1.3.1.2 Endpunkte Forcierte Vitalkapazität

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

Tabelle 4-20: Operationalisierung von FVC

Studie	Operationalisierung
INPULSIS 1 ^a	<u>Jährliche FVC-Abnahme (ml)</u> Definition: Jährliche FVC-Abnahme (ml) errechnet mittels Random-Coefficient-Modell aus der Veränderung der FVC im Zeitverlauf (vom Ausgangswert bis 52 Wochen) Erhebung: - durch Prüfarzt, mittels Spirometer gemäß ATS/ERS/JRS/ALAT-Kriterien zum Ausgangswert, Woche 2, 4, 6, 12, 24, 36, 52 (sowie zu späteren Studienvisiten). - Test wurde dreifach in sitzender Position jeweils zur gleichen Tageszeit (± 60 Minuten) und ohne größere körperliche Belastung innerhalb von 12 Stunden vor der Erhebung durchgeführt. Studienteilnehmer wurden gebeten am Tag der Erhebung nicht zu rauchen, kalte Temperaturen, Staub, Rauch und unangenehme Gerüche (z. B. Parfüm) zu vermeiden. - Kaffee, Tee, Schokolade, Cola und andere koffeinhaltige Getränke, alkoholhaltige Getränke oder Nahrungsmittel und eisgekühlte Getränke waren am Morgen der Erhebung nicht erlaubt. Auswaschphase für Bronchodilatatortherapie von 24 Stunden (für langwirksame Wirkstoffe) bzw. 8 Stunden (für Wirkstoffe mit kurzer Halbwertszeit) vor Testdurchführung. <u>FVC-Responder zu Woche 52 (Schwellenwert 10% bzw. 5%-Punkte^b)</u> Definition: Anteil Studienteilnehmer mit einer absoluten FVC-Abnahme von nicht mehr als 10 Prozentpunkten bzw. 5 Prozentpunkten^b zwischen Studienbeginn und Woche 52. Die FVC wurde jeweils ausgedrückt in % des Sollwertes. Falls zu Woche 52 kein Messwert vorlag, wurde der entsprechende Studienteilnehmer konservativ als Non-Responder bewertet. Um eine sinnvolle Kombination der Daten aus den INPULSIS Studien mit der kürzeren Studie 1199.187 zu ermöglichen, wird eine zusätzliche nicht präspezifizierte Operationalisierung ausgewertet und dargestellt. Diese dient als ergänzende Information zu der für die Nutzenbewertung maßgeblichen Operationalisierung zum Zeitpunkt 52 Wochen. <u>FVC-Responder zu Woche 24 (Schwellenwert 10%-Punkte bzw. 5%-Punkte^b)</u> Definition: Anteil Studienteilnehmer mit einer absoluten FVC-Abnahme von nicht mehr als 10 Prozentpunkten bzw. 5 Prozentpunkten^b zwischen Studienbeginn und Woche 24. Die FVC wurde jeweils ausgedrückt in % des Sollwertes. Falls zu Woche 24 kein Messwert vorlag, wurde der entsprechende Studienteilnehmer konservativ als Non-Responder bewertet.
INPULSIS 2 ^a	Homogenes Studiendesign wie INPULSIS 1

Studie	Operationalisierung
Studie 1199.187 ^a	<u>Jährliche FVC-Abnahme (ml)</u> Definition: Jährliche FVC-Abnahme (ml) errechnet mittels Random-Coefficient-Modell aus der Veränderung der FVC im Zeitverlauf (vom Ausgangswert bis zum Ende der kontrollierten Phase) Erhebung: - durch Prüfarzt, mittels Spirometer gemäß ATS/ERS/JRS/ALAT-Kriterien zum Ausgangswert, Woche 12, 24, 26, 52 (sowie zu späteren Studienvisiten). - Test wurde dreifach in sitzender Position jeweils zur gleichen Tageszeit (± 60 Minuten) und ohne größere körperliche Belastung innerhalb von 12 Stunden vor der Erhebung durchgeführt. Studienteilnehmer wurden gebeten am Tag der Erhebung nicht zu rauchen, kalte Temperaturen, Staub, Rauch und unangenehme Gerüche (z. B. Parfüm) zu vermeiden. Auswaschphase für Bronchodilatatortherapie von 24 Stunden (für langwirksame Wirkstoffe) bzw. 8 Stunden (für Wirkstoffe mit kurzer Halbwertszeit) vor Testdurchführung. <u>FVC-Responder zu Woche 24 (Schwellenwert 10%-Punkte bzw. 5%-Punkte^b)</u> Definition: Anteil Studienteilnehmer mit einer absoluten FVC-Abnahme von nicht mehr als 10 Prozentpunkten bzw. 5 Prozentpunkten^b zwischen Studienbeginn und Woche 24. Die FVC wurde jeweils ausgedrückt in % des Sollwertes. Falls zu Woche 24 kein Messwert vorlag, wurde der entsprechende Studienteilnehmer konservativ als Non-Responder bewertet.
a: Die Operationalisierung für die gepoolten Auswertungen entspricht der Operationalisierung in den Einzelstudien. b: Cut-Off-Wert 10%-Punkte: Hauptanalyse, Cut-Off-Wert 5%-Punkte: supportive Analyse (vgl. Abschnitt 4.5.4) ALAT: Latin American Thoracic Association; ATS: American Thoracic Society; ERS: European Respiratory Society; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); JRS: Japanese Respiratory Society	

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 4-21: Bewertung des Verzerrungspotenzials für den Endpunkt *FVC* 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
INPULSIS 1 ^a	<u>Jährliche FVC-Abnahme (ml)</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>FVC-Responder zu Woche 52 (Schwellenwert von 10%-Punkten und 5%-Punkten^b)</u>					
	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2 ^a	<u>Jährliche FVC-Abnahme (ml)</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>FVC-Responder zu Woche 52 (Schwellenwert von 10%-Punkten und 5%-Punkten^b)</u>					
	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187 ^a	<u>Jährliche FVC-Abnahme (ml)</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>FVC-Responder zu Woche 24 (Schwellenwert von 10%-Punkten und 5%-Punkten^b)</u>					
	niedrig	ja	ja	ja	ja	niedrig
a: Das Verzerrungspotenzial der gepoolten Auswertungen ergibt sich aus dem Verzerrungspotenzial der Einzelstudien. b: Cut-Off-Wert 10%-Punkte: Hauptanalyse, Cut-Off-Wert 5%-Punkte: supportive Analyse (Abschnitt 4.5.4) FVC: Forced Vital Capacity (Forcierte Vitalkapazität); ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für die Endpunkte zur *forcierten Vitalkapazität* erfolgte doppelblind. Die FVC ist zudem objektiv bestimmbar.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und der Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen haben. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip. Für die gepoolten Analysen wurden alle beobachteten Werte in der entsprechenden Beobachtungsperiode herangezogen, unabhängig davon, ob der Patient die Studienmedikation noch einnahm oder die Therapie bereits abgebrochen hatte.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für die FVC waren a priori definiert. Die Ergebnisse aller drei Studien wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. In Bezug auf Studie 1199.187 weichen die im Dossier präsentierten Operationalisierungen vom Studienbericht ab, in dessen Rahmen die absolute und kategorielle Veränderung der FVC gegenüber dem Ausgangswert betrachtet wurden. Die Berechnung der jährlichen Abnahme der FVC sowie der Responderanalyse erfolgt im Dossier aus Gründen der Konsistenz mit den INPULSIS Studien.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP ist transparent. Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die Endpunkte zur FVC ein niedriges Verzerrungspotenzial.

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

Jährliche FVC-Abnahme

Der Endpunkt *jährliche Abnahme der FVC* (ml) wurde in den INPULSIS Studien und in der Studie 1199.187 erhoben. Die jährliche Abnahme der FVC wurde mittels Random-Coefficient-Modell aus der Veränderung der FVC im Zeitverlauf errechnet (vom Ausgangswert bis Woche 52 [INPULSIS Studien] bzw. bis zum Ende der kontrollierten Phase [Studie 1199.187]). Die über die drei Studien gepoolte Analyse wird dabei als relevante Analyse zur Bestimmung des Therapieeffekts von Nintedanib betrachtet.

Tabelle 4-22: Ergebnisse für *jährliche FVC-Abnahme* (ml) aus RCT mit dem zu bewertenden Arzneimittel

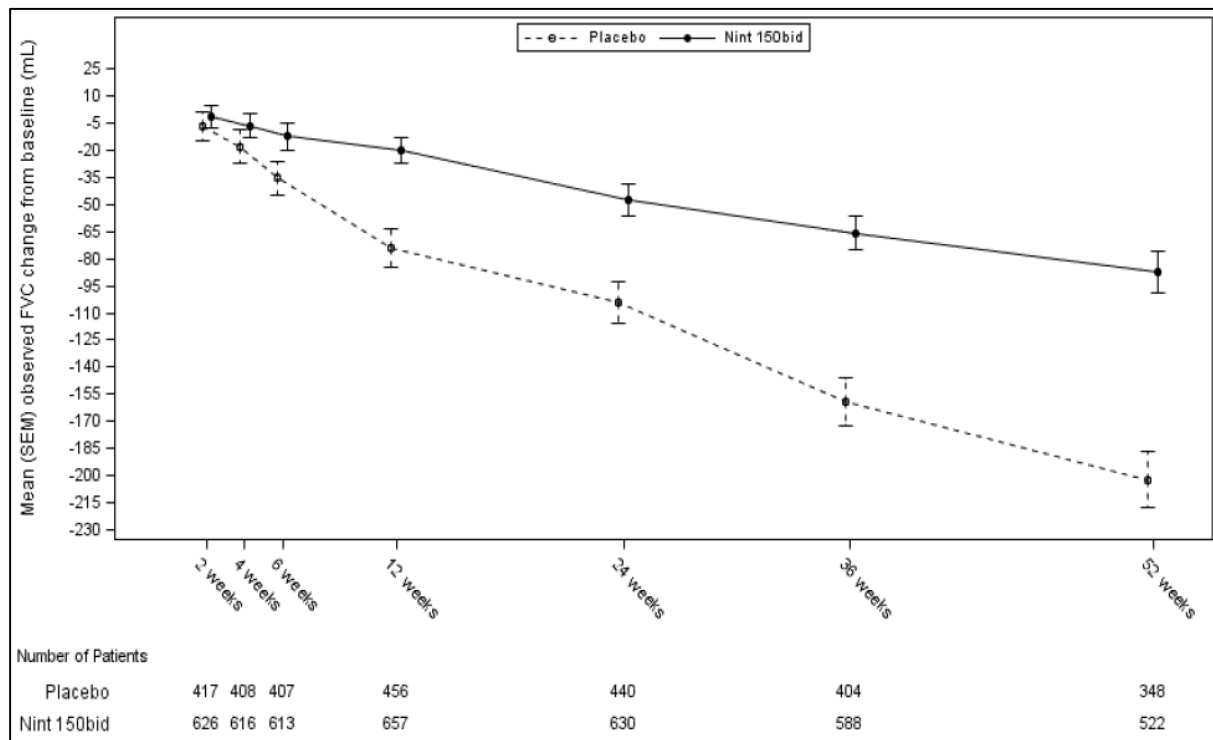
Studie	Behandlung	N ^a	Ausgangswert in ml	Woche 52 in ml		jährliche Abnahme in ml	Behandlungseffekt Nintedanib vs. Placebo ^c in ml		
			MW ^b (SD)	N	MW ^b (SD)	MW ^c (SE)	MWD (SE)	95%-KI	P- Wert
Jährliche Abnahme der FVC (ml)									
INPULSIS + Studie 1199.187 (gepoolt) ^d	Nintedanib	694	2.736,3 (766,5)	-	k. A.	108,88 (10,821)	112,42 (16,998)	[79,06; 145,77]	<0,0001
	Placebo	480	2.750,7 (814,6)	-	k. A.	221,30 (13,106)			
INPULSIS (gepoolt) ^d	Nintedanib	638	2.713,5 (757,03)	519	2.652,6 (792,29)	113,59 (10,984)	109,94 (17,368)	[75,85; 144,03]	<0,0001
	Placebo	423	2.727,7 (810,25)	345	2.585,1 (829,76)	223,53 (13,448)			
INPULSIS 1 ^d	Nintedanib	309	2.756,8 (735,12)	250	2.669,0 (772,04)	114,65 (15,327)	125,26 (24,209)	[77,68; 172,84]	<0,0001
	Placebo	204	2.844,5 (820,11)	165	2.664,4 (834,01)	239,91 (18,709)			
INPULSIS 2 ^d	Nintedanib	329	2.672,8 (775,96)	269	2.637,3 (811,80)	113,59 (15,726)	93,73 (24,907)	[44,78; 142,68]	0,0002
	Placebo	219	2.619,0 (787,35)	180	2.512,5 (821,44)	207,32 (19,309)			
Studie 1199.187 ^{d, e}	Nintedanib	56	2.996,81 (831,25)	-	k. A.	-9,35 ^f (58,504)	195,07 (82,260)	[31,87; 358,26]	0,0196
	Placebo	57	2.921,26 (833,91)	-	k. A.	185,72 (57,359)			
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.1.1, Tab. 15.2.2.2.1.1.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.1.1, Tab. 15.2.2.2.1.1.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c): Tab. 2.1.1.1.1; Tab. 15.2.1.1.1., Tab. 15.2.1.3.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017): Tab. 15.2.2.2.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.28.5.1, Tab. 86.15.1.1, Tab. 1.3.49.2.5.1, Tab. 1.3.65.1.4									

Studie	Behandlung	N ^a	Ausgangswert in ml	Woche 52 in ml		jährliche Abnahme in ml	Behandlungseffekt Nintedanib vs. Placebo ^c in ml		
			MW ^b (SD)	N	MW ^b (SD)	MW ^c (SE)	MWD (SE)	95%-KI	P-Wert
a: Anzahl analysierter Studienteilnehmer									
b: Nicht-adjustierte Mittelwerte									
c: Berechnet mittels Random-Coefficient-Regression mit festen Effekten für Behandlung, Geschlecht, Alter und Größe und einem zufälligen Effekt für Studienteilnehmer und Zeit zum Ausgangswert; für die gepoolten Analysen wurde zusätzlich Studie als fester Effekt einbezogen. Das Modell basiert auf allen verfügbaren Zeitpunkten vom Ausgangswert bis Woche 52; für die Analyse der gepoolten Daten wurden auch Messwerte nach Therapieabbruch berücksichtigt.									
d: Treated Set									
e: Kontrollierte Phase									
f: Negative Werte entsprechen einer Zunahme der FVC									
FVC: Forced Vital Capacity (Forcierte Vitalkapazität); k. A.: keine Angabe; KI: Konfidenzintervall; MW: Mittelwert; MWD: Mittelwertdifferenz; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler)									

Die *jährliche Abnahme der FVC* war sowohl in den betrachteten Einzelstudien als auch in den gepoolten Daten in der Nintedanib-Gruppe geringer als in der Kontrollgruppe. In der über die drei Studien gepoolten Analyse ergab sich unter Nintedanib eine mittlere jährliche FVC-Abnahme von 108,88 ml gegenüber 221,30 ml im Vergleichsarm; für die gepoolten INPULSIS Studien (bei einer Beobachtungsdauer von 52 Wochen) von 113,59 ml vs. 223,53 ml (Tabelle 4-22). Für die IPD Meta-Analyse über die drei Studien beträgt die adjustierte MWD [95%-KI]: 112,42 ml [79,06; 145,77]; $p < 0,0001$; für die beiden INPULSIS Studien 109,94 ml [75,85; 144,03]; $p < 0,0001$. Dies belegt eine geringere *jährliche Abnahme der FVC* im Nintedanib-Arm gegenüber dem Kontrollarm.

In den INPULSIS Studien lag die mittlere jährliche FVC-Abnahme in den Behandlungsarmen bei 114,65 ml vs. 239,91 ml (INPULSIS 1), bzw. 113,59 ml vs. 207,32 ml (INPULSIS 2). In der Studie 1199.187, die eine etwa halb so lange Nachbeobachtungsdauer hatte, zeigten die Patienten in der Nintedanib-Gruppe im Mittel eine jährliche Zunahme der FVC von 9,35 ml gegenüber einer jährlichen Abnahme von 185,72 ml in der Kontrollgruppe. Die adjustierte MWD [95%-KI] beträgt in der INPULSIS 1 Studie 125,26 ml [77,68; 172,84]; $p < 0,0001$, in der INPULSIS 2 Studie 93,73 ml [44,78; 142,68]; $p = 0,0002$ und in Studie 1199.187 195,07 ml [31,87; 358,26]; $p = 0,0196$.

Abbildung 12 zeigt die mittlere beobachtete Änderung der FVC für die gepoolte Auswertung der INPULSIS Studien und Studie 1199.187 im Studienverlauf.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 12: Mittlere FVC-Abnahme (Standardfehler des Mittelwertes, SEM) gegenüber dem Ausgangswert (ml) im Zeitverlauf - aus RCT mit dem zu bewertenden Arzneimittel (gepoolte Analyse der INPULSIS Studien und Studie 1199.187 [kontrollierte Phase])

In allen Studien führte Nintedanib zu einer geringeren *jährlichen FVC-Abnahme* im Vergleich zur zVT; bei einer Beobachtungsdauer von 24 Wochen wurde in Studie 1199.187 im Mittel sogar eine geringe Zunahme der FVC im Nintedanib-Arm beobachtet. Der Unterschied ist in allen Auswertungen statistisch signifikant; bezogen auf den Punktschätzer und den Schwellenwert von 100 ml für einen klinisch relevanten Gruppenunterschied, der mit der Zulassungsbehörde für die Fallzahlplanung der INPULSIS Studien vereinbart wurde, sind die in den Einzelstudien INPULSIS 1 und Studie 1199.187 sowie die in den gepoolten Analysen beobachteten Unterschiede auch klinisch relevant.

Wie in Abschnitt 4.5.4 detailliert ausgeführt, sieht BI hinreichend vorhandene Evidenz, den Endpunkt *jährliche FVC-Abnahme* (ml) als Surrogat für Mortalität zu behandeln. In diesem Abschnitt erfolgt zudem eine konkrete Ableitung der Auswirkungen der jährlichen Reduktion der FVC-Abnahme unter Nintedanib auf die Mortalität. Demnach entspricht eine Reduktion der *jährlichen FVC-Abnahme* von 112,42 ml, wie sie über die gepoolten INPULSIS Studien und die Studie 1199.187 beobachtet wurde (Tabelle 4-22), einer Reduktion des Mortalitätsrisikos um über 20%.

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.

Die Ergebnisse der IPD Meta-Analyse für den Endpunkt *jährliche FVC-Abnahme* wurden bereits in Tabelle 4-22 dargestellt.

Die grafische Darstellung der Meta-Analyse mittels eines Forest-Plots befindet sich in Abbildung 13. Er zeigt die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der IPD Meta-Analyse der Einzelstudien. Der Heterogenitätstest (siehe auch Abschnitt 4.3.1.3.5.1) weist auf eine homogene Datenlage hin (p-Wert des Interaktionstests [Merkmal *Studie*]: $p=0,3867$).

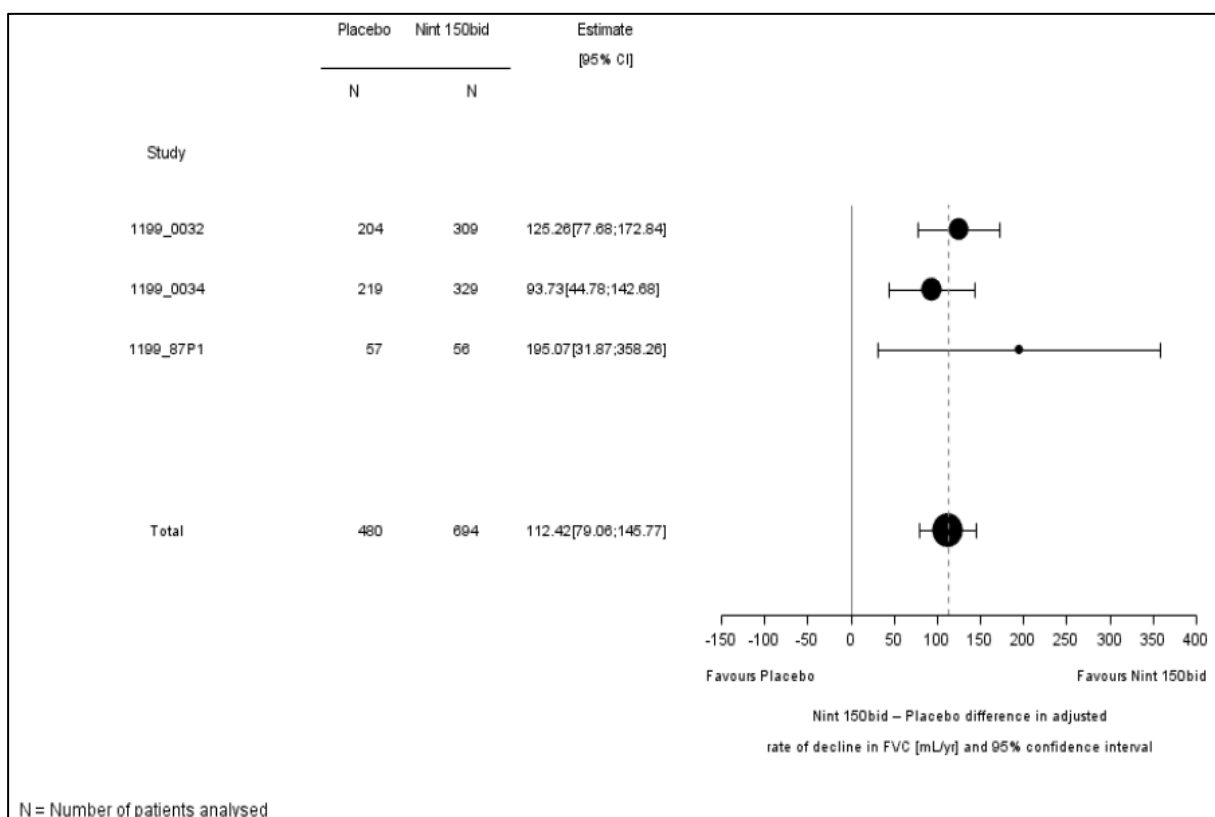


Abbildung 13: Forest-Plot zur Meta-Analyse für Endpunkt *jährliche FVC-Abnahme* - RCT mit dem zu bewertenden Arzneimittel

FVC-Responder

Die Analysen zu *FVC-Respondern zu Woche 52* waren für die INPULSIS Studien präspezifiziert. Für die Responderanalysen wurden zwei unterschiedliche Responsekriterien herangezogen: Das Absinken des FVC-Wertes (ausgedrückt in Prozent des Sollwertes) um nicht mehr als 5 bzw. 10 Prozentpunkte gegenüber dem Ausgangswert. Für Studie 1199.187, mit einer kontrollierten Phase von 24 Wochen, waren FVC-Responderanalysen nicht vorgesehen.

Die Betrachtung der Response nach 52 Wochen war somit nur für die INPULSIS Studien möglich. Die entsprechenden Ergebnisse sind in Tabelle 4-23 dargestellt. Im Kontext der Nutzenbewertung sind im Setting einer chronischen Erkrankung die Analysen zu Woche 52 als relevant zu sehen.

Tabelle 4-23: Ergebnisse für *FVC-Responder zu Woche 52* bezogen auf die Schwellenwerte für die absolute Verschlechterung gegenüber dem Ausgangswert um 10%-Punkte bzw. 5%-Punkte aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
FVC-Responder zu Woche 52 (Schwellenwert 10%-Punkte)							
INPULSIS (gepoolt) ^d	638	447 (70,06)	423	256 (60,52)	1,577	[1,21; 2,05]	0,0007
INPULSIS 1 ^d	309	218 (70,55)	204	116 (56,86)	1,914	[1,32; 2,79]	0,0007
INPULSIS 2 ^d	329	229 (69,60)	219	140 (63,93)	1,286	[0,89; 1,86]	0,1833
FVC-Responder zu Woche 52 (Schwellenwert 5%-Punkte)							
INPULSIS (gepoolt) ^d	638	338 (53,0)	423	164 (38,8)	1,835	[1,43; 2,36]	<0,0001
INPULSIS 1 ^d	309	163 (52,8)	204	78 (38,2)	1,857	[1,28; 2,66]	0,0010
INPULSIS 2 ^d	329	175 (53,2)	219	86 (39,3)	1,794	[1,26; 2,55]	0,0011
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.1.1.3:3, Tab. 15.2.2.2.1.1.3:1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.1.1.3:3, Tab. 15.2.2.2.1.1.3:1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c): Tab. 2.3.2.1.3.3, Tab. 2.3.2.1.3.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.28.6.6, Tab. 1.1.28.7.6 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Response: Studienteilnehmer werden als FVC-Responder betrachtet, wenn sie keinen absoluten Abfall der FVC (% des Sollwertes) von mehr als 10 Prozentpunkten (5 Prozentpunkten) hatten und ein Wert zu Woche 52 bzw. Woche 24 vorliegt. Studienteilnehmer mit fehlenden Daten werden als Non-Responder betrachtet. c: OR [95%-KI] und p-Wert berechnet mittels logistischer Regression mit Termen für Behandlung, Alter, Geschlecht, Größe und Ausgangswert-FVC (%); für die gepoolten Analysen wurde zusätzlich der Term Studie dem Modell hinzugefügt. d: Treated Set FVC: Forced Vital Capacity (Forcierte Vitalkapazität); KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)							

Responderanalyse – Änderung der FVC zu Woche 52

In der gepoolten Analyse der INPULSIS Studien war der Anteil *FVC-Responder* (Absinken des FVC-Wertes um ≤ 10 Prozentpunkte gegenüber dem Ausgangswert) in der Nintedanib-Gruppe größer als in der Kontrollgruppe (70,06% vs. 60,52%) mit einem OR [95%-KI] von 1,577 [1,21; 2,05]; p=0,0007. Die gepoolten Ergebnisse der INPULSIS Studien zeigen somit einen statistisch signifikanten Unterschied zwischen den Studiengruppen und belegen den Vorteil zugunsten von Nintedanib gegenüber der zVT.

Auch in den Einzelstudien INPULSIS 1 (70,55% vs. 56,86%) und INPULSIS 2 (69,60% vs. 63,93%) war der Anteil *FVC-Responder* in der Nintedanib-Gruppe größer als in der Kontrollgruppe. Der entsprechende Effektschätzer ergab ein OR [95%-KI] von 1,914 [1,32; 2,79]; $p=0,0007$ für die Studie INPULSIS 1 bzw. von 1,286 [0,89; 1,86]; $p=0,1833$ für die Studie INPULSIS 2.

Die entsprechende Responderanalyse bezogen auf den Schwellenwert für den absoluten Abfall der FVC (% des Sollwertes) von ≤ 5 Prozentpunkten zwischen Studienbeginn und Woche 52 zeigten vergleichbare Ergebnisse. In allen der Bewertung zugrunde liegenden Studien war der Anteil *FVC-Responder* in der Nintedanib-Gruppe größer als in der Kontrollgruppe. Alle Ergebnisse zum Endpunkt *FVC-Responder* mit Abfall der FVC (% des Sollwertes) von ≤ 5 Prozentpunkten waren statistisch signifikant zugunsten von Nintedanib.

Aus Gründen der Vollständigkeit sind die entsprechenden Analysen für den 24 Wochen-Zeitpunkt, zu dem Daten für alle drei Studien vorlagen, in Tabelle 4-24 dargestellt. Diese Analysen waren in keinem der entsprechenden CTP als Endpunkt präspezifiziert und werden daher lediglich ergänzend dargestellt.

Tabelle 4-24: Ergebnisse für *FVC-Responder zu Woche 24* bezogen auf die Schwellenwerte für die absolute Verschlechterung gegenüber dem Ausgangswert um 10%-Punkte bzw. 5%-Punkte aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
FVC-Responder zu Woche 24 (Schwellenwert 10%-Punkte)							
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	583 (84,01)	480	385 (80,21)	1,319	[0,97; 1,80]	0,0787
INPULSIS (gepoolt) ^d	638	541 (84,80)	423	342 (80,85)	1,370	[0,99; 1,91]	0,0614
INPULSIS 1 ^d	309	264 (85,44)	204	168 (82,35)	1,332	[0,82; 2,17]	0,2491
INPULSIS 2 ^d	329	277 (84,19)	219	174 (79,45)	1,389	[0,88; 2,18]	0,1545
Studie 1199.187 ^{d, e}	56	42 (75,00)	57	43 (75,44)	1,002	[0,41; 2,46]	0,9959
FVC-Responder zu Woche 24 (Schwellenwert 5%-Punkte)							
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	482 (69,45)	480	293 (61,04)	1,466	[1,15; 1,88]	0,0024
INPULSIS (gepoolt) ^d	638	444 (69,59)	423	259 (61,23)	1,468	[1,13; 1,91]	0,0039
INPULSIS 1 ^d	309	216 (69,90)	204	124 (60,78)	1,524	[1,05; 2,22]	0,0273
INPULSIS 2 ^d	329	228 (69,30)	219	135 (61,64)	1,398	[0,97; 2,01]	0,0718
Studie 1199.187 ^{d, e}	56	38 (67,86)	57	34 (59,65)	1,368	[0,61; 3,07]	0,4461

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.28.6.1 - 1.1.28.6.5, Tab. 1.1.28.7.1 - 1.1.28.7.5 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Response: Studienteilnehmer werden als FVC-Responder betrachtet, wenn sie keinen absoluten Abfall der FVC (% des Sollwertes) von mehr als 10 Prozentpunkten (5 Prozentpunkten) hatten und ein Wert zu Woche 52 bzw. Woche 24 vorliegt. Studienteilnehmer mit fehlenden Daten werden als Non-Responder betrachtet. c: OR [95%-KI] und p-Wert berechnet mittels logistischer Regression mit Termen für Behandlung, Alter, Geschlecht, Größe und Ausgangswert-FVC (%); für die gepoolten Analysen wurde zusätzlich der Term Studie dem Modell hinzugefügt. d: Treated Set e: Kontrollierte Phase FVC: Forced Vital Capacity (Forcierte Vitalkapazität); KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)							

Responderanalyse – Änderung der FVC zu Woche 24

Sowohl in der über die drei Studien gepoolten Analyse (84,0% vs. 80,2%) als auch in den Einzelstudien INPULSIS 1 (85,4% vs. 82,4%) und INPULSIS 2 (84,2% vs. 79,5%), war zu Woche 24 der Anteil der FVC-Responder (Schwellenwert <10 Prozentpunkte des Sollwertes) in der Nintedanib-Gruppe und der Kontrollgruppe vergleichbar mit einem numerischen Trend zugunsten von Nintedanib. Der Gruppenunterschied war weder auf Einzelstudien- noch in den IPD Meta-Analysen statistisch signifikant.

Für die Responderanalyse bezogen auf den Schwellenwert von <5 Prozentpunkten des Sollwertes lag der Anteil der Responder in der Nintedanib- bzw. Kontrollgruppe in der IPD Meta-Analyse über die drei Studien bei 69,5% vs. 61,0% und über die INPULSIS Studien bei 69,6% vs. 61,2%. Die IPD Meta-Analyse aller drei Studien zeigte ebenso wie die der INPULSIS Studien einen statistisch signifikanten Effekt zugunsten von Nintedanib (OR [95%-KI]: 1,466 [1,15; 1,88]; p=0,0024, bzw. 1,468 [1,13; 1,91]; p=0,0039).

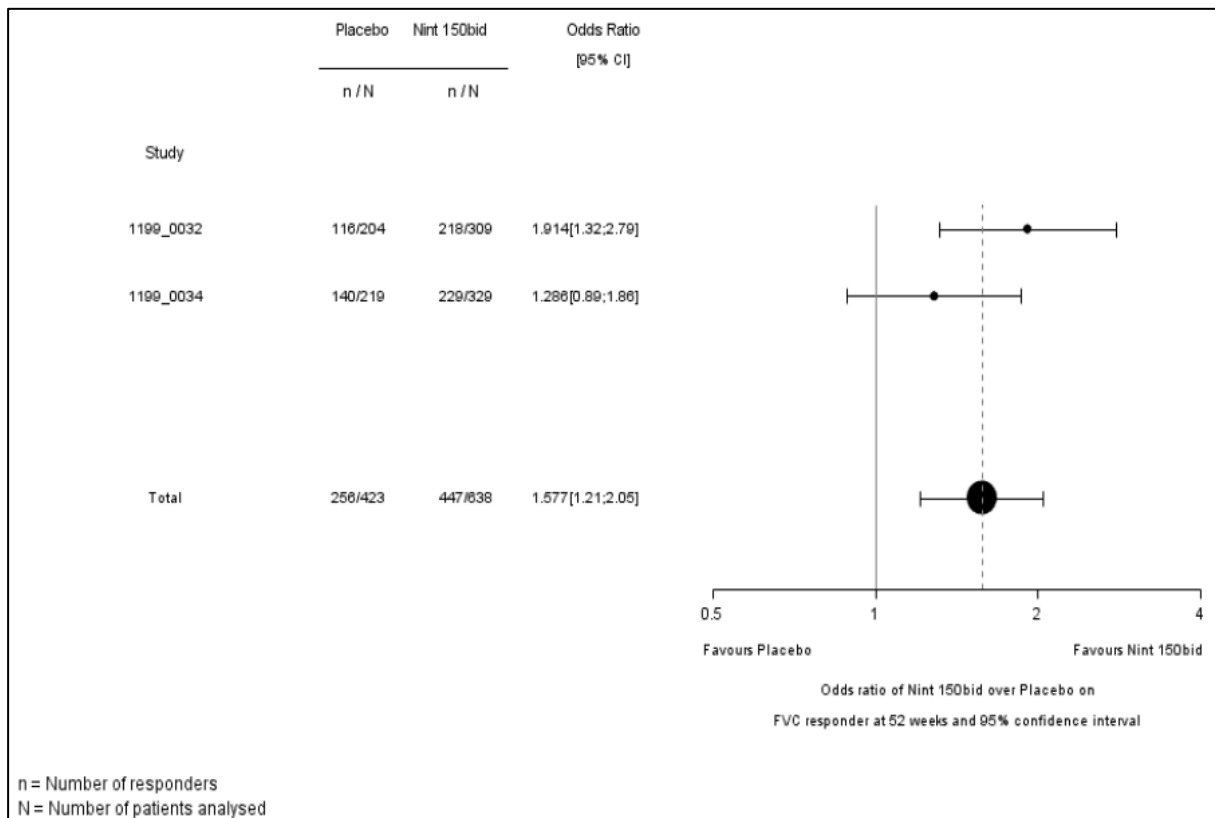
In den Einzelstudien lag der Anteil der Responder in der INPULSIS 1 Studie bei 69,9% vs. 60,8%, in der INPULSIS 2 Studie bei 69,3% vs. 61,6% und in Studie 1199.187 bei 67,9% vs. 59,7%. Für die Studie INPULSIS 1 war der Behandlungsunterschied statistisch signifikant zugunsten von Nintedanib. Das OR [95%-KI] lag bei 1,524 [1,05; 2,22]; p=0,0273. Für die Studien INPULSIS 2 und Studie 1199.187 war der Unterschied jeweils nicht statistisch signifikant (OR [95%-KI]: 1,398 [0,97; 2,01]; p=0,0718, bzw. 1,368 [0,61; 3,07]; p=0,4461), wies jedoch in beiden Fällen einen numerischen Trend zugunsten Nintedanib auf.

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.

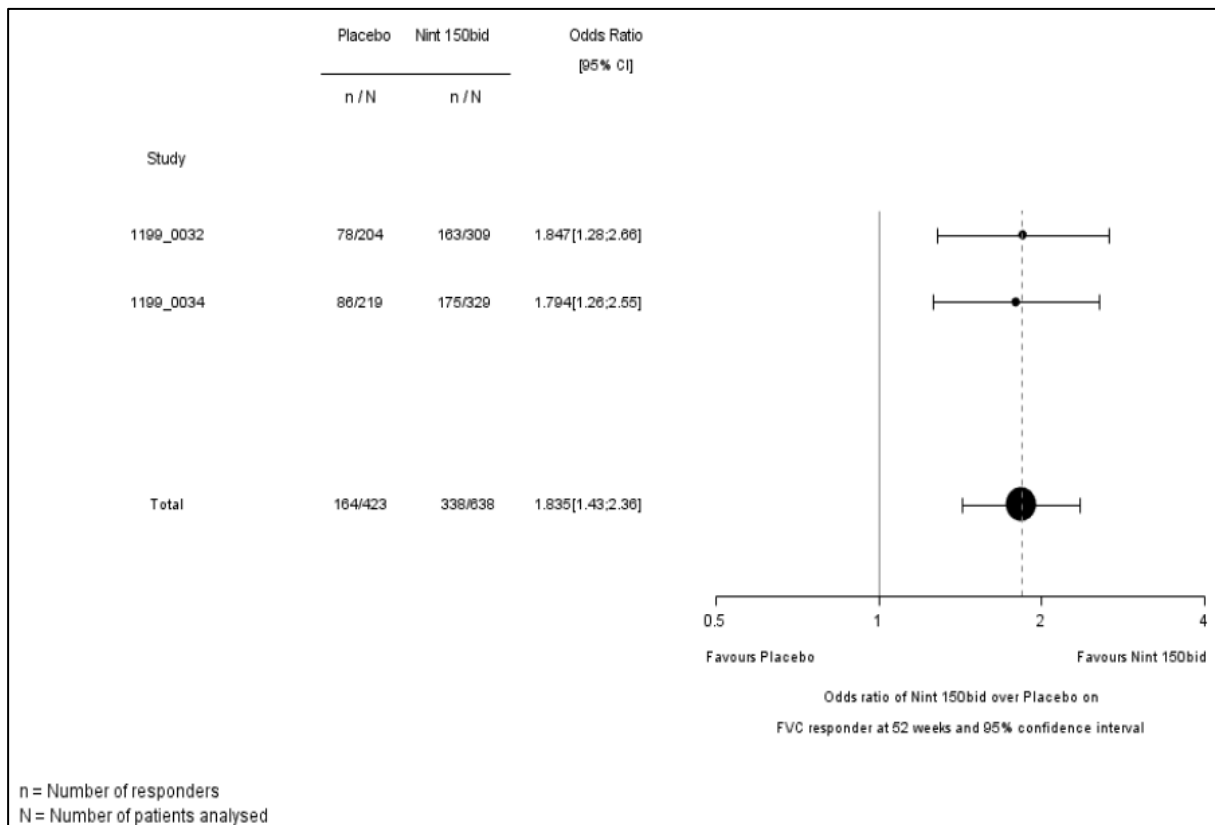
Im Falle der Responderanalysen wird für die Ableitung des Zusatznutzens die meta-analytische Zusammenfassung der längstmöglichen Studiendauer (über 52 Wochen), die für die zwei großen Zulassungsstudien INPULSIS 1 und INPULSIS 2 möglich war, herangezogen (Tabelle 4-23). Die Ergebnisse der meta-analytischen Zusammenfassung aller drei Studien, die nur über den kürzeren Zeitraum (24 Wochen) möglich war (Tabelle 4-24), wird supportiv verwendet; auf eine grafische Darstellung der Ergebnisse wird verzichtet.

Die grafische Darstellung der Meta-Analyse mittels Forest-Plots für die Endpunkte *FVC-Responder zu Woche 52 (Schwellenwert 10 Prozentpunkte)* und *FVC-Responder zu Woche 52 (Schwellenwert 5 Prozentpunkte)* befindet sich in Abbildung 14 und Abbildung 15. Sie zeigen die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der jeweiligen IPD Meta-Analyse der Einzelstudien. Die Heterogenitätstests (siehe auch Abschnitt 4.3.1.3.1) weisen auf eine homogene Datenlage hin (p-Werte der Interaktionstests für das Merkmal *Studie*: $p=0,1432$ bzw. $p=0,9313$).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 14: Forest-Plot zur Meta-Analyse für Endpunkt *FVC-Responder zu Woche 52 (Schwellenwert 10 Prozentpunkte)* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

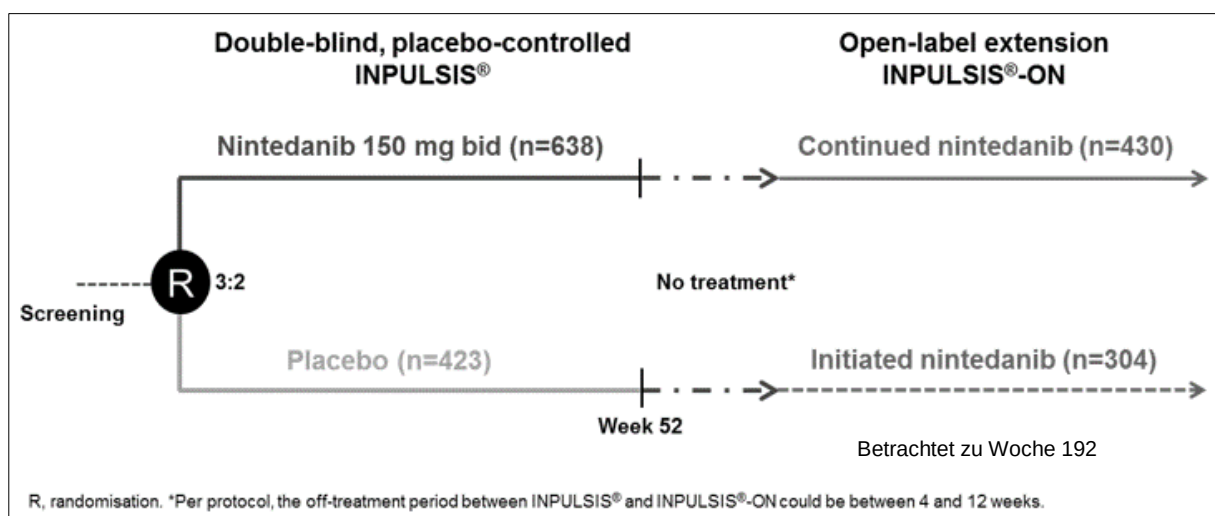
Abbildung 15: Forest-Plot zur Meta-Analyse für Endpunkt *FVC-Responder zu Woche 52* (Schwellenwert 5 Prozentpunkte) - RCT mit dem zu bewertenden Arzneimittel

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Die deutsche S2k-Leitlinie sieht in der longitudinalen Veränderung physiologischer Parameter wie der FVC einen der wichtigsten Prädiktoren für den weiteren Krankheitsverlauf (Behr, 2013). Die Bestimmung der FVC wurde in den berücksichtigten Studien anhand eines standardisierten spirometrischen Verfahrens nach den Kriterien der ATS/ERS/JRS/ALAT durchgeführt. Die Endpunkte *jährliche Abnahme der FVC* und *Anteil FVC-Responder* mit den Schwellenwerten 5% bzw. 10% des Sollwertes sind eindeutig definiert und objektiv zu erheben. Eine Abnahme der FVC um 5% bzw. 10% ist mit einem deutlich erhöhten Sterberisiko verknüpft, die gewählten Responsekriterien entsprechen daher geeigneten Schwellenwerten zur Beurteilung der Mortalität (Abschnitt 4.5.4). Für die drei Endpunkte kann somit von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ausgegangen werden.

Endpunkt *jährliche Abnahme der FVC* über den Zeitraum von einem Jahr hinaus (INPULSIS-ON Studie)

Als chronische Erkrankung bedarf die IPF einer langfristigen Behandlung. Um eine Aussage über die Wirksamkeit und Sicherheit von Nintedanib über einen Zeitraum von 52 Wochen hinaus treffen zu können, wurde die an die INPULSIS Studien anknüpfende Open-label-Extensionsstudie INPULSIS-ON (Studie 1199.33) über 192 Wochen durchgeführt. In dieser wurden 734 Patienten aus beiden Armen der INPULSIS Studien, wovon 430 Patienten Nintedanib und 304 Patienten Placebo erhalten hatten, eingeschlossen und mit Nintedanib 150 mg bid behandelt (Abbildung 16). Die mediane Behandlungsdauer mit Nintedanib über die Studien betrug 44,7 Monate (mind. 11,9 Monate, max. 68,3 Monate), in der INPULSIS-ON alleine 31,5 Monate (0,0-56,3 Monate).

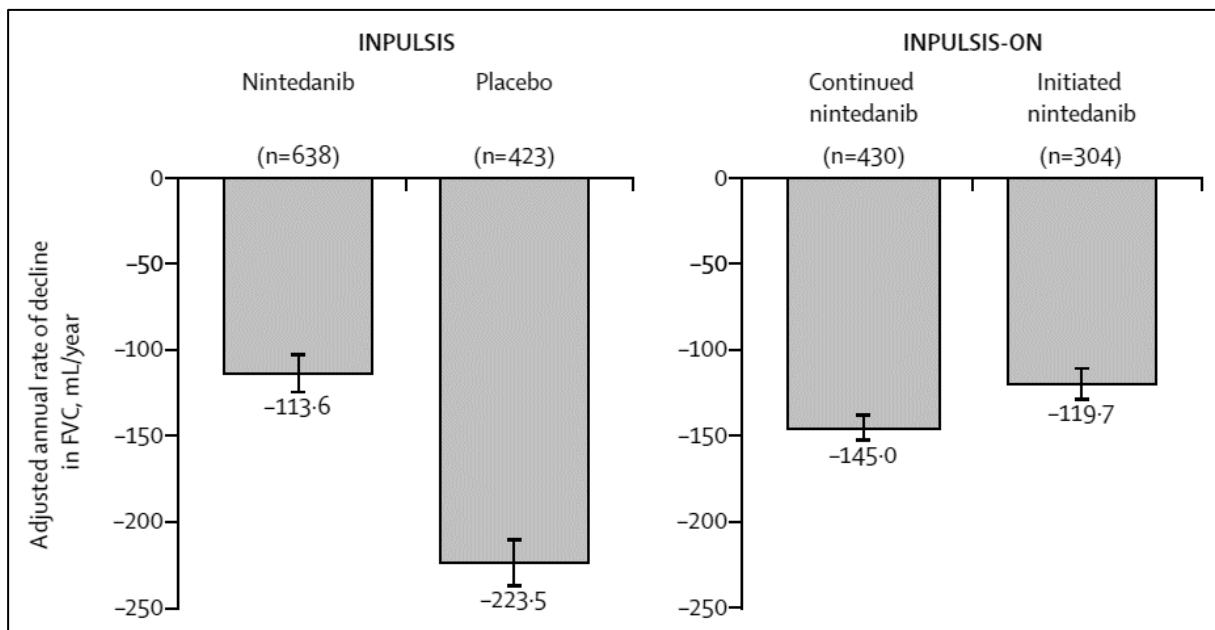


Quelle: (Crestani et al., 2017)

Abbildung 16: Studiendesign der INPULSIS Studien und der Studie INPULSIS-ON

Die Verringerung der FVC-Abnahme unter einer Therapie mit Nintedanib zeigt sich auch über die im vorliegenden Dossier abgebildete Studiendauer von bis zu einem Jahr hinaus. So lag in der Extensionsstudie INPULSIS-ON über 192 Wochen die *jährliche Abnahme der FVC* deutlich unterhalb des Verlustes des Lungenvolumens der Placebo-Gruppe der INPULSIS Studien. Abbildung 17 zeigt die *jährliche Abnahme der FVC* (ml) in den gepoolten INPULSIS Studien und der Extensionsstudie INPULSIS-ON. In den gepoolten INPULSIS Studien war unter Behandlung mit Nintedanib eine deutlich geringere FVC-Abnahme nach 52 Wochen (113,6 ml/Jahr) zu verzeichnen als der mittlere Verlust der FVC bei IPF-Patienten ohne Nintedanib-Behandlung, welcher bei 150-200 ml pro Jahr liegt (Ley et al., 2011; Behr et al., 2013). Dieser Effekt war ebenso in INPULSIS-ON zu beobachten: Die *jährliche FVC-Abnahme* berechnet über 192 Wochen betrug in allen Patienten 135,1 ml/Jahr, 145,0 ml/Jahr in Patienten, die die Nintedanib-Therapie in INPULSIS-ON fortsetzten und 119,7 ml/Jahr in Patienten, die die Nintedanib-Therapie in INPULSIS-ON begannen. Die Werte für die *jährliche FVC-Abnahme* der INPULSIS Studien über 52 Wochen waren konsistent mit denen der Extensionsstudie über 192 Wochen. Daraus ist erkennbar, dass eine

Therapie mit Nintedanib einen Langzeiteffekt auf den Erhalt des Lungenvolumens über mehr als vier Jahre aufweist.



Quelle: (Crestani et al., 2018)

Abbildung 17: Jährliche Abnahme der FVC (ml) in den gepoolten INPULSIS Studien über 52 Wochen (links) und in der Extensionsstudie INPULSIS-ON über 192 Wochen (rechts). Fehlerbalken stellen Standardfehler dar.

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 Morbidität – 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.

4.3.1.3.2.1 Endpunkte *akute Exazerbationen*

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

Tabelle 4-25: Operationalisierung von *akute Exazerbationen*

Studie	Operationalisierung
INPULSIS 1 ^a	<u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u> Definition: Zeit von Randomisierung bis zur ersten akuten Exazerbation, die nachträglich durch ein IAC bestätigt wurde Zensierung von Patienten ohne Ereignis: 372 Tage nach Randomisierung (52 Wochen + 7 Tage Toleranz) bzw. beim letzten Patientenkontakt, definiert als Datum des letzten im CRF dokumentierten Ereignisses Erhebung: - kontinuierlich durch Prüfarzt (als UE), gemäß im Studienprotokoll präspezifizierter Definition für Exazerbationen - Studienteilnehmer wurden sensibilisiert dem Prüfarzt auftretende Verschlechterungen in Bezug auf Dyspnoe frühzeitig zu berichten - Dokumentation aller relevanten klinischen Daten zu jeder Exazerbation in spezifischem Fragebogen (klinischer Verlauf, Anzeichen und Symptome, Laborergebnisse, Lungenfunktion, Ergebnisse Bildgebungsverfahren, erforderliche Behandlung) <i>Akute Exazerbationen</i> waren definiert als akute Verschlechterungen des klinischen Zustandes ohne erkennbare andere Ursachen, einschließlich Vorliegen aller folgenden Merkmale: - Unerklärtes Neuauftreten oder Zunahme von Dyspnoe innerhalb der letzten 30 Tage - Neue diffuse pulmonale Infiltrate auf Röntgenbild des Thorax und/oder neue mittels HRCT festgestellte parenchymale Anomalien ohne Pneumothorax oder Pleuraergüsse (neue Milchglasverschattungen) seit der letzten Visite - Ausschluss von Infektionen gemäß üblicher klinischer Praxis oder mikrobiologischen Studien - Ausschluss anderer möglicher Erklärungen gemäß üblicher klinischer Praxis, einschließlich Linksherzversagen, Lungenembolie und akuter Lungenverletzung Nachträgliche Adjudizierung durch das IAC^b Die Adjudizierung erfolgte verblindet, wobei alle berichteten Exazerbationen basierend auf den vorliegenden Unterlagen anhand der oben genannten Kriterien wie folgt kategorisiert wurden: - bestätigte Exazerbation - vermutete Exazerbation - keine Exazerbation Sowohl <i>bestätigte</i> als auch <i>vermutete Exazerbationen</i> wurden bei der Analyse als Ereignis betrachtet. <u>Zeit bis zur ersten akuten Exazerbation</u> Definition: Zeit von Randomisierung bis zur ersten akuten Exazerbation Zensierung von Patienten ohne Ereignis: 372 Tage nach Randomisierung (52 Wochen + 7 Tage Toleranz) bzw. beim letzten Patientenkontakt, definiert als Datum

Studie	Operationalisierung
	des letzten im CRF dokumentierten Ereignisses Erhebung und Methodik erfolgte analog des Endpunkts <i>Zeit bis zur ersten adjudizierten akuten Exazerbation</i> , jedoch ohne Adjudizierung durch das IAC
INPULSIS 2 ^a	Homogenes Studiendesign wie INPULSIS 1
Studie 1199.187 ^a	<u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u> Definition: Zeit von Randomisierung bis zur ersten akuten Exazerbation, die nachträglich durch ein IAC bestätigt wurde Zensierung von Patienten ohne Ereignis: Für Kaplan-Meier-Schätzung: an Tag 176; für die Cox-Regression am End-Datum der kontrollierten Phase Erhebung und Adjudizierung analog zu INPULSIS Studien <u>Zeit bis zur ersten akuten Exazerbation</u> Definition: Zeit von Randomisierung bis zur ersten akuten Exazerbation Zensierung von Patienten ohne Ereignis: Für Kaplan-Meier-Schätzung: an Tag 176; für die Cox-Regression am End-Datum der kontrollierten Phase
a: Die Operationalisierung für die gepoolten Auswertungen entspricht der Operationalisierung in den Einzelstudien. b: Das IAC bestand aus drei Fachärzten für Pneumologie mit spezieller Expertise auf dem Gebiet der IPF. CRF: Case Report Form; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); IAC: Independent Adjudication Committee (Unabhängiges Bewertungskomitee); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); UE: Unerwünschtes Ereignis	

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 4-26: Bewertung des Verzerrungspotenzials für den Endpunkt *akute Exazerbationen* 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
INPULSIS 1 ^a	<u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zur ersten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
INPULSIS 2 ^a	<u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zur ersten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187 ^a	<u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Zeit bis zur ersten akuten Exazerbation</u>					
	niedrig	ja	ja	ja	ja	niedrig
a: Das Verzerrungspotenzial der gepoolten Auswertungen ergibt sich aus dem Verzerrungspotenzial der Einzelstudien. ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für die Endpunkte *Zeit bis zur ersten (adjudizierten) akuten Exazerbation* erfolgte doppelblind. Das Vorliegen einer *akuten Exazerbation* wurde durch den Prüfarzt festgestellt. Zusätzlich wurden andere Ursachen als die zugrundeliegende IPF für die als Exazerbation imponierende akute Verschlechterung des Gesundheitszustandes durch ein verblindetes IAC anhand der Patientenakten ausgeschlossen. Dadurch kann von einer objektiven Einschätzung des Ereignisses Exazerbation ausgegangen werden.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und der Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen haben. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für die Endpunkte *akute Exazerbationen* waren a priori definiert. Die Ergebnisse aller drei Studien wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Abweichungen der vorliegenden Berichterstattung vom Studienbericht sind durch die Anforderungen der Nutzenbewertungen begründet.¹⁷

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war transparent, Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

Exazerbationen sind seltene, aber potenziell lebensbedrohliche Ereignisse bei IPF-Patienten, die hinsichtlich ihrer Klinik manchmal nur schwer von ähnlichen Ereignissen, die mit einer plötzlichen Verschlechterung des Gesundheitszustandes einhergehen, wie sie z. B. im Rahmen respiratorischer Infektionen auftreten, abgegrenzt werden können. Um diesem Umstand Rechnung zu tragen, wurde in den vorliegenden Studien jeweils eine zusätzliche Überprüfung der durch den Prüfarzt gestellten Diagnose durchgeführt. Die Erhebung der *Zeit bis zur ersten adjudizierten akuten Exazerbation* unterscheidet sich dadurch maßgeblich von der Erhebung der *Zeit bis zur ersten akuten Exazerbation*. Fachärzte für Pneumologie mit spezieller Expertise auf dem Gebiet der IPF haben auf Basis einer strengen Definition für IPF-getriggerte Exazerbationen diesen patientenrelevanten Endpunkt - ebenfalls wie die Prüfarzte verblindet - bewertet. Damit ist der Stellenwert der adjudizierten Ereignisse als präziser einzustufen, als derjenige des Endpunkts *Zeit bis zur ersten akuten Exazerbation* (ohne anschließende Adjudizierung).

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* ein niedriges Verzerrungspotenzial.

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

Die Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* wurden in den INPULSIS Studien sowie in Studie 1199.187 erhoben. Die Analysen der über alle drei Studien bzw. den INPULSIS Studien gepoolten Daten sowie die Daten zur Einzeldarstellung der INPULSIS Studien berichten das 12-Monats-Überleben, die

¹⁷ Abweichend vom Studienbericht wird für die Studie 1199.187 im vorliegenden Nutzendossier für die Analyse der Endpunkte *akute Exazerbationen* die gesamte kontrollierte Phase herangezogen (ohne Einschränkung auf die ersten 24 Wochen), um den üblichen Anforderungen der Nutzenbewertung entsprechend die längste unverzerrt bewertbare Beobachtungsdauer zugrunde zu legen.

Analyse der Einzelstudie 1199.187 das 6-Monats-Überleben. Im Kontext der Nutzenbewertung wird die Analyse der gepoolten Daten über alle drei Studien als relevant betrachtet.

Tabelle 4-27: Ergebnisse für *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	HR ^c	95 %- KI ^c	p- Wert ^c
Zeit bis zur ersten adjudizierten akuten Exazerbation									
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	13 (1,9)	0,979 ^e	480	26 (5,4)	0,939 ^e	0,32	[0,17; 0,63]	0,0010
INPULSIS ^d (gepoolt)	638	12 (1,9)	0,980 ^e	423	24 (5,7)	0,940 ^e	0,32	[0,16; 0,65]	0,0010
INPULSIS 1 ^d	309	7 (2,3)	0,975 ^e	204	8 (3,9)	0,959 ^e	0,55	[0,20; 1,54]	0,3019
INPULSIS 2 ^d	329	5 (1,5)	0,984 ^e	219	16 (7,3)	0,922 ^e	0,20	[0,07; 0,56]	0,0006
Studie 1199.187 ^d	56	1 (1,8)	k. A.	57	2 (3,5)	k. A.	0,39	[0,03; 4,91]	0,5763
Zeit bis zur ersten akuten Exazerbation									
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	32 (4,6)	0,949 ^e	480	34 (7,1)	0,921 ^e	0,63	[0,39; 1,02]	0,0695
INPULSIS ^d (gepoolt)	638	31 (4,9)	0,948 ^e	423	32 (7,6)	0,920 ^e	0,64	[0,39; 1,05]	0,0823
INPULSIS 1 ^d	309	19 (6,1)	0,934 ^e	204	11 (5,4)	0,944 ^e	1,15	[0,54; 2,42]	0,6728
INPULSIS 2 ^d	329	12 (3,6)	0,961 ^e	219	21 (9,6)	0,898 ^e	0,38	[0,19; 0,77]	0,0050
Studie 1199.187 ^d	56	1 (1,8)	0,978 ^f	57	2 (3,5)	0,982 ^f	0,39	[0,03; 4,86]	0,5800
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.1.3; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.1.3; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c): Tab. 2.2.2.1.1, Tab. 2.2.2.2.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017): Tab. 1.3.49.2.3.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 86.17.1.1, Tab. 86.17.1.3, Tab. 86.16.1.1, Tab. 86.16.1.3, Tab. 1.3.49.2.3.2, Tab. 1.3.49.2.4.2, Tab. 1.1.28.4.2, Tab. 1.1.28.3.2 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: HR [95%-KI] berechnet mittels Cox-Regression mit Termen für Behandlung, Geschlecht, Alter und Größe (bei den gepoolten Analysen wurde zusätzlich der Term Studie hinzugefügt); p-Wert berechnet mittels Log-rank-Test; für die									

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	HR ^c	95 %- KI ^c	p- Wert ^c
gepoolte Analyse aller drei Studien stratifiziert für Studie d: Treated Set e: 12-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 372) f: 6-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 176) HR: Hazard Ratio; k. A.: keine Angabe; KI: Konfidenzintervall; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)									

Für die Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* wurde die Wahrscheinlichkeit, keine Exazerbation zu erleiden, mittels Kaplan-Meier-Schätzer pro Behandlungsgruppe dargestellt. Das HR [95%-KI] beschreibt den Effekt der Zeit bis zum Erreichen des Ereignisses *akute Exazerbation* (Tabelle 4-27).

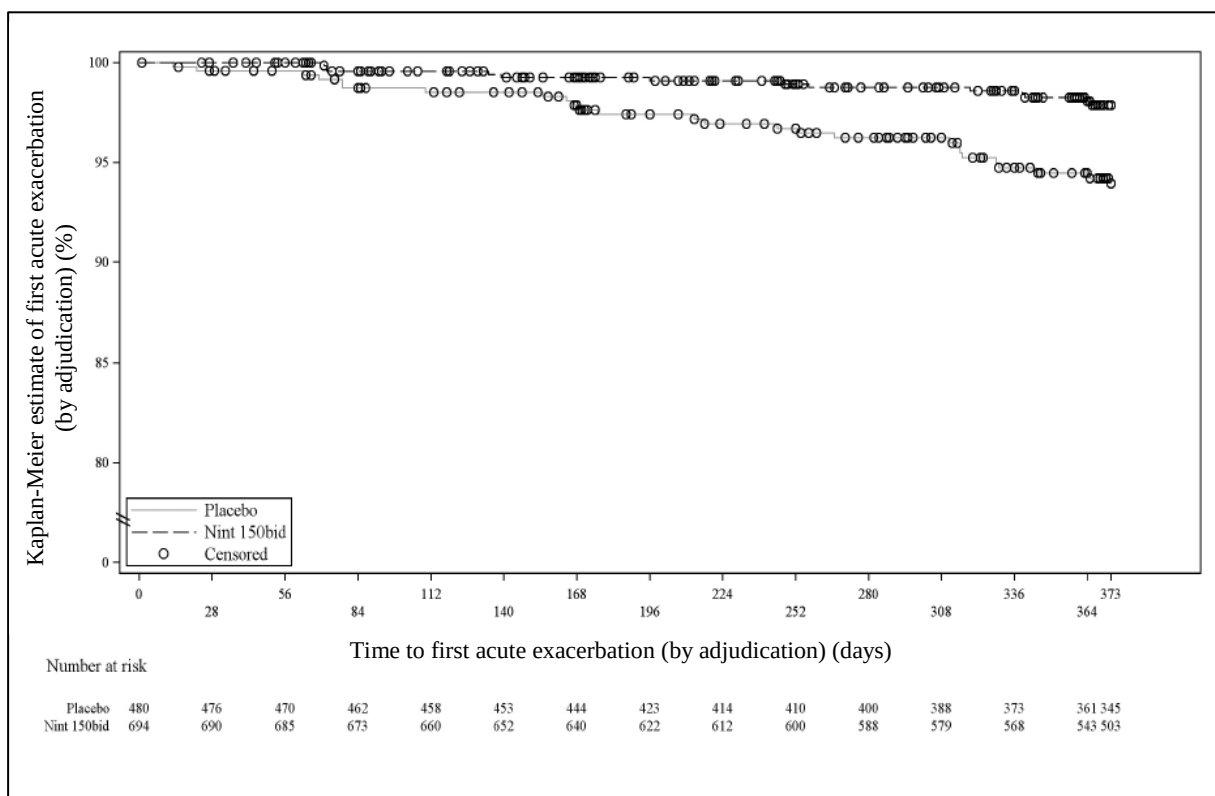
In den gepoolten Analysen trat eine *adjudizierte akute Exazerbation* bei einem geringeren Anteil der mit Nintedanib behandelten Studienteilnehmer auf, als dies in der Kontrollgruppe der Fall war (INPULSIS Studien und Studie 1199.187 (gepoolt): 1,9% vs. 5,4%, INPULSIS Studien (gepoolt): 1,9% vs. 5,7%). Das HR [95%-KI] beträgt für die IPD Meta-Analyse aller drei Studien 0,32 [0,17; 0,63]; p=0,0010, das für die IPD Meta-Analyse der INPULSIS Studien 0,32 [0,16; 0,65]; p=0,0010. Der Unterschied zugunsten von Nintedanib im Auftreten von *adjudizierten akuten Exazerbationen* ist in der gepoolten Auswertung aller drei Studien sowie in der gepoolten Auswertung der INPULSIS Studien statistisch signifikant.

Auch in den Einzelstudien trat eine *adjudizierte akute Exazerbation* unter Nintedanib bei einem geringeren Anteil an Studienteilnehmern auf (INPULSIS 1: 2,3% vs. 3,9%, INPULSIS 2: 1,5% vs. 7,3% und Studie 1199.187: 1,8% vs. 3,5%). Für die Einzelstudien ergab sich ein Effektschätzer von 0,55 [0,20; 1,54]; p=0,3019 für INPULSIS 1, von 0,20 [0,07; 0,56]; p=0,0006 für INPULSIS 2 bzw. von 0,39 [0,03; 4,91]; p=0,5763 für die Studie 1199.187.

Vom Prüfarzt diagnostizierte *akute Exazerbationen (nicht adjudiziert)* traten in den gepoolten Daten aller drei Studien und der INPULSIS Studien seltener im Nintedanib-Arm auf als im Vergleichsarm (INPULSIS Studien und Studie 1199.187 [gepoolt]: 4,6% vs. 7,1%, INPULSIS Studien [gepoolt]: 4,9% vs. 7,6%). Das HR [95%-KI] beträgt für die IPD Meta-Analyse aller drei Studien 0,63 [0,39; 1,02]; p=0,0695, das für die IPD Meta-Analyse der INPULSIS Studien 0,64 [0,39; 1,05]; p=0,0823.

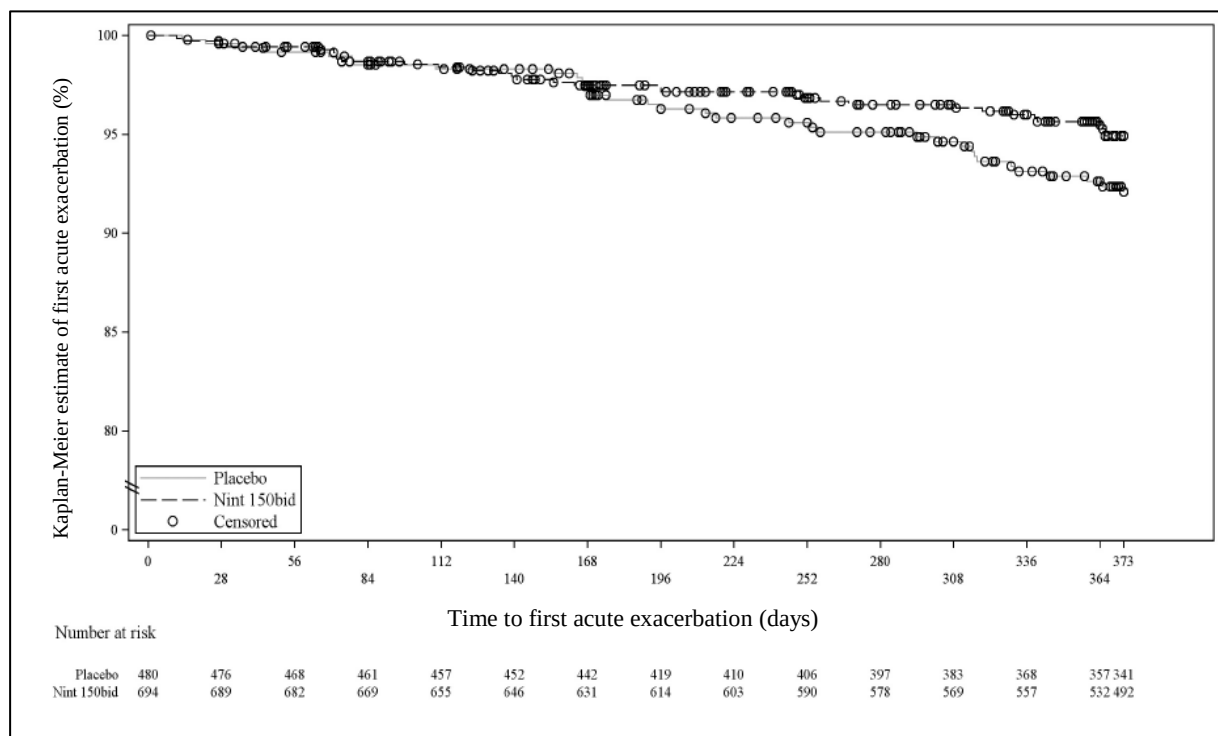
Auch in den einzeln betrachteten Studien INPULSIS 2 und der Studie 1199.187 traten *akute Exazerbationen (nicht adjudiziert)* seltener im Nintedanib-Arm auf als im Vergleichsarm (INPULSIS 2: 3,6% vs. 9,6% und Studie 1199.187: 1,8% vs. 3,5%). In der INPULSIS 1 Studie traten solche Ereignisse in beiden Studienarmen ähnlich häufig auf (INPULSIS 1: 6,1% vs. 5,4%). Für die Studien ergab sich ein Effektschätzer von HR [95%-KI]: 1,15 [0,54; 2,42]; $p=0,6728$ für INPULSIS 1, 0,38 [0,19; 0,77]; $p=0,0050$ für INPULSIS 2 und 0,39 [0,03; 4,86]; $p=0,5800$ für Studie 1199.187. Insgesamt zeichnet sich für den Endpunkt *akute Exazerbationen* eine klare Tendenz hinsichtlich eines reduzierten Risikos unter der Behandlung mit Nintedanib ab.

Die Kaplan-Meier-Kurven zu den Endpunkten *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation (nicht adjudiziert)* auf Basis der gepoolten Daten der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 sind in Abbildung 18 und in Abbildung 19 dargestellt.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 18: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zur ersten adjudizierten akuten Exazerbation* aus RCT mit dem zu bewertenden Arzneimittel; gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 19: Kaplan-Meier-Kurven für Endpunkt *Zeit bis zur ersten akuten Exazerbation* aus RCT mit dem zu bewertenden Arzneimittel; gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

Die entsprechenden Auswertungen zum Auftreten von Exazerbationen, die durch den Prüfarzt als SUE bewertet wurden, sind in Tabelle 4-28 dargestellt. Hier zeigte sich sowohl für die *Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation* (HR [95%-KI]: 0,30 [0,15; 0,62]; p-Wert aus Log-rank-Test stratifiziert nach Studie: p=0,0015; p-Wert aus Wald-Test im adjustierten Cox-Modell: p=0,0012) als auch für die *Zeit bis zur ersten schwerwiegenden akuten Exazerbation* (HR [95%-KI]: 0,55 [0,32; 0,96]; p-Wert aus Log-rank-Test stratifiziert nach Studie: p=0,0545; p-Wert aus Wald-Test im adjustierten Cox-Modell: p=0,0337) ein statistisch signifikanter Vorteil zugunsten der Therapie mit Nintedanib.

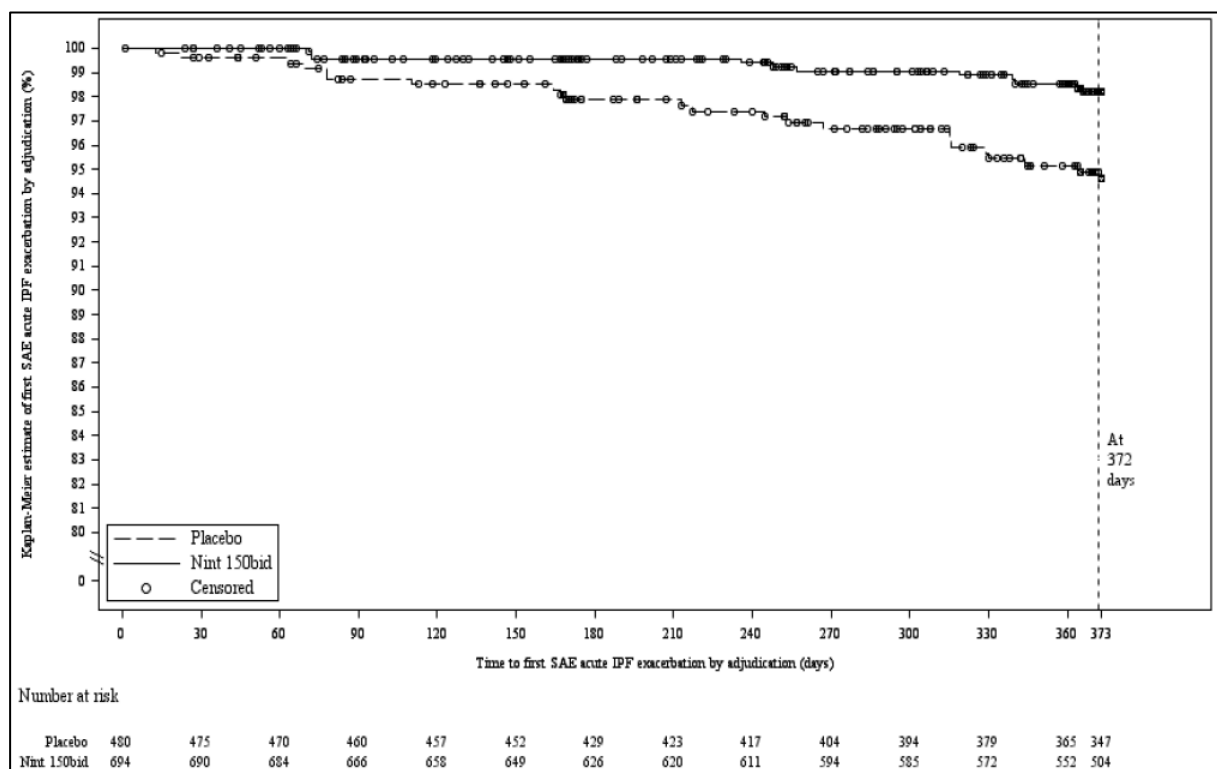
Tabelle 4-28: Sensitivitätsanalyse - Ergebnisse für *Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation* und *Zeit bis zur ersten schwerwiegenden akuten Exazerbation* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	HR ^c	95 %- KI ^c	p- Wert ^c
Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation (Sensitivitätsanalyse)									
IMPULSIS + Studie 1199.187 (gepoolt) ^d	694	11 (1,6)	0,982	480	23 (4,8)	0,946	0,30	[0,15; 0,62]	0,0015
IMPULSIS ^d (gepoolt)	638	10 (1,6)	0,983	423	21 (5,0)	0,948	0,30	[0,14; 0,64]	0,0016
IMPULSIS 1 ^d	309	7 (2,3)	0,975	204	6 (2,9)	0,969	0,74	[0,25; 2,22]	0,6700
IMPULSIS 2 ^d	329	3 (0,9)	0,990	219	15 (6,8)	0,927	0,13	[0,04; 0,44]	0,0001
Studie 1199.187 ^d	56	1 (1,8)	0,952	57	2 (3,5)	0,945	0,41	[0,03; 5,34]	0,6175
Zeit bis zur ersten schwerwiegenden akuten Exazerbation (Sensitivitätsanalyse)									
IMPULSIS + Studie 1199.187 (gepoolt) ^d	694	24 (3,5)	0,962 ^e	480	28 (5,8)	0,934 ^e	0,55	[0,32; 0,96]	0,0545
IMPULSIS ^d (gepoolt)	638	23 (3,6)	0,962	423	26 (6,1)	0,935	0,57	[0,32; 0,99]	0,0632
IMPULSIS 1 ^d	309	15 (4,9)	0,948 ^e	204	7 (3,4)	0,964 ^e	1,35	[0,55; 3,34]	0,4076
IMPULSIS 2 ^d	329	8 (2,4)	0,974 ^e	219	19 (8,7)	0,908 ^e	0,27	[0,12; 0,63]	0,0011
Studie 1199.187 ^d	56	1 (1,8)	0,952 ^f	57	2 (3,5)	0,945 ^f	0,41	[0,03; 5,34]	0,6175

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	N ^a	n ^b (%)	Wahr- scheinlichkeit, keine Exazerbation zu erleiden	HR ^c	95 %- KI ^c	p- Wert ^c

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.71.7.1.1, Tab. 1.3.71.7.1.3, Tab. 1.3.71.7.2.1, Tab. 1.3.71.7.2.3

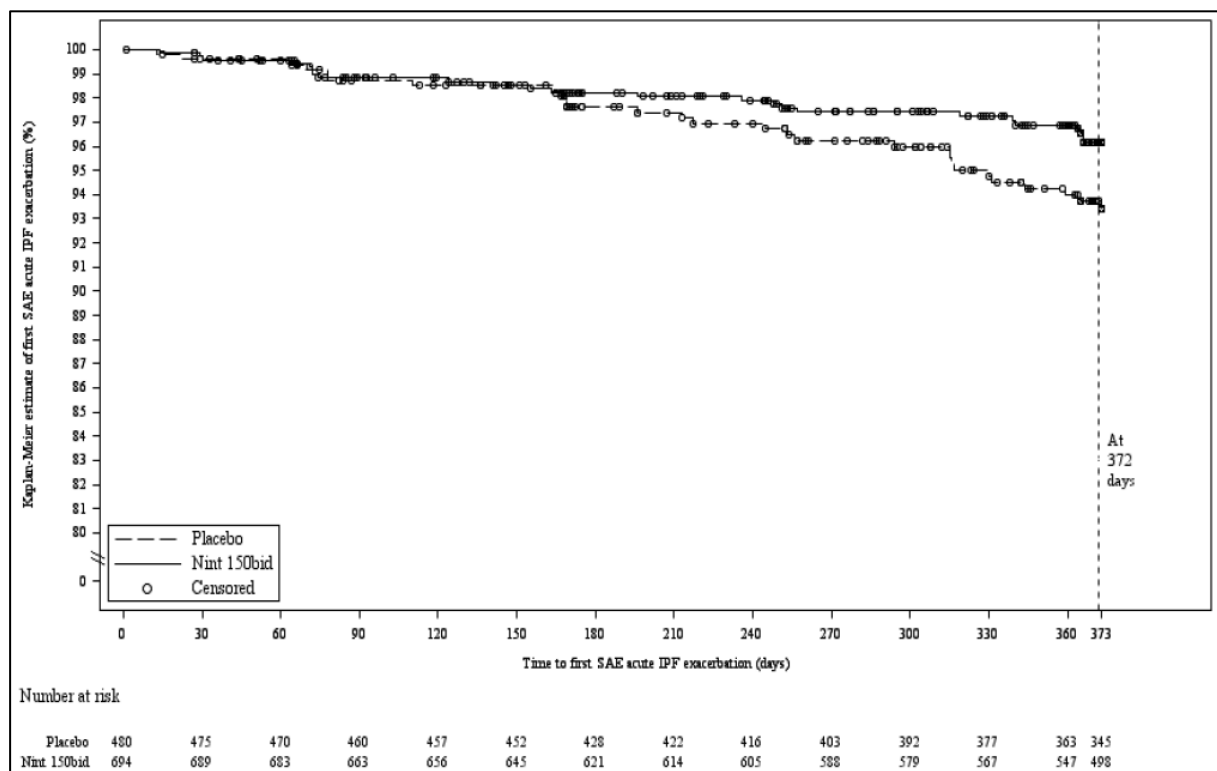
a: Anzahl analysierter Studienteilnehmer
b: Anzahl Studienteilnehmer mit Ereignis
c: HR [95%-KI] berechnet mittels Cox-Regression mit Termen für Behandlung, Geschlecht, Alter und Größe (bei den gepoolten Analysen wurde zusätzlich der Term Studie hinzugefügt); p-Wert berechnet mittels Log-Rank-Test, für die gepoolte Analyse aller drei Studien stratifiziert für Studie; p-Wert aus Wald-Test im adjustierten Cox-Modell für die gepoolte Analyse aller drei Studien: *Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation*: p=0,0012, *Zeit bis zur ersten schwerwiegenden akuten Exazerbation*: p= 0,0337; p-Wert aus Wald-Test im adjustierten Cox-Modell für die gepoolte Analyse der INPULSIS Studien: *Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation*: p=0,0019, *Zeit bis zur ersten schwerwiegenden akuten Exazerbation*: p= 0,0476
d: Treated Set
e: 12-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 372)
f: 6-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 176)
HR: Hazard Ratio; KI: Konfidenzintervall; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 20: Kaplan-Meier-Kurven für Endpunkt Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation aus RCT mit dem zu bewertenden Arzneimittel

(Sensitivitätsanalyse); gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)



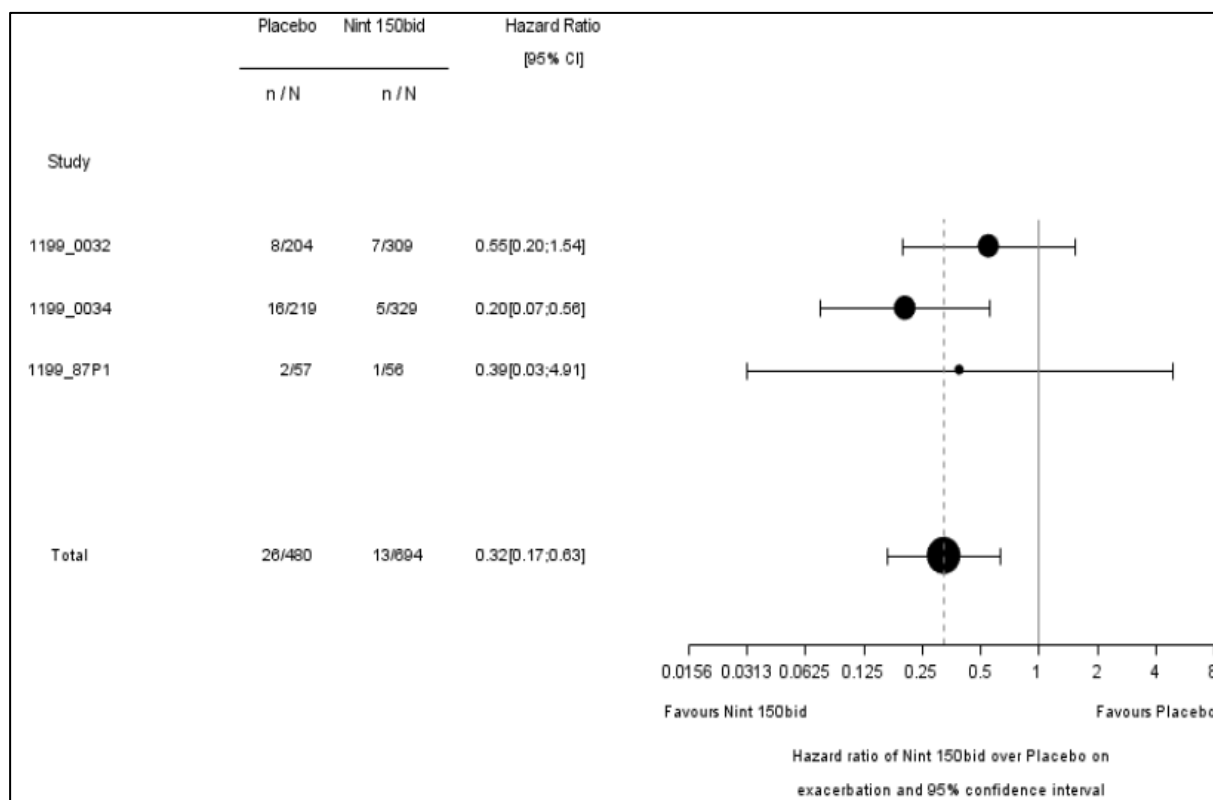
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 21: Kaplan-Meier-Kurven für Endpunkt Zeit bis zur ersten schwerwiegenden akuten Exazerbation aus RCT mit dem zu bewertenden Arzneimittel (Sensitivitätsanalyse); gepoolte Analyse der INPULSIS Studien und Studie 1199.187 (kontrollierte Phase)

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.

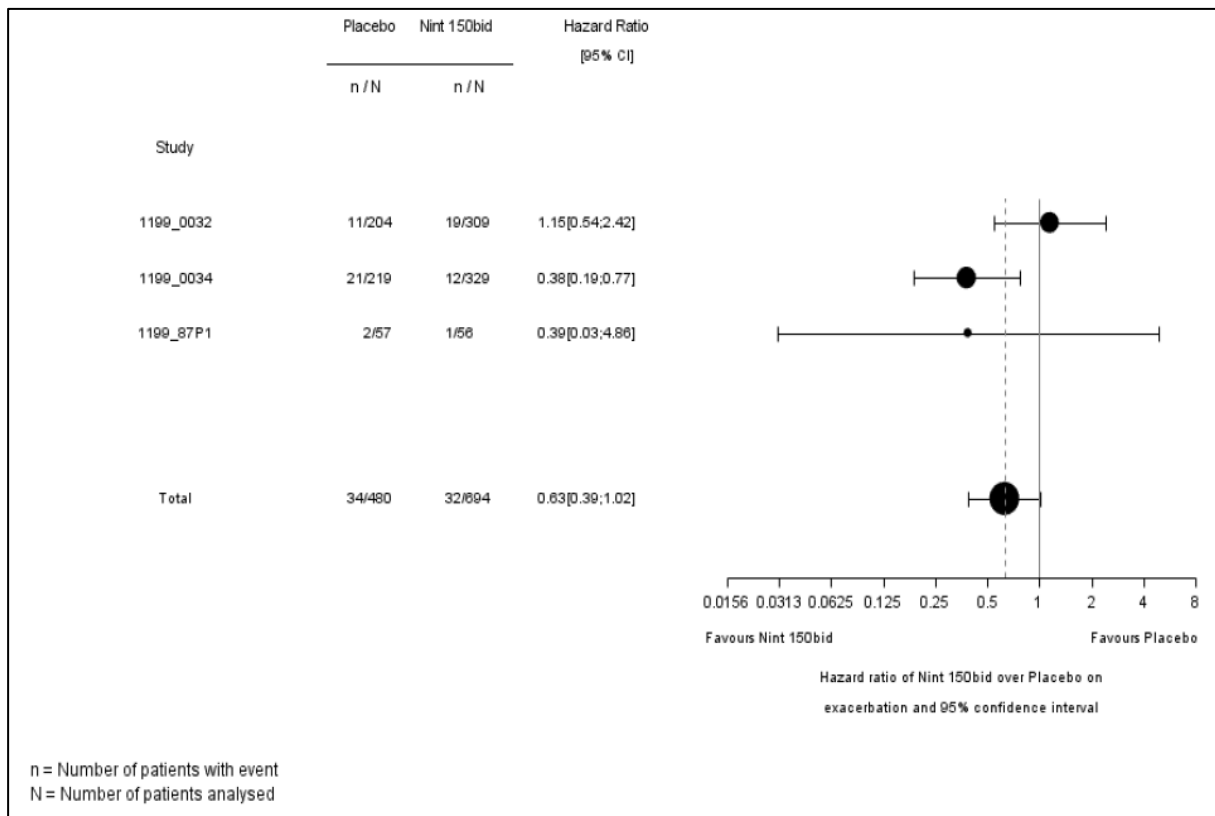
Die Ergebnisse der IPD Meta-Analyse für die Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* und die *Zeit bis zur ersten akuten Exazerbation* wurden bereits in Tabelle 4-27 dargestellt.

Die grafische Darstellung der Meta-Analysen mittels Forest-Plots befinden sich in Abbildung 22 und Abbildung 23. Sie zeigen die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der IPD Meta-Analyse. Die Heterogenitätstests (siehe auch Abschnitt 4.3.1.3.5.2) weisen auf eine homogene Datenlage hin (p-Wert der Interaktionstests [Merkmal Studie]: $p=0,3659$ bzw. $p=0,0948$).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 22: Forest-Plot zur Meta-Analyse für Endpunkt *Zeit bis zur ersten adjudizierten akuten Exazerbation* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 23: Forest-Plot zur Meta-Analyse für den Endpunkt *Zeit bis zur ersten akuten Exazerbation* - RCT mit dem zu bewertenden Arzneimittel

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Akute Exazerbationen der IPF sind respiratorische Dekompensationen des Patientenzustandes ohne erkennbare Ursache, die mit einer Zunahme der Dyspnoe und einer Verschlechterung oder schweren Beeinträchtigung des Gasaustausches einhergehen (Behr, 2013). Diese sind klar abzugrenzen von Verschlechterungen des Patientenzustandes, welche durch bekannte Ursachen wie beispielsweise respiratorische Infektionen oder Lungenembolien auftreten. Dementsprechend waren die Endpunkte *Zeit bis zur ersten akuten Exazerbation* und *Zeit bis zur ersten adjudizierten akuten Exazerbation* definiert als Zeit bis zum Eintreten einer akuten Verschlechterung des klinischen Zustandes ohne erkennbare Ursachen durch andere Faktoren. Das Auftreten einer *akuten Exazerbation* wurde dabei in den berücksichtigten Studien anhand fester Kriterien bestimmt (Tabelle 4-25), welche im Einklang mit den Angaben zur Symptomatik *akuter Exazerbationen* der deutschen S2k-Leitlinie stehen (Behr, 2013).

Bei der Erhebung der Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* wurde demnach die größtmögliche Objektivität angestrebt. Beide Endpunkte waren eindeutig definiert. Von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext kann demnach für beide Endpunkte ausgegangen werden.

Endpunkt *akute Exazerbationen* über den Zeitraum von einem Jahr hinaus (INPULSIS-ON Studie)

Um die Wirksamkeit von Nintedanib über den Studienzeitraum der INPULSIS Studien hinaus zu erheben, konnten Patienten, die an einer der INPULSIS Studien teilgenommen hatten, in eine einarmige Open-Label-Extensionsstudie aufgenommen werden (Studie 1199.33; INPULSIS-ON). Die im Folgenden berichteten Daten decken einen Zeitraum von 144 Wochen bzw. 196 Wochen der Extensionsstudie ab und zeigen, dass die reduzierte Inzidenz auftretender akuter Exazerbationen unter einer Therapie mit Nintedanib auch über die im vorliegenden Dossier abgebildete Studiendauer von bis zu einem Jahr hinaus erhalten bleibt. Tabelle 4-29 zeigt die Anzahl der Patienten, welche während der INPULSIS Studien bzw. der Extensionsstudie INPULSIS-ON (Studie 1199.33) eine *akute Exazerbation* erlitten haben. Der Anteil an IPF-Patienten ohne Behandlung (bzw. mit BSC), welcher jährlich eine *akute Exazerbation* erleidet, beträgt zwischen 5% und 14% (Behr et al., 2013), 3-Jahres-Inzidenzen variieren zwischen 20,7% (Song et al., 2011) und 23,9% (Juarez et al., 2015). In der Extensionsstudie INPULSIS-ON, in der Patienten aus beiden INPULSIS Studien und aus beiden Behandlungsarmen für weitere 144 Wochen mit Nintedanib behandelt wurden, lag der Anteil an Patienten, die mindestens eine akute Exazerbation erlitten hatten, bei 12,6% (Fortführung von Nintedanib) bzw. 11,5% (Beginn von Nintedanib nach Erhalt von Placebo im Zuge der INPULSIS Studien) (Crestani et al., 2017) und damit deutlich unter den oben beschriebenen 3-Jahres-Inzidenzen (20,7-23,9%). Entsprechende Inzidenzen ergaben sich auch über einen Zeitraum von 196 Wochen: 15% (Fortführung von Nintedanib) bzw. 12% (Beginn von Nintedanib) an IPF-Patienten erlitten mindestens eine *akute Exazerbation* (Crestani et al., 2018). Diese deutlich verminderten Raten von *akuten Exazerbationen* zeigen den positiven Langzeiteffekt von Nintedanib auch über einen Zeitraum von drei Jahren hinaus.

Tabelle 4-29: *Akute Exazerbationen* in den gepoolten INPULSIS Studien und der Extensionsstudie INPULSIS-ON

Patienten mit ≥1 akuten Exazerbation, n (%)	INPULSIS (gepoolt)		INPULSIS-ON ^a innerhalb 144 Wochen		INPULSIS-ON ^a innerhalb 196 Wochen	
	Nintedanib n=638	Placebo n=423	Fortführung von Nintedanib n=430	Beginn von Nintedanib n=304	Fortführung von Nintedanib n=430	Beginn von Nintedanib n=304
	31 (5)	32 (8)	54 (12,6)	35 (11,5)	63 (15) ^b	37 (12) ^b
Quelle: nach (Crestani et al., 2017; Crestani et al., 2018) a: Die Extensionsstudie INPULSIS-ON ist eine einarmige, unverblindete Extensionsstudie zu den INPULSIS Studien. Alle eingeschlossenen Patienten erhielten eine Nintedanib-Therapie. b: Basierend auf Kaplan-Meier-Schätzung						

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 Endpunkt HRCT quantitativer Lungenfibrosescore (HRCT-QLF)

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

Tabelle 4-30: Operationalisierung von HRCT-QLF Score

Studie	Operationalisierung
INPULSIS 1	Endpunkt nicht erfasst
INPULSIS 2	Endpunkt nicht erfasst
Studie 1199.187	<u>Veränderung des HRCT-QLF Scores bis Woche 24</u> Definition: Veränderung des HRCT-QLF Scores zwischen Ausgangswert und Wert bis Woche 24 (analysiert jeweils als relative Änderung [Prozentwert] und als absolute Änderung); Berechnung mittels ANCOVA zur Adjustierung für den Ausgangswert Erhebung: - mittels HRCT im Studienzentrum zum Ausgangswert und Woche 24 (sowie späteren Studienvisiten) unter Befolgung der speziellen Vorgaben im Manual für das verwendete Modell - Verwendung desselben Gerätes für alle Untersuchungen eines Studienteilnehmers - zentraler Review aller Scans eines Studienteilnehmers durch denselben Radiologen ANCOVA: Kovarianzanalyse; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore)

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 4-31: Bewertung des Verzerrungspotenzials für den Endpunkt HRCT-QLF Score 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 1199.187	niedrig	ja	nein	ja	ja	hoch
<u>Veränderung des HRCT-QLF Scores bis Woche 24</u> HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); ITT: Intention-to-Treat; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für den Endpunkt *HRCT-QLF Score* erfolgte doppelblind mittels HRCT.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung für den Endpunkt *Veränderung des HRCT-QLF Scores bis Woche 24* erfolgte auf Basis der Studienteilnehmer im Treated Set, für die verfügbare Daten für die Analyse des Endpunkts vorlagen. Das Treated Set in Studie 1199.187 stimmt mit der Population der randomisierten Studienteilnehmer überein. Allerdings konnten insgesamt 26 Studienteilnehmer (23%) aufgrund fehlender Follow-up-Daten nicht bei der Analyse berücksichtigt werden. Das ITT-Prinzip konnte somit nicht vollständig adäquat umgesetzt werden.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für den Endpunkt *Veränderung des HRCT-QLF Scores bis Woche 24* waren a priori definiert. Die Ergebnisse der Studie 1199.187 wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Die im Dossier präsentierten Ergebnisse für diesen Endpunkt bilden die Ergebnisse aus dem Studienbericht ab.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war transparent, Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die *Veränderung des HRCT-QLF Scores bis Woche 24* aufgrund des hohen Anteils an Studienteilnehmern mit fehlenden Daten ein hohes Verzerrungspotenzial.

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

Der Endpunkt *Veränderung des HRCT-QLF Scores bis Woche 24* wurde als primärer Endpunkt der Studie 1199.187 erhoben. Eine Erhebung in den INPULSIS Studien erfolgte nicht, zum Endpunkt liegen daher nur Daten aus der Einzelstudie 1199.187 bis Woche 24 vor. Im Kontext der Nutzenbewertung ergibt sich für den Endpunkt eine untergeordnete Rolle.

Tabelle 4-32: Ergebnisse für *Veränderung des HRCT-QLF Scores bis Woche 24* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	N ^a	Ausgangswert	Woche 24	Veränderung	Behandlungseffekt Nintedanib vs. Placebo ^c		
			MW ^b (SE)	MW ^b (SE)	MW ^c (SE)	MWD (SE)	95%-KI	p-Wert
Relative Veränderung des HRCT-QLF Scores (%)								
Studie 1199.187 ^{d, e}	Nintedanib	41	12,66 (1,226)	13,65 (1,338)	11,091 (4,6086)	-3,511 (6,4482)	[-16,341; 9,319]	-
	Placebo	43	13,59 (1,203)	15,08 (1,419)	14,602 (4,4999)			
Absolute Veränderung des HRCT-QLF Scores								
Studie 1199.187 ^{d, e}	Nintedanib	41	12,66 (1,226)	13,65 (1,338)	0,947 (0,5990)	-0,366 (0,8381)	[-2,033; 1,302]	-
	Placebo	43	13,59 (1,203)	15,08 (1,419)	1,313 (0,5849)			
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.71.1.1, Tab. 1.3.71.1.2 a: Anzahl in MMRM-Auswertung analysierter Studienteilnehmer b: Nicht-adjustierte Mittelwerte. Höhere MW entsprechen dabei einer stärker ausgeprägten Lungenfibrose. c: Berechnet mittels MMRM-Analyse mit festen Effekten für Behandlung und Ausgangswert und Patient als zufälliger Effekt; unstrukturierte Kovarianzmatrix d: Treated Set mit Einschränkung auf Patienten mit verfügbaren Daten für den vorliegenden Endpunkt e: Kontrollierte Phase HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); SE: Standard Error (Standardfehler)								

Im Rahmen des globalen Amendments zum CTP der Studie 1199.187 wurde aufgrund der Verkürzung der Dauer der kontrollierten Phase von 52 auf 24 Wochen und der Reduktion der Stichprobenzahl von 275 auf ungefähr 150 Patienten, sowie der daraus resultierenden verringerten statistischen Power von statistischer Hypothesetestung abgesehen und lediglich eine deskriptive Auswertung mittels Punktschätzer mit 95%-KI durchgeführt (Tabelle 4-32). Auf die Angabe von p-Werten wird aus diesem Grunde verzichtet.

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.

Daten zum Endpunkt *HRCT-QLF Score* wurden nur in der Studie 1199.187 erfasst, weshalb keine Meta-Analyse durchgeführt wurde.

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Als ein standardisiertes computergestütztes Verfahren zur Quantifizierung des Fortschritts der Fibrosierung war der Endpunkt *Veränderung des HRCT-QLF Scores* eindeutig definiert und wurde objektiv erhoben. Für den Endpunkt kann somit von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ausgegangen werden.

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.

4.3.1.3.2.3 Endpunkt *Veränderung der Gehstrecke im 6-Minuten-Gehtest (6 MWT)*

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

Tabelle 4-33: Operationalisierung von 6 MWT

Studie	Operationalisierung
INPULSIS 1	Endpunkt nicht erfasst
INPULSIS 2	Endpunkt nicht erfasst
Studie 1199.187	<u>Veränderung der Gehstrecke im 6 MWT bis Woche 24:</u> Definition: Veränderung zu Woche 24 der im Rahmen des 6 MWT gelaufenen Distanz (in Metern); Betrachtung mittels MMRM basierend auf den Daten des 6 MWT im Zeitverlauf (Ausgangswert, Woche 12, Woche 24) Erhebung: Durchführung des Tests im Studienzentrum zum Ausgangswert, Monat 3, Woche 24 (sowie zu späteren Studienvisiten)
6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); MMRM: Mixed Effect Model for Repeated Measures	

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 4-34: Bewertung des Verzerrungspotenzials für den Endpunkt 6 MWT 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 1199.187	<u>Veränderung der Gehstrecke im 6 MWT bis Woche 24</u>					
	niedrig	ja	ja	ja	ja	niedrig
6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für den Endpunkt *Veränderung der Gehstrecke im 6 MWT* erfolgte doppelblind im Studienzentrum.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung für den Endpunkt erfolgte auf Basis des Treated Set, welches in der Studie 1199.187 mit der Population der randomisierten Patienten übereinstimmt. Allerdings konnten sechs Studienteilnehmer (5%) aufgrund fehlender Follow-up-Daten nicht bei der Analyse berücksichtigt werden. Aufgrund des relativ geringen Anteils nichtberücksichtigter Patienten wird dennoch von einer adäquaten Umsetzung des ITT-Prinzips ausgegangen.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für den Endpunkt *Veränderung der Gehstrecke im 6 MWT bis Woche 24* waren a priori definiert. Die Ergebnisse der Studie 1199.187 wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Die im Dossier präsentierten Ergebnisse für diesen Endpunkt bilden die Ergebnisse aus dem Studienbericht ab.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war transparent, Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die *Veränderung der Gehstrecke im 6 MWT bis Woche 24* ein niedriges Verzerrungspotenzial.

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

Zum Endpunkt *Veränderung der Gehstrecke im 6 MWT bis Woche 24* liegen Daten aus der Studie 1199.187 vor. Im Kontext der Nutzenbewertung ergibt sich für den Endpunkt eine untergeordnete Rolle.

Tabelle 4-35: Ergebnisse für *Veränderung der Gehstrecke im 6 MWT bis Woche 24* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	N ^a	Ausgangswert ^b (Meter)	Woche 24 ^c (Meter)	Veränderung (Meter)	Behandlungseffekt Nintedanib vs. Placebo ^d		
			MW ^e (SD)	MW ^e (SD)	MW ^e (SE)	MWD (SE)	95%-KI	P-Wert
Veränderung der Gehstrecke im 6 MWT bis Woche 24								
Studie 1199.187 ^{f, g}	Nintedanib	55	345,46 (140,712)	360,70 (150,423)	3,688 (11,4443)	17,932 (16,2091)	[-14,257; 50,121]	-
	Placebo	52	347,69 (146,262)	332,12 (161,138)	-14,243 (11,5310)			

Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.71.2.1

a: Anzahl in MMRM-Auswertung analysierter Studienteilnehmer (Patienten mit Ausgangswert und zumindest einer weiteren Messung in der kontrollierten Phase)

b: Vorliegende Messungen bei Ausgangswert: N=57 im Nintedanib-Arm, N=56 im Placebo-Arm

c: Vorliegende Messungen zu Woche 24: N=46 im Nintedanib-Arm, N=46 im Placebo-Arm

d: Kontrast der Behandlungen zu Woche 24, berechnet mittels MMRM mit festen Effekten für Behandlung, Visite, Behandlung×Visite-Interaktion, Ausgangswert und Ausgangswert×Visite-Interaktion sowie einem zufälligen Effekt für Studienteilnehmer; unstrukturierte Kovarianzmatrix.

e: Nicht-adjustierte Mittelwerte. Ein geringerer MW zum 6 MWT entspricht einer im Mittel geringeren Gehstrecke.

f: Treated Set

g: Kontrollierte Phase

6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standardabweichung; SE: Standard Error (Standardfehler)

Im Rahmen des globalen Amendments zum CTP der Studie 1199.187 wurde aufgrund der Verkürzung der Dauer der kontrollierten Phase von 52 auf 24 Wochen und der Reduktion der Stichprobenzahl von 275 auf ungefähr 150 Patienten, sowie der daraus resultierenden verringerten statistischen Power von statistischer Hypothesetestung abgesehen und lediglich eine deskriptive Auswertung mittels Punktschätzer mit 95%-KI durchgeführt (Tabelle 4-35). Auf die Angabe von p-Werten wird aus diesem Grunde verzichtet.

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.

Daten zum Endpunkt *Änderung im 6 MWT* wurden nur in der Studie 1199.187 erfasst, weshalb keine Meta-Analyse durchgeführt wurde.

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Der 6 MWT ist ein standardisiertes und im Anwendungsgebiet IPF validiertes Verfahren zur Beurteilung der Belastungseinschränkung (du Bois et al., 2011c; Behr et al., 2013). Als solches ist der Endpunkt *Änderung der Gehstrecke im 6 MWT* eindeutig definiert und kann objektiv erhoben werden. Für den Endpunkt kann somit von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ausgegangen werden.

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.

4.3.1.3.2.4 Endpunkt Zeit bis zur ersten respiratorisch bedingten Hospitalisierung

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

Tabelle 4-36: Operationalisierung von *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung*

Studie	Operationalisierung
INPULSIS 1	Endpunkt nicht erfasst
INPULSIS 2	Endpunkt nicht erfasst
Studie 1199.187	<u>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</u> Definition: Zeit von erster Verabreichung der Studienmedikation bis zum Zeitpunkt der respiratorisch bedingten Hospitalisierung Zensierung von Patienten ohne Ereignis: 176 Tage nach erstmaliger Einnahme der Studienmedikation Erhebung: im Rahmen der Erfassung unerwünschter Ereignisse

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 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* 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 1199.187	<u>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung über 24 Wochen</u>					
	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für den Endpunkt *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* erfolgte doppelblind im Studienzentrum.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung für den Endpunkt erfolgte auf Basis des Treated Set, welches in der Studie 1199.187 mit der Population der randomisierten Patienten übereinstimmt. Es wird von einer adäquaten Umsetzung des ITT-Prinzips ausgegangen. Da keine Daten über die gesamte kontrollierte Studienphase vorlagen, wurde für diesen Endpunkt die im Studienbericht angegebene Auswertung herangezogen. Diese zensierte alle Daten an Tag 176 nach der ersten Gabe der Studienmedikation. Da der Zensierungsgrund als nicht informativ aufgefasst werden kann, ist nicht von einer Verzerrung aufgrund der Zensierung auszugehen.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für den Endpunkt *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* waren a priori definiert. Die Ergebnisse der Studie 1199.187 wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Die im Dossier präsentierten Ergebnisse für diesen Endpunkt bilden die Ergebnisse aus dem Studienbericht ab.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war transparent, Hinweise auf ein erhöhtes Verzerrungspotenzial liegen diesbezüglich nicht vor.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für die *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* ein niedriges Verzerrungspotenzial.

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

Zum Endpunkt Zeit bis zur ersten respiratorisch bedingten Hospitalisierung liegen Daten aus der Studie 1199.187 vor. Im Kontext der Nutzenbewertung ergibt sich für den Endpunkt eine untergeordnete Rolle.

Tabelle 4-38: Ergebnisse für *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung* mit dem zu bewertenden Arzneimittel

Studie	Nintedanib			Placebo			Behandlungseffekt Nintedanib vs. Placebo		
	N ^a	n ^b (%)	Überlebens- wahrschein- lichkeit	N ^a	n ^b (%)	Überlebens- wahrschein- lichkeit	HR	95%-KI	P-Wert
Zeit bis zur ersten respiratorisch bedingten Hospitalisierung									
Studie 1199.187 ^d	56	0 (0,0)	1,000 ^e	57	4 (7,0)	0,922 ^e	k. A. ^c	k. A. ^c	k. A. ^c
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017): Tab. 15.2.2.7: 1 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: Die Berechnung von Effektschätzern waren im SAP nicht vorgesehen d: Treated Set e: 6-Monats-Überleben (Kaplan-Meier-Schätzer an Tag 176) HR: Hazard Ratio; k. A.: keine Angabe; KI: Konfidenzintervall; SAP: Statistischer Analyseplan									

Im Rahmen des globalen Amendments zum CTP der Studie 1199.187 wurde aufgrund der Verkürzung der Dauer der kontrollierten Phase von 52 auf 24 Wochen und der Reduktion der Stichprobenzahl von 275 auf ungefähr 150 Patienten, sowie der daraus resultierenden verringerten statistischen Power von statistischer Hypothesetestung abgesehen und lediglich eine deskriptive Auswertung mittels Punktschätzer mit 95%-KI durchgeführt (Tabelle 4-35). Auf die Angabe von p-Werten wird aus diesem Grunde verzichtet.

4.3.1.3.3 Gesundheitsbezogene Lebensqualität – 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.

4.3.1.3.3.1 Endpunkte zum SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI- C und EQ-5D

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

Tabelle 4-39: Operationalisierung zu SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI-C und EQ-5D

Studie	Operationalisierung
INPULSIS 1	<u>Veränderung SGRQ-Gesamtscore bis Woche 52 bzw. Woche 24</u> Definition: Veränderung des SGRQ-Gesamtscores über 52 Wochen bzw. über 24 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Der SGRQ-Gesamtscore (Wertebereich 0-100) setzt sich zusammen aus den drei Domänen Symptome, Aktivität und Impact. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 52 Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar. <u>SGRQ-(≤-4 Punkte)-Responder zu Woche 52 bzw. Woche 24</u> Definition: Anteil Studienteilnehmer mit einer absoluten Veränderung des SGRQ-Gesamtscores von ≤-4 Punkten zwischen Ausgangswert und Woche 52 bzw. Woche 24. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 24 bzw. Woche 52. <u>Veränderung SGRQ-I-Gesamtscore bis Woche 52 bzw. Woche 24</u> Definition: Veränderung des IPF-spezifischen SGRQ-I-Gesamtscores über 52 Wochen bzw. über 24 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Beim SGRQ-I Fragebogen (Wertebereich 0-100) handelt es sich um eine IPF-spezifische SGRQ-Fragebogenversion. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 52. Der SGRQ-I-Gesamtscore wurde aus den Antworten im SGRQ-Fragebogen nach einem von den Entwicklern des SGRQ-I definierten Algorithmus ermittelt. Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar. <u>SGRQ-I-(≤-4 Punkte)-Responder zu Woche 52 bzw. Woche 24</u> Definition: Anteil Studienteilnehmer mit einer absoluten Veränderung des SGRQ-I-Gesamtscores von ≤-4 Punkten zwischen Ausgangswert und Woche 52 bzw. Woche 24. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 24 bzw. Woche 52. Der SGRQ-I-Gesamtscore wurde aus den Antworten im SGRQ-Fragebogen nach einem von den Entwicklern des SGRQ-I definierten Algorithmus ermittelt. <u>Veränderung SOBQ-Gesamtscore bis Woche 52 bzw. Woche 24</u> Definition: Veränderung des SOBQ-Gesamtscores über 52 Wochen bzw. über 24 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Der SOBQ-Gesamtscore setzt sich zusammen aus 24 Items, die jeweils auf einer Skala von null bis fünf bewertet werden. Die Summe ergibt den SOBQ-Gesamtscore (Wertebereich 0-120). Erhebung: durch Studienteilnehmer mittels SOBQ-Fragebogen, vom Ausgangswert bis Woche 52 Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar.

Studie	Operationalisierung
	<u>Veränderung CASA-Q Husten-Symptome bis Woche 52</u> Definition: Veränderung der CASA-Q Husten-Symptome über 52 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Erhebung: durch Studienteilnehmer, mittels CASA-Q-Fragebogen (Likert-Skala 1-5), vom Ausgangswert bis Woche 52 Interpretation: Ein größerer Score ist mit weniger Husten-Symptomen assoziiert. <u>Veränderung CASA-Q Husten-Impact bis Woche 52</u> Definition: Veränderung des CASA-Q Husten-Impact über 52 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Erhebung: durch Studienteilnehmer, mittels CASA-Q-Fragebogen (Likert-Skala 1-5), vom Ausgangswert bis Woche 52 Interpretation: Ein größerer Score ist mit weniger Husten-Symptomen assoziiert. <u>PGI-C-Responder zu Woche 52</u> Definition: Anteil Studienteilnehmer, die ihre gesundheitsbezogene Lebensqualität gegenüber dem Befinden bei Studienbeginn als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „ohne Veränderung“ einschätzten. Die Einschätzung erfolgte auf einer 7-Punkte-Likert-Skala von eins (sehr viel besser) bis sieben (sehr viel schlechter). Erhebung: durch Studienteilnehmer, mittels PGI-C-Fragebogen (7-Punkte-Likert-Skala) zu Woche 52 <u>Veränderung EQ-5D-VAS bis Woche 52</u> Definition: Veränderung der EQ-5D-VAS (Skala 0-100) über 52 Wochen gegenüber dem Ausgangswert; Auswertung über ein MMRM. Erhebung: durch Studienteilnehmer, mittels EQ-5D-Fragebogen, vom Ausgangswert bis Woche 52 Interpretation: Ein größerer Score steht für einen besseren subjektiv empfundenen allgemeinen Gesundheitszustand. Um eine sinnvolle gepoolte Analyse mit der kontrollierten Phase der Studie 1199.187 zu ermöglichen werden Änderungen der Scores SGRQ, SGRQ-I und SOBQ zusätzlich zu Woche 24 dargestellt.
INPULSIS 2	Homogenes Studiendesign wie INPULSIS 1
Studie 1199.187	<u>Veränderung SGRQ-Gesamtscore bis Woche 24</u> Definition: Veränderung des SGRQ-Gesamtscores bis Woche 24. Der SGRQ-Gesamtscore (Wertebereich 0-100) setzt sich zusammen aus den drei Domänen Symptome, Aktivität und Impact. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis zum Ende der kontrollierten Phase Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar.

Studie	Operationalisierung
	<u>SGRQ-($\leq$-4 Punkte)-Responder zu Woche 24</u> Definition: Anteil Studienteilnehmer mit einer absoluten Veränderung des SGRQ-Gesamtscores von $\leq$-4 Punkten zwischen Ausgangswert und Woche 24. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 24. <u>Veränderung SGRQ-I-Gesamtscore bis Woche 24</u> Definition: Veränderung des IPF-spezifischen SGRQ-I-Gesamtscores (Wertebereich 0-100) bis Woche 24. Beim SGRQ-I Fragebogen handelt es sich um eine IPF-spezifische SGRQ-Fragebogenversion. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis zum Ende der kontrollierten Phase. Der SGRQ-I-Gesamtscore wurde aus den Antworten im SGRQ-Fragebogen nach einem von den Entwicklern des SGRQ-I definierten Algorithmus ermittelt. Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar. <u>SGRQ-I-($\leq$-4 Punkte)-Responder zu Woche 24</u> Definition: Anteil Studienteilnehmer mit einer absoluten Veränderung des SGRQ-I-Gesamtscores von $\leq$-4 Punkten zwischen Ausgangswert und Woche 24. Erhebung: durch Studienteilnehmer mittels SGRQ-Fragebogen, vom Ausgangswert bis Woche 24. Der SGRQ-I-Gesamtscore wurde aus den Antworten im SGRQ-Fragebogen nach einem von den Entwicklern des SGRQ-I definierten Algorithmus ermittelt. <u>Veränderung SOBQ-Gesamtscore bis Woche 24</u> Definition: Veränderung des SOBQ-Gesamtscores über 24 Wochen. Der SOBQ-Gesamtscore setzt sich zusammen aus 24 Items, die jeweils auf einer Skala von null bis fünf bewertet werden. Die Summe ergibt den SOBQ-Gesamtscore (Wertebereich 0-120). Erhebung: durch Studienteilnehmer mittels SOBQ-Fragebogen, vom Ausgangswert bis zum Ende der kontrollierten Phase Interpretation: Eine negative absolute Veränderung gegenüber dem Ausgangswert stellt eine Verbesserung der Lebensqualität dar. <u>Veränderung CASA-Q Husten-Symptome</u> Nicht erhoben <u>Veränderung CASA-Q Husten-Impact</u> Nicht erhoben <u>PGI-C-Responder</u> Nicht erhoben <u>Veränderung EQ-5D-VAS</u> Nicht erhoben
CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); MMRM: Mixed Effect Model for Repeated Measures; PGI-	

Studie	Operationalisierung
	C: Patient's Global Impression of Change; SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; VAS: Visual Analog Scale

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 4-40: Bewertung des Verzerrungspotenzials für die Endpunkte SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI-C und EQ-5D 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
INPULSIS 1	<u>Veränderung SGRQ-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>SGRQ-(≤4 Punkte)-Responder</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung SGRQ-I-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>SGRQ-I-(≤4 Punkte)-Responder</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung SOBQ-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung CASA-Q Husten-Symptome</u>					
	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	<u>Veränderung CASA-Q Husten-Impact</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>PGI-C-Responder</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung EQ-5D-VAS</u>					
	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	Homogenes Studiendesign wie INPULSIS 1					

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Studie 1199.187	<u>Veränderung SGRQ-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>SGRQ-($\leq$-4 Punkte)-Responder</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung SGRQ-I-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>SGRQ-I-($\leq$-4 Punkte)-Responder</u>					
	niedrig	ja	ja	ja	ja	niedrig
	<u>Veränderung SOBQ-Gesamtscore</u>					
	niedrig	ja	ja	ja	ja	niedrig
CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); ITT: Intention-to-Treat; PGI- C: Patient's Global Impression of Change; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; VAS: Visual Analog Scale						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für die Endpunkte SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI- C und EQ-5D erfolgte doppelblind.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 erfolgte auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen haben. Das Treated Set wick lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung ($<1\%$ der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip. In der MMRM-Auswertung wurden alle Patienten des Treated Set eingeschlossen, für die neben dem Ausgangswert zumindest ein weiterer Wert vorlag.

Der Anteil der Studienteilnehmer, die die Fragebögen zu Woche 24 bzw. Woche 52 ausreichend ausfüllten, lag in den INPULSIS Studien in den Endpunkten zur gesundheitsbezogenen Lebensqualität annähernd bei 90% bezogen auf die Zahl der randomisierten Studienteilnehmer. In Studie 1199.187 lag dieser Anteil zu Woche 24 bei etwa 80%. Dabei unterschieden sich die Nichtberücksichtigungsanteile zwischen den Behandlungsgruppen innerhalb der Studien jeweils um weniger als 5%. Somit liegt ein niedriges Verzerrungspotenzial vor. Studienteilnehmer, für die zu Woche 24 bzw. 52 keine Messwerte vorlagen, wurden bei den Responderanalysen konservativ als Non-Responder bewertet.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien für die Endpunkte zur gesundheitsbezogenen Lebensqualität waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Prüfplan ausgewertet und transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden. Die zusätzliche Betrachtung des Auswertungszeitpunkts nach 24 Wochen für die INPULSIS Studien wurde dargestellt, um eine sinnvolle gepoolte Analyse mit der kontrollierten Phase der Studie 1199.187 zu ermöglichen.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war für alle Endpunkte zur gesundheitsbezogenen Lebensqualität transparent.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für alle Endpunkte zur gesundheitsbezogenen Lebensqualität ein niedriges Verzerrungspotenzial.

Stellen Sie die Ergebnisse für die Endpunkte SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI- C und EQ-5D für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

SGRQ und SGRQ-I

Die Endpunkte *Veränderung SGRQ-Gesamtscore* und *Veränderung SGRQ-I-Gesamtscore* wurden in den INPULSIS Studien und in Studie 1199.187 erhoben. Dabei liegen aus den INPULSIS Studien Daten bis Woche 52 vor, aus Studie 1199.187 bis zum Ende der kontrollierten Phase. Die Darstellung der Ergebnisse für Studie 1199.187 alleine erfolgt daher in Form der Analyse bis Woche 24. Die IPD Meta-Analyse für den 52 Wochen Zeitraum basiert jeweils auf den beiden INPULSIS Studien; für den 24 Wochen Zeitraum wurden jeweils alle drei Studien einbezogen. Die Analysen basieren auf einem MMRM. Im Kontext der Nutzenbewertung sind im Setting einer chronischen Erkrankung die Ergebnisse bis Woche 52 als relevant zu sehen, Ergebnisse bis Woche 24 werden ergänzend dargestellt.

Für den SGRQ fand sich auf Einzelstudienenebene über den Zeitraum von 52 Wochen ein statistisch signifikanter Vorteil zugunsten von Nintedanib: Die Verschlechterung der gesundheitsbezogenen Lebensqualität war in der INPULSIS 2 Studie im Nintedanib-Arm

weniger stark als im Vergleichsarm. Auf Ebene der IPD Meta-Analyse beider INPULSIS Studien auf Basis eines MMRM ergab sich hingegen kein statistisch signifikanter Unterschied. Für die gesundheitsbezogene Lebensqualität gemessen über den IPF-spezifischen Fragebogen *SGRQ-I* zeigte sich in der IPD Meta-Analyse der INPULSIS Studien über den Zeitraum von 52 Wochen ein statistisch signifikanter Vorteil zugunsten von Nintedanib im Vergleich zur zVT.

Generell sank die gesundheitsbezogene Lebensqualität gemessen über den *SGRQ* und den IPF-spezifischen Fragebogen *SGRQ-I* zwischen Ausgangswert und Woche 52 sowohl im Kontrollarm als auch im Nintedanib-Arm beider INPULSIS Studien. In Studie 1199.187 verbesserte sich die Lebensqualität der Studienteilnehmer zwischen Ausgangswert und Woche 24 im Mittel in beiden Behandlungsarmen.

Die Ergebnisse zu der *Veränderung des SGRQ-Gesamtscores* und *des SGRQ-I-Gesamtscores* über den Zeitraum von 52 Wochen sind in Tabelle 4-41 dargestellt und nachfolgend genauer beschrieben. Die entsprechenden Analysen über den Zeitraum von 24 Wochen, denen angesichts des chronischen Charakters der IPF weniger Relevanz für die Nutzenbewertung beizumessen ist, sind in Tabelle 4-42 ergänzend dargestellt und im Anschluss daran zusammenfassend beschrieben.

Tabelle 4-41: Ergebnisse für *Veränderung SGRQ-Gesamtscore* und *Veränderung SGRQ-I-Gesamtscore bis Woche 52* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95% KI	p-Wert
Veränderung SGRQ-Gesamtscore (Punkte) bis Woche 52										
INPULSIS ^d (gepoolt)	Nintedanib	624	39,51 (19,151)	507	41,44 (21,160)	609	3,53 (0,540)	-1,43 (0,848)	[-3,09; 0,23]	0,0923
	Placebo	419	39,58 (18,545)	342	42,30 (21,215)	413	4,96 (0,654)			
INPULSIS 1 ^d	Nintedanib	298	39,55 (17,628)	240	42,67 (20,187)	289	4,34 (0,799)	-0,05 (1,248)	[-2,50; 2,40]	0,9657
	Placebo	202	39,79 (18,478)	163	42,36 (21,391)	200	4,39 (0,960)			
INPULSIS 2 ^d	Nintedanib	326	39,46 (20,471)	267	40,33 (21,977)	320	2,80 (0,730)	-2,69 (1,151)	[-4,95; -0,43]	0,0197
	Placebo	217	39,39 (18,647)	179	42,24 (21,113)	213	5,48 (0,891)			

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95% KI	p-Wert
Veränderung SGRQ-I-Gesamtscore (Punkte) bis Woche 52										
INPULSIS ^d (gepoolt)	Nintedanib	625	41,92 (19,932)	503	43,62 (21,789)	610	3,48 (0,559)	-1,99 (0,876)	[-3,71; -0,27]	0,0230
	Placebo	419	41,94 (19,412)	342	45,08 (21,534)	413	5,47 (0,676)			
INPULSIS 1 ^d	Nintedanib	299	41,95 (18,377)	239	44,74 (20,894)	290	4,30 (0,824)	-0,78 (1,289)	[-3,31; 1,75]	0,5446
	Placebo	202	42,06 (19,233)	163	45,44 (21,768)	200	5,08 (0,992)			
INPULSIS 2 ^d	Nintedanib	326	41,88 (21,286)	264	42,61 (22,560)	320	2,72 (0,757)	-3,12 (1,192)	[-5,46; -0,79]	0,0089
	Placebo	217	41,83 (19,620)	179	44,75 (21,375)	213	5,84 (0,921)			
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.1.2, Tab. 15.2.2.2.2.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.1.2, Tab. 15.2.2.2.2.2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 86.6.1.1, Tab. 86.5.1.1, Tab. 1.1.28.8.2, Tab. 1.1.28.11.2 a: Anzahl analysierter Studienteilnehmer b: Nicht-adjustierte Mittelwerte c: MWD [95%-KI] und p-Wert für die Einzelstudien berechnet mittels MMRM mit festen Effekten für Behandlung, Behandlung Visite, Ausgangswert SGRQ-Gesamtscore bzw. Ausgangswert SGRQ-I-Gesamtscore, Ausgangswert SGRQ-Gesamtscore×Visite bzw. Ausgangswert SGRQ-I-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer; MWD [95%-KI] und p-Wert für IPD Meta-Analyse: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Behandlung Visite, Ausgangswert SGRQ-Gesamtscore bzw. Ausgangswert SGRQ-I-Gesamtscore, Ausgangswert SGRQ-Gesamtscore×Visite bzw. Ausgangswert SGRQ-I-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer. d: Treated Set; OC Ein niedrigerer SGRQ- oder SGRQ-I-Gesamtscore steht für eine bessere subjektiv empfundene krankheitsspezifische Lebensqualität. Eine größere (positive) Veränderung gegenüber dem Ausgangswert ist als eine stärkere Abnahme der Lebensqualität zu werten; demzufolge ist eine negative MWD (Nintedanib – Placebo) als Vorteil für Nintedanib zu werten. IPD: Individual Participant Data (Individuelle Patientendaten); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler); SGRQ: St. George’s Respiratory Questionnaire; SGRQ-I: St. George’s Respiratory Questionnaire IPF-specific version										

Der *SGRQ-Gesamtscore* stieg in den INPULSIS Studien vom Ausgangswert bis Woche 52 jeweils in beiden Untersuchungsgruppen an, was als eine Verschlechterung der gesundheitsbezogenen Lebensqualität zu interpretieren ist. In den gepoolten Daten der INPULSIS Studien fiel der Anstieg im *SGRQ-Gesamtscore* unter der Behandlung mit Nintedanib geringer aus als in der Kontrollgruppe (3,53 Punkte vs. 4,96 Punkte). Dies deutet auf eine geringere Verschlechterung der gesundheitsbezogenen Lebensqualität unter

Nintedanib im Vergleich zur zVT in dieser Zeitspanne hin. Für die Veränderung bis zu Woche 52 zeigte die adjustierte MWD [95%-KI] der zwei Behandlungsgruppen in der gepoolten Analyse der beiden INPULSIS Studien unter Verwendung eines MMRM keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen bezüglich der gesundheitsbezogenen Lebensqualität der Patienten gemessen anhand des SGRQ-Gesamtscores über 52 Wochen; die adjustierte MWD [95%-KI] betrug -1,43 Punkte [-3,09; 0,23]; $p=0,0923$ (Tabelle 4-41).

Ein Anstieg des *SGRQ-Gesamtscores* war bis Woche 52 auch in den Einzelstudien zu beobachten (Nintedanib vs. zVT: INPULSIS 1: 4,34 Punkte vs. 4,39 Punkte und INPULSIS 2: 2,80 Punkte vs. 5,48 Punkte). Die adjustierte MWD [95%-KI] der zwei Behandlungsgruppen betrug in der INPULSIS 1 Studie -0,05 Punkte [-2,50; 2,40]; $p=0,9657$. In der INPULSIS 2 Studie zeigte sich ein statistisch signifikanter Unterschied zugunsten von Nintedanib; die adjustierte MWD betrug hier -2,69 Punkte [-4,95; -0,43]; $p=0,0197$.

Der *SGRQ-I-Gesamtscore* stieg vom Ausgangswert bis Woche 52 in den gepoolten Daten der Studien INPULSIS 1 und 2 in der Nintedanib-Gruppe weniger stark an als in der Kontrollgruppe (3,48 Punkte vs. 5,47 Punkte) (Tabelle 4-41). Dies entspricht einer geringeren Abnahme der gesundheitsbezogenen Lebensqualität unter Nintedanib im Vergleich zur zVT. Die adjustierte MWD [95%-KI] der gepoolten Analyse der beiden INPULSIS Studien unter Verwendung eines MMRM zeigt einen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen bezüglich der Lebensqualität der Patienten gemessen anhand des SGRQ-I-Gesamtscores über 52 Wochen; die adjustierte MWD [95%-KI] betrug -1,99 Punkte [-3,71; -0,27]; $p=0,0230$ (Tabelle 4-41). Diese negative MWD zeigt eine bessere subjektiv empfundene gesundheitsbezogene Lebensqualität zugunsten von Nintedanib an. Der Unterschied ist statistisch signifikant.

Auch auf Ebene der Einzelstudien war unter Nintedanib ein geringerer Anstieg des SGRQ-I-Gesamtscores vom Ausgangswert bis Woche 52 in den Studien INPULSIS 1 (4,30 Punkte vs. 5,08 Punkte) und INPULSIS 2 (2,72 Punkte vs. 5,84 Punkte) zu beobachten. Die adjustierte MWD [95%-KI] für die Behandlungsgruppen beträgt in der Studie INPULSIS 1 -0,78 Punkte [-3,31; 1,75]; $p=0,5446$ und in der Studie INPULSIS 2 -3,12 Punkte [-5,46; -0,79]; $p=0,0089$. Der Effekt für die Studie INPULSIS 2 ist statistisch signifikant.

Tabelle 4-42: Ergebnisse für *Veränderung SGRQ-Gesamtscore* und *Veränderung SGRQ-I-Gesamtscore bis Woche 24* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	Ausgangswert		Woche 24		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	P-Wert
Veränderung SGRQ-Gesamtscore (Punkte) bis Woche 24										
INPULSIS + Studie 1199.187 (gepoolt) ^d	Nintedanib	680	39,20 (19,036)	618	39,85 (19,635)	664	0,04 (0,533)	-0,60 (0,714)	[-2,00; 0,80]	0,3981
	Placebo	476	40,16 (18,584)	438	40,51 (19,599)	466	0,65 (0,596)			
INPULSIS ^d (gepoolt)	Nintedanib	624	39,51 (19,151)	573	40,49 (19,700)	609	1,258 (0,4822)	-0,742 (0,7554)	[-2,223; 0,740]	0,3263
	Placebo	419	39,58 (18,545)	394	40,64 (19,737)	413	1,999 (0,5817)			
INPULSIS 1 ^d	Nintedanib	298	39,55 (17,628)	277	41,18 (18,488)	289	1,459 (0,7184)	0,482 (1,1221)	[-1,720; 2,685]	0,6674
	Placebo	202	39,79 (18,478)	190	40,09 (20,414)	200	0,976 (0,8620)			
INPULSIS 2 ^d	Nintedanib	326	39,46 (20,471)	296	39,84 (20,780)	320	1,067 (0,6473)	-1,883 (1,0169)	[-3,879; 0,112]	0,0643
	Placebo	217	39,39 (18,647)	204	41,14 (19,121)	213	2,951 (0,7843)			
Studie 1199.187 ^d	Nintedanib	56	35,75 (17,491)	45	31,71 (16,976)	55	-2,441 (1,5400)	0,312 (2,2327)	[-4,097; 4,721]	0,8891
	Placebo	57	44,39 (18,488)	44	39,34 (18,491)	53	-2,753 (1,5546)			
Veränderung SGRQ-I-Gesamtscore (Punkte) bis Woche 24										
INPULSIS + Studie 1199.187 (gepoolt) ^d	Nintedanib	681	41,62 (19,814)	616	42,09 (20,613)	665	-0,04 (0,556)	Hedges' g : -1,56 (0,746) ; -0,130	Hedges' g : [-3,03; -0,10] ; [-0,25; -0,01]	0,0360
	Placebo	475	42,52 (19,356)	438	43,65 (20,378)	464	1,52 (0,622)			
INPULSIS ^d (gepoolt)	Nintedanib	625	41,92 (19,932)	571	42,72 (20,703)	610	1,087 (0,5022)	Hedges' g : -1,609 (0,7873) ; -0,1331	Hedges' g : [-3,153; -0,065] ; [-0,261; -0,005]	0,0411
	Placebo	419	41,94 (19,412)	393	43,61 (20,461)	412	2,696 (0,6066)			
INPULSIS 1 ^d	Nintedanib	299	41,95 (18,377)	276	43,33 (19,613)	290	1,180 (0,7410)	-0,371 (1,1614)	[-2,650; 1,909]	0,7496
	Placebo	202	42,06 (19,233)	188	43,06 (20,989)	199	1,551 (0,8943)			

Studie	Behandlung	Ausgangswert		Woche 24		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	p-Wert
INPULSIS 2 ^d	Nintedanib	326	41,88 (21,286)	295	42,14 (21,691)	320	1,017 (0,6813)	-2,731 (1,0683)	[-4,827; -0,634]	0,0107
	Placebo	217	41,83 (19,620)	205	44,13 (20,002)	213	3,747 (0,8229)	Hedges'g -0,2311	Hedges'g [-0,409; -0,054]	
Studie 1199.187 ^d	Nintedanib	56	38,31 (18,293)	45	34,11 (17,790)	55	-2,375 (1,6430)	-1,038 (2,3863)	[-5,750; 3,674]	0,6642
	Placebo	56	46,84 (18,534)	45	43,96 (19,862)	52	-1,337 (1,6653)			

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.71.3.1 – 1.3.71.3.4, Tab. 1.3.71.4.1 – 1.3.71.4.4, Tab. 1.1.28.8.1, Tab. 1.1.28.11.1

a: Anzahl analysierter Studienteilnehmer

b: Nicht-adjustierte Mittelwerte

c: MWD [95%-KI] und p-Wert für Einzelstudien: berechnet mittels MMRM mit festen Effekten für Behandlung, Behandlung Visite, Ausgangswert SGRQ-Gesamtscore bzw. Ausgangswert SGRQ-I-Gesamtscore, Ausgangswert SGRQ-Gesamtscore×Visite bzw. Ausgangswert SGRQ-I-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer; MWD [95%-KI] und p-Wert für IPD Meta-Analyse: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Behandlung Visite, Ausgangswert SGRQ-Gesamtscore bzw. Ausgangswert SGRQ-I-Gesamtscore, Ausgangswert SGRQ-Gesamtscore×Visite bzw. Ausgangswert SGRQ-I-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer.

d: Treated Set; OC

Ein niedrigerer SGRQ- oder SGRQ-I-Gesamtscore steht für eine bessere subjektiv empfundene krankheitsspezifische Lebensqualität.

IPD: Individual Participant Data (Individuelle Patientendaten); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version

In der supportiv dargestellten Analyse über 24 Wochen zeigte sich in der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187, den gepoolten INPULSIS Studien sowie auf Ebene der Einzelstudien für die *Veränderung des SGRQ-Gesamtscores* bis zur Woche 24 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Für die *Veränderung des SGRQ-I-Gesamtscores* ergab sich in der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 ein statistisch signifikanter Vorteil zugunsten von Nintedanib gegenüber der Vergleichstherapie (MWD [95%-KI]: -1,56 [-3,03; -0,10]; p=0,0360); die klinische Relevanz dieses Unterschieds ist allerdings fraglich (Hedges' g [95%-KI]: -0,130 [-0,25; -0,01]). Auch in der gepoolten Analyse der INPULSIS Studien zeigte sich ein statistisch signifikanter Vorteil zugunsten von Nintedanib (MWD [95%-KI]: -1,609 [-3,153; -0,065]; p=0,0411; Hedges' g [95%-KI]: -0,1331 [-0,261; -0,005]).

SGRQ-Responder und SGRQ-I-Responder

Die Endpunkte *SGRQ-(≤ 4 Punkte)-Responder* und *SGRQ-I-(≤ 4 Punkte)-Responder* wurden in den INPULSIS Studien und in Studie 1199.187 erhoben. Dabei liegen aus den INPULSIS Studien Daten zu Woche 24 und Woche 52 vor, aus Studie 1199.187 Daten zu Woche 24. Die Darstellung der Ergebnisse zu Woche 52 erfolgt daher für die gepoolte Analyse der INPULSIS Studien sowie auf Ebene der Einzelstudien INPULSIS 1 und 2. Für Woche 24 erfolgt zusätzlich die Darstellung der gepoolten Analyse über die INPULSIS Studien und Studie 1199.187, sowie die Darstellung der Ergebnisse der Studie 1199.187 alleine. Im Kontext der Nutzenbewertung sind im Setting einer chronischen Erkrankung die Ergebnisse zu Woche 52 als relevant zu sehen, Ergebnisse zu Woche 24 werden ergänzend dargestellt.

Die Ergebnisse der in den CTP der INPULSIS Studien präspezifizierten Responderanalyse zum SGRQ-Gesamtscore und SGRQ-I-Gesamtscore sind in Tabelle 4-43 sowie Tabelle 4-44 zusammengefasst. Berichtet wird der Anteil der Studienteilnehmer mit einer relevanten Verbesserung der gesundheitsbezogenen Lebensqualität über den Beobachtungszeitraum von 52 Wochen (bzw. 24 Wochen).

Tabelle 4-43: Ergebnisse für *SGRQ-(≤ 4 Punkte)-Responder* und *SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 52* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
SGRQ-(≤ 4 Punkte)-Responder zu Woche 52							
INPULSIS (gepoolt) ^d	638	146 (22,88)	423	86 (20,33)	1,186	[0,88; 1,61]	0,2715
INPULSIS 1 ^d	309	63 (20,39)	204	49 (24,02)	0,840	[0,55; 1,29]	0,4298
INPULSIS 2 ^d	329	83 (25,23)	219	37 (16,89)	1,664	[1,08; 2,57]	0,0218
SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 52							
INPULSIS (gepoolt) ^d	638	150 (23,51)	423	83 (19,62)	1,284	[0,95; 1,74]	0,1091
INPULSIS 1 ^d	309	66 (21,36)	204	49 (24,02)	0,886	[0,58; 1,36]	0,5821
INPULSIS 2 ^d	329	84 (25,53)	219	34 (15,53)	1,877	[1,20; 2,93]	0,0056
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.1.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 86.1.1.1, Tab. 86.1.1.2, Tab. 79.1.1.1; Tab. 79.1.1.2 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: OR [95%-KI] und p-Wert: berechnet mittels logistischer Regression mit Termen für Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore OR [95%-KI] und p-Wert: berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore. Studienteilnehmer, für die zum Auswertungszeitpunkt keine Messwerte vorlagen, wurden als Non-Responder bewertet. d: Treated Set; OC							

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version							

Für die Langzeitbetrachtung über 52 Wochen, die nur für die INPULSIS Studien möglich war, ergab sich unter Nintedanib in der gepoolten Analyse ein größerer Anteil an Studienteilnehmern, der eine Abnahme um mindestens 4 Punkte im *SGRQ-Gesamtscore* und damit eine klinisch relevante Verbesserung der subjektiv empfundenen gesundheitsbezogenen Lebensqualität erreicht hatten (22,88% vs. 20,33% Responder). Das OR [95%-KI] lag für die gepoolte Auswertung bei 1,186 [0,88; 1,61]; p=0,2715 (Tabelle 4-43).

Auf Ebene der Einzelstudien ergaben sich für die Studie INPULSIS 1 (Nintedanib vs. zVT) 20,39% vs. 24,02%, bzw. für die Studie INPULSIS 2 25,23% vs. 16,89% Responder. Das OR [95%-KI] lag in der INPULSIS 1 Studie bei 0,840 [0,55; 1,29]; p=0,4298 und in der INPULSIS 2 Studie bei 1,664 [1,08; 2,57]; p=0,0218. Ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen für den Endpunkt *SGRQ-(≤4 Punkte)-Responder* fand sich nur im Falle der Studie INPULSIS 2 zugunsten von Nintedanib.

Der Anteil an Studienteilnehmern, die eine Abnahme im *SGRQ-I-Gesamtscore* um mindestens 4 Punkte und damit eine relevante Verbesserung ihrer gesundheitsbezogenen Lebensqualität erreichten, war in der gepoolten Analyse der INPULSIS Studien unter Nintedanib größer als in der Kontrollgruppe (23,51% vs. 19,62). Das OR [95%-KI] lag für die gepoolte Auswertung bei 1,284 [0,95; 1,74]; p=0,1091 (Tabelle 4-43).

Auf Ebene der Einzelstudien ergaben sich für die Studie INPULSIS 1 (Nintedanib vs. zVT) 21,36% vs. 24,02%, bzw. für die Studie INPULSIS 2 25,53% vs. 15,53% Responder. Das OR [95%-KI] lag in der INPULSIS 1 Studie bei 0,886 [0,58; 1,36]; p=0,5821 und in der INPULSIS 2 Studie bei 1,877 [1,20; 2,93]; p=0,0056. Ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen für den Endpunkt *SGRQ-I-(≤4 Punkte)-Responder* fand sich nur im Falle der Studie INPULSIS 2 zugunsten von Nintedanib.

Tabelle 4-44: Ergebnisse für SGRQ-(≤ 4 Punkte)-Responder und SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 24 aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
SGRQ-(≤ 4 Punkte)-Responder zu Woche 24							
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	199 (28,67)	480	147 (30,63)	0,947	[0,73; 1,23]	0,6813
INPULSIS (gepoolt) ^d	638	185 (29,00)	423	125 (29,55)	0,989	[0,75; 1,31]	0,9373
INPULSIS 1 ^d	309	88 (28,48)	204	65 (31,86)	0,885	[0,59; 1,32]	0,5483
INPULSIS 2 ^d	329	97 (29,48)	219	60 (27,40)	1,103	[0,75; 1,62]	0,6212
Studie 1199.187	56	14 (25,00)	57	22 (38,60)	0,612	[0,27; 1,41]	0,2484
SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 24							
INPULSIS + Studie 1199.187 (gepoolt) ^d	694	197 (28,39)	480	133 (27,71)	1,070	[0,82; 1,40]	0,6197
INPULSIS (gepoolt) ^d	638	182 (28,53)	423	115 (27,19)	1,087	[0,82; 1,44]	0,5586
INPULSIS 1 ^d	309	88 (28,48)	204	60 (29,41)	0,992	[0,66; 1,48]	0,9684
INPULSIS 2 ^d	329	94 (28,57)	219	55 (25,11)	1,192	[0,80; 1,77]	0,3825
Studie 1199.187	56	15 (26,79)	57	18 (31,58)	0,964	[0,41; 2,27]	0,9332
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.1.1; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.58.10.1, Tab. 1.3.58.10.2, Tab. 1.3.58.10.3, Tab. 1.3.58.10.4, Tab. 1.3.58.10.5, Tab. 1.3.58.13.1, Tab. 1.3.58.13.2, Tab. 1.3.58.13.3, Tab. 1.3.58.13.4, Tab. 1.3.58.13.5 a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: OR [95%-KI] und p-Wert für Einzelstudien: berechnet mittels logistischer Regression mit Termen für Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore; OR [95%-KI] und p-Wert für gepoolte Analysen: berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore. Studienteilnehmer, für die zum Auswertungszeitpunkt keine Messwerte vorlagen, wurden als Non-Responder bewertet. d: Treated Set; OC IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version							

In der supportiv dargestellten Auswertung für den 24 Wochen-Zeitpunkt ergaben sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen, weder für den Anteil der Patienten mit einer relevanten Verbesserung des SGRQ (SGRQ-Responder) noch für den Anteil der SGRQ-I-Responder nach 24 Wochen (Tabelle 4-44). Dies galt sowohl für die Einzelstudien als auch für die gepoolten Analysen.

SOBQ

Der Endpunkt *Veränderung SOBQ-Gesamtscore* wurde in den INPULSIS Studien und in Studie 1199.187 erhoben. Dabei liegen aus den INPULSIS Studien Daten bis Woche 52 vor, aus Studie 1199.187 bis zum Ende der kontrollierten Phase. Die Darstellung der Ergebnisse für Studie 1199.187 alleine erfolgt daher in Form der Analyse bis Woche 24. Die IPD Meta-Analyse für den 52 Wochen Zeitraum basiert auf den beiden INPULSIS Studien; für den 24 Wochen Zeitraum wurden alle drei Studien einbezogen. Die Analysen basieren auf einem MMRM. Im Kontext der Nutzenbewertung sind im Setting einer chronischen Erkrankung die Ergebnisse zu Woche 52 als relevant zu sehen, Ergebnisse bis Woche 24 werden ergänzend dargestellt.

Tabelle 4-45: Ergebnisse für *Veränderung SOBQ-Gesamtscore bis Woche 52* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	p-Wert
Veränderung SOBQ-Gesamtscore (Punkte) nach 52 Wochen										
INPULSIS (gepoolt) ^d	Nintedanib	587	32,85 (24,429)	481	37,51 (27,384)	382	6,71 (0,773)	-1,69 (1,220)	[-4,09; 0,70]	0,1648
	Placebo	392	32,93 (23,721)	322	39,38 (26,361)	569	8,40 (0,945)			
INPULSIS 1 ^d	Nintedanib	278	32,58 (22,978)	226	37,50 (26,631)	267	6,73 (1,113)	-0,88 (1,770)	[-4,35; 2,60]	0,6203
	Placebo	182	32,24 (23,349)	150	38,85 (26,878)	178	7,61 (1,376)			
INPULSIS 2 ^d	Nintedanib	309	33,10 (25,700)	255	37,53 (28,087)	302	6,69 (1,073)	-2,38 (1,685)	[-5,68; 0,93]	0,1587
	Placebo	210	33,53 (24,078)	172	39,84 (25,972)	204	9,07 (1,300)			
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.3; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.2.3; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 86.10.1.1, Tab. 1.1.28.14.2 a: Anzahl analysierter Studienteilnehmer b: Nicht-adjustierte Mittelwerte c: MWD [95%-KI] und p-Wert der Einzelstudien: berechnet mittels MMRM mit festen Effekten für Behandlung, Behandlung×Visite, Ausgangswert SOBQ-Gesamtscore, Ausgangswert SOBQ-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer MWD [95%-KI] und p-Wert der gepoolten Daten: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Behandlung×Visite, Ausgangswert SOBQ-Gesamtscore, Ausgangswert SOBQ-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer. d: Treated Set; OC Ein niedrigerer SOBQ-Gesamtscore steht für eine geringere Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch Kurzatmigkeit.										

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	p-Wert

KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler); SOBQ: Shortness of Breath Questionnaire

Für die Langzeitbetrachtung bis Woche 52 ergab sich unter Nintedanib in der gepoolten Analyse der INPULSIS Studien eine geringere Veränderung vom Ausgangswert für den SOBQ-Gesamtscore (INPULSIS [gepoolt]: 6,71 Punkte vs. 8,40 Punkte) (Tabelle 4-45). Die errechnete adjustierte MWD [95%-KI] für die Behandlungsgruppen betrug in der gepoolten Analyse der INPULSIS Studien -1,69 Punkte [-4,09; 0,70]; p=0,1648. Damit hatte sich die subjektiv empfundene krankheitsspezifische Lebensqualität hinsichtlich der Kurzatmigkeit unter der Behandlung mit Nintedanib weniger stark verschlechtert als im Vergleichsarm.

Auf Ebene der Einzelstudien ergaben sich für die Studie INPULSIS 1 Veränderungen vom Ausgangswert bis Woche 52 (Nintedanib vs. zVT) von 6,73 Punkte vs. 7,61 Punkte, bzw. für die Studie INPULSIS 2 6,69 Punkte vs. 9,07 Punkte. Die errechnete adjustierte MWD [95%-KI] für die Behandlungsgruppen betrug in der INPULSIS 1 Studie -0,88 Punkte [-4,35; 2,60]; p=0,6203 und in der INPULSIS 2 Studie -2,38 Punkte [-5,68; 0,93]; p=0,1587.

Interessant ist vor diesem Hintergrund die Beobachtung, dass es in beiden INPULSIS Studien während der Beobachtungsperiode zu Veränderungen des Bedarfs an einer Begleittherapie mit Bronchodilatoren kam (siehe die diesbezüglichen detaillierten Ausführungen im Abschnitt 4.3.1.3.5, der Ergebnisse der Subgruppenanalysen zusammenfasst). Unter der Behandlung mit Studienmedikation fand sich eine Entwicklung hin zu einem höheren Bedarf an Bronchodilatator-Begleittherapie in den Kontrollgruppen sowie eine gegenläufige Entwicklung in den Nintedanib-Gruppen. Dies kann als indirekter Hinweis auf eine höhere subjektive Beeinträchtigung der Patienten durch Dyspnoe/Luftnot unter der zVT verglichen mit Nintedanib interpretiert werden, da Bronchodilatoren üblicherweise zur Behandlung des Symptoms Dyspnoe/Kurzatmigkeit eingesetzt werden. Es ist vorstellbar, dass durch diese Veränderungen in der Begleittherapie in den konkreten Studien möglicherweise die Sensitivität des SOBQ-Fragebogens, eine Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch Kurzatmigkeit nachzuweisen, reduziert wurde.

Tabelle 4-46: Ergebnisse für *Veränderung SOBQ-Gesamtscore bis Woche 24* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behand- lung	Ausgangswert		Woche 24		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%- KI	p- Wert
Veränderung SOBQ-Gesamtscore (Punkte) bis Woche 24										
INPULSIS + Studie 1199.187 (gepoolt) ^d	Nintedanib	641	32,22 (24,156)	577	35,21 (25,924)	622	3,20 (0,762)	0,58 (1,037)	[-1,45; 2,61]	0,5761
	Placebo	447	34,08 (23,993)	408	36,09 (25,121)	431	2,62 (0,860)			
INPULSIS (gepoolt) ^d	Nintedanib	587	32,85 (24,429)	533	35,84 (26,081)	569	4,037 (0,7024)	0,158 (1,1048)	[-2,009; 2,325]	0,8861
	Placebo	392	32,93 (23,721)	366	35,89 (25,346)	381	3,879 (0,8541)			
INPULSIS 1 ^d	Nintedanib	278	32,58 (22,978)	251	36,59 (25,141)	267	4,795 (1,0195)	0,777 (1,6156)	[-2,394; 3,949]	0,6305
	Placebo	182	32,24 (23,349)	171	34,78 (26,200)	177	4,018 (1,2532)			
INPULSIS 2 ^d	Nintedanib	309	33,10 (25,700)	282	35,17 (26,917)	302	3,355 (0,9691)	-0,427 (1,5172)	[-3,404; 2,551]	0,7785
	Placebo	210	33,53 (24,078)	195	36,86 (24,600)	204	3,782 (1,1672)			
Studie 1199.187	Nintedanib	54	25,39 (19,892)	44	27,55 (22,841)	53	3,417 (2,0656)	5,899 (3,0668)	[-0,153, 11,952]]	0,0560
	Placebo	55	42,25 (24,547)	42	37,81 (23,277)	50	-2,482 (2,1030)			

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a); Tab. 1.3.71.5.1 – 1.3.71.5.4, Tab. 1.1.28.14.1

a: Anzahl analysierter Studienteilnehmer

b: Nicht-adjustierte Mittelwerte

c: MWD [95%-KI] und p-Wert der Einzelstudien: berechnet mittels MMRM mit festen Effekten für Behandlung, Behandlung×Visite, Ausgangswert SOBQ-Gesamtscore, Ausgangswert SOBQ-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer

MWD [95%-KI] und p-Wert der gepoolten Daten: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Behandlung×Visite, Ausgangswert SOBQ-Gesamtscore, Ausgangswert SOBQ-Gesamtscore×Visite und zufälligem Effekt für Studienteilnehmer

d: Treated Set; OC

Ein niedrigerer SOBQ-Gesamtscore steht für eine geringere Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch Kurzatmigkeit.

KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler); SOBQ: Shortness of Breath Questionnaire

In der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 sowie in der gepoolten Analyse der INPULSIS Studien zeigte sich für die *Veränderung des SOBQ-Gesamtscores* bis zur Woche 24 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Dies gilt ebenso für den Gruppenvergleich der Einzelstudien.

CASA-Q

Die Endpunkte *Veränderung CASA-Q Husten-Symptome* und *Veränderung CASA-Q Husten-Impact* wurden nur in den INPULSIS Studien erhoben. Die Darstellung der Ergebnisse erfolgt jeweils als Veränderung des Scores vom Ausgangswert bis Woche 52. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien gepoolten Analyse beizumessen.

Tabelle 4-47: Ergebnisse für *Veränderung CASA-Q (Husten-Symptome bzw. Husten-Impact) bis Woche 52* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	p-Wert
Veränderung CASA-Q Husten-Symptome bis Woche 52										
INPULSIS ^c (gepoolt)	Nintedanib	638	60,16 (23,757)	519	60,63 (24,902)	625	-0,57 (0,785)	0,95 (1,241)	[-1,49; 3,38]	0,4450
	Placebo	423	59,52 (22,320)	345	59,32 (23,226)	417	-1,52 (0,962)			
INPULSIS 1 ^c	Nintedanib	309	58,63 (23,592)	251	58,50 (24,455)	302	-0,76 (1,136)	-0,24 (1,803)	[-3,78; 3,30]	0,8942
	Placebo	204	56,29 (22,861)	164	57,22 (24,829)	202	-0,52 (1,400)			
INPULSIS 2 ^c	Nintedanib	329	61,60 (23,858)	268	62,62 (25,197)	323	-0,33 (1,087)	2,05 (1,713)	[-1,31; 5,41]	0,2326
	Placebo	219	62,52 (21,423)	181	61,23 (21,562)	215	-2,38 (1,325)			
Veränderung CASA-Q Husten-Impact bis Woche 52										
INPULSIS ^d (gepoolt)	Nintedanib	637	74,91 (23,495)	520	74,50 (24,212)	624	-2,48 (0,705)	1,73 (1,115)	[-0,46; 3,92]	0,1212
	Placebo	423	75,66 (22,125)	345	73,42 (24,011)	417	-4,21 (0,865)			

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^c		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SE)	MWD (SE)	95%-KI	p-Wert
INPULSIS 1 ^d	Nintedanib	309	74,22 (22,836)	251	73,82 (23,077)	302	-2,36 (1,006)	1,64 (1,596)	[-1,49; 4,77]	0,3042
	Placebo	204	74,18 (22,344)	164	71,91 (25,920)	202	-4,00 (1,240)			
INPULSIS 2 ^d	Nintedanib	328	75,55 (24,116)	269	75,14 (25,252)	322	-2,58 (0,991)	1,81 (1,564)	[-1,26; 4,88]	0,2475
	Placebo	219	77,04 (21,881)	181	74,79 (22,123)	215	-4,39 (1,209)			

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.4:1, Tab. 15.2.2.2.2.4:2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.2.4:1, Tab. 15.2.2.2.2.4:2; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 86.7.1.1, Tab. 86.8.1.1, Tab. 1.1.28.16.1, Tab. 1.1.28.18.1

a: Anzahl analysierter Studienteilnehmer

b: Nicht-adjustierter Mittelwert

c: MWD [95%-KI] und p-Wert der Einzelstudien: berechnet mittels MMRM mit festen Effekten für Behandlung, Behandlung×Visite, Ausgangswert CASA-Q Husten-Symptome bzw. CASA-Q Husten-Impact, Ausgangswert CASA-Q Husten-Symptome×Visite bzw. CASA-Q Husten-Impact×Visite und zufälligem Effekt für Studienteilnehmer

MWD [95%-KI] und p-Wert der gepoolten Daten: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Behandlung×Visite, Ausgangswert CASA-Q Husten-Symptome bzw. CASA-Q Husten-Impact, Ausgangswert CASA-Q Husten-Symptome×Visite bzw. CASA-Q Husten-Impact×Visite und zufälligem Effekt für Studienteilnehmer.

d: Treated Set; OC

Ein höherer Score ist mit weniger Husten-Symptomen bzw. weniger Husten-Impact assoziiert.

CASA-Q: Cough and Sputum Assessment Questionnaire; KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler)

Für den Endpunkt *CASA-Q Husten-Symptome* war die Veränderung des Scores vom Ausgangswert bis Woche 52 in der gepoolten Analyse der INPULSIS Studien unter Nintedanib im Vergleich zur zVT geringer (-0,57 Punkte vs. -1,52 Punkte) (Tabelle 4-47). Die MWD [95%-KI] für die Behandlungsgruppen beträgt für die gepoolten Daten 0,95 Punkte [-1,49; 3,38]; p=0,4450. Dies spiegelt eine geringere Verschlechterung der gesundheitsbezogenen Lebensqualität hinsichtlich der Husten-Symptome unter Nintedanib wider.

Auf Ebene der Einzelstudien ergaben sich für die Studie INPULSIS 1 Veränderungen vom Ausgangswert bis Woche 52 (Nintedanib vs. zVT) von -0,76 Punkte vs. -0,52 Punkte) Hingegen war in der INPULSIS 2 Studie (-0,33 Punkte vs. -2,38 Punkte) die Veränderung des Scores unter Nintedanib im Vergleich zur zVT geringer. Die jeweilige MWD [95%-KI] für die Behandlungsgruppen beträgt für die INPULSIS 1 Studie -0,24 Punkte [-3,78; 3,30]; $p=0,8942$ und für die INPULSIS 2 Studie 2,05 Punkte [-1,31; 5,41]; $p=0,2326$.

Wie für den Endpunkt *CASA-Q Husten-Symptome*, der die gesundheitsbezogene Lebensqualität hinsichtlich der Husten-Symptome berichtet, war auch für den Endpunkt *CASA-Q Husten-Impact* die Veränderung vom Ausgangswert bis Woche 52 in der gepoolten Analyse der INPULSIS Studien unter Nintedanib geringer als in der Kontrollgruppe (-2,48 Punkte vs. -4,21 Punkte) (Tabelle 4-47). Die errechnete MWD [95%-KI] für die Behandlungseffekte zwischen Nintedanib und der zVT beträgt für die gepoolten Daten 1,73 Punkte [-0,46; 3,92]; $p=0,1212$. Die Analyse zeigt somit die Tendenz, dass die Behandlung mit Nintedanib im Vergleich zur zVT zu einer geringeren Verschlechterung der gesundheitsbezogenen Lebensqualität in Bezug auf den *Husten-Impact* führt.

Auf Ebene der Einzelstudien ergaben sich für die Studie INPULSIS 1 Veränderungen vom Ausgangswert bis Woche 52 (Nintedanib vs. zVT) von -2,36 Punkte vs. -4,00 Punkte, bzw. für die INPULSIS 2 -2,58 Punkte vs. -4,39 Punkte. Die errechnete MWD [95%-KI] für die Behandlungseffekte zwischen Nintedanib und der zVT beträgt in der INPULSIS 1 Studie 1,64 Punkte [-1,49; 4,77]; $p=0,3042$ und in der INPULSIS 2 Studie 1,81 Punkte [-1,26; 4,88]; $p=0,2475$.

PGI-C

Der Endpunkt *PGI-C-Responder* wurde nur in den INPULSIS Studien erhoben. Als Responder wurden solche Patienten gewertet, welche zu Woche 52 ihre gesundheitsbezogene Lebensqualität gegenüber dem Befinden bei Studienbeginn mindestens als nicht schlechter bewerteten. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien gepoolten Analyse beizumessen.

Tabelle 4-48: Ergebnisse für *PGI-C-Responder* zu Woche 52 aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
PGI-C-Responder zu Woche 52							
INPULSIS (gepoolt) ^d	638	391 (61,29)	423	230 (54,37)	1,328	[1,04; 1,70]	0,0254
INPULSIS 1 ^d	309	188 (60,84)	204	112 (54,90)	1,276	[0,89; 1,83]	0,1818
INPULSIS 2 ^d	329	203 (61,70)	219	118 (53,88)	1,379	[0,98; 1,95]	0,0690
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.5; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.2.5; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.28.20.4							

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c		
	N ^a	n ^b (%)	N ^a	n ^b (%)	OR	95%-KI	p-Wert
a: Anzahl analysierter Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: OR [95%-KI] und p-Werte der Einzelstudien: berechnet mittels logistischer Regression mit Term Behandlung INPULSIS (gepoolt): OR [95%-KI] und p-Wert berechnet mittels logistischer Regression mit Termen Studien und Behandlung d: Treated Set, OC PGI-C-Responder definiert als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „keine Veränderung“ der subjektiv empfundenen gesundheitsbezogenen Lebensqualität ermittelt durch den PGI-C-Fragebogen. KI: Konfidenzintervall; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); OR: Odds Ratio; PGI-C: Patient's Global Impression of Change; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)							

In der gepoolten Analyse der INPULSIS Studien war der Anteil der *PGI-C-Responder* zu Woche 52 in der Nintedanib-Gruppe größer als in der Kontrollgruppe (61,29% vs. 54,37% Responder) (Tabelle 4-48). Das entsprechende OR [95%-KI] für den Behandlungseffekt von Nintedanib vs. zVT beträgt 1,328 [1,04; 1,70]; p=0,0254. Dieser Unterschied ist statistisch signifikant. Der Anteil der Studienteilnehmer, die ihre gesundheitsbezogene Lebensqualität in Woche 52 im Vergleich zum Ausgangswert als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „ohne Veränderung“ einschätzen, ist demnach unter Nintedanib höher als im Vergleichsarm mit statistisch signifikantem Effekt.

Auf Ebene der Einzelstudien wurde zu Woche 52 ein Anteil an PGI-C-Respondern von 60,84% vs. 54,90% in der Studie INPULSIS 1, bzw. von 61,70% vs. 53,88% Respondern in der Studie INPULSIS 2 beobachtet. Das OR [95%-KI] für den Behandlungseffekt von Nintedanib vs. zVT beträgt in der INPULSIS 1 Studie 1,276 [0,89; 1,83]; p=0,1818 und in der INPULSIS 2 Studie 1,379 [0,98; 1,95]; p=0,0690.

EQ-5D-VAS

Der Endpunkt *EQ-5D-VAS* wurde nur in den INPULSIS Studien erhoben. Die Darstellung der Ergebnisse erfolgt jeweils als Veränderung des Scores vom Ausgangswert bis Woche 52. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien gepoolten Analyse beizumessen.

Tabelle 4-49: Ergebnisse für *Veränderung EQ-5D-VAS bis Woche 52* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Behandlung	Ausgangswert		Woche 52		Veränderung		Behandlungseffekt Nintedanib vs. Placebo ^d		
		N ^a	MW ^b (SD)	N ^a	MW ^b (SD)	N ^a	MWD ^c (SD)	MWD (SE)	95%-KI	p-Wert
Veränderung EQ-5D-VAS bis Woche 52										
INPULSIS (gepoolt) ^e	Nintedanib	630	68,28 (18,22)	517	66,61 (18,23)	605	-2,70 (0,655)	3,81 (1,037)	[1,78; 5,85]	0,0002
	Placebo	421	67,88 (16,39)	340	62,61 (19,25)	408	-6,51 (0,804)	Hedges‘ g 0,252	[0,12; 0,39]	
INPULSIS 1 ^e	Nintedanib	306	66,71 (17,42)	249	65,04 (17,65)	247	-2,95 (0,943)	3,09 (1,500)	[0,14; 6,03]	0,0399
	Placebo	203	68,02 (16,34)	161	62,12 (20,11)	160	-6,04 (1,167)	Hedges‘ g 0,203	[0,01; 0,40]	
INPULSIS 2 ^e	Nintedanib	324	69,77 (18,84)	268	68,06 (18,66)	265	-2,50 (0,908)	4,39 (1,431)	[1,59; 7,20]	0,0022
	Placebo	218	67,75 (16,47)	179	63,04 (18,48)	178	-6,90 (1,106)	Hedges‘ g 0,292	[0,11; 0,48]	
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a): Tab. 15.2.2.2.2.6; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b): Tab. 15.2.2.2.2.6; (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab: 1.1.28.21.1, Tab. 1.1.28.21.2, Tab. Tab. 1.1.28.21.3										
a: Anzahl analysierter Studienteilnehmer										
b: Nicht-adjustierte Mittelwerte										
c: Nicht-adjustierte Mittelwertdifferenz										
d: MWD [95%-KI] und p-Wert der gepoolten Daten: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Visite, Behandlung×Visite, Ausgangswert EQ-5D-VAS, Ausgangswert EQ-5D-VAS×Visite und zufälligem Effekt für Studienteilnehmer.										
e: Treated Set; OC										
Ein höherer Score steht für eine bessere subjektiv empfundene gesundheitsbezogene Lebensqualität.										
EQ-5D: European Quality of Life-5 Dimensions; KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MW: Mittelwert; MWD: Mittelwertdifferenz; OC: Observed Case (Auswertung nur der tatsächlich beobachteten Werte); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SD: Standard Deviation (Standardabweichung); SE: Standard Error (Standardfehler); VAS: Visual Analog Scale										

Die Verschlechterung des Gesundheitszustandes der Patienten gemessen anhand der Veränderung auf der VAS des EQ-5D zwischen Ausgangswert und Woche 52 war in der gepoolten Analyse der INPULSIS Studien in der Nintedanib-Gruppe geringer als in der Kontrollgruppe (-2,70 Punkte vs. -6,51 Punkte) (Tabelle 4-49). Im Verlauf der Studien war bei beiden Behandlungsgruppen der Gesundheitszustand, gemessen anhand der EQ-5D-VAS, (aufgrund des progressiven Charakters der Erkrankung erklärbar) schlechter als zu Studienbeginn. Jedoch gaben die mit Nintedanib behandelten Studienteilnehmer gegenüber den Studienteilnehmern der Vergleichsgruppe eine geringere Verschlechterung des Gesundheitszustandes an. Der beobachtete Unterschied zwischen den Behandlungsgruppen ist statistisch

signifikant (adjustierte MWD [95%-KI]: 3,81 [1,78; 5,85]; $p=0,0002$), jedoch von unklarer klinischer Relevanz (Hedges' g [95%-KI]: 0,252 [0,12; 0,39]).

Auf Ebene der Einzelstudien ergab sich in der Studie INPULSIS 1 eine Veränderung der VAS des EQ-5D zwischen Ausgangswert und Woche 52 (Nintedanib vs. zVT) von -2,95 Punkte vs. -6,04 Punkte, bzw. von -2,50 Punkte vs. -6,90 Punkte in der Studie INPULSIS 2.

Die deskriptiv dargestellten Ergebnisse der EQ-5D-Domänen *Mobilität, Fähigkeit für sich selbst zu sorgen, alltägliche Aktivitäten, Schmerzen/körperliche Beschwerden* und *Angst/Niedergeschlagenheit* befinden sich in Modul 5.

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.

Die Ergebnisse der IPD Meta-Analyse für die Endpunkte zur gesundheitsbezogenen Lebensqualität und zum Gesundheitszustand wurden bereits in Abschnitt 4.3.1.3.3.1 dargestellt.

Die grafische Darstellung der Meta-Analysen mittels Forest-Plots befindet sich in nachfolgenden Abbildungen. Sie zeigen die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der IPD Meta-Analyse.

Die Heterogenitätstests (siehe auch Abschnitt 4.3.1.3.5.3) weisen für die Endpunkte *Veränderung des SGRQ-Gesamtscores*, *Veränderung des SGRQ-I-Gesamtscores* sowie *Veränderung des SOBQ-Gesamtscores* – für die Ergebnisse aus allen drei Studien vorlagen – auf eine homogene Datenlage hin (p -Wert der Interaktionstests [Merkmal Studie]: $p=0,2081$, $p=0,3285$ bzw. $p=0,6483$).

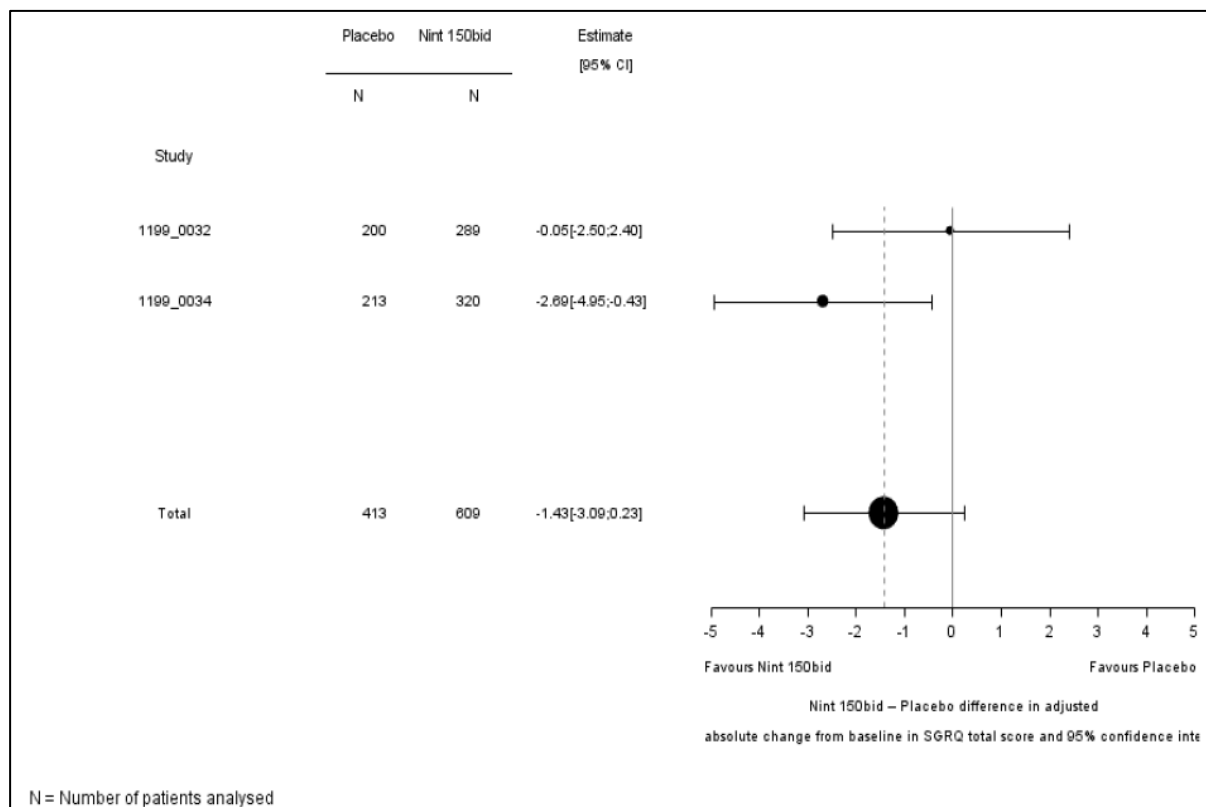
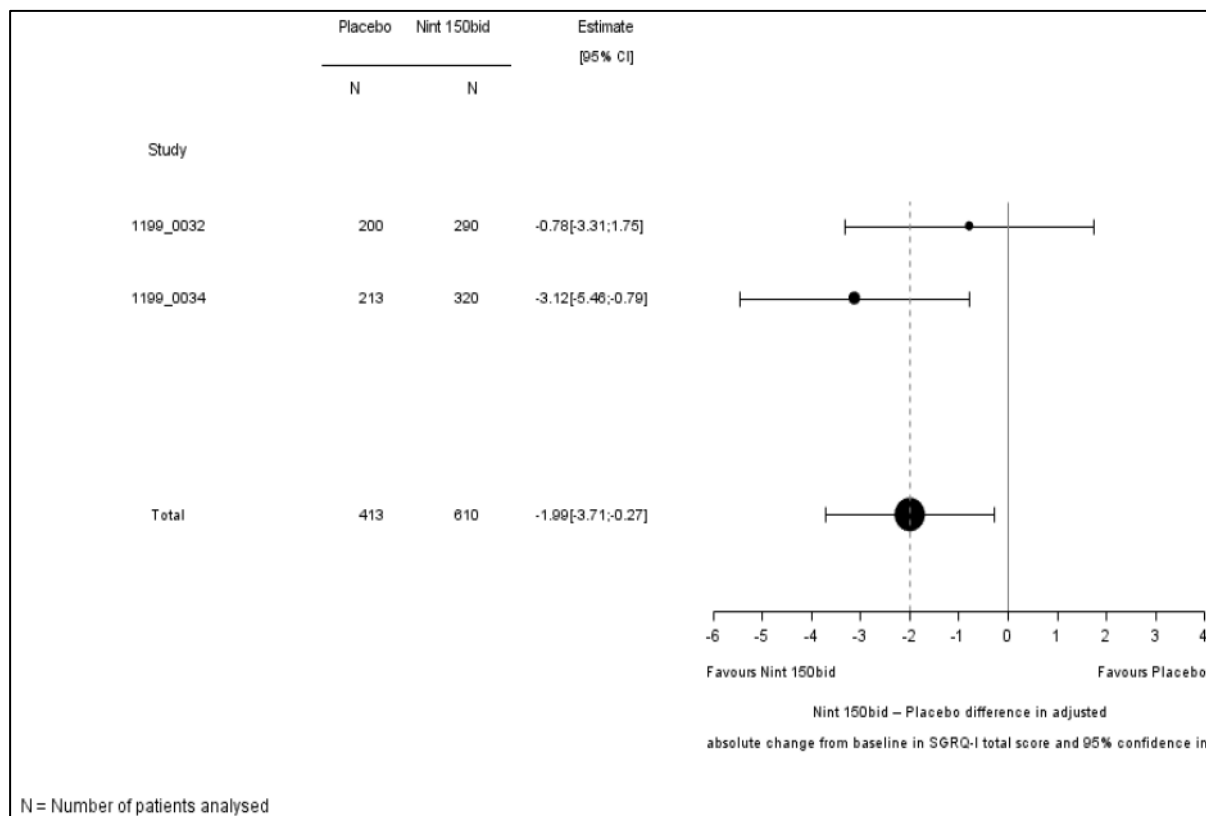
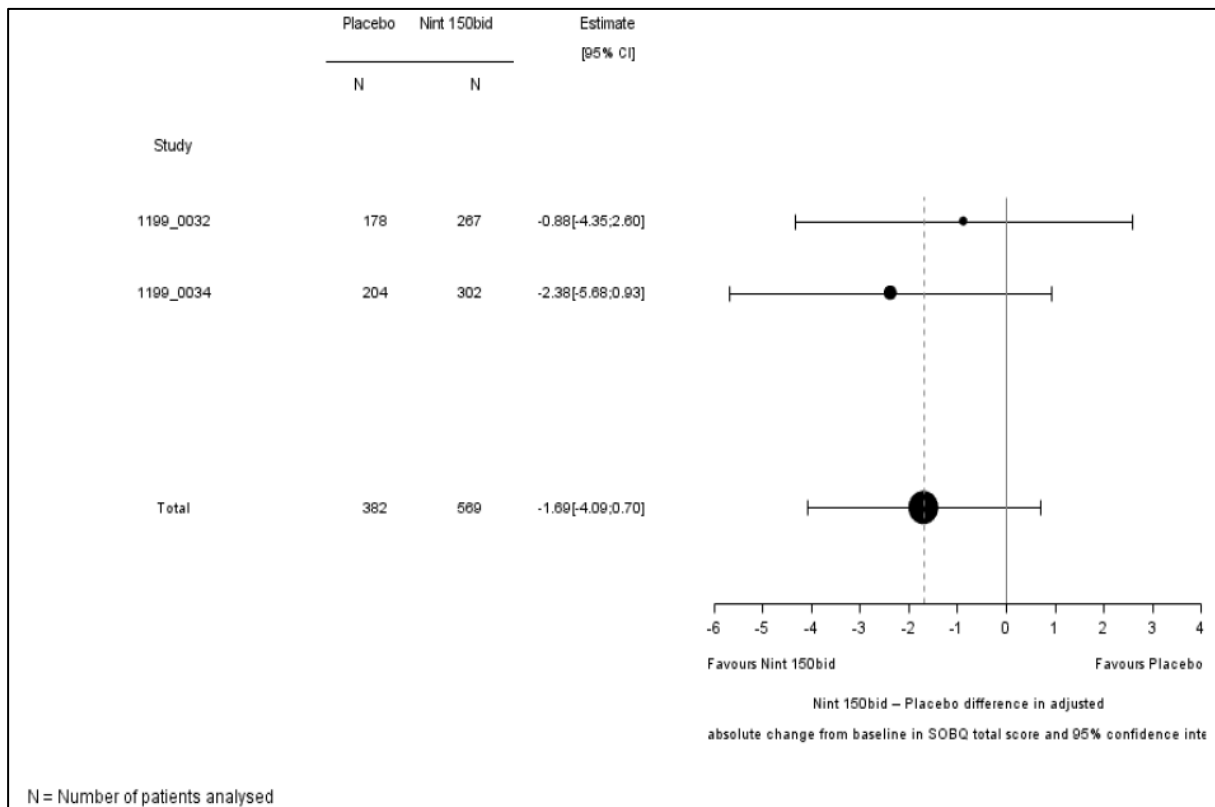


Abbildung 24: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung des SGRQ-Gesamtscores* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

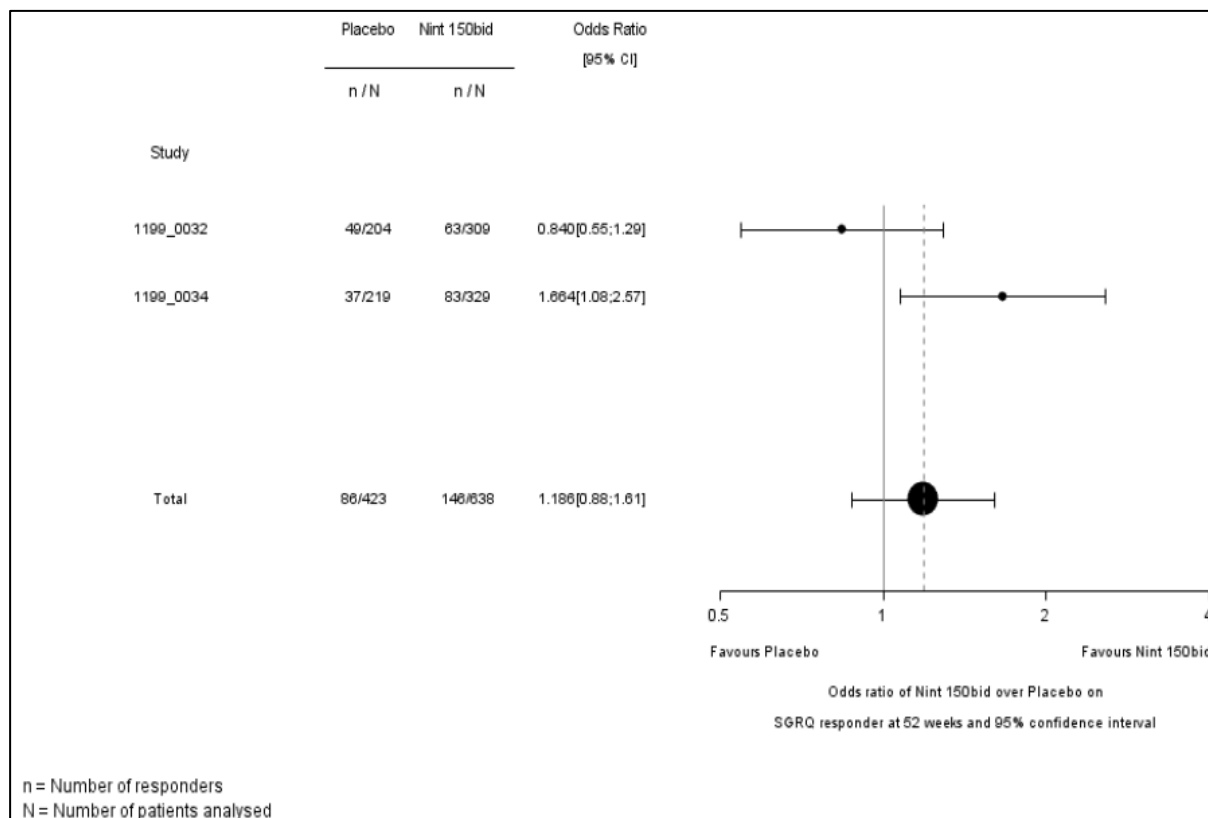
Abbildung 25: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung des SGRQ-I-Gesamtscores* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

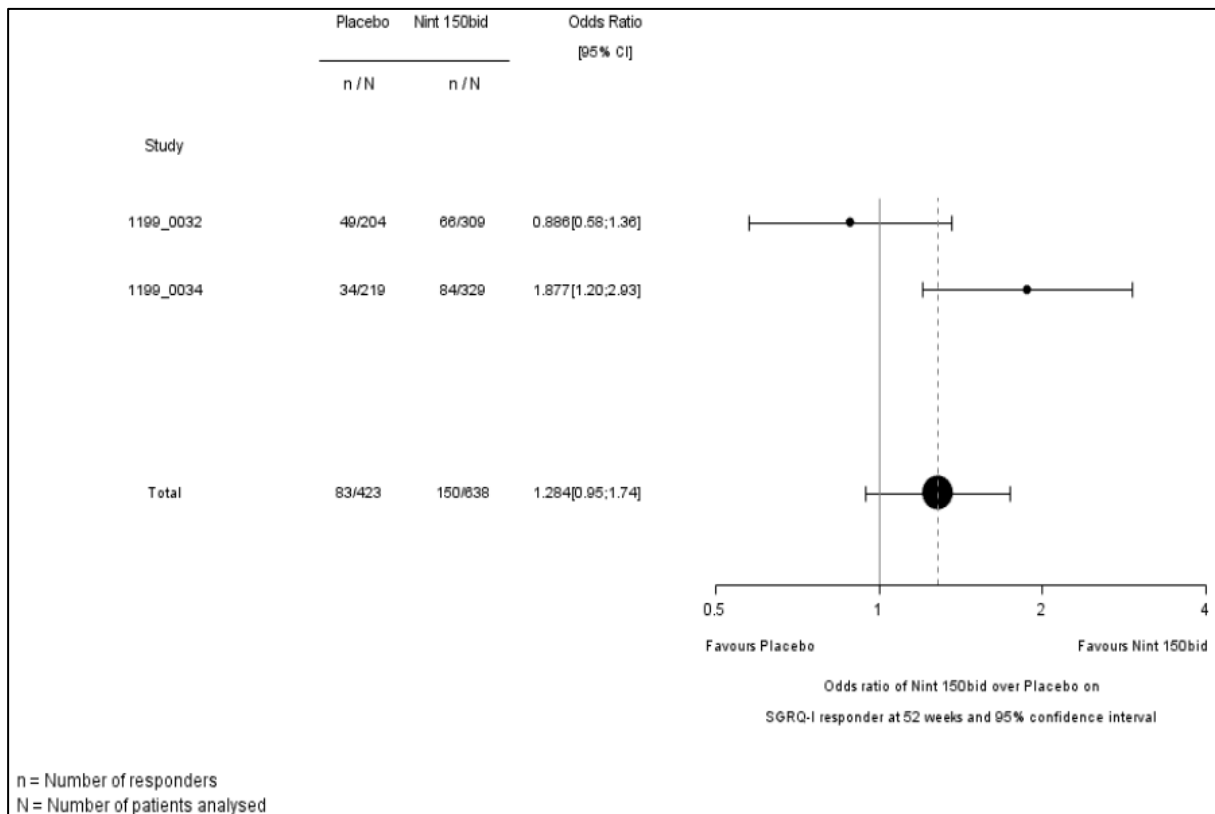
Abbildung 26: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung des SOBQ-Gesamtscores* - RCT mit dem zu bewertenden Arzneimittel

Für die Responderanalysen zum *SGRQ-Gesamtscore zu Woche 52* und zum *SGRQ-I-Gesamtscore zu Woche 52*, die ebenfalls nur für die INPULSIS Studien vorlagen, ist die Datenlage heterogen (Interaktionstest: $p=0,0277$ bzw. $p=0,0164$). Die in den Einzelstudien beobachteten Effekte sind gegenläufig.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 27: Forest-Plot zur Meta-Analyse für den Endpunkt *SGRQ-(≤ 4 Punkte)-Responder zu Woche 52* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 28: Forest-Plot zur Meta-Analyse für den Endpunkt *SGRQ-I-(≤ 4 Punkte)-Responder zu Woche 52* - RCT mit dem zu bewertenden Arzneimittel

Mit Ausnahme des Endpunktes *Veränderung des EQ-5D-VAS* (p-Wert der Interaktionstests [Merkmal *Studie*]: $p=0,0200$), für den die Datenlage heterogen war, ist die Datenlage für die Endpunkte *Veränderung des CASA-Q Husten-Symptom-Scores*, *Veränderung des CASA-Q Husten-Impact-Scores*, *PGI-C-Responder zu Woche 52*, die nur in den INPULSIS Studien erhoben wurden, homogen (p-Wert der Interaktionstests [Merkmal *Studie*]: $p=0,3627$, $p=0,9667$ bzw. $p=0,7607$).

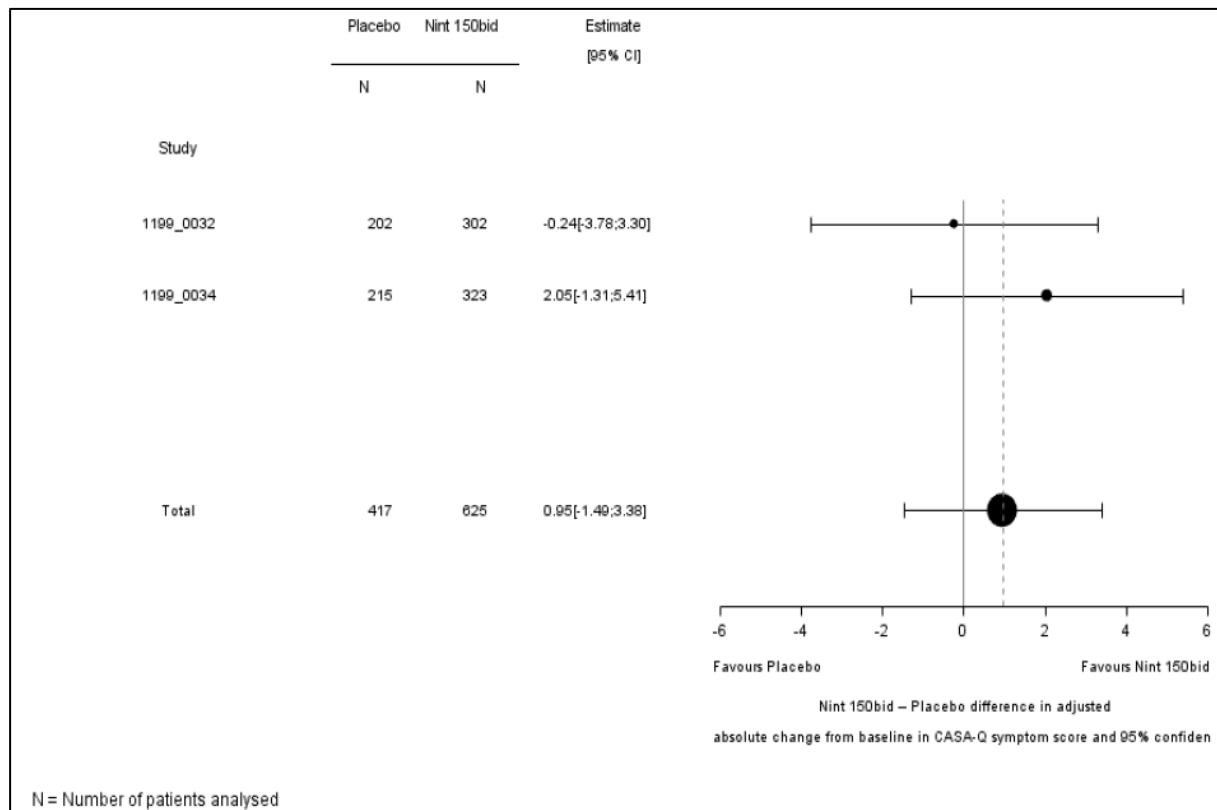


Abbildung 29: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung des CASA-Q Husten-Symptom-Scores* - RCT mit dem zu bewertenden Arzneimittel

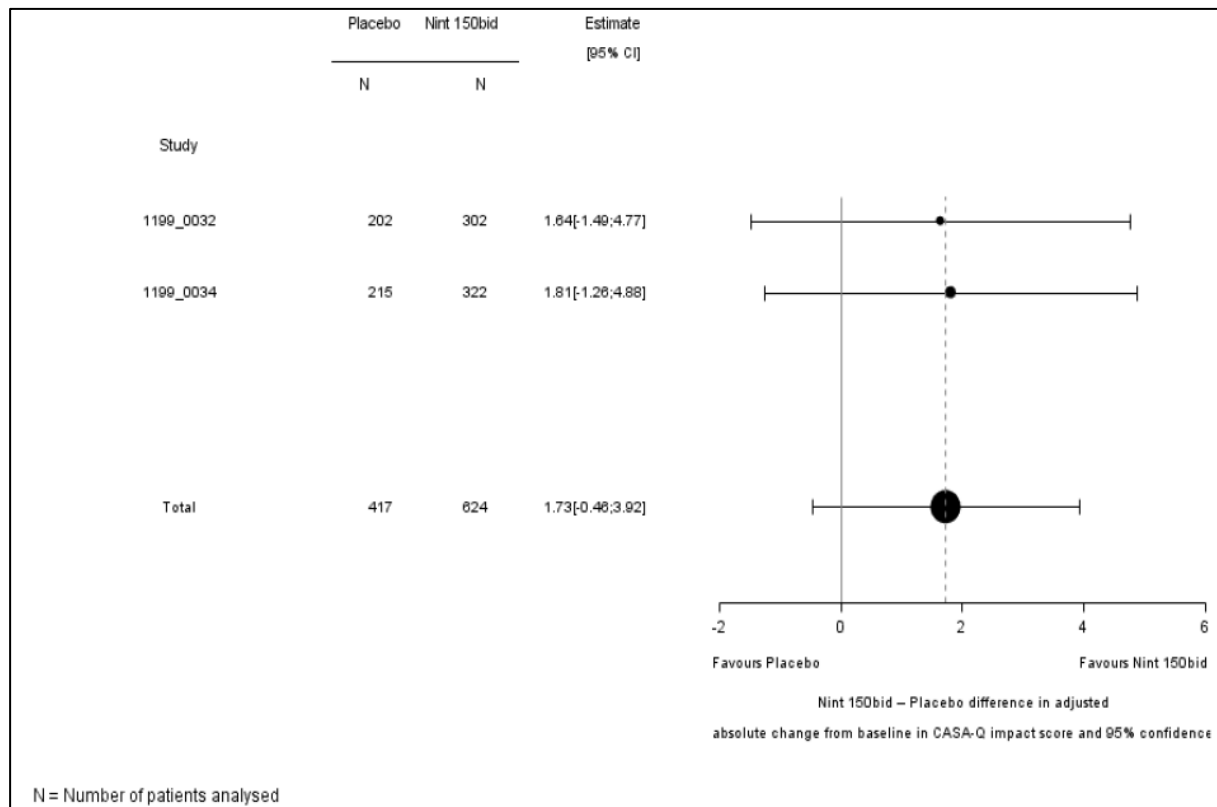
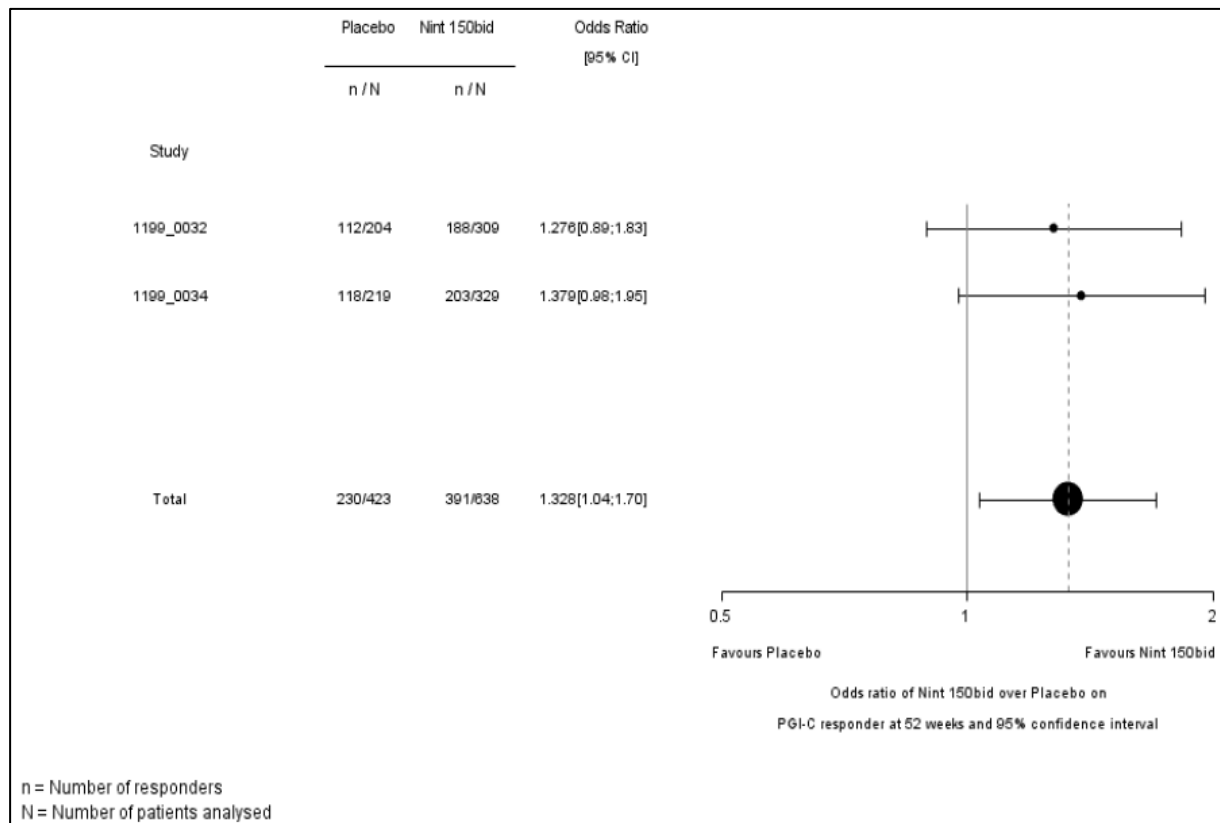
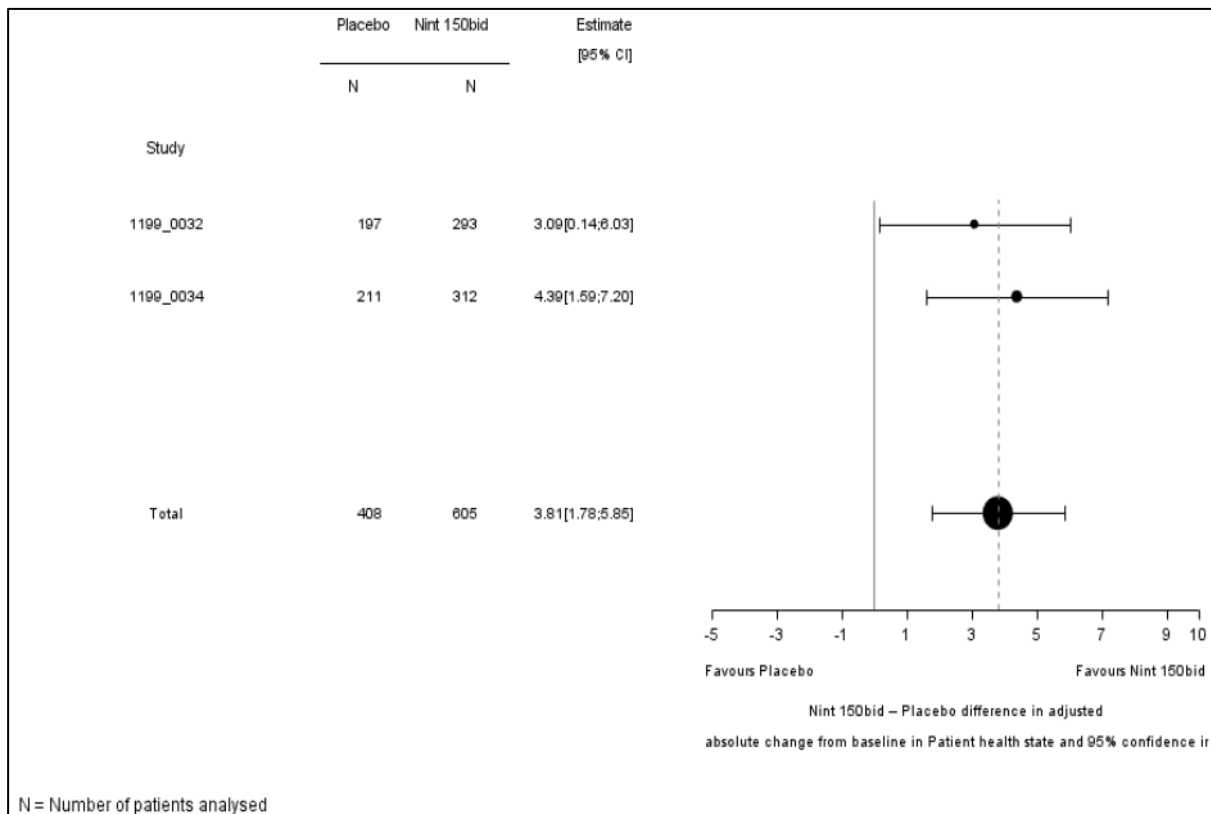


Abbildung 30: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung des CASA-Q Husten-Impact-Scores* - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 31: Forest-Plot zur Meta-Analyse für den Endpunkt *PGI-C-Responder* zu Woche 52 - RCT mit dem zu bewertenden Arzneimittel



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 32: Forest-Plot zur Meta-Analyse für den Endpunkt *Veränderung EQ-5D-VAS* - RCT mit dem zu bewertenden Arzneimittel

Analog zum Vorgehen bei den Responderanalysen für die *Abnahme der FVC*, ist auch im Fall der Responderanalysen zum *SGRQ* und *SGRQ-I* die Auswertung für die längstmögliche Beobachtungsdauer (52 Wochen) diejenige, die für die Bewertung des Zusatznutzens herangezogen wurde. Die Auswertung aller drei Studien, die nur über den kürzeren Zeitraum (24 Wochen) möglich war, wurde supportiv dargestellt (Tabelle 4-44); auf eine grafische Darstellung der Meta-Analyse der Ergebnisse zum Zeitpunkt von 24 Wochen wurde hierbei verzichtet.

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Den in Abschnitt 4.3.1.3.3 dargestellten Instrumenten zur Beurteilung der krankheitsspezifischen und generischen gesundheitsbezogenen Lebensqualität und des allgemeinen Gesundheitszustandes liegen weit verbreitete und validierte Fragebögen zugrunde. Als solche sind die davon abgeleiteten Endpunkte eindeutig definiert und können objektiv erhoben werden. Von einer Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext kann somit ausgegangen werden.

4.3.1.3.4 Unerwünschte Ereignisse – 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.

4.3.1.3.4.1 Endpunkt Gesamtzahl UE

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

Tabelle 4-50: Operationalisierung von Gesamtzahl UE

Studie	Operationalisierung
INPULSIS 1	Definition: Ein UE war definiert als jegliches unerwünschtes medizinisches Vorkommnis (einschließlich einer Verschlechterung des bestehenden Krankheitszustandes) bei einem Studienteilnehmer, der Studienmedikation erhielt. Dabei musste das Ereignis nicht unbedingt in einem kausalen Zusammenhang mit dem angewandten Arzneimittel stehen. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme.
INPULSIS 2	Homogenes Studiendesign zu INPULSIS 1
Studie 1199.187	Definition: Ein UE war definiert als jegliches unerwünschtes medizinisches Vorkommnis (einschließlich einer Verschlechterung des bestehenden Krankheitszustandes) bei einem Studienteilnehmer, der Studienmedikation erhielt. Dabei musste das Ereignis nicht unbedingt in einem kausalen Zusammenhang mit dem angewandten Arzneimittel stehen. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme. Für die vorliegende Nutzenbewertung wurden die Analysen der unerwünschten Ereignisse der kontrollierten Studienphase (im Studienbericht wird diese Phase als P1 bezeichnet) berücksichtigt.
UE: Unerwünschtes Ereignis	

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 4-51: Bewertung des Verzerrungspotenzials für den Endpunkt *Gesamtzahl UE* 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
INPULSIS 1	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für UE erfolgte in den INPULSIS Studien doppelblind. Für Studie 1199.187 wurde die kontrollierte Studienphase herangezogen.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der beiden INPULSIS Studien und der Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen hatten. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien aller Endpunkte zu den unerwünschten Ereignissen waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Analyseplan ausgewertet und im Studienbericht transparent und vollständig berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war für die *Gesamtzahl UE* transparent.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für den Endpunkt *Gesamtzahl UE* ein niedriges Verzerrungspotenzial.

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

Der Endpunkt Gesamtzahl UE wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimiteleinnahe erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Dargestellt wird die Gesamtzahl UE sowohl als gepoolte Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-52: Ergebnisse für die *Gesamtzahl UE* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Gesamtzahl UE								
INPULSIS + Studie 1199.187 (gepoolt)	694	664 (95,7)	480	431 (89,8)	1,07 [1,03; 1,10]	2,56 [1,59; 4,10]	5,97 [2,86; 9,08]	<0,001
INPULSIS (gepoolt)	638	609 (95,5)	423	379 (89,6)	1,07 [1,03; 1,11]	2,44 [1,50; 3,96]	5,86 [2,53; 9,18]	<0,001
INPULSIS 1	309	298 (96,4)	204	181 (88,7)	1,09 [1,03; 1,15]	3,44 [1,64; 7,23]	7,71 [2,91; 12,52]	<0,001
INPULSIS 2	329	311 (94,5)	219	198 (90,4)	1,05 [0,99; 1,10]	1,83 [0,95; 3,53]	4,12 [-0,49; 8,73]	0,066
Studie 1199.187	56	55 (98,2)	57	52 (91,2)	1,08 [0,99; 1,18]	5,29 [0,60; 46,79]	6,99 [-1,14; 15,11]	0,098
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.5.1.1, Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 1.3.49.2.6.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]; Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein UE zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

Der Anteil der Patienten mit mindestens einem UE war in den gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) höher in der Nintedanib-Gruppe im Vergleich zur Placebo-Gruppe (95,7% vs. 89,8% für die gepoolte Analyse aller drei eingeschlossenen Studien) (Tabelle 4-52). Der Unterschied war statistisch signifikant zuungunsten von Nintedanib mit einem RR [95%-KI] von 1,07 [1,03; 1,10]; $p < 0,001$ für die über die drei Studien gepoolte Analyse bzw. von 1,07 [1,03; 1,11]; $p < 0,001$ für die gepoolten INPULSIS Studien. Der beobachtete Unterschied zwischen den Behandlungsarmen ist jedoch allenfalls als geringfügig einzustufen.

Auf Ebene der Einzelstudien fand sich in der Studie INPULSIS 1 ein statistisch signifikant höherer Anteil an Patienten mit mindestens einem UE (Tabelle 4-52).

Die entsprechenden Sensitivitätsanalysen für den Endpunkt *Gesamtzahl UE* ohne Berücksichtigung von adjudizierten Exazerbationen sind in Tabelle 4-53 dargestellt. Die Ergebnisse unterscheiden sich nicht relevant von den Ergebnissen der Hauptanalyse.

Tabelle 4-53: Sensitivitätsanalyse: Ergebnisse für die *Gesamtzahl UE ohne Berücksichtigung von adjudizierten Exazerbationen* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Gesamtzahl UE								
INPULSIS + Studie 1199.187 (gepoolt)	694	664 (95,7)	480	431 (89,8)	1,07 [1,03; 1,10]	2,56 [1,59; 4,10]	5,97 [2,86; 9,08]	<0,001
Gesamtzahl UE ohne Berücksichtigung von adjudizierten Exazerbationen								
INPULSIS + Studie 1199.187 (gepoolt)	694	662 (95,4)	480	428 (89,2)	1,07 [1,03; 1,11]	2,56 [1,62; 4,05]	6,33 [3,13; 9,52]	<0,001
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.13.1.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein UE zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

In Tabelle 4-54 sind alle UE nach MedDRA Systemorganklasse (SOC) und Preferred term (PT) dargestellt, die mit einer Häufigkeit von $\geq 5\%$ in mindestens einem der Studienarme auftraten. Datenbasis sind die gepoolten Safety Daten der drei Studien INPULSIS 1 und 2 sowie Studie 1199.187. Die Ereignisse wurden für alle Studien einheitlich mit dem Medical Dictionary for Regulatory Activities (MedDRA) Dictionary Version 20.1 kodiert und klassifiziert.

Tabelle 4-54: Häufige UE ($\geq 5\%$ in mindestens einem Studienarm) nach MedDRA SOC und PT aus RCT mit dem zu bewertenden Arzneimittel – INPULSIS Studien und Studie 1199.187 (gepoolt)

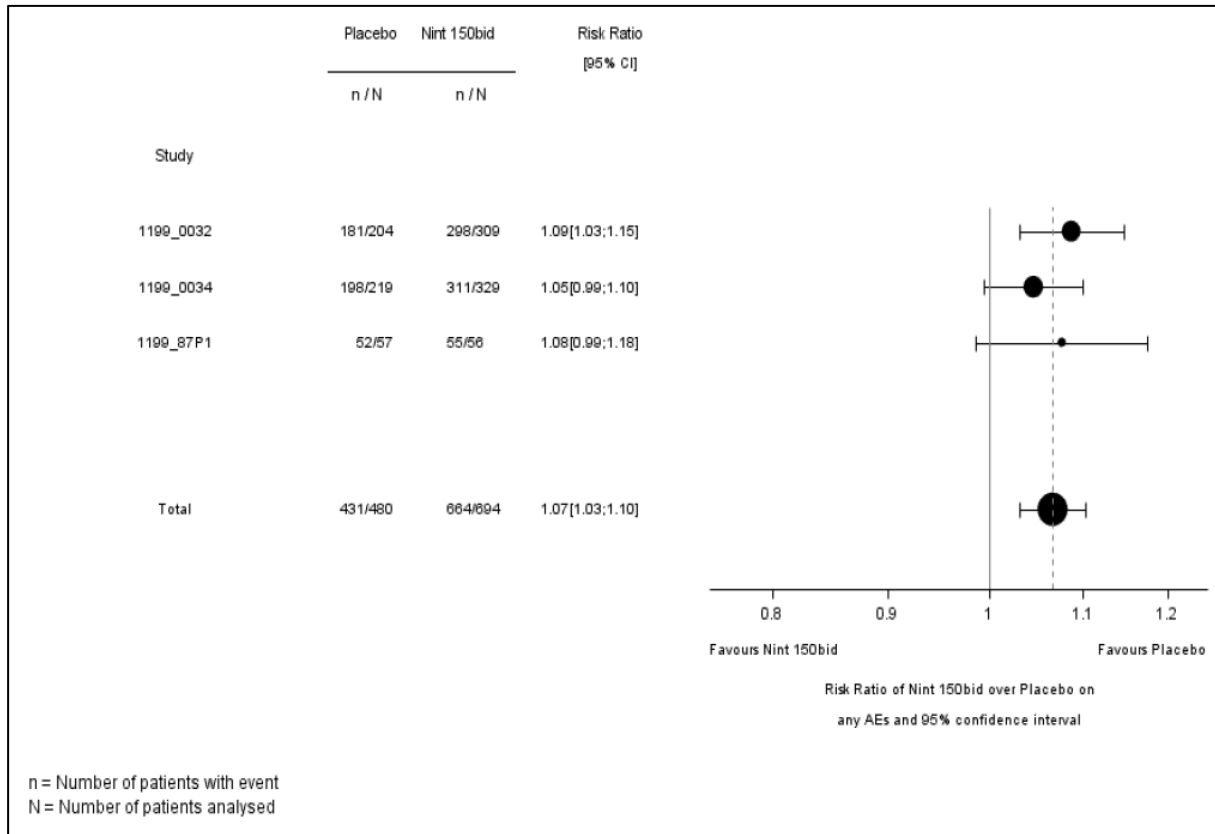
MedDRA SOC Preferred Term	INPULSIS + Studie 1199.187 (gepoolt)	
	Nintedanib N=694	Placebo N=480
Gesamtzahl UE, n (%)	664 (95,7)	431 (89,8)
Erkrankungen des Gastrointestinaltrakts	537 (77,4)	198 (41,3)
Bauchschmerz	59 (8,5)	15 (3,1)
Schmerz im Oberbauch	44 (6,3)	18 (3,8)
Konstipation	39 (5,6)	19 (4,0)
Diarrhö	433 (62,4)	99 (20,6)
Übelkeit	172 (24,8)	41 (8,5)
Erbrechen	83 (12,0)	14 (2,9)
Infektionen und parasitäre Erkrankungen	385 (55,5)	255 (53,1)
Bronchitis	69 (9,9)	48 (10,0)
Nasopharyngitis	90 (13,0)	72 (15,0)
Pneumonie	34 (4,9)	28 (5,8)
Infektionen der oberen Atemwege	60 (8,6)	47 (9,8)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	269 (38,8)	201 (41,9)
Husten	90 (13,0)	66 (13,8)
Dyspnoe	57 (8,2)	54 (11,3)
IPF	67 (9,7)	66 (13,8)
Untersuchungen	203 (29,3)	77 (16,0)
Gewichtsverlust	70 (10,1)	15 (3,1)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	166 (23,9)	118 (24,6)
Brustschmerz	35 (5,0)	23 (4,8)
Fatigue	50 (7,2)	39 (8,1)
Stoffwechsel- und Ernährungsstörungen	129 (18,6)	64 (13,3)
Appetit vermindert	80 (11,5)	28 (5,8)

	INPULSIS + Studie 1199.187 (gepoolt)	
MedDRA SOC Preferred Term	Nintedanib N=694	Placebo N=480
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen	125 (18,0)	109 (22,7)
Rückenschmerz	37 (5,3)	32 (6,7)
Erkrankungen des Nervensystems	119 (17,1)	75 (15,6)
Kopfschmerz	50 (7,2)	26 (5,4)
Erkrankungen der Haut und des Unterhautzellgewebes	103 (14,8)	67 (14,0)
Verletzungen, Vergiftungen und Behandlungskomplikationen	58 (8,4)	52 (10,8)
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.5.1.2 IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); MedDRA: Medical Dictionary for Regulatory Activities; PT: Preferred Term; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SOC: System Organ Class; UE: Unerwünschtes Ereignis		

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.

Das Ergebnis der IPD Meta-Analyse für den Endpunkt *Gesamtzahl UE* wurde bereits in Tabelle 4-52 dargestellt.

Die grafische Darstellung der Meta-Analyse mittels Forest-Plot befindet sich in Abbildung 33. Der Forest-Plot zeigt die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, den Schätzer der IPD Meta-Analyse. Der entsprechende Heterogenitätstest (siehe auch Abschnitt 4.3.1.3.5.4) weist auf eine homogene Datenlage hin (p-Wert der Interaktionstests [Merkmal *Studie*]: p=0,356).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 33: Forest-Plot zur Meta-Analyse für den Endpunkt *Gesamtzahl UE* - RCT mit dem zu bewertenden Arzneimittel**Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext**

Die Erhebung der unerwünschten Ereignisse ist international standardisiert (International Conference on Harmonisation, 1996) und wird einheitlich dokumentiert. Von einer Übertragbarkeit der Studienergebnisse zum Endpunkt *Gesamtzahl UE* auf den deutschen Versorgungskontext kann somit ausgegangen werden.

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.

4.3.1.3.4.2 Endpunkt schwerwiegende UE

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

Tabelle 4-55: Operationalisierung von SUE

Studie	Operationalisierung
INPULSIS 1	Definition: Ein SUE war definiert als jegliches UE, welches entweder zum Tode oder zu einer unmittelbar lebensbedrohlichen Situation führte, eine dauerhafte oder erhebliche Behinderung/Invalidität hervorrief, einen Krankenhausaufenthalt erforderlich machte oder verlängerte, eine angeborene Anomalie/einen Geburtsdefekt darstellte, aus einem anderen Grund für schwerwiegend erachtet wurde oder einen medizinischen oder chirurgischen Eingriff erforderlich machte, um eine der oben genannten Konsequenzen zu verhindern. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme.
INPULSIS 2	Homogenes Studiendesign zu INPULSIS 1
Studie 1199.187	Definition: Ein SUE war definiert als jegliches UE, welches entweder zum Tode oder zu einer unmittelbar lebensbedrohlichen Situation führte, eine dauerhafte oder erhebliche Behinderung/Invalidität hervorrief, einen Krankenhausaufenthalt erforderlich machte oder verlängerte, eine angeborene Anomalie/einen Geburtsdefekt darstellte, aus einem anderen Grund für schwerwiegend erachtet wurde oder einen medizinischen oder chirurgischen Eingriff erforderlich machte, um eine der oben genannten Konsequenzen zu verhindern. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme. Für die vorliegende Nutzenbewertung wurden die Analysen der unerwünschten Ereignisse der kontrollierten Studienphase (im Studienbericht wird diese Phase als P1 bezeichnet) berücksichtigt.
SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis	

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 4-56: Bewertung des Verzerrungspotenzials für die Endpunkte zu den SUE 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
INPULSIS 1	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SUE: Schwerwiegendes unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für UE erfolgte in den INPULSIS Studien doppelblind. Für Studie 1199.187 wurde die kontrollierte Studienphase herangezogen.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen hatten. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien aller Endpunkte zu den unerwünschten Ereignissen waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Analyseplan ausgewertet und transparent und vollständig im Studienbericht berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war für die SUE transparent.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für den Endpunkt *SUE* ein niedriges Verzerrungspotenzial.

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

Der Endpunkt SUE wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimittleinnahme erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-57: Ergebnisse für die *SUE* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
SUE								
INPULSIS + Studie 1199.187 (gepoolt)	694	202 (29,1)	480	136 (28,3)	1,01 [0,84; 1,21]	1,01 [0,78; 1,31]	0,20 [-5,05; 5,45]	0,941
INPULSIS (gepoolt)	638	194 (30,4)	423	127 (30,0)	1,01 [0,84; 1,22]	1,02 [0,78; 1,33]	0,38 [-5,26; 6,02]	0,894
INPULSIS 1	309	96 (31,1)	204	55 (27,0)	1,15 [0,87; 1,53]	1,22 [0,83; 1,81]	4,11 [-3,87; 12,09]	0,318
INPULSIS 2	329	98 (29,8)	219	72 (32,9)	0,91 [0,70; 1,17]	0,87 [0,60; 1,25]	-3,09 [-11,03; 4,86]	0,444
Studie 1199.187	56	8 (14,3)	57	9 (15,8)	0,90 [0,38; 2,18]	0,89 [0,32; 2,50]	-1,50 [-14,68; 11,67]	0,823
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 1.3.49.2.6.1, Tab. 1.5.5.2.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]; Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); SUE: Schwerwiegendes unerwünschtes Ereignis								

SUE traten in der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 bei 29,1% der Studienteilnehmer, die mit Nintedanib behandelt wurden, auf, und somit mit vergleichbarer Häufigkeit wie im Vergleichsarm (28,3%). Der Unterschied zwischen den Gruppen ist nicht statistisch signifikant (RR [95%-KI]: 1,01 [0,84; 1,21]; $p=0,941$) (Tabelle 4-57).

Auch auf Ebene der Einzelstudien ergab sich kein statistisch signifikant häufigeres Auftreten von *SUE* in einer der Behandlungsgruppen.

Die entsprechenden Sensitivitätsanalysen für den Endpunkt *Gesamtzahl SUE ohne Berücksichtigung von adjudizierten Exazerbationen* der IPF-Erkrankung sind in Tabelle 4-58 dargestellt. Die Ergebnisse unterscheiden sich nicht relevant von den Ergebnissen der Hauptanalyse.

Tabelle 4-58: Sensitivitätsanalyse: Ergebnisse für die *Gesamtzahl SUE ohne Berücksichtigung von adjudizierten Exazerbationen* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Gesamtzahl SUE								
INPULSIS + Studie 1199.187 (gepoolt)	694	202 (29,1)	480	136 (28,3)	1,01 [0,84; 1,21]	1,01 [0,78; 1,31]	0,20 [-5,05; 5,45]	0,941
Gesamtzahl SUE ohne Berücksichtigung von adjudizierten Exazerbationen								
INPULSIS + Studie 1199.187 (gepoolt)	694	190 (27,4)	480	129 (26,9)	1,00 [0,83; 1,21]	1,00 [0,77; 1,30]	-0,04 [-5,20; 5,12]	0,987
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.13.1.2 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein UE zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis								

Die nachfolgende Tabelle zeigt eine Übersicht über *SUE*, die mit einer Häufigkeit von $\geq 1\%$ in mindestens einem Behandlungsarm auftraten nach MedDRA SOC und PT.

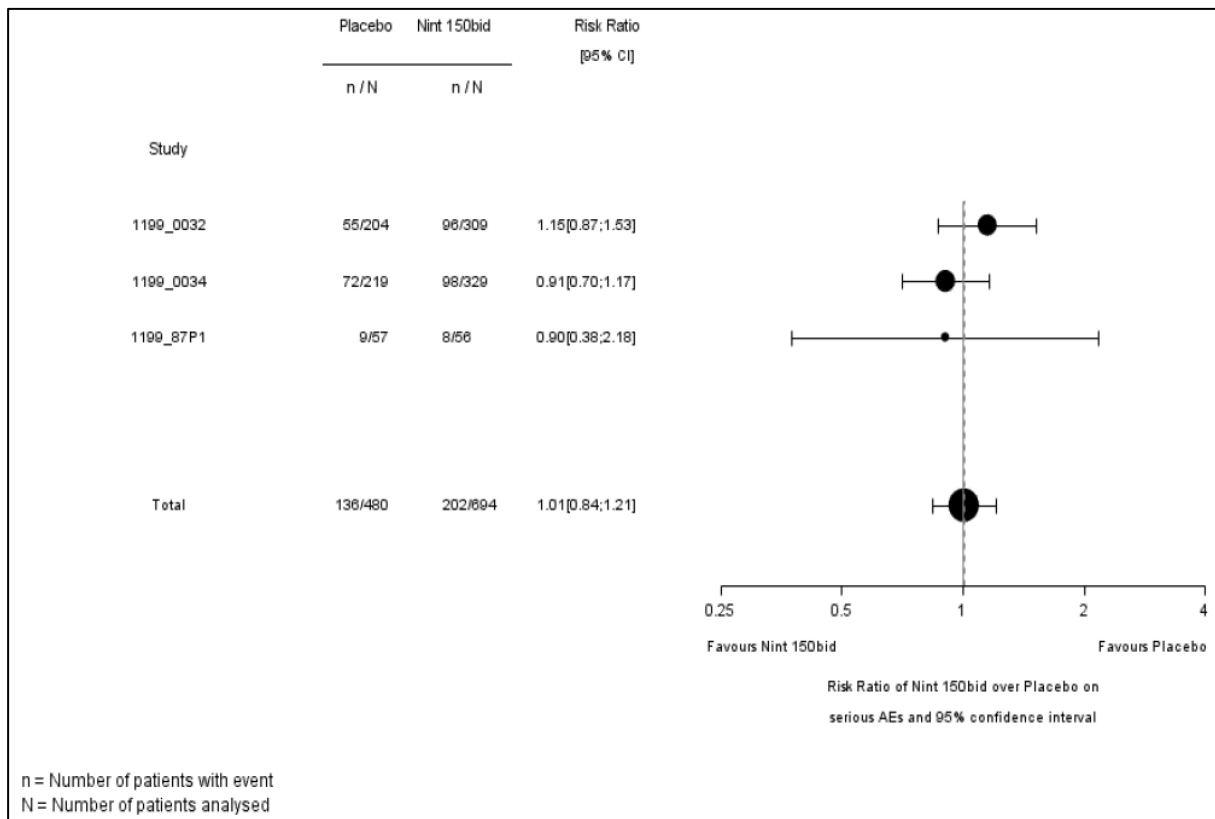
Tabelle 4-59: Häufige SUE ($\geq 1\%$ in mindestens einem Studienarm) nach MedDRA SOC und PT aus RCT mit dem zu bewertenden Arzneimittel – INPULSIS Studien und Studie 1199.187 (gepoolt)

MedDRA SOC Preferred Term	INPULSIS + Studie 1199.187 (gepoolt)	
	Nintedanib N=694	Placebo N=480
Gesamtzahl SUE, n (%)	202 (29,1)	136 (28,3)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	80 (11,5)	66 (13,8)
Dyspnoe	3 (0,4)	6 (1,3)
IPF	43 (6,2)	42 (8,8)
Pulmonale Hypertension	12 (1,7)	9 (1,9)
Respiratorisches Versagen	2 (0,3)	9 (1,9)
Infektionen und parasitäre Erkrankungen	56 (8,1)	38 (7,9)
Bronchitis	8 (1,2)	2 (0,4)
Infektionen der unteren Atemwege	1 (0,1)	5 (1,0)
Pneumonie	25 (3,6)	20 (4,2)
Herzerkrankungen	32 (4,6)	23 (4,8)
Myokardinfarkt	7 (1,0)	2 (0,4)
Neoplasmen (gutartig, bösartig und unspezifisch inkl. Zysten und Polypen)	20 (2,9)	16 (3,3)
Plattenepithelkarzinom	2 (0,3)	5 (1,0)
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.5.2.2 IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); MedDRA: Medical Dictionary for Regulatory Activities; PT: Preferred Term; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); SOC: System Organ Class; SUE: Schwerwiegendes Unerwünschtes Ereignis		

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.

Die Ergebnisse der IPD Meta-Analyse für den Endpunkt *Gesamtzahl SUE* wurde bereits in Tabelle 4-52 dargestellt.

Die grafische Darstellung der Meta-Analyse mittels Forest-Plot befindet sich in Abbildung 34. Der Forest-Plot zeigt die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, das Resultat der IPD Meta-Analyse. Der entsprechende Heterogenitätstest (siehe auch Abschnitt 4.3.1.3.5.4) weist auf eine homogene Datenlage hin (p-Wert des Interaktionstest [Merkmal Studie]: $p=0,443$).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 34: Forest-Plot zur Meta-Analyse für den Endpunkt *Gesamtzahl SUE* - RCT mit dem zu bewertenden Arzneimittel

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Die Erhebung der unerwünschten Ereignisse ist international standardisiert (International Conference on Harmonisation, 1996) und wird einheitlich dokumentiert. Von einer Übertragbarkeit der Studienergebnisse zum Endpunkt *SUE* auf den deutschen Versorgungskontext kann somit ausgegangen werden.

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.

4.3.1.3.4.3 Endpunkt UE, die zum Therapieabbruch führten

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

Tabelle 4-60: Operationalisierung von UE, die zum Therapieabbruch führten in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
INPULSIS 1	Definition: Jegliches berichtete UE, das zum permanenten Absetzen der Studienmedikation führte (keine Re-Exposition in voller oder reduzierter Dosierung).
INPULSIS 2	Homogenes Studiendesign zu INPULSIS 1
Studie 1199.187	Definition: Jegliches berichtete UE, das zum permanenten Absetzen der Studienmedikation führte (keine Re-Exposition in voller oder reduzierter Dosierung). Für die vorliegende Nutzenbewertung wurden die Analysen der unerwünschten Ereignisse der kontrollierten Studienphase (im Studienbericht wird diese Phase als P1 bezeichnet) berücksichtigt.
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis	

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 4-61: Bewertung des Verzerrungspotenzials für den Endpunkt UE, die zum Therapieabbruch führten, 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
INPULSIS 1	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für UE erfolgte in den INPULSIS Studien doppelblind. Für Studie 1199.187 wurde die kontrollierte Studienphase herangezogen.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen hatten. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien aller Endpunkte zu den unerwünschten Ereignissen waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Analyseplan ausgewertet und transparent und vollständig im Studienbericht berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war für den Endpunkt *UE, die zum Therapieabbruch führten* transparent.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für den Endpunkt *UE, die zum Therapieabbruch führten* ein niedriges Verzerrungspotenzial.

Stellen Sie die Ergebnisse für den Endpunkt UE, die zum Therapieabbruch führten, für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Der Endpunkt *UE, die zum Therapieabbruch führten* wurde in den drei berichteten Studien erhoben und erfasst jegliches UE, welches zum permanenten Absetzen der Studienmedikation führte. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-62: Ergebnisse für *UE, die zum Therapieabbruch führten* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
UE, die zum Therapieabbruch führten								
INPULSIS + Studie 1199.187 (gepoolt)	694	131 (18,9)	480	58 (12,1)	1,54 [1,15; 2,04]	1,66 [1,19; 2,32]	6,55 [2,41; 10,69]	0,003
INPULSIS (gepoolt)	638	123 (19,3)	423	55 (13,0)	1,48 [1,11; 1,99]	1,60 [1,13; 2,26]	6,28 [1,84; 10,71]	0,007
INPULSIS 1	309	65 (21,0)	204	22 (10,8)	1,95 [1,24; 3,06]	2,20 [1,31; 3,71]	10,25 [4,02; 16,48]	0,002
INPULSIS 2	329	58 (17,6)	219	33 (15,1)	1,17 [0,79; 1,73]	1,21 [0,76; 1,92]	2,56 [-3,72; 8,84]	0,430
Studie 1199.187	56	8 (14,3)	57	3 (5,3)	2,71 [0,76; 9,71]	3,00 [0,75; 11,96]	9,02 [-1,82; 19,87]	0,106
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.5.3.1, Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 1.3.49.2.6.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

UE, die zum Therapieabbruch führten traten in der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 bei 18,9% der Studienteilnehmer, die mit Nintedanib behandelt wurden, auf, verglichen mit 12,1% der Studienteilnehmer im Kontrollarm. Der Unterschied zwischen den Gruppen ist statistisch signifikant (RR [95%-KI]: 1,54 [1,15; 2,04]; p=0,003). Dies traf auch für die Analyse der gepoolten INPULSIS Studien zu. Somit brachen in der Nintedanib-Gruppe häufiger Studienteilnehmer die Behandlung aufgrund eines UE ab, als in der Kontrollgruppe (Tabelle 4-62).

Auf Ebene der Einzelstudien ergab sich ein statistisch signifikant häufigerer Therapieabbruch aufgrund von UE in der Studie INPULSIS 1.

Die entsprechenden Sensitivitätsanalysen für den Endpunkt *UE, die zum Therapieabbruch führten* ohne Berücksichtigung von adjudizierten Exazerbationen der IPF-Erkrankung sind in Tabelle 4-63 dargestellt. Die Ergebnisse unterscheiden sich nicht relevant von den Ergebnissen der Hauptanalyse.

Tabelle 4-63: Sensitivitätsanalyse: Ergebnisse für *UE, die zum Therapieabbruch führten* ohne Berücksichtigung von adjudizierten Exazerbationen aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
UE, die zum Therapieabbruch führten								
INPULSIS + Studie 1199.187 (gepoolt)	694	131 (18,9)	480	58 (12,1)	1,54 [1,15; 2,04]	1,66 [1,19; 2,32]	6,55 [2,41; 10,69]	0,003
UE, die zum Therapieabbruch führten ohne Berücksichtigung von adjudizierten Exazerbationen								
INPULSIS + Studie 1199.187 (gepoolt)	694	126 (18,2)	480	49 (10,2)	1,74 [1,28; 2,37]	1,91 [1,34; 2,72]	7,68 [3,73; 11,64]	<0,001
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.13.1.3 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko, eine RD größer null ein höheres Risiko, ein UE zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

Tabelle 4-64 zeigt eine Übersicht der *UE, die zum Therapieabbruch führten* und die mit einer Häufigkeit von $\geq 1\%$ in mindestens einem Behandlungsarm auftraten.

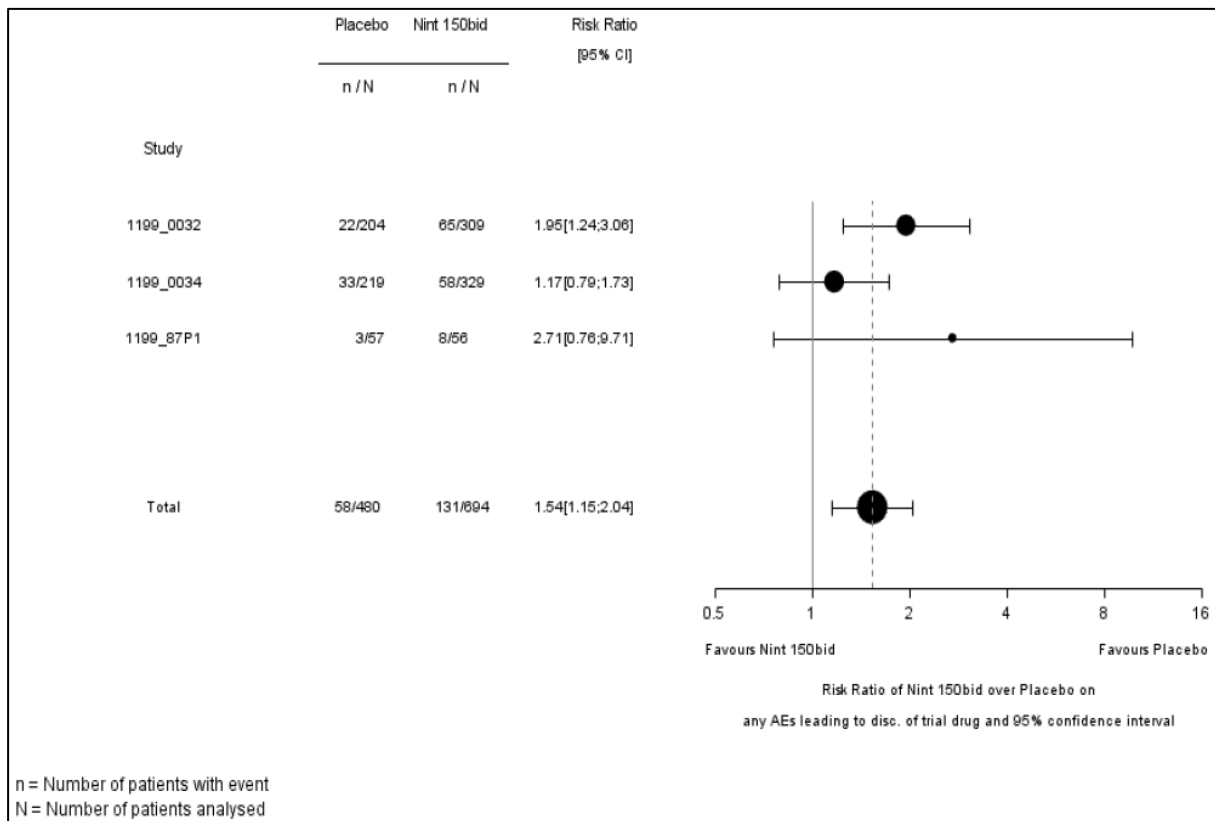
Tabelle 4-64: Ergebnisse für *UE*, die zum Therapieabbruch führten mit Inzidenz $\geq 1\%$ aus RCT mit dem zu bewertenden Arzneimittel - INPULSIS Studien und Studie 1199.187 (gepoolt)

MedDRA SOC Preferred Term	INPULSIS + Studie 1199.187 (gepoolt)	
	Nintedanib N=694	Placebo N=480
Gesamtzahl Therapieabbrüche aufgrund UE, n (%)	131 (18,9)	58 (12,1)
Erkrankungen des Gastrointestinaltrakts	53 (7,6)	5 (1,0)
Diarrhö	32 (4,6)	1 (0,2)
Übelkeit	14 (2,0)	0
Erkrankungen der Atemwege, des Brustraums und Mediastinums	21 (3,0)	31 (6,5)
IPF	14 (2,0)	24 (5,0)
Stoffwechsel- und Ernährungsstörungen	9 (1,3)	2 (0,4)
Appetit vermindert	9 (1,3)	1 (0,2)
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.5.3.2 IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); MedDRA: Medical Dictionary for Regulatory Activities; SOC: System Organ Class; PT: Preferred Term; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis		

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.

Die Ergebnisse der IPD Meta-Analyse für den Endpunkt *UE*, die zum Therapieabbruch führten wurden bereits in Tabelle 4-64 dargestellt.

Die grafische Darstellung der Meta-Analyse mittels Forest-Plot befindet sich in Abbildung 35. Der Forest-Plot zeigt die Effektschätzer und zugehörigen KI der Einzelstudien sowie, als zusammenfassenden Schätzer, das Resultat der IPD Meta-Analyse. Der Heterogenitätstest (siehe auch Abschnitt 4.3.1.3.5.1) weist auf eine homogene Datenlage hin (p-Wert der Interaktionstests [Merkmal Studie]: p=0,159).



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a)

Abbildung 35: Forest-Plot zur Meta-Analyse für den Endpunkt *UE, die zum Therapieabbruch führten* - RCT mit dem zu bewertenden Arzneimittel

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Die Erhebung der unerwünschten Ereignisse ist international standardisiert (International Conference on Harmonisation, 1996) und wird einheitlich dokumentiert. Von einer Übertragbarkeit der Studienergebnisse zum Endpunkt *UE, die zum Therapieabbruch führten* auf den deutschen Versorgungskontext kann somit ausgegangen werden.

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.

4.3.1.3.4.4 Endpunkt UE von besonderem Interesse

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

Tabelle 4-65: Operationalisierung von UE von besonderem Interesse

Studie	Operationalisierung
INPULSIS 1	Definition: Als UE von besonderem Interesse wurden gastrointestinale Störungen (Diarrhö, Übelkeit, Erbrechen, Appetitlosigkeit, Gewichtsverlust, gastrointestinale Perforation), Leber- und Gallenerkrankungen (Leberenzymerrhöhung, Bilirubinerhöhung), kardiale Ereignisse (darunter Myokardinfarkt), arterielle thromboembolische Ereignisse, respiratorische Ereignisse (Progression der IPF, Husten, Dyspnoe, Bronchitis, Nasopharyngitis), alle weiteren Störungen, die in mindestens einem Studienarm mit einer Häufigkeit von >10% der Studienteilnehmer auftraten, sowie Blutungsereignisse betrachtet. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme.
INPULSIS 2	Homogenes Studiendesign zu INPULSIS 1
Studie 1199.187	Definition: Als UE von besonderem Interesse wurden gastrointestinale Störungen (Diarrhö, Übelkeit, Erbrechen, Appetitlosigkeit, Gewichtsverlust, gastrointestinale Perforation), Leber- und Gallenerkrankungen (Leberenzymerrhöhung, Bilirubinerhöhung), kardiale Ereignisse (darunter Myokardinfarkt), arterielle thromboembolische Ereignisse, respiratorische Ereignisse (Progression der IPF, Husten, Dyspnoe, Bronchitis, Nasopharyngitis), alle weiteren Störungen, die in mindestens einem Studienarm mit einer Häufigkeit von >10% der Studienteilnehmer auftraten, sowie Blutungsereignisse betrachtet. Ereignisse wurden registriert von Studienbeginn bis 28 Tage nach der letzten Arzneimittleinnahme. Für die vorliegende Nutzenbewertung wurden die Analysen der unerwünschten Ereignisse der kontrollierten Studienphase (im Studienbericht wird diese Phase als P1 bezeichnet) berücksichtigt.
IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); UE: Unerwünschtes Ereignis	

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 4-66: Bewertung des Verzerrungspotenzials für den Endpunkt *UE von besonderem Interesse* 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
INPULSIS 1	niedrig	ja	ja	ja	ja	niedrig
INPULSIS 2	niedrig	ja	ja	ja	ja	niedrig
Studie 1199.187	niedrig	ja	ja	ja	ja	niedrig
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

- Verblindung Endpunkterheber

Die Datenerhebung für UE erfolgte in den INPULSIS Studien doppelblind. Für Studie 1199.187 wurde die kontrollierte Studienphase herangezogen.

- Adäquate Umsetzung des ITT-Prinzips

Die Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187 erfolgte jeweils auf Basis des Treated Set, definiert als randomisierte Studienteilnehmer, die mindestens eine Dosis der Studienmedikation eingenommen hatten. Das Treated Set wich lediglich um zwei Studienteilnehmer in der INPULSIS 1 Studie und drei Studienteilnehmer in der INPULSIS 2 Studie, die keine Behandlung erhielten, von allen randomisierten Studienteilnehmern ab. Damit ist die Abweichung (<1% der randomisierten Population) als nicht ergebnisrelevant einzuschätzen. Die Auswertung erfolgte in der durch Randomisierung zugeordneten Behandlungsgruppe und erfüllte daher das ITT-Prinzip.

- Ergebnisunabhängige Berichterstattung

Die Auswertungskriterien aller Endpunkte zu den unerwünschten Ereignissen waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Analyseplan ausgewertet und transparent und vollständig berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

- Sonstige Aspekte

Die Beschreibung der Methodik und Ergebnisse von Analysen im Studienbericht, SAP und CTP war für den Endpunkt *UE von besonderem Interesse* transparent.

- Verzerrungspotenzial auf Endpunktebene

Abschließend ergibt sich für den Endpunkt *UE von besonderem Interesse* ein niedriges Verzerrungspotenzial.

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

Als *UE von besonderem Interesse* wurden *gastrointestinale Störungen (Diarrhö, Übelkeit, Erbrechen, Appetitlosigkeit, Gewichtsverlust, gastrointestinale Perforation), Leber- und Gallenerkrankungen (Leberenzymerrhöhung, Bilirubinerhöhung), kardiale Ereignisse (darunter Myokardinfarkt), respiratorische Ereignisse (Progression der IPF, Husten, Dyspnoe, Bronchitis, Nasopharyngitis), Blutungsereignisse* sowie alle weiteren Störungen, die in mindestens einem Studienarm mit einer Häufigkeit von >10% der Studienteilnehmer auftraten, betrachtet.

Unerwünschte gastrointestinale Ereignisse

Der Endpunkt *UE von besonderem Interesse (gastrointestinale Ereignisse)* wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimiteleinnahe erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-67: Ergebnisse für UE – gastrointestinale Ereignisse aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS + Studie 1199.187 (gepoolt)	Diarrhö (Gesamtzahl)							
	694	433 (62,4)	480	99 (20,6)	3,09 [2,56; 3,72]	6,70 [5,09; 8,81]	42,30 [37,24; 47,37]	<0,001
	Diarrhö (Intensität leicht^d)^e							
	433	249 (57,5)	99	77 (77,8)	0,75 [0,66; 0,86]	0,41 [0,24; 0,68]	-19,02 [-28,57; -9,47]	<0,001
	Diarrhö (Intensität moderat^d)^e							
	433	162 (37,4)	99	20 (20,2)	1,82 [1,20; 2,75]	2,30 [1,35; 3,91]	16,51 [7,27; 25,75]	0,002
	Diarrhö (Intensität schwer^d)^e							
	433	21 (4,8)	99	2 (2,0)	2,08 [0,50; 8,68]	2,14 [0,49; 9,27]	2,29 [-1,19; 5,78]	0,299
	Diarrhö (angegeben als SUE)^e							
	694	2 (0,3)	480	1 (0,2)	1,33 [0,12; 14,63]	1,33 [0,12; 14,72]	0,08 [-0,56; 0,71]	0,816
	Diarrhö (angegeben als UE, die zum Therapieabbruch führten)^e							
	694	32 (4,6)	480	1 (0,2)	21,95 [3,09; 156,12]	22,88 [3,16; 165,65]	4,45 [2,83; 6,08]	<0,001
	Übelkeit							
	694	172 (24,8)	480	41 [8,5]	3,02 [2,18; 4,18]	3,72 [2,57; 5,39]	16,63 [12,59; 20,67]	<0,001
	Erbrechen							
	694	83 (12,0)	480	14 (2,9)	4,20 [2,40; 7,35]	4,65 [2,59; 8,32]	9,18 [6,33; 12,03]	<0,001
	Appetitlosigkeit							
	694	80 (11,5)	480	28 (5,8)	2,02 [1,34; 3,05]	2,15 [1,38; 3,37]	5,93 [2,74; 9,12]	<0,001
	Gewichtsverlust							
	694	72 (10,4)	480	15 (3,1)	3,27 [1,91; 5,61]	3,49 [1,99; 6,12]	7,27 [4,48; 10,05]	<0,001
	Gastrointestinale Perforation							
	694	2 (0,3)	480	0	nicht berechenbar	nicht berechenbar	0,31 [-0,12; 0,75]	0,250

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS (gepoolt)	Diarrhö (Gesamtzahl)							
	638	398 (62,4)	423	78 (18,4)	3,38 [2,74; 4,17]	7,33 [5,47; 9,84]	43,94 [38,67; 49,21]	<0,001
	Diarrhö (Intensität leicht^d)							
	638	226 (35,4)	423	60 (14,2)	2,50 [1,93; 3,23]	3,32 [2,42; 4,56]	21,24 [16,26; 26,22]	<0,001
	Diarrhö (Intensität moderat^d)							
	638	150 (23,5)	423	16 (3,8)	6,22 [3,77; 10,3]	7,82 [4,59; 13,3]	19,73 [15,97; 23,49]	<0,001
	Diarrhö (Intensität schwer^d)							
	638	21 (3,3)	423	2 (0,5)	6,96 [1,64; 29,5]	7,16 [1,67; 30,7]	2,82 [1,29; 4,35]	0,002
	Diarrhö (angegeben als SUE)							
	638	2 (0,3)	423	1 (0,2)	1,33 [0,12; 14,6]	1,33 [0,12; 14,7]	0,08 [-0,56; 0,71]	0,817
	Diarrhö (angegeben als UE, die zum Therapieabbruch führten)							
	638	28 (4,4)	423	1 (0,2)	18,6 [2,54; 136]	19,4 [2,63; 143]	4,15 [2,50; 5,81]	<0,001
	Übelkeit							
	638	156 (24,5)	423	28 (6,6)	3,69 [2,52; 5,42]	4,57 [2,99; 6,97]	17,83 [13,74; 21,92]	<0,001
	Erbrechen							
	638	74 (11,6)	423	11 (2,6)	4,46 [2,40; 8,30]	4,91 [2,58; 9,38]	9,00 [6,09; 11,91]	<0,001
	Appetitlosigkeit							
	638	68 (10,7)	423	24 (5,7)	1,88 [1,20; 2,94]	1,98 [1,22; 3,21]	4,98 [1,73; 8,24]	0,005
	Gewichtsverlust							
	638	62 (9,7)	423	15 (3,5)	2,74 [1,58; 4,75]	2,93 [1,64; 5,22]	6,17 [3,28; 9,07]	<0,001
	Gastrointestinale Perforation							
	638	2 (0,3)	423	0 (0)	nicht berechenbar	nicht berechenbar	0,31 [-0,12; 0,75]	0,2491

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS 1	Diarrhö (Gesamtzahl)							
	309	190 (61,5)	204	38 (18,6)	3,30 [2,45; 4,46]	6,97 [4,58; 10,6]	42,86 [35,25; 50,48]	<0,001
	Diarrhö (Intensität leicht^d)							
	309	103 (33,3)	204	29 (14,2)	2,34 [1,62; 3,40]	3,02 [1,91; 4,77]	19,12 [12,00; 26,23]	<0,001
	Diarrhö (Intensität moderat^d)							
	309	75 (24,3)	204	9 (4,4)	5,50 [2,82; 10,70]	6,94 [3,39; 14,20]	19,86 [14,31; 25,41]	<0,001
	Diarrhö (Intensität schwer^d)							
	309	11 (3,6)	204	0 (0,0)	nicht berechenbar	nicht berechenbar	3,56 [1,49; 5,63]	0,006
	Diarrhö (angegeben als SUE)							
	309	1 (0,3)	204	0 (0,0)	nicht berechenbar	nicht berechenbar	0,32 [-0,31; 0,96]	0,416
	Diarrhö (angegeben als UE, das zum Therapieabbruch führte)							
	309	14 (4,5)	204	0 (0,0)	nicht berechenbar	nicht berechenbar	4,53 [2,21; 6,85]	0,002
	Übelkeit							
	309	70 (22,7)	204	12 (5,9)	3,85 [2,14; 6,92]	4,69 [2,47; 8,90]	16,77 [11,10; 22,45]	<0,001
	Erbrechen							
	309	40 (12,9)	204	4 (2,0)	6,60 [2,40; 18,2]	7,43 [2,62; 21,1]	10,98 [6,79; 15,18]	<0,001
	Appetitlosigkeit							
	309	26 (8,4)	204	14 (6,9)	1,23 [0,66; 2,29]	1,25 [0,63; 2,45]	1,55 [-3,10; 6,20]	0,521
	Gewichtsverlust							
	309	25 (8,1)	204	13 (6,4)	1,27 [0,67; 2,42]	1,29 [0,65; 2,59]	1,72 [-2,81; 6,24]	0,467

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS 2	Diarrhö (Gesamtzahl)							
	329	208 (63,2)	219	40 (18,3)	3,46 [2,58; 4,64]	7,69 [5,11; 11,60]	44,96 [37,65; 52,26]	<0,001
	Diarrhö (Intensität leicht^d)							
	329	123 (37,4)	219	31 (14,2)	2,64 [1,85; 3,77]	3,62 [2,33; 5,63]	23,23 [16,26; 30,21]	<0,001
	Diarrhö (Intensität moderat^d)							
	329	75 (22,8)	219	7 (3,2)	7,13 [3,35; 15,20]	8,94 [4,04; 19,80]	19,60 [14,50; 24,70]	<0,001
	Diarrhö (Intensität schwer^d)							
	329	10 (3,0)	219	2 (0,9)	3,33 [0,74; 15,00]	3,40 [0,74; 15,70]	2,13 [-0,12; 4,37]	0,096
	Diarrhö (angegeben als SUE)							
	329	1 (0,3)	219	1 (0,5)	0,67 [0,04; 10,60]	0,66 [0,04; 10,70]	-0,15 [-1,23; 0,92]	0,772
	Diarrhö (angegeben als UE, das zum Therapieabbruch führte)							
	329	14 (4,3)	219	1 (0,5)	9,32 [1,23; 70,40]	9,69 [1,26; 74,20]	3,80 [1,44; 6,16]	0,008
	Übelkeit							
	329	86 (26,1)	219	16 (7,3)	3,58 [2,16; 5,93]	4,49 [2,55; 7,90]	18,83 [12,97; 24,70]	<0,001
	Erbrechen							
	329	34 (10,3)	219	7 (3,2)	3,23 [1,46; 7,16]	3,49 [1,52; 8,02]	7,14 [3,11; 11,17]	0,002
	Appetitlosigkeit							
	329	42 (12,8)	219	10 (4,6)	2,80 [1,43; 5,45]	3,06 [1,50; 6,24]	8,20 [3,66; 12,74]	0,001
	Gewichtsverlust							
	329	37 (11,2)	219	2 (0,9)	12,3 [3,00; 50,6]	13,7 [3,28; 57,70]	10,33 [6,69; 13,97]	<0,001

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Studie 1199.187	Diarrhö (Gesamtzahl)							
	56	40 (71,4)	57	21 (36,8)	1,94 [1,33; 2,83]	4,29 [1,94; 9,45]	34,59 [17,36; 51,81]	<0,001
	Diarrhö (Intensität leicht^d)							
	56	28 (50,0)	57	17 (29,8)	1,68 [1,04; 2,70]	2,35 [1,09; 5,09]	20,18 [2,50; 37,85]	0,028
	Diarrhö (Intensität moderat^d)							
	56	12 (21,4)	57	4 (7,0)	3,05 [1,05; 8,90]	3,61 [1,09; 12,00]	14,41 [1,78; 27,04]	0,028
	Diarrhö (Intensität schwer^d)							
	56	0	57	0	nicht berechenbar	nicht berechenbar	nicht berechenbar	nicht berech enbar
	Diarrhö (angegeben als SUE)							
	56	0	57	0	nicht berechenbar	nicht berechenbar	nicht berechenbar	nicht berech enbar
	Diarrhö (angegeben als UE, das zum Therapieabbruch führte)							
	56	4 (7,1)	57	0 (0,0)	nicht berechenbar	nicht berechenbar	7,14 [0,40; 13,89]	0,040
	Übelkeit							
	56	16 (28,6)	57	13 (22,8)	1,25 [0,67; 2,36]	1,35 [0,58; 3,16]	5,76 [-10,32; 21,85]	0,483
	Erbrechen							
	56	9 (16,1)	57	3 (5,3)	3,05 [0,87; 10,70]	3,45 [0,88; 13,48]	10,81 [-0,42; 22,04]	0,062
	Appetitlosigkeit							
	56	12 (21,4)	57	4 (7,0)	3,05 [1,05; 8,90]	3,61 [1,09; 12,00]	14,41 [1,78; 27,04]	0,028
	Gewichtsverlust							
	56	8 (14,3)	57	0 (0,0)	nicht berechnet	nicht berechnet	14,29 [5,12; 23,45]	0,003

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a) Tab. 1.5.5.5.1, Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 1.5.5.1.3, Tab. 1.5.5.3.1, Tab. 1.5.5.1.1, Tab. Additional AEs, Tab. 1.3.49.2.6.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. d: Für ein Ereignis in der Nintedanib-Gruppe lag keine Bewertung der Intensität vor. Intensität leicht: Anzeichen oder Symptome sind spürbar, aber leicht zu tolerieren. Intensität moderat: Anzeichen oder Symptome führen zu Beeinträchtigungen im Alltag, ziehen aber keine Arbeitsunfähigkeit oder Einschränkung von Alltagsaktivitäten nach sich. Intensität schwer: Beeinträchtigungen führen zu Fatigue und Arbeitsunfähigkeit bzw. Einschränkung der täglichen Aktivitäten. e: Prozentuierung basiert auf der Zahl der Patienten mit mindestens einem Ereignis; der Effektschätzer auf der Gesamtzahl aller Patienten im Treated Set p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. AE: Adverse Event (Unerwünschtes Ereignis); KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis								

Unter den *gastrointestinalen Störungen* waren die Ereignisse *Diarrhö*, *Übelkeit*, *Erbrechen*, *Appetitlosigkeit*, *Gewichtsverlust* und *gastrointestinale Perforation* von besonderem Interesse. *Diarrhö* kam in der gepoolten Analyse der INPULSIS Studien und der Studie 1199.187 am häufigsten vor (62,4% in der Nintedanib-Gruppe vs. 20,6% in der Kontrollgruppe). Der Unterschied ist statistisch signifikant (RR [95%-KI]: 3,09 [2,56; 3,72]; $p < 0,001$) (Tabelle 4-67).

Diarrhöen traten erstmalig überwiegend in den ersten drei Monaten der Behandlung auf und waren bei den meisten Studienteilnehmern von leichter bis moderater Intensität. Zu einer Einschränkung in der Ausübung von Alltagsaktivitäten oder zu Arbeitsunfähigkeit (Intensität „schwer“) kam es bei 4,8% der Nintedanib-Studienteilnehmer mit *Diarrhö* (entsprechend 3,0% aller Studienteilnehmer der Nintedanib-Gruppe). Das CTP der INPULSIS Studien und der Studie 1199.187 gab Empfehlungen zum Management von *Diarrhöen*, die eine pharmakologische Intervention (z. B. mit dem Antidiarrhoikum Loperamid) und bei Persistenz der Beschwerden eine Dosisreduktion oder Unterbrechung der Behandlung vorsahen. Nach Abklingen der Beschwerden konnte die Arzneimittelgabe in normaler oder reduzierter Dosis wieder aufgenommen werden. Auf diese Weise ließ sich das *UE Diarrhö* bei den meisten Studienteilnehmern der Nintedanib-Gruppe gut tolerieren: Über alle drei eingeschlossenen Studien hinweg konnten 78,5% der Studienteilnehmer mit *Diarrhö* die Therapie mit der Normaldosis fortsetzen, weitere 14,1% der Studienteilnehmer mit reduzierter Dosis. Lediglich bei 7,4% der Studienteilnehmer mit *Diarrhö* (entsprechend 4,6% aller Studienteilnehmer der Nintedanib-Gruppe) kam es aufgrund dieses *UE* zu einem endgültigen Therapieabbruch

(Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d; Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

Übelkeit wurde in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 im Nintedanib-Arm (24,8%) häufiger dokumentiert als im Kontrollarm (8,5%). Der Unterschied ist statistisch signifikant: (RR [95%-KI]: 3,02 [2,18; 4,18]; $p < 0,001$). Auch *Erbrechen* wurde im Nintedanib-Arm (12,0%) häufiger berichtet als im Vergleichsarm (2,9%). Der beobachtete Effekt ist statistisch signifikant (RR [95%-KI]: 4,20 [2,40; 7,35]; $p < 0,001$) (Tabelle 4-67).

Die Fachinformation für das Produkt Ofev[®] gibt Empfehlungen zu Maßnahmen beim Auftreten von *Diarrhöen*, *Übelkeit* und *Erbrechen*, die auf dem Vorgehen in den beiden Phase III-Studien beruhen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Appetitlosigkeit und *Gewichtsverlust* werden gemäß den beim Erfassen von *UE* in den INPULSIS Studien zugrunde gelegten „preferred terms“ des MedDRA formal nicht der Kategorie der gastrointestinalen Ereignisse zugerechnet, könnten jedoch durch gastrointestinale Beschwerden beeinflusst sein und werden daher hier dargestellt. *Appetitlosigkeit* wurde in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 im Nintedanib-Behandlungsarm (11,5%) häufiger dokumentiert als im Vergleichsarm (5,8%). Der Unterschied ist statistisch signifikant: (RR [95%-KI]: 2,02 [1,34; 3,05]; $p < 0,001$). *Gewichtsverlust* trat im Nintedanib-Behandlungsarm (10,4%) häufiger auf als im Vergleichsarm (3,1%). Der beobachtete Effekt ist statistisch signifikant (RR [95%-KI]: 3,27 [1,91; 5,61]; $p < 0,001$).

Bedingt durch den Wirkmechanismus von Nintedanib könnten Patienten ein erhöhtes Risiko für eine *gastrointestinale Perforation* haben. In der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 traten *gastrointestinale Perforationen* sehr selten auf (0% in der Kontrollgruppe versus 0,3% in der Nintedanib-Gruppe), sodass keine Aussage zum Risiko einer *gastrointestinalen Perforation* unter einer Behandlung mit Nintedanib getroffen werden kann. Es wurde ein entsprechender Warnhinweis in die Fachinformation von Ofev[®] aufgenommen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Zusammenfassend traten die *gastrointestinalen UE Diarrhö*, *Übelkeit* und *Erbrechen* sowie die weiteren *UE Appetitlosigkeit* und *Gewichtsverlust* unter der Behandlung mit Nintedanib signifikant häufiger auf als in der Kontrollgruppe. Häufigstes gastrointestinales Ereignis war die *Diarrhö*. Die Mehrzahl der *Diarrhöen* war von leichter bis moderater Intensität und ließ sich in der Regel mithilfe der Empfehlungen im Prüfplan zum Vorgehen beim Auftreten von *Diarrhö* kontrollieren.

Unerwünschte Leberenzym-und Bilirubinveränderungen

Der Endpunkt *UE von besonderem Interesse (Leber- und Gallenerkrankungen)* wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimittelaufnahme erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch

auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-68: Ergebnisse für *UE – Leber- und Gallenerkrankungen* (Leberenzym-, Bilirubin-, AP-Erhöhung) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS + Studie 1199.187 (gepoolt)	Leberenzymerrhöhung (AST, ALT)							
	≥3x ULN ^{d,e}							
	694	34 (4,9)	480	4 (0,8)	5,99 [2,10; 17,07]	6,24 [2,18; 17,87]	4,06 [2,26; 5,85]	<0,001
	≥5x ULN ^{d,e}							
	694	14 (2,0)	480	1 (0,2)	9,30 [1,23; 70,53]	9,51 [1,24; 72,75]	1,76 [0,65; 2,88]	0,008
	≥8x ULN ^d							
	694	5 (0,7)	480	1 (0,2)	3,32 [0,39; 28,36]	3,34 [0,39; 28,78]	0,49 [-0,25; 1,24]	0,244
	Bilirubinerhöhung							
	≥1,5x ULN ^{d,e}							
	694	16 (2,3)	480	3 (0,6)	3,60 [1,06; 12,17]	3,67 [1,07; 12,61]	1,66 [0,33; 2,98]	0,027
	≥2x ULN ^d							
	694	3 (0,4)	480	2 (0,4)	1,00 [0,17; 5,94]	1,00 [0,17; 6,00]	-0,00 [-0,76; 0,76]	0,997
	AP-Erhöhung							
	≥1,5x ULN ^{d,e}							
	694	39 (5,6)	480	5 (1,0)	5,44 [2,13; 13,87]	5,72 [2,22; 14,75]	4,55 [2,62; 6,49]	<0,001
	≥2x ULN ^d							
	694	17 (2,4)	480	1 (0,2)	11,29 [1,51; 84,64]	11,58 [1,53; 87,50]	2,19 [0,97; 3,40]	0,003

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p- Wert
INPULSIS (gepoolt)*	Leberenzymerrhöhung (AST, ALT)							
	≥3x ULN ^{d,e}							
	638	32 (5,0)	423	3 (0,7)	7,07 [2,18; 22,9]	7,39 [2,25; 24,3]	4,31 [2,43; 6,18]	<0,001
	≥5x ULN ^{d,e}							
	638	14 (2,2)	423	1 (0,2)	9,28 [1,23; 70,3]	9,47 [1,24; 72,3]	1,96 [0,73; 3,19]	0,008
	≥8x ULN ^d							
	638	5 (0,8)	423	1 (0,2)	3,32 [0,39; 28,3]	3,33 [0,39; 28,6]	0,55 [-0,28; 1,37]	0,244
	Bilirubinerhöhung							
	≥1,5x ULN ^{d,e}							
	638	15 (2,4)	423	3 (0,7)	3,32 [0,97; 11,4]	3,37 [0,97; 11,7]	1,64 [0,22; 3,06]	0,043
	≥2x ULN ^d							
	638	3 (0,5)	423	2 (0,5)	0,99 [0,17; 5,93]	0,99 [0,17; 5,98]	-0,00 [-0,84; 0,84]	0,995
	AP-Erhöhung							
	≥1,5x ULN ^{d,e}							
	638	37 (5,8)	423	4 (0,9)	6,13 [2,20; 17,1]	6,45 [2,28; 18,2]	4,85 [2,82; 6,89]	<0,001
	≥2x ULN ^d							
	638	17 (2,7)	423	1 (0,2)	11,3 [1,51; 84,4]	11,6 [1,53; 87,1]	2,43 [1,10; 3,76]	0,003
Studie 1199.187	Leberenzymerrhöhung (AST, ALT)							
	≥3x ULN ^{d,e}							
	56	2 (3,6)	57	1 (1,8)	2,04 [0,19; 21,82]	2,07 [0,18; 23,54]	1,82 [-4,12; 7,75]	0,548
	≥5x ULN ^{d,e}							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.
	≥8x ULN ^d							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p- Wert
	Bilirubinerhöhung							
	≥1,5x ULN ^{d,e}							
	56	1 (1,8)	57	0 (0,0)	n. b.	n. b.	1,79 [-1,68; 5,25]	0,311
	≥2x ULN ^d							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.
	AP-Erhöhung							
	≥1,5x ULN ^{d,e}							
	56	2 (3,6)	57	1 (1,8)	2,04 [0,19; 21,82]	2,07 [0,18; 23,54]	1,82 [-4,12; 7,75]	0,548
	≥2x ULN ^d							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 23.1.1, Tab. 1.3.49.2.6.1, Tab. 1.5.5.6.1 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. d: ULN: Upper Limits of Normal (Oberer Normgrenzwert). e: Kategorien sind kumulativ, d. h. Studienteilnehmer, die in die höheren Kategorien fallen, sind bei den niedrigeren Kategorien mit angegeben. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. *Aufgrund zu geringer Zellenbesetzung werden die Ergebnisse auf Einzelstudienbene im Dossier nicht dargestellt, sondern lediglich in der Anlage für Modul 5 geliefert. Grundsätzlich zeigen beide Studien ähnliche Ereignisraten in den jeweiligen Behandlungsarmen. ALT: Alanin-Aminotransferase; AP: Alkalische Phosphatase; AST: Aspartat-Aminotransferase; KI: Konfidenzintervall; n. b.: Nicht berichtet; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis; ULN: Upper Limits of Normal (Oberer Normgrenzwert)								

UE, welche das Leber- und Gallenwegssystem betrafen, wurden in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 häufiger unter Nintedanib als unter der zVT dokumentiert: Die Behandlung mit Nintedanib war im Vergleich zur zVT häufiger verbunden mit einer Erhöhung der Leberenzyme (ALT [Alanin-Aminotransferase], AST [Aspartat-Aminotransferase], AP [Alkalische Phosphatase]) oder des Bilirubin-Wertes (Tabelle 4-68).

Erhöhungen der AST- und ALT-Werte um mindestens 3x über dem oberen Normgrenzwert (Upper Limits of Normal, ULN) wurden bei 4,9% der Nintedanib-behandelten Studienteilnehmer (gegenüber 0,8% in der Kontrollgruppe) registriert (RR [95%-KI]: 5,99 [2,10; 17,07]; p<0,001). Bei 2% (versus 0,2%) erreichte die Erhöhung mindestens

5x ULN (RR [95%-KI]: 9,30 [1,23; 70,53]; $p=0,007$). Eine Erhöhung um mindestens 8x ULN trat unter Nintedanib (0,7%) jedoch nicht signifikant häufiger auf als in der Kontrollgruppe (0,2%) (RR [95%-KI]: 3,32 [0,39; 28,36]; $p=0,226$).

Eine *Erhöhung des Bilirubinwertes* um mindestens 1,5x ULN fand sich in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 2,3% der Nintedanib-Studienteilnehmer (verglichen mit 0,6% der Studienteilnehmer des Vergleichsarms) (RR [95%-KI]: 3,60 [1,06; 12,17]; $p=0,025$). Eine *Erhöhung des Bilirubinwertes* um mindestens 2x ULN trat in der Nintedanib-Gruppe jedoch nicht häufiger auf als in der Kontrollgruppe. Eine *Erhöhung der AP* um mindestens 1,5x ULN wurde in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 5,6% der Nintedanib-Studienteilnehmer (verglichen mit 1,0% der Studienteilnehmer des Vergleichsarms) dokumentiert (RR [95%-KI]: 5,44 [2,13; 13,87]; $p<0,001$). Eine *Erhöhung der AP* um mindestens 2,0x ULN wurde in der Nintedanib-Gruppe (2,4%) ebenfalls häufiger registriert als in der Kontrollgruppe (0,2%) (RR [95%-KI]: 11,29 [1,51; 84,64]; $p=0,002$).

UE aus der Gruppe der *Leber- und Gallenerkrankungen* waren überwiegend nicht-schwerwiegend. Als *SUE* wurden *Leber- und Gallenerkrankungen* bei sieben Studienteilnehmern der Nintedanib-Gruppe (1,0% aller Studienteilnehmer in der gepoolten Auswertung der INPULSIS Studien und Studie 1199.187) angegeben. Das CTP der INPULSIS Studien legte ein standardisiertes Vorgehen bei einer Erhöhung von Leberenzymen fest, welches sich an der Höhe des Transaminasen-Anstiegs in Relation zum ULN sowie dem Vorhandensein von Zeichen einer schweren Leberschädigung ausrichtete. Unter Beachtung dieser Vorgaben führte ein *UE* aus der Gruppe der *Leber- und Gallenerkrankungen* bei sieben Studienteilnehmern der Nintedanib-Gruppe (1,0% aller Studienteilnehmer) zu einem vorzeitigen Therapieabbruch. Zu einer dauerhaften Dosisreduktion auf 100 mg 2x täglich kam es aufgrund eines *UE* aus der Gruppe der *Leber- und Gallenerkrankungen* bei weiteren sechs Studienteilnehmern (0,9% aller Studienteilnehmer).

Die Fachinformation für das Produkt Ofev® gibt Empfehlungen zu Maßnahmen bei Leberenzym erhöhungen und zum Monitoring der Leberenzymwerte und enthält darüber hinaus einen Warnhinweis, wonach die Behandlung mit Nintedanib bei Studienteilnehmern mit bekannten mittelschweren oder schweren Leberfunktionsstörungen nicht empfohlen wird (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Unerwünschte kardiale Ereignisse

Der Endpunkt *UE von besonderem Interesse (kardiale Ereignisse)* wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimittelleinnahme erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-69: Ergebnisse für *UE - kardiale Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS + Studie 1199.187 (gepoolt)	Unerwünschte kardiale Ereignisse (SOC)							
	694	67 (9,7)	480	46 (9,6)	0,98 [0,69; 1,40]	0,98 [0,66; 1,45]	-0,18 [-3,62; 3,25]	0,916
	Schwerwiegende unerwünschte kardiale Ereignisse (SOC)							
	694	32 (4,6)	480	23 (4,8)	0,92 [0,55; 1,55]	0,92 [0,53; 1,59]	-0,42 [-3,17; 2,33]	0,763
	Tödliche kardiale Ereignisse							
	694	3 (0,4)	480	6 (1,3)	0,33 [0,08; 1,32]	0,33 [0,08; 1,32]	-0,95 [-2,19; 0,30]	0,100
	Arterielle thromboembolische Ereignisse							
	694	17 (2,4)	480	3 (0,6)	3,82 [1,13; 12,85]	3,89 [1,14; 13,30]	1,80 [0,44; 3,15]	0,020
INPULSIS (gepoolt)	Myokardinfarkt/akuter Myokardinfarkt							
	694	11 (1,6)	480	2 (0,4)	3,65 [0,81; 16,36]	3,69 [0,81; 16,74]	1,25 [0,05; 2,45]	0,070
	Unerwünschte kardiale Ereignisse							
	638	64 (10,0)	423	45 (10,6)	0,94 [0,66; 1,35]	0,94 [0,63; 1,40]	-0,61 [-4,36; 3,14]	0,750
	Schwerwiegende unerwünschte kardiale Ereignisse							
	638	32 (5,0)	423	23 (5,4)	0,92 [0,55; 1,55]	0,92 [0,53; 1,59]	-0,42 [-3,17; 2,32]	0,762
	Tödliche kardiale Ereignisse							
	638	3 (0,5)	423	6 (1,4)	0,33 [0,08; 1,32]	0,33 [0,08; 1,32]	-0,95 [-2,19; 0,30]	0,099
INPULSIS 1	Arterielle thromboembolische Ereignisse							
	638	16 (2,5)	423	3 (0,7)	3,54 [1,04; 12,06]	3,60 [1,04; 12,44]	1,8 [0,35; 3,25]	0,0305
	Myokardinfarkt/akuter Myokardinfarkt							
	638	10 (1,6)	423	2 (0,5)	3,32 [0,73; 15,05]	3,35 [0,73; 15,37]	1,09 [-0,07; 2,26]	0,0988
	Unerwünschte kardiale Ereignisse							
INPULSIS 1	309	30 (9,7)	204	19 (9,3)	1,04 [0,60; 1,80]	1,05 [0,57; 1,92]	0,40 [-4,78; 5,57]	0,882
	Schwerwiegende unerwünschte kardiale Ereignisse							
	309	14 (4,5)	204	11 (5,4)	0,84 [0,39; 1,81]	0,83 [0,37; 1,87]	-0,86 [-4,73; 3,01]	0,657

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
	Tödliche kardiale Ereignisse							
	309	1 (0,3)	204	2 (1,0)	0,33 [0,03; 3,62]	0,33 [0,03; 3,64]	-0,66 [-2,15; 0,84]	0,340
	Arterielle thromboembolische Ereignisse							
	309	8 (2,6)	204	1 (0,5)	5,25 [0,67; 41,91]	5,40 [0,67; 43,47]	2,1 [0,09; 4,11]	0,0764
	Myokardinfarkt/akuter Myokardinfarkt							
	309	5 (1,6)	204	1 (0,5)	3,30 [0,39; 28,05]	3,34 [0,39; 28,79]	1,13 [-0,57; 2,83]	0,2449
INPULSIS 2	Unerwünschte kardiale Ereignisse							
	329	34 (10,3)	219	26 (11,9)	0,87 [0,54; 1,41]	0,86 [0,50; 1,47]	-1,54 [-6,94; 3,86]	0,572
	Schwerwiegende unerwünschte kardiale Ereignisse							
	329	18 (5,5)	219	12 (5,5)	1,00 [0,49; 2,03]	1,00 [0,47; 2,12]	-0,01 [-3,90; 3,88]	0,997
	Tödliche kardiale Ereignisse							
	329	2 (0,6)	219	4 (1,8)	0,33 [0,06; 1,80]	0,33 [0,06; 1,81]	-1,22 [-3,18; 0,74]	0,179
	Arterielle thromboembolische Ereignisse							
	329	8 (2,4)	219	2 (0,9)	2,66 [0,57; 12,42]	2,70 [0,57; 12,86]	1,52 [-0,57; 3,61]	0,1933
	Myokardinfarkt/akuter Myokardinfarkt							
	329	5 (1,5)	219	1 (0,5)	3,33 [0,39; 28,29]	3,36 [0,39; 28,99]	1,06 [-0,53; 2,66]	0,2414
Studie 1199.187	Unerwünschte kardiale Ereignisse (SOC)							
	56	3 (5,4)	57	1 (1,8)	3,05 [0,33; 28,48]	3,17 [0,32; 31,43]	3,60 [-3,21; 10,41]	0,300
	Schwerwiegende unerwünschte kardiale Ereignisse (SOC)							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.
	Tödliche kardiale Ereignisse (SOC)							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.
	Ischämische Herzerkrankung (SMQ)							
	56	3 (5,4)	57	0 (0,0)	n. b.	n. b.	5,36 [-0,54; 11,25]	0,077
	Myokardinfarkt/akuter Myokardinfarkt							
	56	0 (0,0)	57	0 (0,0)	n. b.	n. b.	n. b.	n. b.

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 1.5.5.2.1, Tab. 1.5.5.1.1, Tab. Additional AEs, Tab. 1.3.49.2.6.1, Tab. 1.5.5.1.3 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. AE: Adverse Event (Unerwünschtes Ereignis); KI: Konfidenzintervall; n. b.: nicht berichtet; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); SMQ: Standardised MedDRA Queries; SOC: System Organ Class; UE: Unerwünschtes Ereignis								

Bedingt durch den Wirkmechanismus von Nintedanib besteht die Möglichkeit, dass Patienten ein erhöhtes Risiko für *thromboembolische Ereignisse* haben. *Arterielle thromboembolische Ereignisse* wurden selten berichtet; in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 0,6% der Patienten in der Kontrollgruppe und bei 2,4% in der mit Nintedanib behandelten Gruppe (RR [95%-KI]: 3,82 [1,13; 12,85]; p=0,020) (Tabelle 4-69). *UE*, die auf eine *ischämische Herzerkrankung* deuten, waren zwar zwischen der Nintedanib- und der Kontrollgruppe ausgeglichen, aber in der Nintedanib-Gruppe erlitt ein numerisch höherer Prozentsatz der Patienten (1,6%) einen *Myokardinfarkt* als dies in der Kontrollgruppe (0,4%) der Fall war (RR [95%-KI]: 3,65 [0,81; 16,36]; p=0,060). Es wurde ein entsprechender Warnhinweis in die Fachinformation von Ofev[®] aufgenommen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Bei Betrachtung aller *unerwünschten kardialen Ereignisse* wurden diese in den gepoolten INPULSIS Studien und Studie 1199.187 vergleichbar häufig unter Nintedanib und zVT (9,7% vs. 9,6%) beobachtet. Außer den bereits oben genannten *Herzinfarkten* waren hierunter beispielsweise folgende Ereignisse subsummiert: Vorhofflimmern, koronare Herzerkrankung, Herzversagen, Elektrokardiogramm (EKG)- Auffälligkeiten wie Arrhythmien oder Schenkelblockbilder oder Cor pulmonale. Auch *schwerwiegende unerwünschte kardiale Ereignisse* traten in der Nintedanib-Gruppe (4,6%) und Kontrollgruppe (4,8%) vergleichbar häufig auf (RR [95%-KI]: 0,92 [0,55; 1,55]; p=0,763). *Tödliche kardiale Ereignisse* traten numerisch seltener in der Nintedanib-Gruppe (0,4%) als in der Kontrollgruppe (1,3%) auf (RR [95%-KI]: 0,33 [0,08; 1,32]; p=0,100).

Unerwünschte respiratorische Ereignisse

Der Endpunkt *UE von besonderem Interesse (respiratorische Ereignisse)* wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimittelaufnahme erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der

Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-70: Ergebnisse für *UE - respiratorische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS + Studie 1199.187 (gepoolt)	Progression der IPF							
	694	67 (9,7)	480	66 (13,8)	0,69 [0,50; 0,95]	0,66 [0,46; 0,94]	-4,29 [-8,08; -0,50]	0,023
	Husten							
	694	90 (13,0)	480	66 (13,8)	0,94 [0,70; 1,27]	0,93 [0,66; 1,31]	-0,83 [-4,80; 3,14]	0,683
	Dyspnoe							
	694	57 (8,2)	480	54 (11,3)	0,74 [0,52; 1,05]	0,72 [0,48; 1,06]	-2,92 [-6,43; 0,58]	0,093
	Bronchitis							
	694	69 (9,9)	480	48 (10,0)	0,97 [0,68; 1,38]	0,97 [0,65; 1,43]	-0,30 [-3,78; 3,18]	0,866
	Nasopharyngitis							
	694	90 (13,0)	480	72 (15,0)	0,84 [0,63; 1,12]	0,82 [0,59; 1,15]	-2,36 [-6,42; 1,70]	0,249
INPULSIS (gepoolt)	Progression der IPF							
	638	64 (10,0)	423	61 (14,4)	0,70 [0,50; 0,97]	0,66 [0,45; 0,96]	-4,39 [-8,47; -0,31]	0,030
	Husten							
	638	85 (13,3)	423	57 (13,5)	0,99 [0,72; 1,35]	0,99 [0,69; 1,42]	-0,15 [-4,34; 4,04]	0,943
	Dyspnoe							
	638	49 (7,7)	423	48 (11,3)	0,68 [0,46; 0,99]	0,65 [0,43; 0,99]	-3,67 [-7,33; -0,01]	0,042
	Bronchitis							
	638	67 (10,5)	423	45 (10,6)	0,99 [0,69; 1,41]	0,99 [0,66; 1,47]	-0,14 [-3,92; 3,64]	0,943
	Nasopharyngitis							
	638	87 (13,6)	423	68 (16,1)	0,85 [0,63; 1,14]	0,82 [0,58; 1,16]	-2,44 [-6,84; 1,96]	0,271

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
INPULSIS 1	Progression der IPF							
	309	31 (10,0)	204	21 (10,3)	0,97 [0,58; 1,65]	0,97 [0,54; 1,74]	-0,26 [-5,61; 5,09]	0,923
	Husten							
	309	47 (15,2)	204	26 (12,7)	1,19 [0,76; 1,86]	1,23 [0,73; 2,06]	2,47 [-3,62; 8,55]	0,434
	Dyspnoe							
	309	22 (7,1)	204	23 (11,3)	0,63 [0,36; 1,10]	0,60 [0,33; 1,11]	-4,15 [-9,36; 1,05]	0,104
	Bronchitis							
	309	36 (11,7)	204	28 (13,7)	0,85 [0,54; 1,35]	0,83 [0,49; 1,41]	-2,08 [-8,00; 3,85]	0,486
	Nasopharyngitis							
	309	39 (12,6)	204	34 (16,7)	0,76 [0,50; 1,16]	0,72 [0,44; 1,19]	-4,05 [-10,36; 2,27]	0,199
INPULSIS 2	Progression der IPF							
	329	33 (10,0)	219	40 (18,3)	0,55 [0,36; 0,84]	0,50 [0,30; 0,82]	-8,23 [-14,29; -2,17]	0,005
	Husten							
	329	38 (11,6)	219	31 (14,2)	0,82 [0,52; 1,27]	0,79 [0,48; 1,32]	-2,61 [-8,37; 3,16]	0,368
	Dyspnoe							
	329	27 (8,2)	219	25 (11,4)	0,72 [0,43; 1,20]	0,69 [0,39; 1,23]	-3,21 [-8,36; 1,94]	0,209
	Bronchitis							
	329	31 (9,4)	219	17 (7,8)	1,21 [0,69; 2,14]	1,24 [0,67; 2,29]	1,66 [-3,09; 6,41]	0,501
	Nasopharyngitis							
	329	48 (14,6)	219	34 (15,5)	0,94 [0,63; 1,41]	0,93 [0,58; 1,50]	-0,94 [-7,06; 5,19]	0,764
Studie 1199.187	Progression der IPF							
	56	3 (5,4)	57	5 (8,8)	n. b.	n. b.	n. b.	n. b.
	Husten							
	56	5 (8,9)	57	9 (15,8)	0,57 [0,20; 1,58]	0,52 [0,16; 1,67]	-6,86 [-18,92; 5,20]	0,268

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
	Dyspnoe							
	56	8 (14,3)	57	6 (10,5)	1,36 [0,50; 3,66]	1,42 [0,46; 4,38]	3,76 [-8,38; 15,90]	0,544
	Bronchitis							
	56	2 (3,6)	57	3 (5,3)	n. b.	n. b.	n. b.	n. b.
	Nasopharyngitis							
	56	3 (5,4)	57	4 (7,0)	n. b.	n. b.	n. b.	n. b.
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 23.1.1, Tab. 23.1.2, Tab. 23.1.3, Tab. 12.1.2.1: 1, Tab. 1.5.5.1.1, Tab. 1.5.5.1.2 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; n. b.: nicht berichtet, OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

Husten und Dyspnoe sind typische Symptome der Grunderkrankung IPF. Daher wurden Therapieeffekte auf die Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch diese Symptome mittels der spezifischen Fragebögen *CASA-Q* und *SOBQ* analysiert. Diese Messinstrumente erfragen explizit die durchschnittliche Beeinträchtigung der gesundheitsbezogenen Lebensqualität durch die genannten Symptome über die Zeitspanne der letzten sieben Tage vor der jeweiligen Studienvisite und bilden somit chronische Effekte der Therapie auf die Symptomintensität ab.

Husten und Dyspnoe können jedoch auch als einzeln berichtenswerte akute medizinische Vorkommnisse auftreten, die gegenüber der vom Studienteilnehmer als normal empfundenen Tagesvariabilität hervorstechen. Dies kann zudem auch außerhalb der Zeitspanne der letzten sieben Tage vor einer Studienvisite der Fall gewesen sein. Daher lassen sich aus der Registrierung der Symptome *Husten* oder *Dyspnoe* als *UE* gegenüber den Fragebogeninstrumenten zusätzliche patientenrelevante Informationen gewinnen.

Als *respiratorische Ereignisse* wurden eine *Progression der IPF*, *Husten*, *Dyspnoe*, *Bronchitis* und *Nasopharyngitis* im Rahmen der Erfassung von *UE* angegeben (Tabelle 4-70).

Eine *Progression der IPF* wurde in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 in der Nintedanib-Gruppe von 9,7% der Studienteilnehmer als *UE* angegeben,

verglichen mit 13,8% in der Kontrollgruppe. Der Unterschied ist statistisch signifikant zugunsten von Nintedanib (RR [95%-KI]: 0,69 [0,50; 0,95]; $p=0,023$). *Husten* als UE trat in beiden Behandlungsarmen vergleichbar häufig auf (13,0% vs. 13,8%; RR [95%-KI]: 0,94 [0,70; 1,27]; $p=0,683$). *Bronchitis* und *Nasopharyngitis* als UE fanden sich in den gepoolten INPULSIS Studien und Studie 1199.187 in beiden Behandlungsarmen vergleichbar häufig: *Bronchitis* in der Nintedanib-Gruppe bei 9,9% der Studienteilnehmer vs. 10,0% der Studienteilnehmer in der Kontrollgruppe (RR [95%-KI]: 0,97 [0,68; 1,38]; $p=0,866$); *Nasopharyngitis* in der Nintedanib-Gruppe bei 13,0% der Studienteilnehmer vs. 15,0% der Studienteilnehmer in der Kontrollgruppe (RR [95%-KI]: 0,84 [0,63; 1,12]; $p=0,249$).

Unerwünschte Blutungsereignisse

Aufgrund theoretischer – den Wirkmechanismus betreffender – Erwägungen ist eine Hemmung des Vaskulären endothelialen Wachstumsfaktorrezeptors (VEGFR) möglicherweise mit einem erhöhten *Blutungsrisiko* verbunden (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b). Daher wurden Blutungsereignisse als UE von besonderem Interesse betrachtet.

Der Endpunkt *UE von besonderem Interesse (Blutungsereignisse)* wurde in den drei berichteten Studien von Studienbeginn bis 28 Tage nach letztmaliger Arzneimittleinnahme erhoben. Für die Studie 1199.187 wurden die Analysen der kontrollierten Phase berücksichtigt. Der Endpunkt wird dargestellt für die gepoolten Analysen (INPULSIS Studien und Studie 1199.187 [gepoolt] und INPULSIS Studien [gepoolt]) als auch auf Ebene der Einzelstudien. Im Kontext der Nutzenbewertung ist die größte Relevanz der über die INPULSIS Studien und Studie 1199.187 gepoolten Analyse beizumessen.

Tabelle 4-71: Ergebnisse für *UE - Blutungsereignisse* aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Blutungsereignisse								
INPULSIS + Studie 1199.187 (gepoolt)	694	68 (9,8)	480	35 (7,3)	1,31 [0,89; 1,94]	1,35 [0,88; 2,06]	2,30 [-0,92; 5,51]	0,171
INPULSIS (gepoolt)	638	66 (10,3)	423	33 (7,8)	1,33 [0,89; 1,98]	1,36 [0,88; 2,11]	2,54 [-0,94; 6,02]	0,1631
INPULSIS 1	309	29 (9,4)	204	17 (8,3)	1,13 [0,64; 2,00]	1,14 [0,61; 2,13]	1,05 [-3,94; 6,05]	0,6832
INPULSIS 2	329	37 (11,2)	219	16 (7,3)	1,54 [0,88; 2,70]	1,61 [0,87; 2,97]	3,94 [-0,91; 8,79]	0,1264
Studie 1199.187	56	2 (3,6)	57	2 (3,5)	n. b.	n. b.	n. b.	n. b.

Studie	Nintedanib		Placebo		Behandlungseffekt Nintedanib vs. Placebo ^c			
	N ^a	n ^b (%)	N ^a	n ^b (%)	RR [95%-KI]	OR [95%-KI]	RD [95%-KI]	p-Wert
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. Additional AEs, Tab. 1.5.5.1.3 a: Anzahl Studienteilnehmer b: Anzahl Studienteilnehmer mit Ereignis c: RR [95%-KI], OR [95%-KI], RD [95%-KI]: Ein OR bzw. RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Eine RD kleiner null zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. p-Wert für Analysen der INPULSIS Studien (gepoolt): berechnet mittels Chi-Quadrat-Test; p-Wert und Effektschätzer für INPULSIS + Studie 1199.187 (gepoolt): berechnet mittels Cochran-Mantel-Haenszel Methode. AE: Adverse Event (Unerwünschtes Ereignis); KI: Konfidenzintervall; n. b.: nicht berichtet; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); RD: Risk Difference (Absolute Risikodifferenz); RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis								

In der gepoolten Auswertung der INPULSIS Studien und Studie 1199.187 traten *Blutungsereignisse* im Nintedanib-Behandlungsarm numerisch häufiger (9,8%) als im Vergleichsarm (7,3%) auf (RR [95%-KI]: 1,31 [0,89; 1,94]; p=0,171) (Tabelle 4-71). Nicht schwerwiegende Epistaxis und Kontusionen stellten die häufigsten Blutungsereignisse dar. Die Häufigkeit *schwerwiegender Blutungsereignisse* war in beiden Behandlungsgruppen gering und vergleichbar (Placebo: 1,5%; Nintedanib: 1,2%). Da jedoch mit der anti-angiogenen Aktivität von VEGFR-Inhibitoren ein mögliches Risiko von Blutungsereignissen assoziiert wurde, wurde ein entsprechender Warnhinweis in die Fachinformation von Ofev[®] aufgenommen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

Unerwünschte Ereignisse mit einer Häufigkeit von >10%

Im Sinne der Vollständigkeit wurde geprüft, welche *UE* in den gepoolten INPULSIS Studien und der Studie 1199.187 in mindestens einer der Gruppen mit einer Häufigkeit >10% auftraten. Hierbei wurden keine *UE* auf MedDRA PT-Level identifiziert, die nicht bereits in einer der oben genannten Kategorien nach Organsystemen berücksichtigt wurden.

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.

Übertragbarkeit des Endpunkts auf den deutschen Versorgungskontext

Die Erhebung der unerwünschten Ereignisse ist international standardisiert (International Conference on Harmonisation, 1996) und wird einheitlich dokumentiert. Von einer Übertragbarkeit der Studienergebnisse zum Endpunkt *UE* von besonderem Interesse auf den deutschen Versorgungskontext kann somit ausgegangen werden.

Im vorliegenden Dossier werden UE berichtet, die im Rahmen der nutzenbewertungsrelevanten Studien aufgetreten waren. Weitere Hinweise auf UE, welche nach erfolgter Zulassung am 15.01.2015 dokumentiert wurden, sind in Modul 3 Abschnitt 3.4 sowie in der aktuellen Fachinformation enthalten (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018b).

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.

4.3.1.3.5 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.

Generell sollte die Aussagekraft von Subgruppenanalysen, insbesondere auf Studienebene, kritisch beurteilt werden. Geringe Studienteilnehmerzahlen innerhalb der Subgruppenkategorien, ein Ungleichgewicht zwischen den Subgruppenkategorien und die steigende Anzahl statistischer Tests erhöhen die Wahrscheinlichkeit für falsch-positive signifikante Ergebnisse (IQWiG, 2017). Diese Aspekte müssen bei der Bewertung von Ergebnissen aus Subgruppenanalysen berücksichtigt werden.

Basis der Subgruppenbetrachtungen ist, wenn nicht anders vermerkt, die gepoolte Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187. Die verwendete Methodik ist in Abschnitt 4.2.5.5 beschrieben. Tabelle 4-72 zeigt nochmals die untersuchten Subgruppenmerkmale.

Tabelle 4-72: Subgruppenmerkmale, die als potenzielle Effektmodifikatoren untersucht wurden

Subgruppmerkmal	Subgruppenkategorien
Geschlecht	männlich; weiblich
Alter	<65 Jahre; ≥65 Jahre
Ethnie	kaukasisch; asiatisch
Region	Europa; Amerika; Asien/Australien
FVC (% des Sollwertes)	≤70%; >70%
SGRQ-Gesamtscore zu Studienbeginn	≤40; >40
Raucherstatus	Nichtraucher; Ex-Raucher/Raucher
Systemische Kortikosteroide zu Studienbeginn	ja; nein
Bronchodilatator zu Studienbeginn	ja; nein
Studie	INPULSIS 1; INPULSIS 2; Studie 1199.187
Niereninsuffizienz ^a	ja; nein
Körpergewicht zu Studienbeginn	<65 kg; ≥65 kg
a: Definiert als Kreatinin-Clearance von <90 ml/min. Berechnung der Kreatinin-Clearance mithilfe der Cockcroft–Gault-Formel.	
FVC: Forced Vital Capacity (Forcierte Vitalkapazität); SGRQ: St. George's Respiratory Questionnaire	

Verzerrungspotenzial des Subgruppenmerkmals Bronchodilatator

Eine der präspezifizierten Subgruppenanalysen betraf das Merkmal Begleittherapie mit Bronchodilatator (ja/nein). Die Subgruppenzuordnung erfolgte gemäß CTP anhand der bei Studienbeginn angegebenen Begleittherapie. Dieses Vorgehen erwies sich jedoch als nicht zielführend: Die Analyse der Begleittherapie während der Studie ergab, dass die individuelle Ausprägung des Merkmals Bronchodilatator über die Beobachtungsdauer der Studie bei einem relevanten Anteil der Studienteilnehmer nicht konstant war. In der INPULSIS 1 Studie erhielten bei Studienbeginn n=61 Studienteilnehmer im Nintedanib-Arm (19,7%) und n=34 Studienteilnehmer im Vergleichsarm (16,7%) begleitend Bronchodilatoren. Während der Beobachtungsdauer der Studie kehrte sich das Verhältnis jedoch um: Eine Bronchodilatator-Begleittherapie erhielten n=40 Studienteilnehmer im Nintedanib-Arm (12,9%) bzw. n=38 im Kontrollarm (18,6%). In der INPULSIS 2 Studie zeigten sich dazu analoge Verschiebungen der Merkmalsausprägung im Studienverlauf: Bei Studienbeginn erhielten n=68 Studienteilnehmer im Nintedanib-Arm (20,7%) und n=38 Studienteilnehmer im Vergleichsarm (17,4%) begleitend Bronchodilatoren. Während der Beobachtungsdauer der Studie kehrte sich das Verhältnis um: Mindestens einmal erhielten Bronchodilatoren n=49 Studienteilnehmer im Nintedanib-Arm (14,9%) bzw. n=48 im Kontrollarm (21,9%).

Konsequenz dieser Änderungen bezüglich der individuellen Ausprägung des Subgruppenmerkmals Bronchodilatator (ja/nein) war, dass die anhand der Dokumentation des Ausgangswertes (Boehringer Ingelheim Pharma GmbH & Co. KG, 2015) erfolgte Zuordnung

eines relevanten Anteils der Studienteilnehmer im weiteren Studienverlauf nicht den tatsächlichen Gegebenheiten entsprach. Daher muss von einer erheblichen Verzerrung der Ergebnisse der Subgruppenanalysen zum Merkmal Bronchodilatator (ja/nein) ausgegangen werden: Eine Begleittherapie kann sich nur insofern als potenzieller Effektmodifikator auswirken, wie sie auch tatsächlich besteht.

Die Tatsache, dass diese Veränderungen der Ausprägung für das Subgruppenmerkmal Bronchodilatator (ja/nein) analog in beiden Studien beobachtet wurden, weist dabei auf einen systematischen Effekt hin. Es ist auffällig, dass die Veränderungen des Bronchodilatator-Bedarfs unter der Behandlung mit Studienmedikation eine Entwicklung hin zu einem höheren Bedarf an Bronchodilatator-Begleittherapie in den Kontrollgruppen sowie eine gegenläufige Entwicklung in den Nintedanib-Gruppen anzeigen. Dies kann als indirekter Hinweis auf eine höhere subjektive Beeinträchtigung der Studienteilnehmer durch Dyspnoe/Luftnot unter Behandlung mit der zVT verglichen mit Nintedanib interpretiert werden.

Aus den genannten Gründen werden nachfolgend Ergebnisse der Subgruppenanalyse zum Merkmal Bronchodilatator (ja/nein) zwar dargestellt; diese werden von BI jedoch aufgrund des hohen Verzerrungspotenzials als nicht schlussfolgerungsrelevant betrachtet.

Ergebnisse der Subgruppenanalysen

Im Folgenden (Abschnitt 4.3.1.3.5.1 bis Abschnitt 4.3.1.3.5.4) werden die Ergebnisse der Subgruppenanalysen zusammengefasst. Im Sinne der Übersichtlichkeit werden zunächst jeweils die p-Werte der Interaktionstests in einer Übersichtstabelle für die einzelnen Endpunkte zur Mortalität, Morbidität, gesundheitsbezogenen Lebensqualität und zu UE dargestellt. Einzelergebnisse der untersuchten Subgruppen werden für diejenigen Endpunkte und Subgruppen tabellarisch berichtet, bei denen sich Belege für Effektmodifikationen (p-Wert des Interaktionstests $< 0,05$) durch den subgruppenbildenden Faktor ergeben haben.

Wenn nicht anders angegeben, wurden die Subgruppenanalysen für die gepoolten Daten des gesamten Studienpools durchgeführt und dargestellt. Wie bereits in Abschnitt 4.3.1.2 dargelegt, eignen sich die drei Studien für eine gepoolte Auswertung. Auf eine zusätzliche Darstellung der Subgruppenergebnisse auf Einzelstudienebene wird verzichtet. Im vorliegenden Fall handelt es sich um IPD Meta-Analysen über drei Studien. Die statistischen Modelle wurden wie in den Einzelstudien mit dem dazugehörigen Effektmaß gewählt und auf die Darstellung von Forest-Plots wurde verzichtet. Alle Ergebnisse der durchgeführten Subgruppenanalysen sind zur Gewährleistung der notwendigen Transparenz in Modul 5 dargestellt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

4.3.1.3.5.1 Mortalität

Tabelle 4-73: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur Mortalität basierend auf der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187

Subgruppe	Zeit bis zum Tod, unabhängig von der Ursache	Zeit bis zum respiratorisch bedingten Tod	Jährliche Abnahme der FVC (ml)	FVC-Responder zu Woche 52 (Schwellenwert >10%)	FVC-Responder zu Woche 52 (Schwellenwert >5%)
Geschlecht	0,2279	0,9468	0,3855	0,1746	0,8984
Alter	0,7515	0,6425	0,7357	0,3193	0,6963
Ethnie	0,9874	0,8969	0,4213	0,8936	0,9310
Region	0,9512	0,5627	0,4615	0,9460	0,9842
FVC (% des Sollwertes)	0,1041	0,0710	0,8513	0,3835	0,8662
SGRQ-Gesamtscore	0,4861	0,2534	0,4410	0,4478	0,5704
Raucherstatus	0,7571	0,9526	0,2861	0,8275	0,0166
Systemische Kortikosteroide	0,3155	0,7054	0,5757	0,7904	0,6808
Bronchodilatator	0,8210	0,8579	0,9997	0,2814	0,8043
Studie	0,5795	0,7910	0,3867	0,1432	0,9313
Niereninsuffizienz	0,0681	0,0831	0,3950	0,6427	0,4218
Körpergewicht	0,3957	0,1279	0,4558	0,6318	0,2661
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.69.1.1 - 1.12, Tab. 1.3.69.2.1 - 2.12, Tab. 1.3.69.5.1 - 5.12, Tab. 1.3.69.6.1.1 - 6.12.1, Tab. 1.3.69.7.1.1 - 7.12.1 Mit Ausnahme der Analysen der FVC-Responder zu Woche 52 basieren alle Interaktionstests auf der gepoolten Analyse aller drei Studien. Die statistischen Modelle verwenden die gesamte kontrollierte Phase aller drei eingeschlossenen Studien. Die Subgruppenuntersuchungen für die Endpunkte FVC-Responder zu Woche 52 basieren auf der gepoolten Analyse der INPULSIS Studien. FVC: Forced Vital Capacity (Forcierte Vitalkapazität); SGRQ: St. George's Respiratory Questionnaire					

4.3.1.3.5.1.1 Endpunkt *Zeit bis zum Tod, unabhängig von der Ursache, Zeit bis zum respiratorisch bedingten Tod*

Die Subgruppenuntersuchungen für die Endpunkte *Zeit bis zum Tod, unabhängig von der Ursache* und *Zeit bis zum respiratorisch bedingten Tod* auf Basis der gepoolten Analyse aller drei Studien ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale (Tabelle 4-73). Für alle drei Studien wurden für die Subgruppenanalyse jeweils die Daten der gesamten kontrollierten Phase herangezogen.

4.3.1.3.5.1.2 Endpunkt *Forcierte Vitalkapazität*

Entsprechend der Ergebnisse der Interaktionstests für die Subgruppenanalysen zu Endpunkten der Endpunktkategorie Mortalität (Tabelle 4-73) werden im Folgenden für den Endpunkt *FVC-Responder zu Woche 52* (Abfall der FVC [% des Sollwertes] um <5 Prozentpunkte gegenüber dem Ausgangswert) die Subgruppen nach *Raucherstatus* berichtet. Die Subgruppenanalyse basiert auf der gepoolten Analyse der INPULSIS Studien. Für Studie 1199.187 lagen für diesen Beobachtungszeitpunkt keine Daten vor.

Tabelle 4-74: Subgruppenergebnisse für *FVC-Responder* (Schwellenwert 5% des Sollwertes) aus RCT für die gepoolte Auswertung der INPULSIS Studien

Endpunkt/ Subgruppe	Behandlung	N ^a	Anzahl Ereignisse N (%)	Behandlungseffekt Nintedanib vs. Placebo ^b	
				OR	95%-KI
FVC-Responder zu Woche 52 (Schwellenwert 5%-Punkte)					
Raucherstatus				p-Wert Interaktionsterm ^c : 0,0166	
Nichtraucher	Nintedanib	174	77 (44,25)	1,122	[0,70; 1,80]
	Placebo	122	51 (41,80)		
Ex-Raucher/ Raucher	Nintedanib	464	261 (56,25)	2,216	[1,64; 2,99]
	Placebo	301	113 (37,54)		
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.69.7.6.1					
a: Anzahl Studienteilnehmer in Subgruppe					
b: OR [95%-KI]: berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Alter, Geschlecht, Größe und Ausgangswert FVC (% des Sollwertes)					
c: Berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Alter, Geschlecht, Größe, Subgruppe und Wechselwirkung Behandlung×Subgruppe					
Studienteilnehmer waren FVC-Responder, wenn sie keinen absoluten Abfall der FVC von mehr als 10%-Punkten oder 5%-Punkten der FVC (% des Sollwertes) hatten und ein Wert zu Woche 52 vorlag.					
FVC: Forced Vital Capacity (Forcierte Vitalkapazität); KI: Konfidenzintervall; OR: Odds Ratio; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)					

Für den Endpunkt *FVC-Responder zu Woche 52* (bewertet anhand des Schwellenwertes von 5 Prozentpunkte) ergab sich in der gepoolten Analyse der INPULSIS Studien ein Beleg für eine quantitative Effektmodifikation durch das Merkmal *Raucherstatus*. In der Subgruppe der Ex-Raucher und Raucher zeigte sich ein statistisch signifikanter Effekt zugunsten von Nintedanib (OR [95%-KI]: 2,216 [1,64; 2,99]), in der Subgruppe der Nichtraucher zeigte sich ein numerischer Trend zugunsten von Nintedanib; der Unterschied war jedoch in dieser Subgruppe nicht statistisch signifikant (OR [95%-KI]: 1,122 [0,70; 1,80]).

4.3.1.3.5.2 Morbidität

Tabelle 4-75: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur Morbidität

	INPULSIS Studien + Studie 1199.187 (gepoolt)		Studie 1199.187 ^a	
Subgruppe	Zeit bis zur ersten adjudizierten akuten Exazerbation	Zeit bis zur ersten akuten Exazerbation	Änderung im HRCT-QLF Score bis Woche 24	Veränderung der Gehstrecke im 6 MWT bis Woche 24
Geschlecht	0,6068	0,9631	0,6638	0,1360
Alter	0,3673	0,3102	0,3735	0,2501
Ethnie	0,7930	0,8293	0,3009	0,2758
Region	0,4354	0,4346	0,8463	0,6728
FVC (% des Soll- wertes)	0,4596	0,2362	0,8299	0,3520
SGRQ- Gesamtscore	0,6755	0,6422	0,4807	0,2847
Raucherstatus	0,8631	0,7332	0,3487	0,4660
Systemische Kortikosteroide	nicht berechenbar ^b	0,2975	0,4795	0,9385
Bronchodilatator	0,1312	0,1358	0,6646	0,6537
Studie	0,3659	0,0948	-	-
Niereninsuffizienz	0,5358	0,8320	0,9226	0,1484
Körpergewicht	0,3913	0,5439	0,6530	0,6734
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.69.3.1-3.12, Tab. 1.3.69.4.1-4.12, Tab. 1.3.69.18.1-18.12, Tab. 1.3.69.19.1-19.12 a: Aufgrund der geringen Zahl an Ereignissen in der Gesamtpopulation (0 vs. 4 Patienten mit Ereignissen im Nintedanib- bzw. Placebo-Arm) wurde für den Endpunkt <i>Zeit bis zur ersten respiratorisch bedingten Hospitalisierung</i> auf Subgruppenuntersuchungen verzichtet b: statistisches Modell konvergierte nicht 6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); SGRQ: St. George's Respiratory Questionnaire				

4.3.1.3.5.2.1 Endpunkte zu akuten Exazerbationen

Die Subgruppenuntersuchungen für die Endpunkte *Zeit bis zur ersten adjudizierten akuten Exazerbation* und *Zeit bis zur ersten akuten Exazerbation* auf Basis der gepoolten Analyse aller drei Studien ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale (Tabelle 4-75). Für alle drei Studien wurden für die Analyse jeweils die Daten der gesamten kontrollierten Phase herangezogen.

4.3.1.3.5.2.2 Endpunkt *HRCT-QLF Score*

Die Subgruppenuntersuchungen für den Endpunkt *Änderung des HRCT-QLF Scores bis Woche 24* basieren auf der Studie 1199.187; in den INPULSIS Studien wurde dieser Endpunkt nicht erhoben. Es ergab sich kein Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale (Tabelle 4-75).

4.3.1.3.5.2.3 Endpunkt *Veränderung der Gehstrecke im 6 MWT*

Die Subgruppenuntersuchungen für den Endpunkt *Veränderung der Gehstrecke im 6 MWT bis Woche 24* basieren auf der Studie 1199.187; in den INPULSIS Studien wurde dieser Endpunkt nicht erhoben. Es ergab sich kein Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale (Tabelle 4-75).

4.3.1.3.5.3 Gesundheitsbezogene Lebensqualität

Tabelle 4-76: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zur gesundheitsbezogenen Lebensqualität

	INPULSIS 1 und 2 (gepoolt)								
Subgruppe	SGRQ-Gesamt-score	SGRQ-I-Gesamt-score	SOBQ	SGRQ-Responder zu Woche 52	SGRQ-I-Responder zu Woche 52	CASA-Q Husten-Symptome	CASA-Q Husten-Impact	PGI-C Responder zu Woche 52	EQ-5D-VAS
Geschlecht	0,0591	0,1267	0,6121	0,9095	0,9617	0,6729	0,1151	0,9424	0,6307
Alter	0,2232	0,1848	0,2029	0,0839	0,3068	0,2318	0,0994	0,5522	0,5983
Ethnie	0,7292	0,4855	0,8864	0,8388	0,7679	0,3784	0,7846	0,2534	0,3126
Region	0,1511	0,1597	0,7570	0,1016	0,0956	0,0462	0,3282	0,2771	0,0177
FVC (% des Sollwertes)	0,0631	0,0256	0,2132	0,0293	0,1058	0,3644	0,2234	0,0757	0,7897
SGRQ-Gesamt-score	0,4785	0,8811	0,7292	0,7346	0,5631	0,7429	0,4575	0,5437	0,4334
Raucher-status	0,9753	0,6958	0,6280	0,9076	0,4283	0,8299	0,5731	0,9953	0,1377
Systemische Kortikosteroide	0,7503	0,4765	0,7742	0,6980	0,9119	0,1573	0,4959	0,8655	0,5095
Bronchodilatator	0,1116	0,1962	0,9918	0,0002	0,0030	0,1639	0,0595	0,8758	0,5742
Studie	0,2081	0,3285	0,6483	0,0277	0,0164	0,3627	0,9667	0,7607	0,0200

	INPULSIS 1 und 2 (gepoolt)								
Subgruppe	SGRQ-Gesamt-score	SGRQ-I-Gesamt-score	SOBQ	SGRQ-Responder zu Woche 52	SGRQ-I-Responder zu Woche 52	CASA-Q Husten-Symptome	CASA-Q Husten-Impact	PGI-C Responder zu Woche 52	EQ-5D-VAS
Nieren-insuffizienz	0,3095	0,1962	0,3074	0,2571	0,7498	0,9046	0,1975	0,7845	0,2058
Körpergewicht	0,5082	0,5431	0,0589	0,3229	0,2223	0,9661	0,0685	0,4399	0,8054
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.29.8.1 – 12, Tab. 1.1.29.9.1.1 – 12.1, Tab. 1.1.29.10.1 – 12, Tab. 1.1.29.11.1.1 – 12.1, Tab. 1.1.29.12.1 – 12, Tab. 1.1.29.13.1 – 12, Tab. 1.1.29.14.1 – 12, Tab. 1.1.29.15.1.1 – 12.1, Tab. 1.1.29.16.1 – 12 CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); PGI-C: Patient's Global Impression of Change; SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; VAS: Visual Analog Scale									

4.3.1.3.5.3.1 Endpunkte zum SGRQ, SGRQ-I, SOBQ, CASA-Q, PGI-C und EQ-5D-VAS

Entsprechend der Ergebnisse der Interaktionstests (Tabelle 4-76) wird in Tabelle 4-77 für den Endpunkt *Veränderung des SGRQ-I-Gesamtscores* die Subgruppe nach *FVC (% des Sollwertes)* zu Studienbeginn berichtet.

Tabelle 4-77: Subgruppenergebnisse für die Veränderung des SGRQ-I-Gesamtscores – basierend auf der gepoolten Analyse der Studien INPULSIS 1 und INPULSIS 2

Endpunkt/ Subgruppe	Behandlung	N ^a	N ^b	MWD ^c (SE)	Behandlungseffekt Nintedanib vs. Placebo ^c	
					MWD (SE)	95%-KI
Veränderung SGRQ-I-Gesamtscore						
FVC (% des Sollwertes)				p-Wert Interaktionsterm ^d : 0,0256		
≤70%	Nintedanib	207	198	3,33 (0,996)	-3,86 (1,542)	[-6,89; -0,83]
	Placebo	154	149	7,19 (1,178)	Hedges' g: -0,295 [-0,53; -0,06]	
>70%	Nintedanib	431	412	3,57 (0,670)	-0,92 (1,059)	[-3,00; 1,15]
	Placebo	269	264	4,50 (0,820)		

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.29.10.4, Tab. 1.1.31.1.1

a: Anzahl Studienteilnehmer in Subgruppe

b: Anzahl analysierter Studienteilnehmer in Subgruppe

c: MWD [95%-KI] anhand IPD Meta-Analyse der Studien INPULSIS 1 und INPULSIS 2: berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Visite, Behandlung×Visite, (Ausgangswert SGRQ-I-Gesamtscore), (Ausgangswert SGRQ-I-Gesamtscore) ×Visite und zufälligem Effekt für Studienteilnehmer

d: Berechnet mittels MMRM mit festen Effekten für Studie, Behandlung, Visite, Behandlung×Visite, (Ausgangswert SGRQ-I-Gesamtscore), (Ausgangswert SGRQ-I-Gesamtscore) ×Visite, Subgruppe, Behandlung×Subgruppe und zufälligem Effekt für Studienteilnehmer.

Endpunkt/ Subgruppe	Behandlung	N ^a	N ^b	MWD ^c (SE)	Behandlungseffekt Nintedanib vs. Placebo ^c	
					MWD (SE)	95 %-KI
FVC: Forced Vital Capacity (Forcierte Vitalkapazität), IPD: Individual Participant Data (Individuelle Patientendaten); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MWD: Mittelwertdifferenz; SE: Standard Error (Standardfehler); SGRQ-I: St. George’s Respiratory Questionnaire IPF-specific version						

Für den Endpunkt *Veränderung des SGRQ-I-Gesamtscores bis Woche 52* fand sich in der gepoolten Analyse der beiden INPULSIS Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *FVC ($\leq 70\%$ des Sollwertes)* (Tabelle 4-77). Bei Patienten mit einer zu Studienbeginn niedrigeren FVC ($\leq 70\%$ des Sollwertes), zeigte sich ein statistisch signifikanter Behandlungseffekt zugunsten von Nintedanib gegenüber der Vergleichstherapie (adjustierte MWD [95%-KI]: -3,86 [-6,89; -0,83]). Die klinische Relevanz des beobachteten Unterschieds ist jedoch unklar (Hedges' g [95%-KI]: -0,295 [-0,53; -0,06]). Bei Patienten mit einer höheren FVC ($> 70\%$ des Sollwertes) zu Studienbeginn, zeigte sich ein gleichgerichteter Effekt, der beobachtete Unterschied zwischen den Behandlungsgruppen ist jedoch nicht statistisch signifikant (adjustierte MWD [95%-KI]: -0,92 [-3,00; 1,15]).

Die Subgruppenuntersuchungen für die Endpunkte *Veränderung des SGRQ-Gesamtscores bis Woche 52* und *Veränderung des SOBQ bis Woche 52* auf Basis der gepoolten INPULSIS Studien ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale.

Entsprechend der Ergebnisse der Interaktionstests (Tabelle 4-76) werden im Folgenden für den Endpunkt *SGRQ-Responder zu Woche 52* die Subgruppen nach *FVC (% des Sollwertes) zu Studienbeginn*, nach *Bronchodilatator zu Studienbeginn* und nach *Studie* sowie für den Endpunkt *SGRQ-I-Responder zu Woche 52* die Subgruppen nach *Bronchodilatator zu Studienbeginn* und nach *Studie* berichtet.

Tabelle 4-78: Subgruppenergebnisse für *SGRQ-Responder* und *SGRQ-I-Responder zu Woche 52* basierend auf der gepoolten Analyse der INPULSIS Studien

Endpunkt/ Subgruppe	Behandlung	N ^a	Responder N (%)	Behandlungseffekt Nintedanib vs. Placebo ^b	
				OR	95%-KI
SGRQ-(≤4 Punkte)-Responder zu Woche 52					
FVC (% des Sollwertes)				p-Wert Interaktionsterm ^c : 0,0293	
≤70%	Nintedanib	207	51 (24,64)	1,934	[1,11; 3,38]
	Placebo	154	23 (14,94)		
>70%	Nintedanib	431	95 (22,04)	0,917	[0,63; 1,33]
	Placebo	269	63 (23,42)		

Endpunkt/ Subgruppe	Behandlung	N ^a	Responder N (%)	Behandlungseffekt Nintedanib vs. Placebo ^b	
				OR	95 %-KI
Bronchodilatator				p-Wert Interaktionsterm ^c : 0,0002	
ja	Nintedanib	131	32 (24,43)	7,236	[2,12; 24,68]
	Placebo	72	3 (4,17)		
nein	Nintedanib	507	114 (22,49)	0,976	[0,70; 1,35]
	Placebo	351	83 (23,65)		
Studie				p-Wert Interaktionsterm ^c : 0,0277	
INPULSIS 1	Nintedanib	309	63 (20,39)	0,840	[0,55; 1,29]
	Placebo	204	49 (24,02)		
INPULSIS 2	Nintedanib	329	83 (25,23)	1,664	[1,08; 2,57]
	Placebo	219	37 (16,89)		
SGRQ-I (≤-4 Punkte)-Responder zu Woche 52					
Bronchodilatator				p-Wert Interaktionsterm ^c : 0,0030	
ja	Nintedanib	131	37 (28,24)	4,232	[1,68; 10,67]
	Placebo	72	6 (8,33)		
nein	Nintedanib	507	113 (22,29)	1,066	[0,76; 1,49]
	Placebo	351	77 (21,94)		
Studie ^d				p-Wert Interaktionsterm ^e : 0,0164	
INPULSIS 1	Nintedanib	309	66 (21,36)	0,886	[0,58; 1,36]
	Placebo	204	49 (24,02)		
INPULSIS 2	Nintedanib	329	84 (25,53)	1,877	[1,20; 2,93]
	Placebo	219	34 (15,53)		

Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.3.69.9.4.1, Tab. 1.3.69.9.8.1, Tab. 1.3.69.9.9.1, Tab. 1.3.69.11.8.1, Tab. 1.3.69.11.9.1

a: Anzahl Studienteilnehmer in Subgruppe

b: OR [95%-KI] gepoolte Daten: berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore

c: Berechnet mittels logistischer Regression mit Termen für Studie, Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore, Subgruppe, Wechselwirkung Behandlung×Subgruppe

d: Bei den Subgruppenanalysen für die gepoolten Daten wurde die Kovariate „Studie“ bei den Analysemodellen berücksichtigt, um ggf. vorhandene Studieneffekte auszugleichen.

e: Berechnet mittels logistischer Regression mit Termen für Behandlung, Ausgangswert SGRQ-Gesamtscore bzw. SGRQ-I-Gesamtscore, Subgruppe, Wechselwirkung Behandlung×Subgruppe

FVC: Forced Vital Capacity (Forcierte Vitalkapazität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; OR: Odds Ratio; SGRQ: St. George’s Respiratory Questionnaire; SGRQ-I: St. George’s Respiratory Questionnaire IPF-specific version

Subgruppenergebnisse – SGRQ-Responder zu Woche 52

Für den Endpunkt *SGRQ-Responder zu Woche 52* fand sich in der gepoolten Analyse der INPULSIS Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *FVC (% des Sollwertes) zu Studienbeginn* (Tabelle 4-78). Bei Patienten mit einer höheren FVC (>70% des Sollwertes) zu Studienbeginn zeigte sich kein Unterschied zwischen den Behandlungsarmen im Anteil der *SGRQ-Responder zu Woche 52* (OR [95%-KI]: 0,917 [0,63; 1,33]). Bei Patienten mit einer zu Studienbeginn niedrigeren FVC (≤70% des Sollwertes) lag der Anteil der Patienten mit einer relevanten Verbesserung gemessen anhand des SGRQ-Gesamtscores in der Nintedanib-Gruppe höher als in der Kontrollgruppe (24,64% vs. 14,94%). Der Unterschied ist statistisch signifikant (OR [95%-KI]: 1,934 [1,11; 3,38]).

Weiterhin ergab sich ein Beleg für eine Effektmodifikation durch das Merkmal *Bronchodilatator zu Studienbeginn*. Bei Patienten mit *Bronchodilatator zu Studienbeginn* zeigte sich ein statistisch signifikanter Unterschied in den Anteilen der *SGRQ-Responder zu Woche 52* zugunsten von Nintedanib (24,43% vs. 4,17%) gegenüber der Vergleichstherapie (OR [95%-KI]: 7,236 [2,12; 24,68]). Bei Patienten ohne eine solche Begleittherapie zu Studienbeginn zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 0,976 [0,70; 1,35]).

Überdies zeigte sich in der gepoolten Analyse der INPULSIS Studien eine Effektmodifikation durch den Faktor *Studie*, was auf eine heterogene Datenlage hindeutet (Tabelle 4-78). Die in den Einzelstudien beobachteten Effekte sind gegenläufig. Während der Behandlungseffekt in Studie INPULSIS 1 nicht statistisch signifikant war (OR [95%-KI]: 0,840 [0,55; 1,29]), war der Anteil der *SGRQ-Responder zu Woche 52* in Studie INPULSIS 2 im Nintedanib-Arm höher als im Kontrollarm (OR [95%-KI]: 1,664 [1,08; 2,57]).

Subgruppenergebnisse - SGRQ-I-Responder zu Woche 52

Für den Endpunkt *SGRQ-I-Responder zu Woche 52* fand sich in der gepoolten Analyse der INPULSIS Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *Bronchodilatator zu Studienbeginn* (Tabelle 4-78). Bei Patienten mit *Bronchodilatator zu Studienbeginn* zeigte sich ein statistisch signifikanter Unterschied in den Anteilen der *SGRQ-I-Responder zu Woche 52* zugunsten von Nintedanib (28,2% vs. 8,3%) gegenüber der Vergleichstherapie (OR [95%-KI]: 4,232 [1,68; 10,67]). Bei Patienten ohne eine solche Begleittherapie zu Studienbeginn zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 1,066 [0,76; 1,49]).

Weiterhin zeigte sich in der gepoolten Analyse der INPULSIS Studien eine Effektmodifikation durch den Faktor *Studie*, was auf eine heterogene Datenlage hindeutet (Tabelle 4-78). Die in den Einzelstudien beobachteten Effekte sind gegenläufig. Während der Behandlungseffekt in Studie INPULSIS 1 nicht statistisch signifikant war (OR [95%-KI]: 0,886 [0,58; 1,36]), war der Anteil der *SGRQ-I-Responder zu Woche 52* in Studie INPULSIS 2 im Nintedanib-Arm signifikant höher als im Kontrollarm (OR [95%-KI]: 1,877 [1,20; 2,93]).

Die Subgruppenuntersuchungen für den Endpunkt Veränderung des *CASA-Q Husten-Impact-Scores* und *PGI-C-Responder* zu Woche 52 basierend auf der gepoolten Analyse der INPULSIS Studien ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale (Tabelle 4-76).

Tabelle 4-79: Subgruppenergebnisse für die Veränderung des *CASA-Q Husten-Symptom Scores* und die Veränderung auf der *EQ-5D-VAS* – basierend auf der gepoolten Analyse der INPULSIS Studien

Endpunkt/ Subgruppe	Behandlung	N ^a	N ^b	MWD ^c (SE)	Behandlungseffekt Nintedanib vs. Placebo ^c	
					MWD (SE)	95%-KI
Veränderung CASA-Q Husten-Symptom-Score						
Region				p-Wert Interaktionsterm ^d : 0,0462		
Europa	Nintedanib	312	309	−1,21 (1,138)	0,46 (1,781)	[-3,03; 3,95]
	Placebo	216	212	−1,67 (1,377)		
Amerika	Nintedanib	120	119	1,74 (1,758)	5,11 (2,781)	[-0,35; 10,58]
	Placebo	72	72	−3,37 (2,186)		
Asien/Austra- lien	Nintedanib	206	197	−0,87 (1,395)	-0,99 (2,207)	[-5,32; 3,34]
	Placebo	135	133	0,12 (1,731)		
Veränderung EQ-5D-VAS						
Region				p-Wert Interaktionsterm ^d : 0,0177		
Europa	Nintedanib	312	301	−2,67 (0,915)	3,24 (1,427)	[0,44; 6,04]
	Placebo	216	208	−5,91 (1,103)	Hedges‘ g: 0,217 [0,03; 0,40]	
Amerika	Nintedanib	120	115	−1,29 (1,606)	10,44 (2,557)	[5,41; 15,47]
	Placebo	72	69	−11,73 (2,013)	Hedges‘ g: 0,663 [0,34; 0,98]	
Asien/Austra- lien	Nintedanib	206	189	−3,23 (1,165)	0,80 (1,844)	[−2,82; 4,42]
	Placebo	135	131	−4,03 (1,449)		
Studie ^e				p-Wert Interaktionsterm ^d 0,0200		
INPULSIS 1	Nintedanib	309	293	−2,95 (0,943)	3,09 (1,500)	[0,14; 6,03]
	Placebo	204	197	−6,04 (1,167)	Hedges‘ g: 0,203 [0,01; 0,40]	
INPULSIS 2	Nintedanib	329	312	−2,50 (0,908)	4,39 (1,431)	[1,59; 7,20]
	Placebo	219	211	−6,90 (1,106)	Hedges‘ g: 0,292 [0,11; 0,48]	

Endpunkt/ Subgruppe	Behandlung	N ^a	N ^b	MWD ^c (SE)	Behandlungseffekt Nintedanib vs. Placebo ^c	
					MWD (SE)	95%-KI
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.1.29.13.12, Tab. 1.1.29.16.9, Tab. 1.1.29.16.12, Tab. 1.1.31.1.4, Tab. 1.1.31.1.3						
a: Anzahl Studienteilnehmer in Subgruppe						
b: Anzahl analysierter Studienteilnehmer in Subgruppe						
c: MWD [95%-KI] anhand IPD Meta-Analyse der Studien INPULSIS 1 und INPULSIS 2: berechnet mittels MMRM über den Zeitraum der doppelblinden Studienperiode mit festen Effekten für (Studie), Behandlung, Visite, Behandlung×Visite, Ausgangswert, Ausgangsért×Visite und zufälligem Effekt für Studienteilnehmer.						
d: Basierend auf einem MMRM mit festem Effekt für (Studie), Behandlung, Visite, Subgruppe, Behandlung×Visite×Subgruppe, Ausgangswert, Ausgangswert ×Visite, und zufälligem Effekt für Studienteilnehmer.						
e: Bei Subgruppenanalysen nach Studie enthielt das statistische Modell „Studie“ nicht als Faktor. Für die Untersuchung des Interaktionsterms wurde „Studie“ als Subgruppenfaktor eingeschlossen.						
CASA-Q: Cough and Sputum Assessment Questionnaire; EQ-5D: European Quality of Life-5 Dimensions; IPD: Individual Participant Data (Individuelle Patientendaten); KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; MWD: Mittelwertdifferenz; SE: Standard Error (Standardfehler); VAS: Visuelle Analogskala						

Subgruppenergebnisse - CASA-Q Husten-Symptom Score

Für den Endpunkt Veränderung des *CASA-Q Husten-Symptom Score* fand sich in der gepoolten Analyse der INPULSIS Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *Region* (Tabelle 4-79). Für keine der Vergleichsgruppen bestehend aus Patienten, die in Europa (adjustierte MWD [95%-KI]: 0,46 [−3,03; 3,95]), Amerika (adjustierte MWD [95%-KI]: 5,11 [−0,35; 10,58]) oder Asien/Australien (adjustierte MWD [95%-KI]: −0,99 [−5,32; 3,34]) eingeschlossen wurden, zeigte sich dabei ein statistisch signifikanter Unterschied.

Subgruppenergebnisse – EQ-5D-VAS

Für den Endpunkt Veränderung des *EQ-5D-VAS* fand sich in der gepoolten Analyse der INPULSIS Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *Region* (Tabelle 4-79). Bei Patienten aus Asien/Australien zeigte sich kein statistisch signifikanter Unterschied zwischen den Vergleichsgruppen (adjustierte MWD [95%-KI]: 0,80 [−2,82; 4,42]). Bei Patienten, die in Europa oder Amerika eingeschlossen wurden, zeigte sich ein statistisch signifikanter Behandlungseffekt zugunsten von Nintedanib gegenüber der Vergleichstherapie (adjustierte MWD [95%-KI]: 3,24 [0,44; 6,04], bzw. 10,44 [5,41; 15,47]). Der Unterschied bei Patienten in Amerika ist als klinisch relevant zu bewerten (Hedges' g: 0,663 [0,34; 0,98]). Unter der Therapie mit Nintedanib verschlechterte sich im Mittel der Gesundheitszustand der Patienten weniger als unter der Vergleichstherapie.

Weiterhin zeigte sich in der gepoolten Analyse der INPULSIS Studien eine Effektmodifikation durch den Faktor *Studie*, was auf eine heterogene Datenlage hindeutet (Tabelle 4-79). Die Behandlungseffekte waren sowohl in der INPULSIS 1 Studie (adjustierte MWD [95%-KI]: 3,09 [0,14; 6,03]) als auch in der INPULSIS 2 Studie (adjustierte MWD [95%-KI]: 4,39 [1,59; 7,20]) statistisch signifikant. Die klinische Relevanz der Unterschiede ist in beiden Einzelstudien fraglich (Hedges' g: 0,203 [0,01; 0,40] bzw. 0,292 [0,11; 0,48]). In beiden Studien verschlechterte sich im Mittel der Gesundheitszustand der Patienten unter Nintedanib weniger als unter der Vergleichstherapie.

4.3.1.3.5.4 Unerwünschte Ereignisse

Im Folgenden werden die Subgruppenanalysen für die Endpunkte zu UE berichtet. Untersucht wurden die in Abschnitt 4.2.5.5 beschriebenen Merkmale. Basis der Subgruppenuntersuchung ist die gepoolte Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187. Für Studie 1199.187 wurde die gesamte kontrollierte Periode berücksichtigt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

Tabelle 4-80: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zu unerwünschten Ereignissen – basierend auf der gepoolten Analyse der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187

Subgruppe	Gesamtzahl UE	Gesamtzahl SUE	UE, die zum Therapieabbruch führten
Geschlecht	0,725	0,555	0,014
Alter	0,528	0,645	0,156
Ethnie	0,708	0,217	0,147
Region	0,648	0,178	0,079
FVC (% des Sollwertes)	0,504	0,284	<0,001
SGRQ-Gesamtscore	0,390	0,913	0,033
Raucherstatus	0,732	0,706	0,119
Systemische Kortikosteroidtherapie	0,365	0,770	0,365
Bronchodilatator	0,101	0,682	0,094
Studie	0,356	0,443	0,159
Niereninsuffizienz	0,681	0,827	0,373
Körpergewicht	0,084	0,386	0,220
Quellen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.6.1.1, Tab. 1.5.6.2.1, Tab. 1.5.6.3.1 FVC: Forced Vital Capacity (Forcierte Vitalkapazität); SGRQ: St. George's Respiratory Questionnaire; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis			

4.3.1.3.5.4.1 Endpunkt Gesamtzahl UE

Die Subgruppenuntersuchungen für den Endpunkt Gesamtzahl UE ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale.

4.3.1.3.5.4.2 Endpunkt Gesamtzahl SUE

Die Subgruppenuntersuchungen für den Endpunkt Gesamtzahl SUE ergaben keinen Beleg für eine Effektmodifikation durch eines der untersuchten Merkmale.

4.3.1.3.5.4.3 Endpunkt UE, die zum Therapieabbruch führten

In Tabelle 4-81 sind die Ergebnisse zu UE, die zum Therapieabbruch führten separat für die Subgruppen nach *Geschlecht*, *FVC (% des Sollwertes) zu Studienbeginn* sowie *SGRQ-Gesamtscore zu Studienbeginn* dargestellt. Für diese Subgruppenmerkmale hatten die durchgeführten Interaktionstests eine signifikante ($p < 0,05$) Interaktion mit der Behandlung ergeben.

Tabelle 4-81: Subgruppenergebnisse für die UE, die zum Therapieabbruch führten für die gepoolte Auswertung der INPULSIS Studien und der Studie 1199.187

Endpunkt/ Subgruppe	Behandlung	N ^a	Anzahl Ereignisse N (%)	Behandlungseffekt Nintedanib vs. Placebo	
				RR ^b	95%-KI
UE, die zum Therapieabbruch führten					
Geschlecht				p-Wert Interaktionsterm ^c :	0,014
männlich	Nintedanib	552	87 (15,8)	1,23	[0,88; 1,71]
	Placebo	371	47 (12,7)		
weiblich	Nintedanib	142	44 (31,0)	3,00	[1,65; 5,47]
	Placebo	109	11 (10,1)		
FVC (% des Sollwertes) zu Studienbeginn				p-Wert Interaktionsterm ^c :	<0,001
≤70%	Nintedanib	229	35 (15,3)	0,75	[0,49; 1,14]
	Placebo	177	36 (20,3)		
>70%	Nintedanib	465	96 (20,6)	2,78	[1,79; 4,32]
	Placebo	303	22 (7,3)		
SGRQ-Gesamtscore zu Studienbeginn				p-Wert Interaktionsterm ^c :	0,033
≤40	Nintedanib	360	58 (16,1)	1,88	[1,18; 3,00]
	Placebo	255	22 (8,6)		
>40	Nintedanib	320	71 (22,2)	1,33	[0,92; 1,93]
	Placebo	221	35 (15,8)		

Endpunkt/ Subgruppe	Behandlung	N ^a	Anzahl Ereignisse N (%)	Behandlungseffekt Nintedanib vs. Placebo	
				RR ^b	95 %- KI
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.6.3.1					
a: Anzahl Studienteilnehmer in Subgruppe					
b: Ein RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an. Berechnet mittels Mantel-Haenszel Methode.					
c: Berechnet mittels Breslow-Day-Test					
FVC: Forced Vital Capacity (Forcierte Vitalkapazität); KI: Konfidenzintervall; RR: Risk Ratio (Relatives Risiko); SGRQ: St. George's Respiratory Questionnaire; UE: Unerwünschtes Ereignis					

Für den Sicherheits-Endpunkt *UE, die zum Therapieabbruch führten* fand sich in der gepoolten Analyse aller drei Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *Geschlecht* (Tabelle 4-81). Das RR für das Auftreten eines solchen UE war bei Frauen (RR [95%-KI]: 3,00 [1,65; 5,47]) höher als bei Männern. Bei isolierter Betrachtung der Subgruppe der Männer fand sich keine statistisch signifikante Zunahme der Wahrscheinlichkeit des Auftretens von *UE, die zum Therapieabbruch führten*. Allerdings lag der Punktschätzer für das RR auch in der Subgruppe der Männer mit 1,23 über 1 (RR [95%-KI]: 1,23 [0,88; 1,71]), sodass sich das Subgruppenmerkmal *Geschlecht* nicht als qualitativer, sondern als quantitativer Effektmodifikator auswirkt. Darüberhinaus zeigte sich ein Beleg für eine qualitative Effektmodifikation durch das Merkmal *FVC (% des Sollwertes) zu Studienbeginn*. Für Patienten mit einer *niedrigeren FVC* zu Studienbeginn (bezogen auf den Sollwert) traten Therapieabbrüche aufgrund unerwünschter Ereignisse im Nintedanib-Arm gegenüber dem Kontrollarm seltener auf; der Unterschied zwischen den Behandlungsarmen ist jedoch nicht statistisch signifikant (RR [95%-KI]: 0,75 [0,49; 1,14]). Für Patienten mit einer höheren FVC (>70%) waren Therapieabbrüche im Kontrollarm seltener als im Nintedanib-Arm. Der Unterschied ist hier statistisch signifikant (RR [95%-KI]: 2,78 [1,79; 4,32]). Weiterhin zeigte sich ein Beleg für eine Effektmodifikation durch das Merkmal *SGRQ-Gesamtscore zu Studienbeginn*. Für Patienten, die zu Studienbeginn einen niedrigeren *SGRQ-Gesamtscore* (≤ 40) hatten, zeigte sich bei *UE, die zum Therapieabbruch führten* ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen (RR [95%-KI]: 1,88 [1,18; 3,00]). Für Patienten mit einem höheren *SGRQ-Gesamtscore zu Studienbeginn* war der Unterschied im Auftreten von *UE, die zum Therapieabbruch führten* gegenüber dem Kontrollarm nicht statistisch signifikant (RR [95%-KI]: 1,33 [0,92; 1,93]).

4.3.1.3.5.4 Endpunkt UE von besonderem InteresseTabelle 4-82: Übersicht zu den p-Werten der Interaktionstests für die Endpunkte zu *UE von besonderem Interesse*

Subgruppe	Diarrhö	Übelkeit	Erbrechen	Appetit- losigkeit	Kardiale Ereignisse	Progression IPF	Husten	Dyspnoe	Bronchitis	Naso- pharyngitis	Blutungen
Geschlecht	0,412	0,076	0,255	0,021	0,759	0,395	0,287	0,751	0,374	0,693	0,807
Alter	0,182	0,278	0,797	0,251	0,456	0,279	0,402	0,163	0,445	0,140	0,839
Ethnie	0,715	0,066	0,700	0,174	0,250	0,683	0,237	0,304	0,309	0,658	0,666
Region	0,644	0,060	0,627	0,038	0,380	0,718	0,451	0,566	0,331	0,444	0,661
FVC (% des Sollwertes)	0,334	0,226	0,772	0,316	0,651	0,654	0,401	0,396	0,601	0,524	0,757
SGRQ	0,066	0,194	0,308	0,066	0,774	0,580	0,582	0,599	0,617	0,355	0,779
Raucherstatus	0,148	0,016	0,563	0,151	0,341	0,684	0,120	0,291	0,341	0,776	0,587
Systemische Kortikosteroide	0,067	0,487	0,692	0,166	0,195	0,183	0,740	0,445	0,715	0,517	0,955
Bronchodilatator	0,298	0,237	0,482	0,190	0,261	0,502	0,694	0,574	0,483	0,217	0,212
Studie	0,482	0,037	0,488	0,120	0,496	0,230	0,291	0,417	0,578	0,768	0,715
Nieren- insuffizienz	0,497	0,029	0,356	0,296	0,823	0,447	0,664	0,753	0,757	0,602	0,269
Körpergewicht	0,789	0,028	0,701	0,242	0,707	0,367	0,423	0,511	0,845	0,656	0,902
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.6.5.1 FVC: Forced Vital Capacity (Forcierte Vitalkapazität); IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); SGRQ: St. George's Respiratory Questionnaire; Unerwünschtes Ereignis											

Tabelle 4-83: Subgruppenergebnisse für die *UE von besonderem Interesse* für die gepoolte Auswertung der Studien INPULSIS 1, INPULSIS 2 und Studie 1199.187

Endpunkt/ Subgruppe	Behandlung	N ^a	Anzahl Ereignisse n (%)	Behandlungseffekt Nintedanib vs. Placebo	
				RR	95%-KI
Übelkeit					
Raucherstatus				p-Wert Interaktionsterm ^c : 0,016	
Nichtraucher	Nintedanib	188	53 (28,2)	4,55	[2,32; 8,94]
	Placebo	139	9 (6,5)		
Ex-Raucher/ Raucher	Nintedanib	506	119 (23,5)	2,60	[1,79; 3,78]
	Placebo	341	32 (9,4)		
Studie				p-Wert Interaktionsterm ^b : 0,037	
INPULSIS 1	Nintedanib	309	70 (22,7)	3,85	[2,14; 6,92]
	Placebo	204	12 (5,9)		
INPULSIS 2	Nintedanib	329	86 (26,1)	3,58	[2,16; 5,93]
	Placebo	219	16 (7,3)		
Studie 1199.187	Nintedanib	56	16 (28,6)	1,25	[0,67; 2,36]
	Placebo	57	13 (22,8)		
Niereninsuffizienz				p-Wert Interaktionsterm ^b : 0,029	
ja	Nintedanib	359	95 (26,5)	3,57	[2,26; 5,63]
	Placebo	269	20 (7,4)		
nein	Nintedanib	334	77 (23,1)	2,58	[1,60; 4,17]
	Placebo	210	21 (10,0)		
Körpergewicht				p-Wert Interaktionsterm ^c : 0,028	
<65kg	Nintedanib	129	33 (25,6)	6,75	[2,20; 20,69]
	Placebo	80	3 (3,8)		
≥65kg	Nintedanib	565	139 (24,6)	2,68	[1,91; 3,78]
	Placebo	400	38 (9,5)		
Appetitlosigkeit					
Geschlecht				p-Wert Interaktionsterm ^c : 0,021	
männlich	Nintedanib	552	56 (10,1)	1,47	[0,94; 2,29]
	Placebo	371	26 (7,0)		
weiblich	Nintedanib	142	24 (16,9)	11,08	[2,38; 51,51]
	Placebo	109	2 (1,8)		

Endpunkt/ Subgruppe	Behandlung	N ^a	Anzahl Ereignisse n (%)	Behandlungseffekt Nintedanib vs. Placebo	
				RR	95 %-KI
Region				p-Wert Interaktionsterm ^c : 0,038	
Europa/ mittlerer Osten	Nintedanib	323	32 (9,9)	2,80	[1,32; 5,97]
	Placebo	227	8 (3,5)		
Amerika	Nintedanib	165	21 (12,7)	1,62	[0,80; 3,26]
	Placebo	118	10 (8,5)		
Asien/ Australien	Nintedanib	206	27 (13,1)	1,79	[0,90; 3,56]
	Placebo	135	10 (7,4)		
Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a): Tab. 1.5.6.5.1					
a: Anzahl Studienteilnehmer in Subgruppe					
b: Ein RR kleiner eins zeigt ein geringeres Risiko ein Ereignis zu erleiden unter Nintedanib im Vergleich zu Placebo an.					
Berechnet mittels Mantel-Haenszel Methode					
c: berechnet mittels Breslow-Day Test					
KI: Konfidenzintervall; RR: Risk Ratio (Relatives Risiko); UE: Unerwünschtes Ereignis					

Für den Endpunkt *Übelkeit* fanden sich in der gepoolten Analyse aller drei Studien Belege für eine Effektmodifikation durch die Subgruppenmerkmale *Raucherstatus*, *Körpergewicht* sowie *Niereninsuffizienz zu Studienbeginn*. Überdies zeigte sich eine Effektmodifikation durch den Faktor *Studie*, was auf eine heterogene Studienlage hindeutet (Tabelle 4-83). Die Behandlungsunterschiede sind in allen Subgruppenkategorien nach *Raucherstatus*, *Körpergewicht* und *Nierenfunktionsstatus zu Studienbeginn* statistisch signifikant zuungunsten von Nintedanib gegenüber Placebo; die Effekte unterscheiden sich in ihrer Größenordnung. Der Effekt ist größer bei Nichtrauchern, Patienten mit geringerem Körpergewicht und bei Patienten, die zu Studienbeginn eine Niereninsuffizienz aufwiesen. Bei der Untersuchung der *Studien-Effekte* zeigte sich jeweils ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen in den beiden INPULSIS Studien, während in Studie 1199.187 kein statistisch signifikanter Unterschied im Auftreten von *Übelkeit* zwischen den Behandlungsgruppen beobachtet wurde.

Für den Endpunkt *Appetitlosigkeit* fand sich in der gepoolten Analyse aller drei Studien ein Beleg für eine Effektmodifikation durch das Subgruppenmerkmal *Geschlecht* und *Region* (Tabelle 4-83). Bei *Männern* und *Frauen* zeigten sich gleichgerichtete Effekte. Der beobachtete Unterschied ist in der Subgruppe der *Frauen* statistisch signifikant, während er in der Subgruppe der *Männer* nicht statistisch signifikant ist. Bei Patienten, die in *Europa* behandelt wurden zeigte sich ein statistisch signifikanter Unterschied zuungunsten der Therapie mit Nintedanib verglichen mit der zVT. Bei Patienten, die in *Asien* oder *Amerika* behandelt wurden, war der zwischen den Behandlungsgruppen beobachtete Unterschied nicht statistisch signifikant.

4.3.1.3.6 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.

Der vorliegenden Nutzenbewertung für Nintedanib (Ofev®) im Anwendungsgebiet IPF liegen drei RCT zugrunde. Im Folgenden werden die Ergebnisse der IPD Meta-Analyse dieser Studien zusammengefasst. In dem Fall, dass die drei Studien nicht sinnvoll zusammengefasst werden konnten, werden Ergebnisse der Einzelstudien separat dargestellt. Ebenso werden in dem Fall, dass es Anzeichen für unterschiedliche Effekte in relevanten Subgruppen gab, Ergebnisse für die Subgruppen getrennt berichtet.

Mortalität

Für den Endpunkt *Zeit bis zum Tod, unabhängig von der Ursache* zeigt sich in der IPD Meta-Analyse der drei eingeschlossenen Studien kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Der Effektschätzer deutet auf ein reduziertes Sterberisiko unter Nintedanib gegenüber BSC hin (HR [95%-KI]: 0,65 [0,41; 1,03]; p=0,0775). Für den Endpunkt *Zeit bis zum respiratorisch bedingten Tod* zeigte sich ebenfalls kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (HR [95%-KI]: 0,71 [0,40; 1,26]; p=0,2960). Trotz der meta-analytischen Zusammenfassung der Studien waren die vorliegenden Daten aufgrund der geringen Ereigniszahl nicht geeignet, signifikante Therapieunterschiede nachzuweisen.

Für den Endpunkt *jährliche Abnahme der FVC (ml)* – als Surrogat für die Mortalität sowie als Maß für die Reduktion der Krankheitsprogression – zeigten sich sowohl in den gepoolten Analysen als auch in den Einzelstudien statistisch signifikante Unterschiede zugunsten von Nintedanib gegenüber der Vergleichstherapie. In der IPD Meta-Analyse der drei Studien war die *jährliche Abnahme der FVC* im Nintedanib-Arm im Mittel um 112,42 ml geringer als im Kontrollarm. Die Abnahme der FVC im Behandlungsarm mit Nintedanib fiel somit durchschnittlich um etwa 50% geringer aus als im Vergleichsarm. Der Unterschied ist - bezogen auf den für die Fallzahlplanung als relevant ausgewiesenen und mit den Zulassungsbehörden abgestimmten Gruppenunterschied von 100 ml - auch als klinisch relevant zu sehen. Überdies ist der beobachtete Effekt auch in seinem Ausmaß vergleichbar über alle untersuchten Subgruppen hinweg; dies gilt insbesondere auch für die Subgruppen nach FVC-Ausgangswert.

Hinsichtlich des Anteils der Patienten, die nach 52 Wochen als Responder gewertet wurden (gemessen an einer absoluten FVC-Abnahme von nicht mehr als 5- bzw. 10 Prozentpunkten gegenüber dem Ausgangswert) ergaben sich in der gepoolten Analyse der INPULSIS Studien statistisch signifikante Effekte zugunsten von Nintedanib (OR [95%-KI]: 1,577 [1,21; 2,05]; $p=0,0007$ bzw. 1,835 [1,43; 2,36]; $p<0,0001$). Die beobachtete Effektmodifikation durch den Raucherstatus der Patienten wurde nur in der Responderanalyse für den Schwellenwert von 5 Prozentpunkten beobachtet, nicht aber für die stetige Betrachtung der FVC oder den Schwellenwert von 10 Prozentpunkten und wird daher als nicht fazitrelevant erachtet.

Für die *FVC-Responderanalysen* nach 24 Wochen (INPULSIS + Studie 1199.187 gepoolt) ergaben sich für den Schwellenwert 10% (Abfall der FVC [% des Sollwertes] um weniger als 10 Prozentpunkte) keine statistisch signifikanten Effekte. Für den Schwellenwert 5% (Abfall der FVC [% des Sollwertes] um weniger als 5 Prozentpunkte) zeigte sich in den gepoolten Analysen ein statistisch signifikanter Unterschied zugunsten von Nintedanib gegenüber der Vergleichstherapie. In den Analysen der Einzelstudien ergaben sich für den Schwellenwert 10 Prozentpunkte keine signifikanten Effekte. Für den Schwellenwert von 5 Prozentpunkten ergab sich für INPULSIS 1 ein statistisch signifikanter Unterschied zugunsten von Nintedanib gegenüber der Vergleichstherapie.

Morbidität

Die Ergebnisse aller Endpunkte zur Morbidität zeigen einen Vorteil oder einen numerischen Trend zugunsten der Nintedanib-Therapie im Vergleich zur zVT. Speziell für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* ist der Unterschied statistisch signifikant zugunsten von Nintedanib.

Exazerbationen

Für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* zeigt sich in der IPD Meta-Analyse der drei eingeschlossenen Studien und der IPD Meta-Analyse der INPULSIS Studien ein statistisch signifikanter Unterschied zugunsten von Nintedanib. Der Effektschätzer der IPD Meta-Analyse der drei eingeschlossenen Studien belegt einen deutlichen Effekt auf die *Zeit bis zur ersten adjudizierten akuten Exazerbation* (HR [95%-KI]: 0,32 [0,17; 0,63]; $p=0,0010$). Dieser ist als relative Risikoreduktion von 68% für das Auftreten adjudizierter akuter Exazerbationen unter der Behandlung mit Nintedanib gegenüber BSC zu werten.

Für die *Zeit bis zur ersten akuten Exazerbation*, bewertet durch den Prüfarzt, zeigte sich für die IPD Meta-Analyse der drei eingeschlossenen Studien ein numerischer Trend zugunsten von Nintedanib gegenüber der Vergleichstherapie; der Unterschied zwischen den Behandlungsgruppen war jedoch nicht statistisch signifikant (HR [95%-KI]: 0,63 [0,39; 1,02]; $p=0,0695$). Auf Ebene der Einzelstudien zeigte sich nur für die Studie INPULSIS 2 ein statistisch signifikanter Effekt zugunsten von Nintedanib (HR [95%-KI]: 0,38 [0,19; 0,77]; $p=0,0050$).

Betrachtet man nur Exazerbationen, die als SUE eingestuft wurden, ergab sich sowohl für die *Zeit bis zur ersten schwerwiegenden adjudizierten akuten Exazerbation* (HR [95%-KI]: 0,30 [0,15; 0,62]; $p=0,0015$) als auch für die *Zeit bis zur ersten schwerwiegenden akuten Exazerbation* (HR [95%-KI]: 0,55 [0,32; 0,96]; p -Wert aus Log rank-Test stratifiziert nach Studie: $p=0,0545$; p -Wert aus Wald-Test im adjustierten Cox-Modell: $p=0,0337$) ein statistisch signifikanter Vorteil zugunsten der Therapie mit Nintedanib.

Gesundheitsbezogene Lebensqualität

Die Ergebnisse aller Endpunkte zur gesundheitsbezogenen Lebensqualität zeigen einen Vorteil oder einen numerischen Trend zugunsten einer Behandlung mit Nintedanib im Vergleich zur zVT in Form einer geringeren Verschlechterung der gesundheitsbezogenen Lebensqualität innerhalb des Studienzeitraums.

Endpunkte zum SGRQ und SGRQ-I

Für die gesundheitsbezogene Lebensqualität gemessen über den SGRQ-Gesamtscore zeigte sich in der IPD Meta-Analyse der beiden INPULSIS Studien über den Zeitraum von 52 Wochen auf Basis einer MMRM-Auswertung ein numerischer Trend zugunsten von Nintedanib gegenüber der Vergleichstherapie; der beobachtete Unterschied ist jedoch nicht statistisch signifikant (MWD [95%-KI]: -1,43 Punkte [-3,09; 0,23]; $p=0,0923$). Für den IPF-spezifischen Fragebogen SGRQ-I zeigte sich in der gepoolten Analyse der INPULSIS Studien ein statistisch signifikanter Unterschied zugunsten von Nintedanib gegenüber der Vergleichstherapie (MWD [95%-KI]: -1,99 Punkte [-3,71; -0,27]; $p=0,0230$) (Tabelle 4-41).

Aufgrund einer potenziellen Effektmodifikation durch das Merkmal *FVC zu Studienbeginn* wurden die Ergebnisse für die Veränderung des SGRQ-I Gesamtscores für die entsprechenden Subgruppen getrennt betrachtet. Für die Subgruppe der Patienten mit einem FVC-Ausgangswert $\leq 70\%$ zu Studienbeginn zeigte sich ein statistisch signifikanter Vorteil in der gesundheitsbezogenen Lebensqualität gemessen anhand des SGRQ-I-Gesamtscores für Nintedanib gegenüber der Vergleichstherapie (adjustierte MWD [95%-KI]: -3,86 [-6,89; -0,83]). Die klinische Relevanz des beobachteten Unterschieds blieb jedoch unklar (Hedges' g [95%-KI]: -0,295 [-0,53; -0,06]). Der Behandlungsunterschied in der Subgruppe der Patienten mit einer höheren FVC zu Studienbeginn ($>70\%$ des Sollwertes) ist nicht statistisch signifikant.

Die Analyse der *SGRQ-(≤ -4 Punkte)-Responder* zu Woche 52 auf Basis der gepoolten INPULSIS Studien zeigte in der Nintedanib-Gruppe einen höheren Anteil an Studienteilnehmern mit einer klinisch relevanten Verbesserung der gesundheitsbezogenen Lebensqualität als im Vergleichsarm (22,88% vs. 20,33% Responder); der Unterschied war nicht statistisch signifikant (OR [95%-KI]: 1,186 [0,88; 1,61]; $p=0,2715$). Aufgrund von Belegen für eine potenzielle Effektmodifikation durch die Merkmale *FVC zu Studienbeginn*, *Bronchodilatator zu Studienbeginn* sowie durch das Merkmal *Studie* wurden die entsprechenden Subgruppen separat betrachtet. Für die Subgruppe der Studienteilnehmer mit einer niedrigeren Vitalkapazität zu Studienbeginn ($FVC \leq 70\%$) zeigte sich ein statistisch

signifikanter Effekt zugunsten von Nintedanib (OR [95%-KI]: 1,934 [1,11; 3,38]), während der Behandlungsunterschied in der Subgruppe der Patienten mit einer höheren Vitalkapazität zu Studienbeginn (*FVC* >70%) nicht signifikant war (OR [95%-KI]: 0,917 [0,63; 1,33]).

Bei Patienten mit *Bronchodilatator* zu Studienbeginn zeigte sich ein statistisch signifikanter Unterschied in den Anteilen der *SGRQ-Responder* zu Woche 52 zugunsten von Nintedanib gegenüber der Vergleichstherapie (OR [95%-KI]: 7,236 [2,12; 24,68]). Bei Patienten ohne eine solche Begleittherapie zu Studienbeginn zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 0,976 [0,70; 1,35]).

Bei separater Betrachtung der INPULSIS Studien zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen in INPULSIS 1 (OR [95%-KI]: 0,840 [0,55; 1,29]). In INPULSIS 2 ergab sich ein statistisch signifikanter Unterschied zugunsten von Nintedanib (OR [95%-KI]: 1,664 [1,08; 2,57]).

Die Analyse der *SGRQ-I* (≤ 4 Punkte)-*Responder* zu Woche 52 auf Basis der gepoolten INPULSIS Daten zeigt, dass unter der Behandlung mit Nintedanib ebenfalls mehr Studienteilnehmer eine klinisch relevante Verbesserung der gesundheitsbezogenen Lebensqualität angeben als unter der Vergleichstherapie (23,51% vs. 19,62%) (OR [95%-KI]: 1,284 [0,95; 1,74]; $p=0,1091$). Aufgrund von Belegen für eine potenzielle Effektmodifikation durch die Merkmale *Bronchodilatator* zu Studienbeginn sowie durch das Merkmal *Studie* wurden die entsprechenden Subgruppen separat betrachtet.

Bei Patienten mit *Bronchodilatator* zu Studienbeginn zeigte sich ein statistisch signifikanter Unterschied in den Anteilen der *SGRQ-I-Responder* zu Woche 52 zugunsten von Nintedanib (28,2% vs. 8,3%) gegenüber der Vergleichstherapie (OR [95%-KI]: 4,232 [1,68; 10,67]). Bei Patienten ohne eine solche Begleittherapie zu Studienbeginn zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (OR [95%-KI]: 1,066 [0,76; 1,49]). Bei separater Betrachtung der INPULSIS Studien zeigten sich gegenläufige Effekte in den Einzelstudien. Während der Behandlungseffekt in Studie INPULSIS 1 nicht statistisch signifikant war (OR [95%-KI]: 0,886 [0,58; 1,36]), war der Anteil der *SGRQ-I-Responder* zu Woche 52 in Studie INPULSIS 2 im Nintedanib-Arm signifikant höher als im Kontrollarm (OR [95%-KI]: 1,877 [1,20; 2,93]).

Weitere Endpunkte zur krankheitsspezifischen Lebensqualität

Für die Veränderung des *SOBQ-Gesamtscores*, der die subjektiv empfundene krankheitsspezifische Lebensqualität hinsichtlich Kurzatmigkeit misst, zeigte sich in der IPD Meta-Analyse der beiden INPULSIS Studien über den Zeitraum von 52 Wochen auf Basis einer MMRM-Auswertung eine geringere Verschlechterung unter Nintedanib im Vergleich zur Kontrollgruppe (6,71 Punkte vs. 8,40 Punkte; ein höherer Wert entspricht einer stärkeren Verschlechterung der Kurzatmigkeit). Der Unterschied ist nicht statistisch signifikant (MWD [95%-KI]: -1,69 Punkte [-4,09; 0,70]; $p=0,1648$) (Tabelle 4-45). In der IPD Meta-Analyse aller drei Studien für die Veränderung des *SOBQ-Gesamtscores* nach

24 Wochen ergab sich ebenfalls kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen (MWD [95%-KI]: 0,58 [-1,45; 2,61]; $p=0,5761$).

Auch die Ergebnisse zur *Veränderung des CASA-Q* (Wertebereich: 0-100; ein hoher Wert entspricht einer geringeren Beeinträchtigung) in der gepoolten Auswertung der INPULSIS Studien zeigten hinsichtlich der *Veränderung des CASA-Q Husten-Impact bis Woche 52* eine geringere Verschlechterung der krankheitsspezifischen Lebensqualität unter der Behandlung mit Nintedanib im Vergleich zur zVT (Impact: MWD [95%-KI]: 1,73 Punkte [-0,46; 3,92]; $p=0,1212$; Symptome: MWD [95%-KI]: 0,95 Punkte [-1,49; 3,38]; $p=0,4450$). Der Unterschied ist nicht statistisch signifikant, bestätigt jedoch in der Tendenz die Ergebnisse der anderen Auswertungen zur krankheitsspezifischen Lebensqualität (Tabelle 4-47). In Studie 1199.187 wurde der CASA-Q nicht verwendet.

Aufgrund eines Belegs für eine Effektmodifikation durch das Subgruppenmerkmal *Region* für den Endpunkt *Veränderung des CASA-Q Husten-Symptom-Scores bis Woche 52* wurden die entsprechenden Subgruppen getrennt betrachtet. In keiner der drei Regionen zeigte sich ein statistisch signifikanter Behandlungseffekt. Während bei Patienten aus den Studienzentren in Asien/Australien eher ein Trend zugunsten der Vergleichstherapie zu beobachten war, zeigte sich für Patienten aus Europa und Amerika ein Trend zugunsten von Nintedanib.

Endpunkte zur allgemeinen gesundheitsbezogenen Lebensqualität

Der Anteil *PGI-C-Responder*, definiert als Studienteilnehmer, die ihre gesundheitsbezogene Lebensqualität in Woche 52 im Vergleich zum Ausgangswert als „sehr viel besser“, „viel besser“, „ein wenig besser“ oder „ohne Veränderung“ einschätzen, ist in der gepoolten Auswertung der INPULSIS Studien unter Nintedanib höher als unter der Vergleichstherapie (61,29% vs. 54,37%). Der Unterschied ist statistisch signifikant (OR [95%-KI]: 1,328 [1,04; 1,70]; $p=0,0254$) (Tabelle 4-48). In Studie 1199.187 wurde das PGI-C Instrument nicht eingesetzt. Es ergaben sich keine Belege für eine Effektmodifikation durch eines der untersuchten Merkmale.

Auch die Veränderung der *EQ-5D-VAS* vom Ausgangswert bis Woche 52 zeigte in den gepoolten INPULSIS Studien eine geringere Verschlechterung des allgemeinen Gesundheitszustandes unter Nintedanib im Vergleich zur zVT (2,70 Punkte vs. 6,51 Punkte) (Tabelle 4-49). Der beobachtete Unterschied ist statistisch signifikant, jedoch von unklarer klinischer Relevanz (adjustierte MWD [95%-KI]: 3,81 [1,78; 5,85]; $p=0,0002$; Hedges' g [95%-KI]: 0,252 [0,12; 0,39]). In Studie 1199.187 wurden keine Daten zum EQ-5D erhoben.

Aufgrund eines Belegs für eine potenzielle Effektmodifikation durch die Merkmale *Region* wurden die Ergebnisse zur EQ-5D-VAS separat für die Regionen Europa, Amerika und Asien/Australien betrachtet. Während sich für Patienten aus den Studienzentren in Asien/Australien kein statistisch signifikanter Unterschied bzgl. der Änderung des Gesundheitszustandes zwischen den Behandlungsgruppen zeigte, ergaben sich für Patienten

aus Europa und Amerika statistisch signifikante Unterschiede zugunsten von Nintedanib gegenüber der Vergleichstherapie (MWD [95%-KI]: 3,24 [0,44; 6,04], Hedges' g: 0,217 [0,03; 0,40] bzw. 10,44 [5,41; 15,47], Hedges' g: 0,663 [0,34; 0,98]) (Tabelle 4-79).

Zudem fand sich ein Hinweis auf eine Effektmodifikation durch das Merkmal *Studie*. In beiden INPULSIS Studien zeigte sich eine statistisch signifikant geringere Verschlechterung des allgemeinen Gesundheitszustandes unter Nintedanib im Vergleich zur zVT (MWD [95%-KI] INPULSIS 1: 3,09 [0,14; 6,03], Hedges' g: 0,203 [0,01; 0,40], bzw. INPULSIS 2: 4,39 [1,59; 7,20], Hedges' g: 0,292 [0,11; 0,48]).

Unerwünschte Ereignisse

Nintedanib ist ein Wirkstoff, welcher neben einer sehr guten Effektivität zur Behandlung der IPF im Allgemeinen ein gutes Verträglichkeitsprofil zeigt. Gastrointestinale Ereignisse traten vermehrt auf, dabei handelte es sich in den meisten Fällen um milde bis moderate Ereignisse, die nur mit geringer bis mäßiger Beeinträchtigung der Studienteilnehmer verbunden waren und durch Dosisreduktion oder medikamentöse Behandlungen gut handhabbar waren. Insgesamt brachen über alle drei Studien hinweg 18,9% der Studienteilnehmer des Nintedanib-Arms im Vergleich zu 12,1% der Studienteilnehmer des Kontrollarms die Behandlung wegen unerwünschter Ereignisse ab (Abschnitt 4.3.1.3.4).

Gesamtzahl UE

UE traten in der gepoolten Analyse der drei Studien in der Nintedanib-Gruppe mit 95,7% der Studienteilnehmer statistisch signifikant häufiger auf als in der Kontrollgruppe (89,8%) (RR [95%-KI]: 1,07 [1,03; 1,10]; $p < 0,001$). Hierbei handelte es sich allerdings um Ereignisse von heterogener klinischer Relevanz. Es ergaben sich keine Belege für eine Effektmodifikation durch eines der untersuchten Merkmale.

Schwerwiegende UE

SUE traten in der gepoolten Analyse der drei Studien bei 29,1% der Studienteilnehmer des Nintedanib-Arms auf und bei 28,3% der Studienteilnehmer des Vergleichsarms (Tabelle 4-57). Der Unterschied zwischen den Gruppen war nicht statistisch signifikant (RR [95%-KI]: 1,01 [0,84; 1,21]; $p = 0,941$). Es ergaben sich keine Belege für eine Effektmodifikation durch eines der untersuchten Merkmale.

UE, die zum Therapieabbruch führten

UE, die zum Therapieabbruch führten traten in der gepoolten Analyse aller drei Studien bei 18,9% der Studienteilnehmer, die mit Nintedanib behandelt wurden verglichen mit 12,1% der Studienteilnehmer der Kontrollgruppe auf (Tabelle 4-62). Der Unterschied zwischen den Gruppen ist statistisch signifikant (RR [95%-KI]: 1,54 [1,15; 2,04]; $p = 0,003$).

Aufgrund von Belegen für eine Effektmodifikation durch die Merkmale *Geschlecht*, *FVC zu Studienbeginn* sowie *SGRQ-Gesamtscore zu Studienbeginn* wurden die Ergebnisse für die entsprechenden Subgruppen getrennt betrachtet. Für die Subgruppe der *Frauen*, der *Patienten mit einer höheren FVC zu Studienbeginn* sowie für die Subgruppe der Patienten mit einem *niedrigeren SGRQ-Gesamtscore zu Studienbeginn* war der Anteil der Patienten mit *UE, die zum Therapieabbruch führten* im Nintedanib-Arm verglichen mit dem Placebo-Arm signifikant höher. In der Subgruppe der *Männer*, der *Patienten mit einem höheren SGRQ-Gesamtscore zu Studienbeginn* und der *Patienten mit einer niedrigen FVC* lag kein statistisch signifikanter Unterschied im Anteil der Patienten mit *UE, die zum Therapieabbruch führten* vor. In der Subgruppe der Patienten mit einer *geringeren FVC* war der Effekt gegenläufig.

UE von besonderem Interesse

Gastrointestinale Störungen

Diarrhö

Häufigste *gastrointestinale UE* waren *Diarrhöen*, die in den gepoolten INPULSIS Studien und Studie 1199.187 unter Nintedanib häufiger auftraten als in der Vergleichstherapie (62,4% vs. 20,6%) mit statistisch signifikantem Unterschied ($p < 0,001$) (Tabelle 4-67). *Diarrhöen* traten erstmalig überwiegend in den ersten drei Monaten der Behandlung auf und waren bei den meisten Studienteilnehmern von leichter bis moderater Intensität. Zu einer Einschränkung in der Ausübung von Alltagsaktivitäten oder zu Arbeitsunfähigkeit (Intensität „schwer“) kam es bei 4,8% der Nintedanib-Studienteilnehmer mit *Diarrhö* (entsprechend 3,0% aller Studienteilnehmer der Nintedanib-Gruppe). 78,5% der Studienteilnehmer mit *Diarrhö* konnten die Therapie mit der Normaldosis fortsetzen, weitere 14,1% der Studienteilnehmer mit reduzierter Dosis. Lediglich bei 7,4% der Studienteilnehmer mit *Diarrhö* (entsprechend 4,6% aller Studienteilnehmer der Nintedanib-Gruppe) kam es aufgrund dieses *UE* zu einem endgültigen Therapieabbruch (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a).

Übelkeit

Übelkeit wurde in den gepoolten INPULSIS Studien und Studie 1199.187 unter der Behandlung mit Nintedanib häufiger berichtet als unter der Vergleichstherapie (24,8% vs. 8,5%). Der Unterschied ist statistisch signifikant (RR [95%-KI]: 3,02 [2,18; 4,18]; $p < 0,001$) (Tabelle 4-67).

Erbrechen

Auch *Erbrechen* wurde im Nintedanib-Behandlungsarm (12,0%) häufiger berichtet als im Kontrollarm (2,9%). Der beobachtete Effekt ist statistisch signifikant (RR [95%-KI]: 4,20 [2,40; 7,35]; $p < 0,001$) (Tabelle 4-67).

Appetitlosigkeit

Appetitlosigkeit wurde in der gepoolten Auswertung der INPULSIS Studien und Studie 1199.187 im Nintedanib-Behandlungsarm (11,5%) häufiger dokumentiert als im

Vergleichsarm (5,8%). Der Unterschied ist statistisch signifikant (RR [95%-KI]: 2,02 [1,34; 3,05]; $p < 0,001$) (Tabelle 4-67).

Gewichtsverlust

Gewichtsverlust trat im Nintedanib-Behandlungsarm (10,4%) häufiger auf als im Kontrollarm (3,1%). Der beobachtete Effekt ist statistisch signifikant (RR [95%-KI]: 3,27 [1,91; 5,61]; $p < 0,001$) (Tabelle 4-67).

Gastrointestinale Perforation

Gastrointestinale Perforationen traten insgesamt so selten auf, dass keine Aussage zu einem möglichen Risiko vorgenommen werden kann (Tabelle 4-67).

Unerwünschte Leberenzym- und Bilirubinerhöhungen

UE, welche das Leber- und Gallenwegssystem betrafen wurden in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 häufiger im Nintedanib-Arm als im Vergleichsarm dokumentiert (Tabelle 4-68). Die Behandlung mit Nintedanib war im Vergleich zur zVT häufiger verbunden mit einer Erhöhung der Leberenzyme (ALT, AST, AP) oder des Bilirubin-Wertes.

- AST- und ALT-Werte um mindestens 3x ULN erhöht (Nintedanib vs. zVT): 4,9% vs. 0,8% (RR [95%-KI]: 5,99 [2,10; 17,07]; $p < 0,001$)
- AST- und ALT-Werte um mindestens 5x ULN erhöht (Nintedanib vs. zVT): 2,0% vs. 0,2% (RR [95%-KI]: 9,30 [1,23; 70,53]; $p = 0,008$)
- AST- und ALT-Werte um mindestens 8x ULN erhöht (Nintedanib vs. zVT): 0,7% vs. 0,2% (RR [95%-KI]: 3,32 [0,39; 28,36]; $p = 0,244$)
- Bilirubinwert um mindestens 1,5x ULN erhöht (Nintedanib vs. zVT): 2,3% vs. 0,6% (RR [95%-KI]: 3,60 [1,06; 12,17]; $p = 0,027$)
- Bilirubinwert um mindestens 2x ULN erhöht (Nintedanib vs. zVT): 0,4% vs. 0,4% (RR [95%-KI]: 1,00 [0,17; 5,94]; $p = 0,997$)
- AP-Wert um mindestens 1,5x ULN erhöht (Nintedanib vs. zVT): 5,6% vs. 1,0% (RR [95%-KI]: 5,44 [2,13; 13,87]; $p < 0,001$)
- AP-Wert um mindestens 2x ULN erhöht (Nintedanib vs. zVT): 2,4% vs. 0,2% (RR [95%-KI]: 11,29 [1,51; 84,64]; $p = 0,003$)

Unerwünschte kardiale Ereignisse

Unerwünschte kardiale Ereignisse wurden in der gepoolten Auswertung der INPULSIS Studien und Studie 1199.187 vergleichbar häufig im Nintedanib-Arm und im Kontrollarm (9,7% vs. 9,6%) beobachtet (Tabelle 4-69). Auch *schwerwiegende unerwünschte kardiale Ereignisse* traten in der Nintedanib-Gruppe (4,6%) und Kontrollgruppe (4,8%) vergleichbar häufig auf. *Tödliche kardiale Ereignisse* traten numerisch seltener in der

Nintedanib-Gruppe (0,4%) als in der Kontrollgruppe (1,3%) auf (RR [95%-KI]: 0,33 [0,08; 1,32]; p=0,100).

Arterielle thromboembolische Ereignisse wurden in den gepoolten INPULSIS Studien und Studie 1199.187 häufiger unter Nintedanib als in der Vergleichstherapie (2,4% vs. 0,6%) beobachtet (Tabelle 4-69). *Myokardinfarkte* traten numerisch häufiger unter Nintedanib als unter der Vergleichstherapie (1,6% vs. 0,4%) auf.

Unerwünschte respiratorische Störungen

Respiratorische Störungen traten in der gepoolten Analyse der INPULSIS Studien seltener unter der Behandlung mit Nintedanib auf im Vergleich zur zVT (Tabelle 4-70). Unter respiratorischen Störungen sind z. B. Progression der IPF, Husten, Dyspnoe, Bronchitis etc. subsummiert. Dabei wurde eine Progression der IPF als häufigstes Ereignis in den der Bewertung zugrunde liegenden Studien dokumentiert.

Unerwünschte Blutungsereignisse

Blutungsereignisse fanden sich bei 9,8% der Patienten im Nintedanib-Arm und bei 7,3% der Patienten im Vergleichsarm (Tabelle 4-71). Der Gruppenunterschied war nicht statistisch signifikant. Nicht schwerwiegende Epistaxis und Kontusionen stellten die häufigsten *Blutungsereignisse* dar. Die Häufigkeit *schwerwiegender Blutungsereignisse* war in beiden Behandlungsgruppen gering und vergleichbar (zVT: 1,5%; Nintedanib: 1,2%).

Unerwünschte Ereignisse >10%

Es wurden keine *UE* auf MedDRA PT-Level identifiziert, die nicht bereits in einer der oben genannten Kategorien nach Organsystemen berücksichtigt wurden.

Die wenigen Effektmodifikationen, die sich auf der Basis von Interaktionstests ergaben, waren vor allem quantitativer Natur. Während sich die beobachteten Effekte meistens in der Größenordnung, nicht aber hinsichtlich der statistischen Signifikanz unterschieden, lag für das Ereignis *Appetitlosigkeit* nur in der Subgruppe der weiblichen Patienten ein statistisch signifikanter Nachteil von Nintedanib gegenüber der Vergleichstherapie vor. Bei Männern war der Unterschied nicht statistisch signifikant.

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 zutreffend

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 zutreffend

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 4-84: 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
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)					

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 4-85: 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 zutreffend

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

Tabelle 4-86: 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 4-87: 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>
ITT: Intention-to-Treat; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)						

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend

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 4-88: 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>	
RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)	

Nicht zutreffend

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 zutreffend

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 4.3.1.3.5.

Nicht zutreffend

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 zutreffend

4.3.2.2.2 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 4-89: Verzerrungsaspekte auf Studienebene – nicht-randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten	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 zutreffend

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 4-90: 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 4-91: 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
ITT: Intention-to-Treat				

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend

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 zutreffend

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 4.3.1.3.5.**

Nicht zutreffend

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 zutreffend

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 zutreffend

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 4-92: 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 zutreffend

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 zutreffend

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

4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.5.***

Nicht zutreffend

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 zutreffend

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.

Der Nutzenbewertung von Nintedanib in IPF liegen drei RCT (INPULSIS 1, INPULSIS 2 und Studie 1199.187) zugrunde. Bei allen Studien handelt es sich jeweils um Nachweise des Evidenzlevels Ib (nach § 11 2. Kapitel 3. Abschnitt der VerfO des G-BA vorgenommene Klassifizierung der Evidenzstufen). Die INPULSIS Studien sowie die Studie 1199.187 ermöglichen den direkten Vergleich von Nintedanib gegenüber der zVT und weisen sowohl endpunktübergreifend als auch in allen patientenrelevanten Endpunkten ein niedriges Verzerrungspotenzial auf. Daher ist die Aussagekraft der Nachweise als „hoch“ einzustufen.

In Absprache mit den Zulassungsbehörden und entsprechend der Forderung nach Replikation der Studienergebnisse sind die Zulassungsstudien INPULSIS 1 und INPULSIS 2 als Zwillingsstudien nach identischen CTP konzipiert und durchgeführt worden. Es war a priori geplant, beide INPULSIS Studien auf Patientenebene gepoolt im Rahmen von IPD Meta-Analysen auszuwerten. Mit der Studie 1199.187 liegt eine weitere RCT mit ähnlichem Studiendesign und gleicher Endpunktoperationalisierung vor. Unter Beachtung der Anforderungen an die Überprüfung hinsichtlich des Vorliegens von Heterogenität werden die drei Studien meta-analytisch auf Basis der IPD zusammengefasst und zur Ableitung des Zusatznutzens herangezogen. Durch die IPD meta-analytische Zusammenfassung liegt Evidenz der Stufe 1a vor. Die vorhandene Evidenz ist daher methodisch geeignet, Belege für einen Zusatznutzen zu generieren.

Die Untersuchung des Zusatznutzens wurde anhand der folgenden Endpunkte durchgeführt:

- Mortalität
 - *Zeit bis zum Tod (unabhängig von der Ursache bzw. respiratorisch bedingt)*
 - *Jährliche FVC-Abnahme (ml) und FVC-Responder*
- Morbidität
 - *Zeit bis zur ersten adjudizierten akuten Exazerbation (bewertet durch ein IAC) und Zeit bis zur ersten akuten Exazerbation (bewertet durch den Prüfarzt)*
 - *HRCT-QLF Score*
 - *6 MWT*
 - *Zeit bis zur ersten respiratorisch bedingten Hospitalisierung*

- Gesundheitsbezogene Lebensqualität
 - SGRQ-Veränderung und $-(\leq 4 \text{ Punkte})$ -Responder
 - SGRQ-I-Veränderung und $-(\leq 4 \text{ Punkte})$ -Responder
 - Veränderung des SOBQ
 - Veränderung des CASA-Q
 - PGI-C-Responder
 - Veränderung der EQ-5D-VAS
- Unerwünschte Ereignisse
 - UE
 - SUE
 - UE, die zum Therapieabbruch führten
 - UE von besonderem Interesse (gastrointestinale Störungen, Leberenzymwerte- und Bilirubinerhöhungen, kardiale Ereignisse, respiratorische Störungen, Blutungsereignisse, weitere häufige UE, >10%)

Mortalität

Die Bestimmung der FVC ist ein validiertes Verfahren, welches die Lungenfunktion der IPF-Patienten analysiert. FVC ist ein wichtiger klinisch relevanter Parameter, der Auskunft hinsichtlich der klinischen Progredienz der IPF-Erkrankung und der Mortalität gibt (Abschnitt 4.5.4).

Morbidität

Die Endpunkte *adjudizierte akute Exazerbationen* und *akute Exazerbationen* sind insofern patientenrelevant, da Exazerbationen mit einem Krankheitsprogress und einer erhöhten Sterblichkeit einhergehen. Kriterien zur Diagnose akuter Exazerbationen waren im CTP definiert.

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität wurde mittels validierter Messinstrumente von den Studienteilnehmern direkt vor jedem Arztbesuch bewertet. Aufgrund der Verblindung der Studienteilnehmer und behandelnden Personen kann von einer unverzerrten Messung der Symptomatik und der HRQoL ausgegangen werden. Der Zusatznutzen für diese Endpunkte wurde zusätzlich anhand der Responderanalysen abgeleitet.

Unerwünschte Ereignisse

Die Auswertungskriterien aller Endpunkte zu den unerwünschten Ereignissen waren im Prüfplan a priori definiert. Die Ergebnisse wurden gemäß Analyseplan ausgewertet und transparent berichtet. Eine ergebnisgesteuerte Publikation kann daher ausgeschlossen werden.

Alle UE wurden in den berücksichtigten Studien routinemäßig bei jeder Studienvisite ab der Randomisierung bis 28 Tage nach der letzten Einnahme der Studienmedikation erfasst.

Subgruppenanalysen

Statistisch signifikante Unterschiede zwischen den Subgruppen wurden im Dossier beschrieben und diskutiert.

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 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.

Die Bewertung des medizinischen Zusatznutzens erfolgte in Bezug auf die patientenrelevanten Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und UE.

Die Phase III-Studien INPULSIS 1 und INPULSIS 2 haben einen identischen Studienaufbau und waren für eine gepoolte Auswertung bei der Studienplanung bereits vorgesehen. Auf Basis der gepoolten Analyse wird daher eine wesentlich höhere Ergebnissicherheit im Vergleich zu den einzelnen Studien erreicht. Mit der Studie 1199.187 liegt eine weitere RCT vor, die aufgrund ihres vergleichbaren Studiendesigns sowie gleicher Endpunktoperationalisierung für die gepoolte Auswertung zusammen mit den INPULSIS Studien geeignet ist. Somit wird die Ergebnissicherheit nochmals erhöht. In der Ergebnisdarstellung und Zusatznutzenbewertung wurde deshalb auf die Ergebnisse der gepoolten Auswertung fokussiert. Im Falle von Endpunkten, für die eine gepoolte Auswertung aller drei Studien über die jeweils maximale Studiendauer nicht möglich war (FVC-, SGRQ-, SGRQ-I- und PGI-C-Responder), wird für die Ableitung des Zusatznutzens die gepoolte Auswertung der INPULSIS Studien über die längste Beobachtungsdauer (52 Wochen) herangezogen.

Mortalität

Die IPF ist eine chronische, rasch progrediente unheilbare Erkrankung. Die Mortalitätsrate ist dabei höher als bei den meisten Krebserkrankungen (Bjoraker et al., 1998; Douglas et al., 2000; Kim et al., 2006). Ein wichtiges Therapieziel ist daher eine Verlängerung des Überlebens der IPF-Patienten bei gleichzeitigem Erhalt der Lebensqualität.

In der gepoolten Analyse war die Überlebenswahrscheinlichkeit unter Nintedanib höher als unter der zVT BSC (Abbildung 2 und Abbildung 8). Das HR [95%-KI] für die *Zeit bis zum Tod, unabhängig von der Ursache* beträgt 0,65 ([0,41; 1,03]; $p=0,0775$), d. h. das Sterberisiko im Verlauf des Beobachtungszeitraums, lag bei Nintedanib-behandelten Studienteilnehmern um 35% niedriger im Vergleich zu BSC. Ein vergleichbares Bild zeigt sich auch bei der *Zeit bis zum respiratorisch bedingten Tod*.

Der Anteil *FVC-Responder* (definiert als Anteil der Studienteilnehmer, deren Abfall der FVC [% des Sollwertes] weniger als 10% bzw. 5% im beobachteten Zeitraum beträgt) war unter Nintedanib deutlich höher als unter BSC. Die gepoolte Analyse der INPULSIS Studien ergab zu Woche 52 ein OR [95%-KI] entsprechend 1,577 ([1,21; 2,05]; $p=0,0007$) bezogen auf den Schwellenwert von 10% bzw. 1,835 ([1,43; 2,36]; $p<0,0001$) bezogen auf den Schwellenwert von 5%. Damit zeigen die gepoolten Ergebnisse der INPULSIS Studien einen statistisch signifikanten Vorteil zugunsten von Nintedanib gegenüber der zVT. Auch in der supportiv herangezogenen Betrachtung des Anteils der *FVC-Responder zu Woche 24* unter Hinzunahme der Studie 1199.187, die den Evidenzkörper der vorliegenden Nutzenbewertung gegenüber der ersten Nutzenbewertung ergänzt, zeigt sich, dass bereits für den kürzeren Betrachtungszeitraum der Vorteil der Behandlung mit Nintedanib gegenüber einer BSC-Therapie zum Tragen kommt (OR [95%-KI] 1,319 [0,97; 1,80]; $p=0,0787$ bzw. 1,466 [1,15; 1,88]; $p=0,0024$). Die Studienergebnisse belegen ebenfalls, dass der Abfall der FVC um ca. 50% geringer unter der Therapie mit Nintedanib ausfällt im Vergleich zu BSC. Somit konnte gezeigt werden, dass bei Studienteilnehmern, die mit Nintedanib behandelt wurden, der Abfall der jährlichen FVC statistisch und klinisch signifikant aufgehalten werden kann (Tabelle 4-93).

Vor dem Hintergrund, dass akute Exazerbationen mit einer hohen Sterblichkeit assoziiert sind (medianes Überleben nach akuter Exazerbation: 2,2 Monate (Song et al., 2011), 10-fach erhöhtes Sterberisiko nach Erleiden einer akuten Exazerbation im Vergleich zu Patienten ohne Exazerbation (Paterniti et al., 2017), wird der patientenrelevante Vorteil einer Behandlung mit Nintedanib in der Nutzendimension Mortalität außerdem durch den statistisch signifikanten Vorteil hinsichtlich der *Zeit bis zur ersten adjudizierten Exazerbation* deutlich.

Der Zusatznutzen für den Endpunkt FVC kann vor diesem Hintergrund als beträchtlich und dessen Wahrscheinlichkeit als Beleg eingestuft werden. Als valides Surrogat für die Mortalität deutet dies auf eine Verlängerung des Gesamtüberlebens bei Behandlung von IPF-Patienten mit Nintedanib hin (Tabelle 4-93).

Für die Mortalität ergibt sich hieraus ein **Beleg für einen beträchtlichen Zusatznutzen** von Nintedanib gegenüber der zVT für IPF-Patienten.

Tabelle 4-93: Bewertung des Zusatznutzens mit Erläuterung - Mortalität

Ergebnisse der gepoolten Analyse der INPULSIS Studien bzw. der INPULSIS Studien und Studie 1199.187	Mortalität <u>Zeit bis zum Tod, unabhängig von der Ursache^a</u> HR ^c [95%-KI]: 0,65 [0,41; 1,03]; p=0,0775 FVC als valider Surrogatendpunkt für Mortalität: <u>Jährliche FVC-Abnahme (ml)^a</u> MWD [95%-KI]: 112,42 [79,06; 145,77]; p< 0,0001 <u>FVC-Responder (10% des Sollwertes) zu Woche 52^b</u> OR ^d [95%-KI]: 1,577 [1,21; 2,05]; p=0,0007 <u>FVC-Responder (5% des Sollwertes) zu Woche 52^b</u> OR ^d [95%-KI]: 1,835 [1,43; 2,36]; p< 0,0001 Adjudizierte akute Exazerbationen als Mortalitäts-assoziiierter Endpunkt: <u>Zeit bis zur ersten adjudizierten Exazerbation^a</u> HR ^c [95%-KI]: 0,32 [0,17; 0,63]; p=0,0010
Erläuterungen	Nach dem SGB V – beispielsweise in der AM-NutzenV operationalisiert – wird ein patientenrelevanter Nutzen durch eine Verlängerung des Überlebens bestimmt.
Bewertung des Zusatznutzens	Insgesamt ergibt sich für den Endpunkt FVC als valides Surrogat für die Mortalität gemäß AM-NutzenV: • ein Beleg für einen beträchtlichen Zusatznutzen von Nintedanib im Vergleich zu BSC aufgrund einer bisher nicht erreichten Verbesserung des Überlebens der mit Nintedanib behandelten Patienten. Angesichts des vorliegenden statistisch signifikanten Effekts der Meta-Analyse von drei Studien mit homogener Datenlage und hoher qualitativer Ergebnissicherheit ist die Aussagesicherheit als Beleg zu werten.
a: Gepoolte Analyse INPULSIS Studien und Studie 1199.187 b: Gepoolte Analyse INPULSIS Studien c: Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib-Therapie an. d: Ein OR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib-Therapie an AM-NutzenV: Arzneimittel-Nutzenbewertungsverordnung; BSC: Best Supportive Care; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); HR: Hazard Ratio; KI: Konfidenzintervall; MWD: Mittelwertdifferenz; OR: Odds Ratio; SGB: Sozialgesetzbuch	

Morbidität

IPF ist eine Erkrankung mit einer hohen Krankheitslast. Die Krankheitssymptomatik ist insbesondere gekennzeichnet durch eine zunehmende Luftnot/Dyspnoe der Patienten mit erheblichen Auswirkungen auf die körperliche Leistungs- und Bewegungsfähigkeit. Diese sind meist schleichend progredient. Hinzu können akute Exazerbationen (respiratorische Dekompensationen) kommen, welche akut lebensbedrohlich sind, einen maßgeblichen Einfluss auf die Progression der Erkrankung haben und zu einer weiteren rapiden Verschlechterung des Gesundheitszustandes der Betroffenen führen. Solche Ereignisse sind mit einer hohen Sterblichkeit assoziiert (medianes Überleben nach akuter Exazerbation 2,2 Monate (Song et al., 2011), 10-fach erhöhtes Sterberisiko nach Erleiden einer akuten Exazerbation im Vergleich zu Patienten ohne Exazerbation (Paterniti et al., 2017)). Mit der in den Studien durchgeführten Adjudizierung durch ein verblindetes Expertenkomitee wurde das Vorliegen einer Exazerbation überprüft und damit die Ergebnissicherheit des Endpunktes erhöht.

Adjudizierte akute Exazerbationen traten in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 unter Nintedanib deutlich seltener auf als unter BSC (1,9% vs. 5,4%). Das HR von 0,32 für die *Zeit bis zur ersten adjudizierten akuten Exazerbation* zeigt mit einer 68%-igen Reduktion der Exazerbationen einen statistisch signifikanten Vorteil zugunsten der Behandlung mit Nintedanib (HR [95%-KI]: 0,32 [0,17; 0,63]; $p=0,0010$) (Tabelle 4-94). Die Kaplan-Meier-Kurven verdeutlichen, dass mit zunehmender Beobachtungszeit die Kurven für den Endpunkt *Zeit bis zur ersten adjudizierten akuten Exazerbation* zwischen den Behandlungsgruppen Nintedanib und zVT kontinuierlich auseinandergehen (Abbildung 4 und Abbildung 18).

Das Ausmaß des Zusatznutzens für den Endpunkt *adjudizierte akute Exazerbationen* ist vor diesem Hintergrund als beträchtlich einzuschätzen. Angesichts des vorliegenden statistisch signifikanten Effekts der Meta-Analyse von drei Studien mit homogener Datenlage und hoher qualitativer Ergebnissicherheit ist die Aussagesicherheit als Beleg zu werten (Tabelle 4-94).

Bezüglich der Morbidität ergibt sich hieraus ein **Beleg für einen beträchtlichen Zusatznutzen** von Nintedanib gegenüber der zVT für IPF-Patienten.

Tabelle 4-94: Bewertung des Zusatznutzens mit Erläuterung - Morbidität

Ergebnisse der gepoolten Analyse der INPULSIS Studien und Studie 1199.187	Morbidität <u>Zeit bis zur ersten adjudizierten akuten Exazerbation</u> HR ^a [95%-KI]: 0,32 [0,17; 0,63]; p=0,0010
Erläuterungen	Nach dem SGB V – beispielsweise in der AM-NutzenV operationalisiert – wird ein patientenrelevanter Nutzen durch eine Verlängerung des Überlebens oder eine Verringerung von Symptomen bzw. Nebenwirkungen bestimmt. Vor dem Hintergrund, dass <i>akute Exazerbationen</i> Ereignisse sind, die mit einer hohen Sterblichkeit assoziiert sind (medianes Überleben nach <i>akuter Exazerbation</i>: 2,2 Monate) wird durch die Vermeidung des Auftretens <i>akuter Exazerbationen</i> implizit eine Verlängerung des Überlebens (Song et al., 2011) sowie explizit eine Verbesserung hinsichtlich schwerwiegender Symptomatik erreicht.
Bewertung des Zusatznutzens	Die Ergebnisse aller Endpunkte zur Morbidität zeigen einen Vorteil der Nintedanib-Therapie im Vergleich zu BSC. Der Zusatznutzen wird vor allem aber basierend auf den Ergebnissen hinsichtlich der Vermeidung <i>adjudizierter akuter Exazerbationen</i> abgeleitet. Insgesamt ergibt sich für den Endpunkt <i>adjudizierte akute Exazerbationen</i> gemäß AM-NutzenV: • ein Beleg für einen beträchtlichen Zusatznutzen von Nintedanib im Vergleich zu BSC bezüglich einer Verbesserung schwerwiegender Symptome bzw. einer weitgehenden Vermeidung schwerwiegender Nebenwirkungen bzw. einer Verlängerung des Überlebens. Dies trifft hier aufgrund der erzielten Vermeidung der potenziell lebensbedrohlichen (<i>adjudizierten</i>) <i>akuten Exazerbationen</i> zu. Angeichts des vorliegenden statistisch signifikanten Effekts der Meta-Analyse von drei Studien mit homogener Datenlage und hoher qualitativer Ergebnissicherheit ist die Aussagesicherheit als Beleg zu werten.
a: Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib-Therapie an. AM-NutzenV: Arzneimittel-Nutzenbewertungsverordnung; BSC: Best Supportive Care; HR: Hazard Ratio; KI: Konfidenzintervall; SGB: Sozialgesetzbuch	

Gesundheitsbezogene Lebensqualität

IPF führt zu erheblichen Einschränkungen in der gesundheitsbezogenen Lebensqualität betroffener Patienten. Symptome wie Dyspnoe, Erschöpfung sowie eingeschränkte Mobilität beeinträchtigen die Teilhabe der Patienten am gesellschaftlichen und sozialen Leben. Aufgrund der kontinuierlichen Progression der IPF besteht ein wichtiges Therapieziel in der Erhaltung der gesundheitsbezogenen Lebensqualität im Krankheitsverlauf.

Die Verschlechterung der Lebensqualität zu Woche 52, ausgedrückt durch eine Zunahme des *SGRQ-I-Gesamtscores*, war in den gepoolten INPULSIS Studien unter Nintedanib signifikant geringer (MWD [95%-KI]: -1,99 [-3,71; -0,27]; $p=0,0230$). Darüber hinaus war der Anteil der Studienteilnehmer, die eine Verbesserung (Abnahme) um mindestens 4 Punkte im *SGRQ-I-Gesamtscore* und damit eine klinisch relevante Verbesserung der gesundheitsbezogenen Lebensqualität erreichten, in der Nintedanib-Gruppe im Vergleich zur zVT in der INPULSIS 2 Studie statistisch signifikant größer (OR [95%-KI] 1,877 [1,20; 2,93]; $p=0,0056$). Des Weiteren fand sich in den gepoolten INPULSIS Studien unter der Behandlung mit Nintedanib verglichen mit BSC ein statistisch signifikant höherer Anteil *PGI-C-Responder* zu Woche 52 (OR [95%-KI] 1,328 [1,04; 1,70]; $p=0,0254$) sowie eine signifikant geringere Verschlechterung des Gesundheitszustandes der Patienten gemessen anhand des *EQ-5D-VAS* zu Woche 52 (MWD [95%-KI]: 3,81 [1,78; 5,85]; $p=0,0002$; Hedges' g : 0,252 [0,12; 0,39]) (Tabelle 4-95).

Die Ergebnisse zur gesundheitsbezogenen Lebensqualität, gemessen anhand weiterer Instrumente wie *SGRQ*, *SOBQ* oder *CASA-Q*, zeigen einen konsistenten Trend hinsichtlich einer besseren gesundheitsbezogenen Lebensqualität zugunsten einer Behandlung mit Nintedanib im Vergleich zur zVT. Insgesamt wird für die patientenrelevante Endpunktkategorie gesundheitsbezogene Lebensqualität ein Hinweis für einen geringen Zusatznutzen abgeleitet (Tabelle 4-95).

Hinsichtlich der gesundheitsbezogenen Lebensqualität ergibt sich hieraus insgesamt ein **Hinweis für einen geringen Zusatznutzen** von Nintedanib gegenüber der zVT für IPF-Patienten.

Tabelle 4-95: Bewertung des Zusatznutzens mit Erläuterung - gesundheitsbezogene Lebensqualität

Ergebnisse der gepoolten Analyse der INPULSIS Studien	Gesundheitsbezogene Lebensqualität <u>Veränderung des SGRQ-I-Gesamtscores</u> MWD [95%-KI]: -1,99 [-3,71; -0,27]; p=0,0230 <u>PGI-C-Responder zu Woche 52</u> OR ^a [95%-KI]: 1,328 [1,04; 1,70]; p=0,0254 <u>Veränderung EQ-5D-VAS-Score</u> MWD [95%-KI]: 3,81 [1,78; 5,85]; p=0,0002; Hedges' g: 0,252 [0,12; 0,39]
Erläuterungen	Zum Nachweis der therapeutischen Verbesserung sind nach VerfO und AM-NutzenV vorrangig klinische Studien zu berücksichtigen, die unter anderem insbesondere die gesundheitsbezogene Lebensqualität als patientenrelevanten Endpunkt darstellen (Bundesministerium der Justiz und für Verbraucherschutz, 2010; Gemeinsamer Bundesausschuss, 2018b).
Bewertung des Zusatznutzens	Insgesamt ergibt sich für die Endpunkte zur gesundheitsbezogenen Lebensqualität gemäß AM-NutzenV: • ein Hinweis für einen geringen Zusatznutzen von Nintedanib im Vergleich zu BSC bezüglich einer höheren gesundheitsbezogenen Lebensqualität.
a: Ein OR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib-Therapie an. AM-NutzenV: Arzneimittel-Nutzenbewertungsverordnung; BSC: Best Supportive Care; EQ-5D: European Quality of Life-5 Dimensions; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); KI: Konfidenzintervall; MWD: Mittelwertdifferenz; OR: Odds Ratio; PGI-C: Patient's Global Impression of Change; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; VAS: Visual Analog Scale; VerfO: Verfahrensordnung	

Unerwünschte Ereignisse

UE traten in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 95,7% der Studienteilnehmer, die mit Nintedanib behandelt wurden, auf, verglichen mit 89,8% der Studienteilnehmer des Vergleichsarms. Der Unterschied zwischen den Gruppen ist statistisch signifikant (RR [95%-KI]: 1,07 [1,03; 1,10]; $p < 0,001$). Daraus ergibt sich ein Hinweis auf einen Nachteil für Nintedanib gegenüber der zVT für diesen Endpunkt.

SUE traten in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 29,1% der Studienteilnehmer, die mit Nintedanib behandelt wurden, auf, und somit mit vergleichbarer Häufigkeit wie bei den Studienteilnehmern der Vergleichstherapie (28,3%). Der Unterschied zwischen den Gruppen war nicht statistisch signifikant (RR [95%-KI]: 1,01 [0,84; 1,21]; $p = 0,941$). Für diesen Endpunkt liegt somit weder ein Zusatznutzen noch ein Nachteil von Nintedanib gegenüber der zVT vor.

UE, die zum Therapieabbruch führten traten in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 bei 18,9% der Studienteilnehmer, die mit Nintedanib behandelt wurden, auf, verglichen mit 12,1% der Studienteilnehmer der Vergleichstherapie. Der Unterschied zwischen den Gruppen ist statistisch signifikant (RR [95%-KI]: 1,54 [1,15; 2,04]; $p = 0,003$). Somit brachen in der Nintedanib-Gruppe häufiger Studienteilnehmer die Behandlung aufgrund eines *UE* ab als in der Kontrollgruppe. Daraus ergibt sich ein Nachteil für Nintedanib gegenüber der zVT für diesen Endpunkt.

Als *UE* von besonderem Interesse wurden *gastrointestinale Störungen (Diarrhö, Übelkeit, Erbrechen, Appetitlosigkeit, Gewichtsverlust, gastrointestinale Perforation), Leber- und Gallenerkrankungen (Leberenzymerrhöhung, Bilirubinerhöhung), kardiale Ereignisse, respiratorische Ereignisse (Progression der IPF, Husten, Dyspnoe, Bronchitis, Nasopharyngitis), Blutungsereignisse* sowie alle weiteren Störungen, die in mindestens einem Studienarm mit einer Häufigkeit von $>10\%$ der Studienteilnehmer auftraten, betrachtet.

Gastrointestinale Ereignisse

Die gastrointestinalen *UE Diarrhö, Übelkeit* und *Erbrechen* sowie die weiteren *UE Appetitlosigkeit* und *Gewichtsverlust* traten unter der Behandlung mit Nintedanib signifikant häufiger auf als in der Kontrollgruppe. *Gastrointestinale Perforationen* traten insgesamt so selten auf, dass keine Aussage zu einem möglichen Risiko vorgenommen werden kann (Tabelle 4-67).

Leberenzym- und Bilirubinerhöhungen

UE, welche das Leber- und Gallenwegssystem betrafen, wurden in der gepoolten Analyse der INPULSIS Studien und Studie 1199.187 häufiger im Nintedanib-Arm als im Vergleichsarm dokumentiert: Die Behandlung mit Nintedanib war im Vergleich zur zVT häufiger verbunden mit einer *Erhöhung der Leberenzyme (ALT, AST, AP)* oder des *Bilirubin-Wertes*.

Es ist jedoch zu beachten, dass sich die symptomatischen *unerwünschten Ereignisse* insgesamt gut behandeln lassen. Eine *Diarrhö* sollte möglichst frühzeitig erkannt, durch pharmakologische Intervention (z. B. Loperamid) behandelt und nachbeobachtet werden. Außerdem lassen sich die unerwünschten Ereignisse durch eine Unterbrechung der Behandlung und Dosisreduktion häufig reduzieren bzw. beseitigen.

Kardiale Ereignisse

Unerwünschte kardiale Ereignisse wurden in den gepoolten INPULSIS Studien und Studie 1199.187 vergleichbar häufig unter Nintedanib und zVT (9,7% vs. 9,6%) beobachtet. Hierunter waren beispielsweise folgende Ereignisse subsummiert: *Vorhofflimmern, Myokardinfarkt, koronare Herzerkrankung, Herzversagen, EKG-Auffälligkeiten wie Arrhythmien oder Schenkelblockbilder oder Cor pulmonale*. Auch *schwerwiegende unerwünschte kardiale Ereignisse* traten in der Nintedanib-Gruppe (4,6%) und Kontrollgruppe (4,8%) vergleichbar häufig auf.

Tödliche kardiale Ereignisse traten numerisch seltener in der Nintedanib-Gruppe (0,4%) als in der Vergleichsgruppe (1,3%) auf (RR [95%-KI]: 0,33 [0,08; 1,32]; p=0,100). *Arterielle thromboembolische Ereignisse* wurden in den gepoolten INPULSIS Studien häufiger unter Nintedanib als unter zVT (2,4% vs. 0,6%) beobachtet (Tabelle 4-69). *Myokardinfarkte* traten numerisch häufiger unter Nintedanib als unter zVT (1,6% vs. 0,4%) auf.

Respiratorische Ereignisse

Eine *Progression der IPF* wurde in der gepoolten Analyse der INPULSIS Studien und der Studie 1199.187 in der Nintedanib-Gruppe von 9,7% der Studienteilnehmer als *UE* angegeben, verglichen mit 13,8% in der Kontrollgruppe. Der Unterschied ist statistisch signifikant zugunsten von Nintedanib (RR [95%-KI]: 0,69 [0,50; 0,95]; p=0,023). *Husten* als *UE* trat in den gepoolten Analyse der INPULSIS Studien und der Studie 1199.187 in beiden Behandlungsarmen vergleichbar häufig auf (13,0% vs. 13,8%) (RR [95%-KI]: 0,94 [0,70; 1,27]; p=0,683). *Dyspnoe* als *UE* wurde in den gepoolten Analysen seltener bei Nintedanib-behandelten Studienteilnehmern (8,2%) als bei Studienteilnehmern der Vergleichsgruppe (11,3%) erfasst, der Unterschied ist nicht statistisch signifikant (RR [95%-KI]: 0,74 [0,52; 1,05]; p=0,093). *Bronchitis* und *Nasopharyngitis* als *UE* fanden sich in der gepoolten Analyse in beiden Behandlungsarmen vergleichbar häufig: *Bronchitis* in der Nintedanib-Gruppe bei 9,9% der Studienteilnehmer vs. 10,0% der Studienteilnehmer in der Vergleichsgruppe (RR [95%-KI]: 0,97 [0,68; 1,38]; p=0,866). *Nasopharyngitis* in der Nintedanib-Gruppe bei 13,0% der Studienteilnehmer vs. 15,0% der Studienteilnehmer in der Kontrollgruppe (RR [95%-KI]: 0,84 [0,63; 1,12]; p=0,249).

Blutungsereignisse

Blutungsereignisse fanden sich bei 9,8% der Patienten im Nintedanib-Arm und 7,3% der Patienten im Kontrollarm. Der Gruppenunterschied war nicht statistisch signifikant (RR [95%-KI]: 1,31 [0,89; 1,94]; p=0,171).

Ein Hinweis auf einen Vorteil gegenüber der zVT ergibt sich für die *UE von besonderem Interesse Progression der IPF*.

Bei den *UE von besonderem Interesse kardiale Ereignisse, schwerwiegende kardiale Ereignisse, tödliche kardiale Ereignisse, Myokardinfarkt, Husten, Bronchitis, Dyspnoe, Nasopharyngitis und Blutungsereignisse* liegen weder ein Vor- noch ein Nachteil von Nintedanib gegenüber der zVT vor.

Ein Nachteil von Nintedanib gegenüber der zVT ergibt sich für die *gastrointestinalen Ereignisse, Leberenzymerrhöhung und Bilirubinerhöhung sowie arterielle thromboembolische Ereignisse*.

Fazit

Ein Zusatznutzen mit beträchtlichem Ausmaß wird nach AM-NutzenV dann ausgewiesen, wenn eine gegenüber der zVT bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens im Sinne von § 2 Absatz 3 erreicht wird, insbesondere eine Abschwächung schwerwiegender Symptome, eine moderate Verlängerung der Lebensdauer, eine für die Patientinnen und Patienten spürbare Linderung der Erkrankung, eine relevante Vermeidung schwerwiegender Nebenwirkungen oder eine bedeutsame Vermeidung anderer Nebenwirkungen. In der Gesamtschau weist Nintedanib gegenüber der zVT BSC in den Endpunktkategorien Mortalität, Morbidität und gesundheitsbezogene Lebensqualität Vorteile auf, welche die Nachteile in der Endpunktkategorie Nebenwirkungen überwiegen.

Die Vorteile zugunsten einer Nintedanib-Behandlung hinsichtlich Mortalität (statistisch signifikante Reduktion der *jährlichen FVC-Abnahme* als valides Surrogat für eine Reduktion des Mortalitätsrisikos, numerischer Vorteil in der *Zeit bis zum Tod, unabhängig von der Ursache*, statistisch signifikanter Vorteil im Mortalitäts-assoziierten Endpunkt *adjudizierte akute Exazerbationen*, Tabelle 4-93) sowie Morbidität (statistische signifikante Reduktion der *Zeit bis zur ersten adjudizierten akuten Exazerbation*, Tabelle 4-94) sind jeweils als ein Beleg für einen beträchtlichen Zusatznutzen und für die gesundheitsbezogene Lebensqualität (signifikant höherer Anteil *PGI-C-Responder*, signifikanter Vorteil in der *Veränderung des SGRQ-I-Gesamtscores* und der *Veränderung der EQ-5D-VAS*, Tabelle 4-95) als ein Hinweis für einen geringen Zusatznutzen einzustufen. Bezüglich der *UE, die zum Therapieabbruch führten* liegt ein geringerer Nutzen vor.

Aus den dargestellten Daten geht hervor, dass bei der Behandlung mit Nintedanib für erwachsene Patienten mit IPF der Nutzen den potenziellen Nachteil überwiegt. Damit ergibt sich insgesamt ein **Hinweis für einen beträchtlichen Zusatznutzen** für Nintedanib gegenüber der zVT für Patienten mit IPF.

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.

Tabelle 4-96: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit IPF	Hinweis für einen beträchtlichen Zusatznutzen
IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose)	

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 zutreffend

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (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 zutreffend

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 zutreffend

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 2005¹⁸, Molenberghs 2010¹⁹). 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 2006²⁰) 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 2006²¹) 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

¹⁸ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁹ 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.

²⁰ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

²¹ Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

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.

Zur vollständigen Beurteilung eines therapeutischen Effekts auf die mit einer IPF assoziierten Mortalität sind langjährige prospektive Studien notwendig, welche die Wirksamkeit der Therapie auf die Patientensterblichkeit untersuchen. Mortalitätsstudien in der Indikation IPF müssten auf einen vergleichsweise langen Zeitraum ausgerichtet sein und eine sehr hohe Anzahl von Patienten umfassen, um statistische Signifikanz zu erreichen. Um die Verfügbarkeit von Nintedanib für IPF-Patienten nicht zu verzögern, waren die vorliegenden Studien daher nicht als solche angelegt. Um indirekt dennoch Aussagen zu den zu erwartenden Mortalitätsunterschieden zwischen den Studienarmen treffen zu können, wurde im vorliegenden Dossier die FVC als Surrogatparameter für die Mortalität herangezogen.

Rationale für die Verwendung der FVC als valides Surrogat für die Mortalität

Bei der IPF handelt es um eine schwerwiegende und in der Regel zum Tod führende Erkrankung (Behr et al., 2013). Dabei ist die Mortalitätsrate höher als bei vielen Krebserkrankungen (Bjoraker et al., 1998; Douglas et al., 2000; Kim et al., 2006). Wie in Abschnitt 4.2.5.2 ausgeführt, kommt es bei der IPF aufgrund der zunehmenden Vernarbung der Lunge zu einer mit der Zeit fortschreitenden Abnahme der FVC, sodass die Belüftung der Lunge zunehmend eingeschränkt wird, bzw. eine erhebliche Atemanstrengung des Patienten erfordert. Wenngleich sich im Verlauf einer IPF-Erkrankung Phasen der rapiden Verschlechterung und Phasen weitgehender Stagnation abwechseln können, kommt es in der Gesamtschau zu einer fortschreitenden und unumkehrbaren Schädigung der Lungenfunktion. Dies kann durch die Messung der FVC in regelmäßigen Intervallen im Krankheitsverlauf dokumentiert werden. FVC ist ein Lungenfunktionstest, der die Menge an Luft misst, die maximal aus der Lunge ausgeatmet werden kann, nachdem so tief wie möglich eingeatmet wurde. Die Lungenfunktionsmessung wird in der klinischen Praxis regelmäßig durchgeführt, um pathophysiologische Veränderungen im Krankheitsgeschehen darzustellen. Ein wesentlicher Vorteil ist dabei die gute Reproduzierbarkeit und standardisierte Erhebung der Daten. Der altersbedingte Verlust an FVC bei Gesunden liegt bei etwa 30-65 ml pro Jahr (Brändli et al., 1996; Griffith et al., 2001; Janssens, 2005; Alexeeff et al., 2007).

Im vorangegangenen Nutzenbewertungsverfahren zu Nintedanib im Anwendungsgebiet IPF hatte der G-BA den Surrogatcharakter der FVC in Bezug auf Mortalität mit Verweis auf eine fehlende Validierung des Endpunkts nicht anerkannt (Gemeinsamer Bundesausschuss, 2015c). In diesem Zusammenhang führt der G-BA aus, dass keine Modelle zur Validierung vorgelegt wurden, in welchen die Intervention als unabhängige Größe betrachtet wurde und zudem der Zusammenhang zwischen Effekten auf das Surrogat Veränderung der FVC und Effekten auf den patientenrelevanten Endpunkt Gesamtmortalität gezeigt wurde.

Seit dem Zeitpunkt der Erstbewertung liegt nun weitere Evidenz vor, welche die Korrelation zwischen Abnahme der FVC und Mortalität bestätigt (Reichmann et al., 2015; Paterniti et al., 2017) und die Kritikpunkte des G-BA adressiert (Karimi-Shah und Chowdhury, 2015).

Basierend auf der Publikation von Karimi-Shah und Chowdhury zum Zusammenhang zwischen FVC und Mortalität, welche bereits im Stellungnahmeverfahren zur Erstbewertung aufgeführt wurde (Karimi-Shah und Chowdhury, 2015), wird nachfolgend eine weiterführende Analyse vorgelegt. Mit dieser Analyse wird der Zusammenhang zwischen Effekten auf das Surrogat FVC und Effekten auf den patientenrelevanten Endpunkt Mortalität unter Betrachtung der Intervention als unabhängige Größe beschrieben und der Behandlungseffekt durch Nintedanib auf das Mortalitätsrisiko abgeleitet. Zudem haben sich seitdem sowohl Zulassungsbehörden als auch Experten im Anwendungsgebiet der IPF für eine Anerkennung des Endpunkts ausgesprochen (Behr et al., 2015; Karimi-Shah und Chowdhury, 2015; Paterniti et al., 2017); diese werden im Folgenden ebenfalls dargestellt.

Auf der Grundlage der vorhandenen und im Folgenden aufgeführten Evidenz sieht BI insgesamt die Validität des Surrogatcharakters des Endpunkts FVC bezüglich Mortalität sowie die medizinische Relevanz der im Dossier präsentierten Responderanalysen zur Abnahme der FVC, insbesondere im Falle eines Cut-off-Wertes von 10% Abnahme über 52 Wochen im Vergleich zum Ausgangswert, als gegeben an.

Methodische Validität der im vorliegenden Dossier präsentierten Evidenz zur Surrogatvalidierung und Ableitung des Behandlungseffekts durch Nintedanib auf das Mortalitätsrisiko

Im Rahmen eines Positionspapiers hatten Mitarbeiter der FDA eine Analyse veröffentlicht, die die Korrelation zwischen Mortalität und Abnahme der FVC basierend auf sechs im Anwendungsgebiet IPF vorliegenden Studien, den Zulassungstudien zu Nintedanib bzw. Pirfenidon, betrachtet (Karimi-Shah und Chowdhury, 2015) (Abbildung 36). Alle sechs Studien waren nicht für eine konfirmatorische Aussage zur Mortalität gepowert. Primärer Wirksamkeitsendpunkt war jeweils der Abfall der FVC über die Zeit. In fünf der sechs Studien war die FVC-Abnahme deutlich geringer in Patienten, die Nintedanib oder Pirfenidon verabreicht bekamen, als im Kontrollarm. In den Verum-Armen gab es dabei jeweils einen numerischen Trend zu einem geringeren Mortalitätsrisiko (HR <1) im Vergleich zum Kontrollarm. Im Gegenzug konnte in einer der sechs Studien, bei der es in den unterschiedlichen Behandlungsgruppen keinen Unterschied in der FVC-Abnahme gab, auch kein Vorteil in der Mortalität festgestellt werden. Die Autoren geben an, dass der in diesen Studien beobachtete Zusammenhang zwischen FVC und dem Trend in der Mortalität das Vertrauen in die FVC als klinisch relevanten Wirksamkeitsparameter stärkt (Karimi-Shah und Chowdhury, 2015).

Analysis of Forced Vital Capacity and All-Cause Mortality.*						
Study	Forced Vital Capacity			All-Cause Mortality		
	Change from Baseline (ml)		Treatment Difference (95% CI)	No. of Deaths (%)		Hazard Ratio for Time to Death (95% CI)
	study drug	placebo		study drug	placebo	
Pirfenidone study 2 (November 2008)	-318	-475	157 (3 to 311)	14 (8.0)	20 (11.5)	0.65 (0.33 to 1.29)
Pirfenidone study 3 (November 2008)	-379	-373	-6 (-178 to 167)	18 (10.5)	17 (9.8)	1.07 (0.55 to 2.08)
Pirfenidone study 1 (re-submitted; February 2014)	-235	-428	193 (96 to 289)	12 (4.3)	21 (7.6)	0.57 (0.28 to 1.16)
Nintedanib study 1 (June 2010)	-60	-191	131 (27 to 235)	7 (8.1)	9 (10.3)	0.73 (0.27 to 1.98)
Nintedanib study 2 (October 2013)	-115	-240	125 (78 to 173)	13 (4.2)	13 (6.4)	0.63 (0.29 to 1.36)
Nintedanib study 3 (October 2013)	-114	-207	94 (45 to 143)	22 (6.7)	20 (9.1)	0.74 (0.40 to 1.35)

* The studies are listed in chronologic order by drug, with study numbers as referenced in the product labels and the months in which enrollment ended. Data for forced vital capacity are the absolute values for the change from baseline to week 52 for pirfenidone study 1, to week 72 for pirfenidone studies 2 and 3, and to week 52 for all nintedanib studies; the change from baseline for pirfenidone studies was based on descriptive statistics, and the change from baseline for nintedanib studies was based on regression analysis. Mortality data are from the vital status analysis (from randomization to the time of death) and include all deaths irrespective of the cause and of whether the patient had continued treatment. Hazard ratios for time to death are based on Cox proportional-hazards regression analysis. CI denotes confidence interval.

Quelle: (Karimi-Shah und Chowdhury, 2015)

Abbildung 36: Ergebnisse zur FVC-Abnahme und Gesamtmortalität der klinischen Studien zu Nintedanib und Pirfenidon

Im Zuge des Stellungnahmeverfahrens zur Erstbewertung wurde bereits eine weiterführende Aufbereitung der von Karimi-Shah und Chowdhury veröffentlichten Arbeit vorgelegt. Im vorliegenden Dossier wird nun auf Basis der vorhandenen Evidenz eine erneute Analyse vorgenommen. Abbildung 37 zeigt eine grafische Aufbereitung der neuen Analyse des zuvor bereits von der FDA beschriebenen Zusammenhangs zwischen der Reduktion des FVC-Abfalls und der Reduktion der Mortalität (Karimi-Shah und Chowdhury, 2015).

Bezogen auf den Behandlungseffekt ergibt sich eine Korrelation zwischen der Mittelwertdifferenz der FVC und dem logarithmierten HR des Gesamtüberlebens von $r = -0,9605$ ([95%-KI]: [-0,9958; -0,6752]). Das Quadrat dieses Korrelationskoeffizienten ergibt einen Wert von $R^2 = 0,9286$, was einer Aufklärung der Varianz des logarithmierten HR durch die FVC von fast 93% entspricht und damit von einem extrem starken Zusammenhang zwischen diesen Endpunkten auf Effektebene zeugt.

Legt man die vom IQWiG angewandten Kriterien zur Validierung von Surrogatendpunkten²² zugrunde, so dürfte die obere Grenze des KI (-0,6752) einen Wert von -0,85 nicht überschreiten. Dass die vorliegende Korrelation zwischen der FVC und der Mortalität dieses Kriterium nicht erfüllt, ist alleine der Tatsache geschuldet, dass das KI eines Korrelationskoeffizienten vom Schätzer selbst sowie von der Anzahl der in die Berechnung eingegangenen Wertepaare abhängt.

²² Rapid Report des IQWiG zur Aussagekraft von Surrogatendpunkten in der Onkologie

Um für den gegebenen Korrelationskoeffizienten eine obere Grenze des KI kleiner als -0,85 zu erhalten, wären -unabhängig von der Anzahl der Patienten in den Einzelstudien- elf Studien notwendig (Gillhaus et al., 2017), was insbesondere in einem Anwendungsgebiet einer seltenen Erkrankung eine fast unerreichbar hohe Zahl ist. Zudem ist die Durchführbarkeit einer weiteren placebo-kontrollierten Studie vor dem Hintergrund, dass zwei wirksame Therapieoptionen zur Behandlung der IPF zugelassen sind und dem Patienten zur Verfügung stehen, sehr eingeschränkt.

Im Folgenden werden daher die Behandlungseffekte einer Therapie mit Nintedanib hinsichtlich des Mortalitätsrisikos auf Basis der vorliegenden Daten abgeleitet.

Beide zur Bewertung des Behandlungseffekts berücksichtigten Maße, die Abnahme der FVC (ml) sowie das logarithmierte HR für Zeit bis zum Tod, sind Differenzen zwischen den Behandlungsarmen. Da im Falle eines ausbleibenden Behandlungseffekts weder ein Unterschied in der Abnahme der FVC zwischen beiden Armen, noch ein sich daraus ergebender Unterschied auf das Mortalitätsrisiko zu erwarten ist, nehmen in diesem Fall beide Werte den Wert 0 an. Eine Regressionsgerade muss daher durch den Ursprung gehen. Um den Zusammenhang weiter zu quantifizieren, ist folglich ein Metaregressionsmodell ohne Konstante verwendet worden. Die benötigten SE für die logarithmierte HR wurden aus den gegebenen KI geschätzt. In Abbildung 37 ist das sich daraus ergebende Konfidenzband dargestellt (graue Linien). Der Schätzer für den Anstieg der Regressionsgeraden beträgt -0,0030 ([95%-KI]: [-0,0052; -0,0007], $p=0,0103$) und ist somit statistisch signifikant. Dass sich die Pirfenidon-Studie 3 nicht innerhalb des Konfidenzbands befindet, ist der Tatsache geschuldet, dass hier kein Behandlungseffekt auftritt, die Schätzer aber dennoch fehlerbehaftet sind und daher nicht genau den Wert 0 annehmen.

In den gepoolten INPULSIS Studien und der Studie 1199.187 ergab sich unter Therapie mit Nintedanib über 52 Wochen gegenüber dem Vergleichsarm ein Behandlungsunterschied in der Abnahme der FVC von 112,42 ml (Tabelle 4-22). Nimmt man diesen Wert als Grundlage zur Berechnung des Unterschiedes im Mortalitätsrisiko, so ergibt sich mit dem Schätzer zur FVC von 112,42 [79,06; 145,77] ein Schätzer zur Mortalität von HR [95%-KI]: 0,72 [0,56; 0,92]. Selbst wenn man im Einklang mit den Anforderungen des IQWiG die Untergrenze des KI von 79,06 zugrunde legt, ergibt sich ein statistisch signifikanter Wert von HR [95%-KI]: 0,79 [0,66; 0,95].

Dies bedeutet, dass die in den nutzenbewertungsrelevanten Studien beobachtete Verringerung der jährlichen Abnahme der FVC um 112,42 ml durch Therapie mit Nintedanib einer Reduktion des Mortalitätsrisikos von über 20% entspricht.

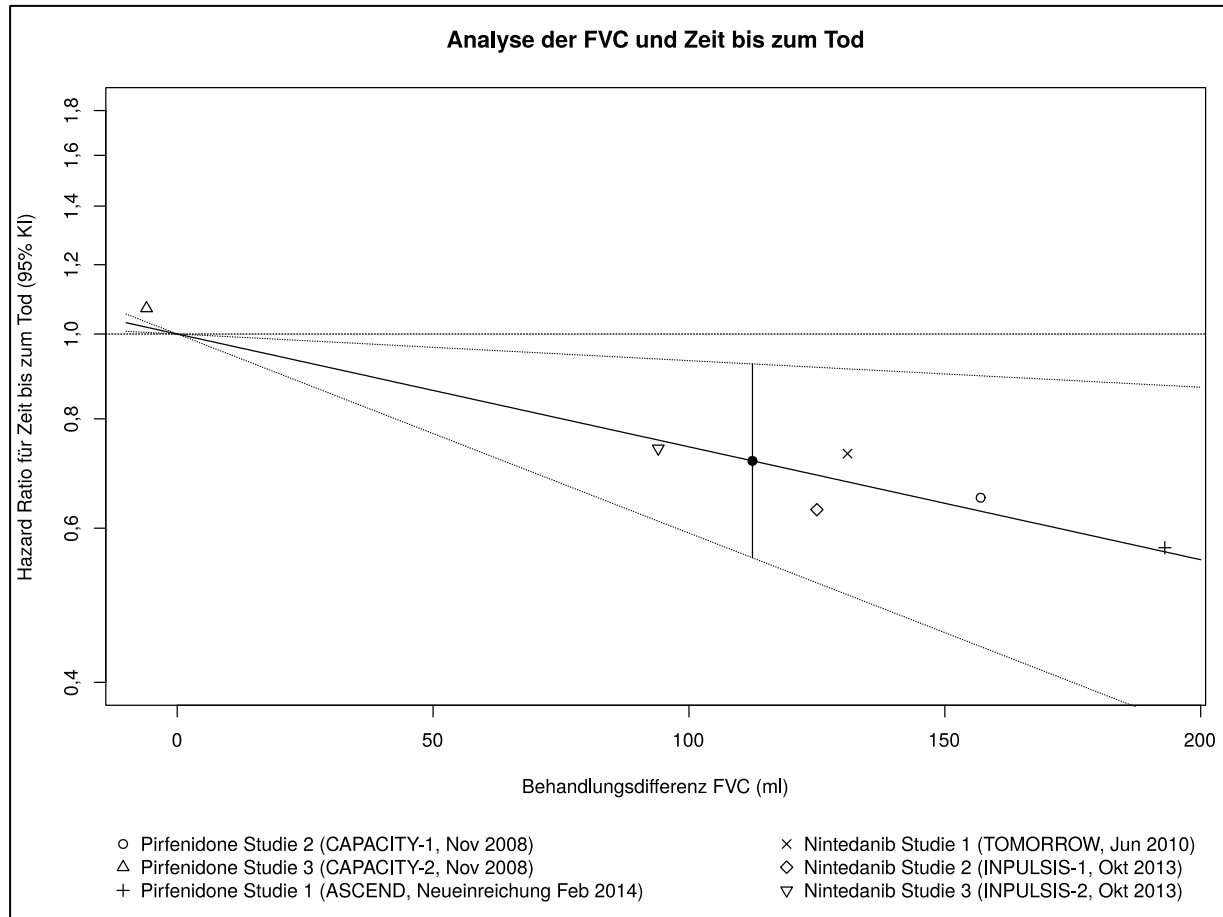


Abbildung 37: Grafische Darstellung der berichteten Behandlungsdifferenz in der Abnahme der FVC (ml) zwischen Vergleichstherapie und aktiver Behandlung und dem HR für die Zeit bis zum Tod.

Gezeigt ist ein Metaregressionsmodell ohne Konstante. Das HR ist inklusive des Konfidenzbandes (graue Linien) abgebildet. Die schwarze Linie bildet den möglichen linearen Zusammenhang ab ($R^2=0,9286$), die senkrechte Linie zeigt die Behandlungsdifferenz von 112,42 ml an.

Weitere Studien zum Zusammenhang von FVC (% des Sollwertes) und Mortalität sowie Beurteilung der Schwellenwerte von 10% und 5%

Der Zusammenhang zwischen FVC und Mortalität wurde des Weiteren im Rahmen verschiedener Studien untersucht, die im Folgenden beschrieben werden. Diese zeigen eine signifikante Korrelation von FVC mit der Patientensterblichkeit auf und bekräftigen die Verwendung der FVC als Surrogatparameter für die Mortalität bei IPF.

Generell kommt angesichts des heterogenen und nicht absehbaren Verlaufs einer IPF prognostischen Faktoren eine besondere Bedeutung zu. Zahlreiche Studien zeigen, dass eine Verschlechterung der FVC um mindestens 10% in einem Follow-up-Zeitraum von sechs bzw. zwölf Monaten als relevanter Prädiktor für Mortalität anzusehen ist. Tritt eine derartige Verschlechterung der FVC ein, steigt die Mortalität um bis zu einem Faktor von acht an, wobei die Zahl zwischen den einzelnen Studien etwas variiert (Flaherty et al., 2003; Latsi et al., 2003; Jegal et al., 2005; Collard et al., 2007a; Zappala et al., 2010; du Bois et al., 2011a; Richeldi et

al., 2012). Dabei erweist sich der Zusammenhang als robust über unterschiedliche Untersuchungsdauern und Patientenpopulationen.

Flaherty et al. konnten in einer retrospektiven Untersuchung an 80 Patienten mit usual interstitial pneumonia (UIP) aufzeigen, dass eine Abnahme der FVC >10% über sechs Monate mit einem zweifachen Anstieg der Mortalität in dem anschließenden 6-Monats-Zeitraum in Verbindung gebracht werden kann (HR [95%-KI]: 2,06 [1,09; 3,89]; p=0,03), adjustiert nach Ausgangswert der FVC, Raucherstatus, Geschlecht und Zeit seit Auftreten der Symptome. Nach zwölf Monaten lag der Wert für das HR (Tod von Patienten mit >10% Rückgang der FVC [% des Sollwertes]) bei 1,70 (HR [95%-KI]: 1,70 [0,76; 3,81]; p=0,20) (Flaherty et al., 2003).

Collard et al. untersuchten das Überleben von 81 Patienten mit IPF (mittleres Alter: 62 Jahre, mittlere FVC: 67% des Sollwertes) im Verhältnis zur absoluten Veränderung der FVC in % des Sollwertes nach zwölf Monaten. Die Überlebenswahrscheinlichkeit nach vier Jahren betrug etwa 40% bei einer Abnahme der FVC (% des Sollwertes) um mindestens 10%, 72% bei stabiler FVC ($\pm 10\%$ des Sollwertes) und 84% bei einer Zunahme der FVC (% des Sollwertes) um mindestens 10%. Das Mortalitätsrisiko bei Patienten mit einer Abnahme der FVC um mindestens 10% des Sollwertes stieg über einen Zeitraum von vier Jahren um das 2,8-fache an im Vergleich zu Patienten mit über zwölf Monaten stabiler FVC ($\pm 10\%$ des Sollwertes) (Collard et al., 2003).

Richeldi et al. untersuchten den Zusammenhang zwischen der Veränderung der FVC (% des Sollwertes) und dem Transplantations-freien Überleben an 142 IPF-Patienten mit einem mittleren Alter von 67 Jahren und mittlerer FVC-Ausgangswert von 2,7 Liter. In dieser Publikation wurden absolute und relative Änderungen unterschieden und die Mortalität über einen Zeitraum von zwei Jahren betrachtet. Eine relative bzw. absolute mindestens 10%-ige Abnahme der FVC nach zwölf Monaten war dabei verbunden mit einem Anstieg des Risikos nach zwei Jahren zu versterben oder eine Transplantation erhalten zu müssen (OR [95%-KI]: 3,39 [1,14; 10,07] bzw. 4,52 [1,27; 16,12]) (Richeldi et al., 2012).

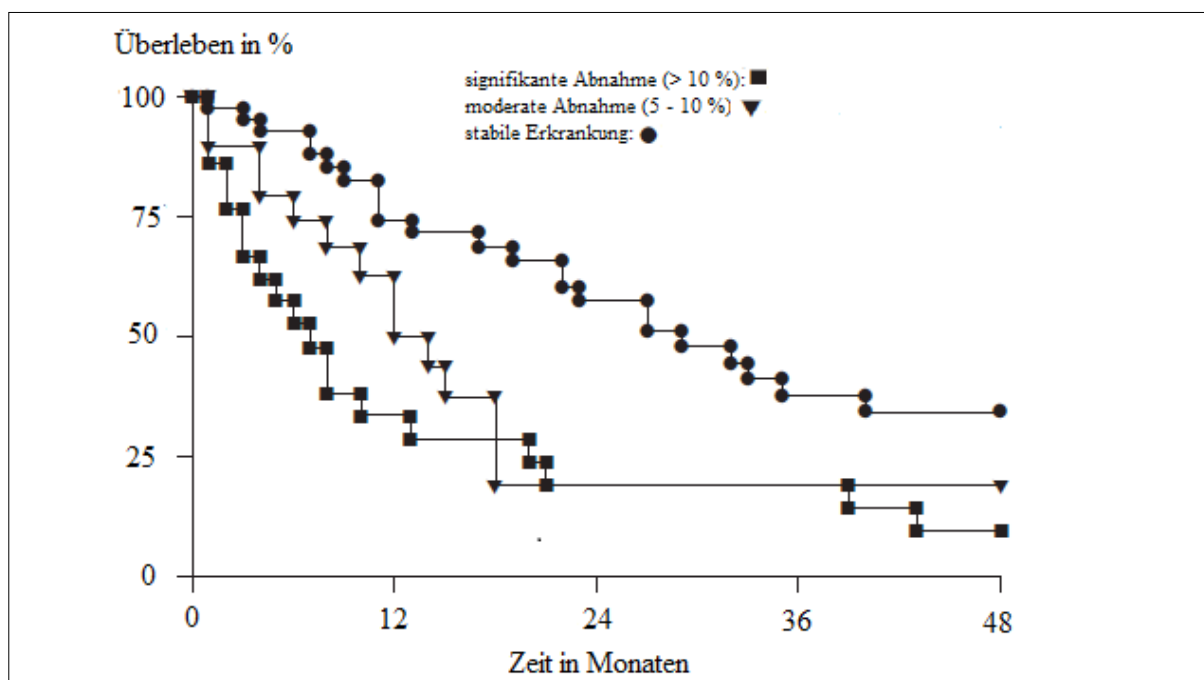
In einer von Du Bois et al. 2011 publizierten Studie mit 1.099 IPF-Patienten erwiesen sich neben dem Alter der Patienten auch eine vorangegangene Hospitalisierung aufgrund respiratorischer Beschwerden, der Ausgangswert der FVC und der DL_{CO} in Prozent des Sollwertes und der Abfall der FVC bzw. der DL_{CO} sowie die gesundheitsbezogene Lebensqualität (gemessen anhand des SGRQ) binnen 24 Wochen als unabhängige, prognostisch relevante Faktoren. Bereits ein absoluter Verlust der FVC von 5-10% des Sollwertes innerhalb 24 Wochen (z. B. von 60% auf 55-50% des Sollwertes) war mit einem mehr als zweifach erhöhten Mortalitätsrisiko (HR [95%-KI]: 2,60 [1,75; 3,85]; p< 0,001), ein Abfall um mehr als 10% des Sollwertes mit einem 8-fach erhöhten Mortalitätsrisiko innerhalb des nächsten Jahres verbunden (HR [95%-KI]: 7,99 [5,26; 12,14]; p< 0,001) (du Bois et al., 2011b).

In einer ebenfalls von Du Bois 2011 publizierten Studie mit 1.156 IPF-Patienten wurde die MCID der FVC mit verschiedenen Methoden ermittelt. Entsprechend dieser Studie liegt dieser Wert bei 2–6% (je nach angewandter Methode; eher niedriger auf Ebene einer Population, eher

höher im Individualfall), d. h. ein Abfall der FVC (in % des Sollwertes) um ca. fünf Prozentpunkte binnen 24 Wochen stellt bereits ein klinisch relevantes Ereignis mit potenziell therapeutischen Implikationen dar (du Bois et al., 2011a).

In einer retrospektiven Untersuchung an 179 Patienten mit fibrotischer interstitieller Pneumonie (UIP/IPF 131; unspezifische interstitielle Pneumonie 48), war eine Veränderung der prozentualen erwarteten FVC (Anstieg von >10%, keine Veränderung, Abnahme von >10%) ein unabhängiger Prädiktor für das Überleben nach sechs Monaten ($p < 0,001$) und hatte einen stärkeren prognostischen Wert als jede andere Messung der Lungenfunktion (Jegal et al., 2005).

Zappala et al. untersuchten das Verhältnis zwischen der Veränderung der FVC (% des Sollwertes) und der 4-Jahres-Mortalität von 84 Patienten mit IPF. Die Auswertung erfolgte in den Kategorien „signifikante Abnahme“ (Abnahme der FVC >10% des Sollwertes), „moderate Abnahme“ (Abnahme der FVC um 5% bis 10% des Sollwertes) und „stabile Erkrankung“. Die Mortalität war bei Patienten mit signifikanter FVC-Abnahme nach sechs Monaten 2,8-fach höher als bei stabiler Erkrankung ($p < 0,001$). Patienten mit moderater FVC-Abnahme hatten im Vergleich zu Patienten mit stabiler Erkrankung ebenfalls ein 2,31-fach erhöhtes Risiko zu versterben ($p = 0,01$). Patienten mit signifikanter Abnahme und mit moderater Abnahme hatten somit eine miteinander vergleichbar stark erhöhte Mortalität (Zappala et al., 2010).



Quelle: (Zappala et al., 2010)

Abbildung 38: 4-Jahres-Überleben in Abhängigkeit vom Ausmaß der FVC-Abnahme in % des Sollwertes über sechs Monate bei Patienten mit IPF (n=84)

In einer retrospektiven Analyse von Reichmann et al. wurden 490 IPF-Patienten nach dem Ausmaß ihres FVC-Abfalls kategorisiert in „stabiler“ (<5%), „geringfügiger“ (≥5% und <10%) und „signifikanter“ Abfall (≥10%) innerhalb von sechs Monaten (Reichmann et al., 2015). Die

Autoren beobachteten ein signifikant erhöhtes Mortalitätsrisiko der Patienten mit signifikantem (HR [95%-KI]: 6,23 [2,96; 13,14]) bzw. geringfügigem Abfall (HR [95%-KI]: 3,22 [1,59; 6,53]) gegenüber solchen mit stabilem Abfall.

Die jüngste Analyse stammt von Paterniti et al. (Paterniti et al., 2017), in welcher 1.132 Placebo-Patienten aus sechs Studien für die klinische Entwicklung von Nintedanib und Pirfenidon eingeschlossen wurden. Auch hier wurden Patienten hinsichtlich des prozentualen FVC-Abfalls im Vergleich zum Ausgangswert in vier Kategorien eingeteilt ($<5\%$, $\geq 5\%$ bis $<10\%$, $\geq 10\%$ bis $<15\%$, $\geq 15\%$). Die mittlere Beobachtungszeit betrug 60 Wochen. Die Autoren fanden bei Patienten mit einer FVC-Abnahme von $\geq 10\%$ bis $<15\%$ ein statistisch signifikant erhöhtes Mortalitätsrisiko (HR [95%-KI]: 2,20 [1,10; 4,37]) verglichen mit Patienten mit einem Abfall unter 5%. Eine entsprechende Korrelation zeigte sich auch für Patienten mit einer FVC-Abnahme von $\geq 15\%$ (HR [95%-KI]: 6,09 [3,14; 11,80]). Daraus lässt sich schließen, dass für eine auf diesen Daten basierende aber von Paterniti et al. nicht berichtete Analyse für Patienten mit einem FVC-Abfall von $\geq 10\%$ ohne Beschränkung auf Patienten mit einem Abfall $<15\%$ ein noch höheres Mortalitätsrisiko zu erwarten wäre. Nach Angaben der Autoren handelt es sich dabei um die bislang umfangreichste diesbezüglich untersuchte Placebo-Kohorte, welche entsprechend valide Analysen zulässt und den im vorliegenden Dossier gewählten Cut-off-Wert von $\leq 10\%$ für Responderanalysen bestätigt. Aussagen zum Cut-off-Wert von $\leq 5\%$ können nicht getroffen werden, da Paterniti et al. hierzu keine Analyse präsentieren.

Sicht der Fachgesellschaften und Zulassungsbehörden

Eine Abnahme der FVC (% des Sollwertes) von 5% bis 10% innerhalb von sechs Monaten gilt bereits als prognostisch relevante Progression (Jegal et al., 2005; Zappala et al., 2010; Raghu et al., 2011). Diese Auffassung wird auch von der deutschen IPF-Konsensus-Konferenz vertreten. Nach Ansicht der Deutschen Gesellschaft für Pneumologie und Beatmungsmedizin (DGP) ist der prognostische Stellenwert durch zahlreiche klinisch kontrollierte Studien belegt. Daher kommt die DGP in ihrer Stellungnahme zur frühen Nutzenbewertung von Nintedanib sowie in der zur frühen Nutzenbewertung von Pirfenidon zum Schluss, dass die FVC den robustesten und sensitivsten Surrogatparameter des Überlebens von IPF-Patienten darstellt (Gemeinsamer Bundesausschuss, 2012; Gemeinsamer Bundesausschuss, 2017).

Auch laut einem Positionspapier führender deutscher Experten im Anwendungsgebiet IPF ist festzuhalten, dass der Nachweis einer Krankheitsprogression in Form eines Abfalls der FVC mit einem erhöhten Mortalitätsrisiko verbunden ist (Behr et al., 2015). Da alle betrachteten Studien, in denen Interventionen einen Effekt auf die Abnahme der FVC hatten, auch einen gleichgerichteten Trend bezüglich der Mortalität aufzeigten, kommen die Experten zum Schluss, dass somit eine zentrale Bedingung erfüllt sei, die FVC als Surrogat für Mortalität anzuerkennen. Behr et al. berichten zudem, dass eine kategorische FVC-Abnahme von 5-10% oder $>10\%$ des Sollwertes innerhalb von 24 Wochen geeignet ist, um prognostisch ein erhöhtes Mortalitätsrisiko zu beschreiben. Die Expertenmeinung stützt daher die im vorliegenden Dossier gewählten Cut-off-Werte der Responderanalysen zur Abnahme der FVC.

Der Parameter FVC erfüllt darüber hinaus die im DIMDI-Bericht 91 (Deutsches Institut für Medizinische Dokumentation und Information) geforderten Validierungsanforderungen für einen patientenrelevanten Endpunkt (biologische Plausibilität, qualitativ und quantitativ gleichgerichtete Veränderung des klinischen Endpunkts, Wirkungsnachweis im Rahmen von RCT) (Mangiapane und Velasco Garrido, 2009).

Wie bereits weiter oben ausgeführt, wird die FVC weltweit von den Zulassungsbehörden als primärer Endpunkt für Studien im Anwendungsgebiet IPF anerkannt. Die allgemeine Akzeptanz der FVC als Prädiktor für die Mortalität spiegelt sich auch in den Richtlinien der EMA wider. So wurde im Jahr 2011 im Rahmen eines Scientific Advice der EMA der primäre Parameter FVC als adäquater Wirksamkeitsendpunkt bestätigt. Zudem hat die EMA den Parameter FVC als prognostisch für das Überleben der Patienten mit IPF akzeptiert (European Medicines Agency (EMA), 2011).

Gegenwärtig ist der Lungenfunktionsparameter FVC der am besten validierte und am breitesten eingesetzte Endpunkt in klinischen Studien im Indikationsgebiet. Du Bois et al. 2012 bekräftigen die Bedeutung dieses Endpunkts als robustes und valides Maß für die Krankheitsprogression bei IPF (du Bois et al., 2012). Der FVC-Wert erfülle alle Anforderungen an einen klinisch aussagekräftigen Endpunkt, mit dessen Hilfe qualitativ hochwertige RCT durchgeführt werden können.

Fazit zur Verwendung der FVC als valides Surrogat für die Mortalität

Nach Beurteilung der in den vorausgehenden Abschnitten beschriebenen Studienlage, unter Berücksichtigung der relevanten Leitlinien und im Einklang mit den Vorgaben der Zulassungsbehörden, wurde in den Studien zur Wirksamkeit und Verträglichkeit von Nintedanib die jährliche Abnahme der FVC als primärer Endpunkt für die Bestimmung des therapeutischen Effekts gewählt. Wie erläutert, steht dieser Parameter in direkter Korrelation mit dem Krankheitsverlauf und stellt den **stärksten klinischen Prädiktor für die Mortalität bei IPF-Patienten dar**. Bereits bei der Erstbewertung von Nintedanib im Anwendungsgebiet IPF lag diesbezüglich umfangreiche Evidenz vor. Seitdem konnte die Korrelation zwischen FVC und Mortalität noch durch weitere umfassende Analysen nachgewiesen werden. **Die FVC wird daher von BI als Surrogatparameter für die Mortalität angesehen und berücksichtigt.**

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).

Tabelle 4-97: Liste der eingeschlossenen Studien

Studie	Quellen
INPULSIS 1	Studienprotokoll: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b) Studienbericht: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a) Zusatzauswertungen:(Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c; Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a) Registerbericht: (PharmNet.bund, 2013b; World Health Organization - International Clinical Trials Registry Platform, 2014; ClinicalTrials.gov, 2016a; EU Clinical Trials Register, 2016b; World Health Organization - International Clinical Trials Registry Platform, 2018b) Publikationen: (Richeldi et al., 2014; Corte et al., 2015; Goyard und Crestani, 2015; Keating, 2015; Costabel et al., 2016; Richeldi et al., 2016; Taniguchi et al., 2016; Wuyts et al., 2016; Azuma et al., 2017; Collard et al., 2017; Hajari Case und Johnson, 2017; Kolb et al., 2017; Raghu et al., 2017; Rinciog et al., 2017; Costabel et al., 2018; Flaherty et al., 2018; Kreuter et al., 2018; Brown et al., 2019)
INPULSIS 2	Studienprotokoll: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012a) Studienbericht: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b) Zusatzauswertungen:(Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c; Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a) Registerbericht: (PharmNet.bund, 2013a; World Health Organization - International Clinical Trials Registry Platform, 2013; ClinicalTrials.gov, 2016b; EU Clinical Trials Register, 2016a; World Health Organization - International Clinical Trials Registry Platform, 2018a) Publikationen: (Richeldi et al., 2014; Corte et al., 2015; Goyard und Crestani, 2015; Keating, 2015; Costabel et al., 2016; Richeldi et al., 2016; Taniguchi et al., 2016; Wuyts et al., 2016; Azuma et al., 2017; Collard et al., 2017; Hajari Case und Johnson, 2017; Kolb et al., 2017; Raghu et al., 2017; Rinciog et al., 2017; Costabel et al., 2018; Flaherty et al., 2018; Kreuter et al., 2018; Brown et al., 2019)
Studie 1199.187	Studienprotokoll: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013) Studienbericht: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2017) Zusatzauswertungen: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2018a) Registerbericht: (ClinicalTrials.gov, 2018; World Health Organization - International Clinical Trials Registry Platform, 2018c)

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte

Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

- [1] Aaronson, N. K. 1993. Assessment of quality of life and benefits from adjuvant therapies in breast cancer. *Recent Results Cancer Res*, 127 201-10.
- [2] Alexeeff, S. E., Litonjua, A. A., Sparrow, D., et al. 2007. Statin use reduces decline in lung function: VA Normative Aging Study. *Am J Respir Crit Care Med*, 176 (8), 742-7.
- [3] Archibald, C. & Guidotti, T. 1987. Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab*, 1 45-54.
- [4] Azuma, A., Nukiwa, T., Tsuboi, E., et al. 2005. Double-blind, placebo-controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 171 (9), 1040-7.
- [5] Azuma, A., Taniguchi, H., Inoue, Y., et al. 2017. Nintedanib in Japanese patients with idiopathic pulmonary fibrosis: A subgroup analysis of the INPULSIS randomized trials. *Respirology*, 22 (4), 750-757.
- [6] Behr, J. 2013. The diagnosis and treatment of idiopathic pulmonary fibrosis. *Dtsch Arztebl Int*, 110 (51-52), 875-81.
- [7] Behr, J., Günther, A., Ammenwerth, W., et al. 2013. S2K-Leitlinie zur Diagnostik und Therapie der idiopathischen Lungenfibrose. *Pneumologie*, 67 81-111.
- [8] Behr, J., Bonella, F., Bonnet, R., et al. 2015. Positionspapier zur Bedeutung der forcierten Vitalkapazität für Patienten mit idiopathischer Lungenfibrose (IPF). *Pneumologie*, 69 (08), 455-458.
- [9] Behr, J., Günther, A., Bonella, F., et al. 2017. S2k-Leitlinie Idiopathische Lungenfibrose – Update zur medikamentösen Therapie 2017. *Pneumologie*, 71 460-474.
- [10] Bender, R., Friede, T., Koch, A., et al. 2018. Methods for evidence synthesis in the case of very few studies. *Research Synthesis Methods*.
- [11] Bjoraker, J. A., Ryu, J. H., Edwin, M. K., et al. 1998. Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 157 (1), 199-203.
- [12] Boehringer Ingelheim Pharma GmbH & Co. KG. 2012a. Studienprotokoll 1199.34 (Phase-III Studie)
- [13] Boehringer Ingelheim Pharma GmbH & Co. KG. 2012b. Studienprotokoll 1199.32 (Phase-III Studie)
- [14] Boehringer Ingelheim Pharma GmbH & Co. KG. 2012c. Charta des unabhängigen Adjudizierungskomitees zu den Studien 1199.32 und 1199.34
- [15] Boehringer Ingelheim Pharma GmbH & Co. KG. 2013. Studienprotokoll 1199.187 (Phase-IIb Studie)
- [16] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014a. Studienbericht 1199.32 (Phase-III Studie)
- [17] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014b. Studienbericht 1199.34 (Phase-III Studie)
- [18] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014c. SCE-Supplement Phase-III Studien 1199.32 und 1199.34 (gepoolt)
- [19] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014d. SCS-Supplement Phase-III Studien 1199.32 und 1199.34

- [20] Boehringer Ingelheim Pharma GmbH & Co. KG. 2015. Baseline-Patientencharakteristika Subgruppen - Nachberechnung
- [21] Boehringer Ingelheim Pharma GmbH & Co. KG. 2016a. Charta des unabhängigen Adjudizierungskomitees zur Studie 1199.187
- [22] Boehringer Ingelheim Pharma GmbH & Co. KG. 2016b. Prozess-Guideline zum Adjudizierungsprozess der Studie 1199.187
- [23] Boehringer Ingelheim Pharma GmbH & Co. KG. 2017. Studienbericht 1199.187 (Phase-IIIb Studie)
- [24] Boehringer Ingelheim Pharma GmbH & Co. KG. 2018a. Weitere Analysen INPULSIS-1, INPULSIS-2, 1199.187
- [25] Boehringer Ingelheim Pharma GmbH & Co. KG. 2018b. Fachinformation Ofev, Stand Oktober 2018
- [26] Boehringer Ingelheim Pharma GmbH & Co. KG 2018c. Übersicht Sensitivitätsanalysen Studien INPULSIS-1, INPULSIS-2 und Studie 1199.187.
- [27] Brändli, O., Schindler, C., Künzli, N., et al. 1996. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. *Thorax*, 51 (3), 277-83.
- [28] Brazier, J., Jones, N. & Kind, P. 1993. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res*, 2 (3), 169-80.
- [29] Brown, A. W., Fischer, C. P., Shlobin, O. A., et al. 2015. Outcomes after hospitalization in idiopathic pulmonary fibrosis: a cohort study. *Chest*, 147 (1), 173-9.
- [30] Brown, K. K., Flaherty, K. R., Cottin, V., et al. 2019. Lung function outcomes in the INPULSIS trials of nintedanib in idiopathic pulmonary fibrosis. *Respiratory Medicine*, 146 42-48.
- [31] Bundesministerium der Justiz und für Verbraucherschutz. 1988. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - Zuletzt geändert durch Artikel 7 des Gesetzes vom 11. Dezember 2018 (BGBl. I S. 2394). Verfügbar: http://www.gesetze-im-internet.de/sgb_5/SGB_5.pdf.
- [32] Bundesministerium der Justiz und für Verbraucherschutz. 2010. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV) - Zuletzt geändert durch Art. 3 G v. 4.5.2017 I 1050. BGBl. I S. 2324. Verfügbar: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf>.
- [33] Caminati, A., Bianchi, A., Cassandro, R., et al. 2009. Walking distance on 6-MWT is a prognostic factor in idiopathic pulmonary fibrosis. *Respir Med*, 103 (1), 117-23.
- [34] ClinicalTrials.gov. 2016a. Safety and Efficacy of BIBF 1120 at High Dose in Idiopathic Pulmonary Fibrosis Patients (NCT01335464). Verfügbar: <https://clinicaltrials.gov/ct2/show/NCT01335464> [Aufgerufen am 17.01.2019].
- [35] ClinicalTrials.gov. 2016b. Safety and Efficacy of BIBF 1120 at High Dose in Idiopathic Pulmonary Fibrosis Patients II (NCT01335477). Verfügbar: <https://clinicaltrials.gov/ct2/show/NCT01335477> [Aufgerufen am 17.01.2019].
- [36] ClinicalTrials.gov. 2018. Nintedanib Twice Daily vs Placebo in Patients Diagnosed With Idiopathic Pulmonary Fibrosis (IPF) (NCT01979952). Verfügbar: <https://clinicaltrials.gov/ct2/show/NCT01979952> [Aufgerufen am 17.01.2019].
- [37] Collard, H. R., King, T. E., Jr., Bartelson, B. B., et al. 2003. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 168 (5), 538-42.

- [38] Collard, H. R., Loyd, J. E., King, T. E., Jr., et al. 2007a. Current diagnosis and management of idiopathic pulmonary fibrosis: a survey of academic physicians. *Respir Med*, 101 (9), 2011-6.
- [39] Collard, H. R., Moore, B. B., Flaherty, K. R., et al. 2007b. Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 176 (7), 636-43.
- [40] Collard, H. R., Richeldi, L., Kim, D. S., et al. 2017. Acute exacerbations in the INPULSIS trials of nintedanib in idiopathic pulmonary fibrosis. *The European respiratory journal*, 49 (5).
- [41] Corte, T., Bonella, F., Crestani, B., et al. 2015. Safety, tolerability and appropriate use of nintedanib in idiopathic pulmonary fibrosis. *Respiratory Research*, 16 (1) (no pagination) (116).
- [42] Costabel, U., Inoue, Y., Richeldi, L., et al. 2016. Efficacy of nintedanib in idiopathic pulmonary fibrosis across prespecified subgroups in INPULSIS. *American Journal of Respiratory and Critical Care Medicine*, 193 (2), 178-185.
- [43] Costabel, U., Behr, J., Crestani, B., et al. 2018. Anti-acid therapy in idiopathic pulmonary fibrosis: Insights from the INPULSIS trials. *Respiratory Research*, 19 (1) (no pagination) (167).
- [44] Cottin, V., Schmidt, A., Catella, L., et al. 2017. Burden of Idiopathic Pulmonary Fibrosis Progression: A 5-Year Longitudinal Follow-Up Study. *PLoS One*, 12 (1), e0166462.
- [45] Crawford, B., Monz, B., Hohlfeld, J., et al. 2008. Development and validation of a cough and sputum assessment questionnaire. *Respir Med*, 102 (11), 1545-55.
- [46] Crestani, B., Quaresma, M., Kaye, M., et al. 2017. Long-term nintedanib treatment in idiopathic pulmonary fibrosis (IPF): new data from INPULSIS®-ON. *European Respiratory Society International Congress, Milan, Italy, 9–13 September 2017*.
- [47] Crestani, B., Huggins, J. T., Kaye, M., et al. 2018. Long-term safety and tolerability of nintedanib in patients with idiopathic pulmonary fibrosis: results from the open-label extension study, INPULSIS-ON. *Lancet Respir Med*.
- [48] de Boer, A., van Lanschott, J., Stalmeier, P., et al. 2004. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Qual Life Res*, 2 (13), 311-320.
- [49] Douglas, W. W., Ryu, J. H. & Schroeder, D. R. 2000. Idiopathic pulmonary fibrosis Impact of oxygen and colchicine, prednisone, or no therapy on survival. *161*, 1172-1178.
- [50] du Bois, R. M., Weycker, D., Albera, C., et al. 2011a. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am J Respir Crit Care Med*, 184 (12), 1382-9.
- [51] du Bois, R. M., Weycker, D., Albera, C., et al. 2011b. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 184 (4), 459-66.
- [52] du Bois, R. M., Weycker, D., Albera, C., et al. 2011c. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *Am J Respir Crit Care Med*, 183 (9), 1231-7.
- [53] du Bois, R. M., Nathan, S. D., Richeldi, L., et al. 2012. Idiopathic pulmonary fibrosis: lung function is a clinically meaningful endpoint for phase III trials. *Am J Respir Crit Care Med*, 186 (8), 712-5.

- [54] Eakin, E. G., Resnikoff, P. M., Prewitt, L. M., et al. 1998. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. *Chest*, 113 (3), 619-24.
- [55] EU Clinical Trials Register. 2016a. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on Forced Vital Capacity decline , in patients with Idiopathic Pulmonary Fibrosis (IPF)(2010-024252-29). Verfügbar: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024252-29 [Aufgerufen am 17.01.2019].
- [56] EU Clinical Trials Register. 2016b. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on annual Forced Vital Capacity decline, in patients with Idiopathic Pulmonary Fibrosis (IPF) (2010-024251-87). Verfügbar: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024251-87 [Aufgerufen am 17.01.2019].
- [57] European Commission (EC). 2015. Note to the Members of the Standing Committee on Medicinal Products for Human Use/Standing Committee on Veterinary Medicinal Products - Adoption of Commission Implementing Decision granting marketing authorisation under Regulation (EC) No 726/2004 of the European Parliament and of the Council for "Ofev - Nintedanib", an orphan medicinal product for human use
- [58] European Medicines Agency (EMA). 2011. Scientific Advice - BIBF 1120
- [59] European Medicines Agency (EMA). 2012. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD) (EMA/CHMP/483572/2012-corr1)
- [60] Flaherty, K. R., Mumford, J. A., Murray, S., et al. 2003. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med*, 168 (5), 543-8.
- [61] Flaherty, K. R., Kolb, M., Vancheri, C., et al. 2018. Stability or improvement in forced vital capacity with nintedanib in patients with idiopathic pulmonary fibrosis. *Eur Respir J*, 52 (2).
- [62] Furukawa, T., Taniguchi, H., Ando, M., et al. 2017. The St. George's Respiratory Questionnaire as a prognostic factor in IPF. *Respir Res*, 18 (1), 18.
- [63] Gemeinsamer Bundesausschuss. 2012. Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-AL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Pirfenidon. Verfügbar: https://www.g-ba.de/downloads/40-268-1915/2012-03-15_AM-RL-XII_Pirfenidon_ZD.pdf.
- [64] Gemeinsamer Bundesausschuss. 2015a. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel -Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Nintedanib. Verfügbar: https://www.g-ba.de/downloads/39-261-2322/2015-09-03_AM-RL-XII_Nintedanib_nAWG_2015-03-15-D-156_BAnz.pdf.
- [65] Gemeinsamer Bundesausschuss. 2015b. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Nintedanib. Verfügbar: https://www.g-ba.de/downloads/40-268-3339/2015-09-03_AM-RL-XII_Nintedanib_nAWG_2015-03-15-D-156_TrG.pdf.
- [66] Gemeinsamer Bundesausschuss. 2015c. Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Bewertung von Arzneimitteln für seltene Leiden nach

- § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 Verfo - Wirkstoff: Nintedanib. Verfügbar: <https://www.g-ba.de/downloads/92-975-813/2015-06-15-D-156-Nintedanib-Nutzenbewertung%20G-BA.pdf>.
- [67] Gemeinsamer Bundesausschuss. 2016a. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Umeclidinium. Verfügbar: <https://www.g-ba.de/downloads/40-268-3895/2016-07-21-AM-RL-XII-Umeclidinium-D-210-TrG.pdf>.
- [68] Gemeinsamer Bundesausschuss. 2016b. Beratungsanforderung 2016-B-049, Niederschrift vom 24.08.2016
- [69] Gemeinsamer Bundesausschuss. 2017. Anlage III - Vorlage zur Abgabe einer schriftlichen Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten- und Kosten-Nutzen-Bewertung nach § 35b SGB V - Stellungnahme der DGB zu Nintedanib/Ofev.
- [70] Gemeinsamer Bundesausschuss. 2018a. Vorsorgliche Aufforderung zur Dossiereinreichung vom 21.06.2018.
- [71] Gemeinsamer Bundesausschuss. 2018b. Verfahrensordnung des Gemeinsamen Bundesausschusses - zuletzt geändert am 16. August 2018, in Kraft getreten am 8. Dezember 2018. Verfügbar: https://www.g-ba.de/downloads/62-492-1614/VerfO_2018-03-16_iK-2018-07-05.pdf.
- [72] Gemeinsamer Bundesausschuss. 2019. Vorsorgliche Aufforderung zur Dossiereinreichung vom 11.01.2019.
- [73] Gillhaus, J., Goertz, R., Jeratsch, U., et al. 2017. Surrogatvalidierung durch Korrelation und Surrogate Threshold Effect – Ergebnisse von Simulationsstudien. *GMS Medizinische Informatik, Biometrie und Epidemiologie*, 13.
- [74] Goldin, J., Kim, G., Lu, P., et al. 2017. Short term Quantitative Lung Fibrosis (QLF) change predicts rapid rate of FVC decline in patients with IPF. *European Respiratory Journal*, 50 (suppl 61).
- [75] Goyard, C. & Crestani, B. 2015. The effect nintedanib on lung function decline according to the initial FVC in idiopathic pulmonary fibrosis: Results of INPULSIS studies. [French]. *Revue des Maladies Respiratoires Actualites*, 7 (2), 151-153.
- [76] Gries, K. S., Esser, D. & Wiklund, I. 2013. Content validity of CASA-Q cough domains and UCSD-SOBQ for use in patients with Idiopathic Pulmonary Fibrosis. *Glob J Health Sci*, 5 (6), 131-41.
- [77] Griffith, K. A., Sherrill, D. L., Siegel, E. M., et al. 2001. Predictors of loss of lung function in the elderly: the Cardiovascular Health Study. *Am J Respir Crit Care Med*, 163 (1), 61-8.
- [78] Hajari Case, A. & Johnson, P. 2017. Clinical use of nintedanib in patients with idiopathic pulmonary fibrosis. *BMJ Open Respiratory Research*, 4 (1) (no pagination) (e000192).
- [79] International Conference on Harmonisation. 1996. Guideline for Good Clinical Practice E6(R1). Verfügbar: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf.
- [80] IQWiG. 2011. IQWiG-Berichte - Nr. 108 - Pirfenidon - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: <https://www.g-ba.de/downloads/92-975-16/2011-09-15-D-020-Pirfenidon-IQWiG-Nutzenbewertung.pdf>.

- [81] IQWiG. 2012. IQWiG-Berichte - Nr. 148 - Acridiniumbromid - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-152/2012-12-21_A12-13_Acridiniumbromid_Nutzenbewertung%20%C2%A735a%20SGB%20V.pdf.
- [82] IQWiG. 2014. IQWiG-Berichte - Nr. 205 - Indacaterol/Glycopyrronium - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-408/2014-02-12_A13-40_Indacaterol-Glycopyrronium_Nutzenbewertung-35a-SGB-V.pdf.
- [83] IQWiG. 2015a. IQWiG-Berichte - Nr. 337 - Tiotropium/Olodaterol - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-914/2015-11-12_Nutzenbewertung-IQWiG_Tiotropium-Olodaterol.pdf.
- [84] IQWiG. 2015b. IQWiG-Berichte - Nr. 298 - Acridiniumbromid/Formoterol - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-767/2015-05-04_Nutzenbewertung-IQWiG_A15-06_Acridiniumbromid-Formoterol.pdf.
- [85] IQWiG. 2016a. IQWiG-Berichte - Nr. 355 - Acridiniumbromid - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-1206/2016-01-15_Nutzenbewertung-IQWiG_Acridiniumbromid.pdf.
- [86] IQWiG. 2016b. IQWiG-Berichte - Nr. 384 - Umeclidinium - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-1326/2016-04-28_Nutzenbewertung-IQWiG_Umeclidinium.pdf.
- [87] IQWiG. 2017. Allgemeine Methoden Version 5.0 [Online]. Adresse: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf Zuletzt geändert.
- [88] Janssens, J. P. 2005. Aging of the respiratory system: impact on pulmonary function tests and adaptation to exertion. *Clin Chest Med*, 26 (3), 469-84, vi-vii.
- [89] Jegal, Y., Kim, D. S., Shim, T. S., et al. 2005. Physiology is a stronger predictor of survival than pathology in fibrotic interstitial pneumonia. *Am J Respir Crit Care Med*, 171 (6), 639-44.
- [90] Johansson, K. & Collard, H. R. 2013. Acute Exacerbation of Idiopathic Pulmonary Fibrosis: A Proposal. *Curr Respir Care Rep*, 2 (4).
- [91] Jones, P. W., Quirk, F. H., Baveystock, C. M., et al. 1992. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis*, 145 (6), 1321-7.
- [92] Jones, P. W., Donohue, J. F., Nedelman, J., et al. 2011. Correlating changes in lung function with patient outcomes in chronic obstructive pulmonary disease: a pooled analysis. *Respir Res*, 12 161.
- [93] Jones, P. W. 2014. Estimation and application of the minimum clinically important difference in COPD. *Lancet Respir Med*, 2 (3), 167-9.
- [94] Juarez, M. M., Chan, A. L., Norris, A. G., et al. 2015. Acute exacerbation of idiopathic pulmonary fibrosis - a review of current and novel pharmacotherapies. *Journal of Thoracic Disease*, 7 (3), 499-519.
- [95] Karimi-Shah, B. A. & Chowdhury, B. A. 2015. Forced vital capacity in idiopathic pulmonary fibrosis--FDA review of pirfenidone and nintedanib. *N Engl J Med*, 372 (13), 1189-91.
- [96] Keating, G. M. 2015. Nintedanib in idiopathic pulmonary fibrosis: a guide to its use. *Drugs and Therapy Perspectives*, 31 (10), 329-333.

- [97] Kim, D. S., Park, J. H., Park, B. K., et al. 2006. Acute exacerbation of idiopathic pulmonary fibrosis: frequency and clinical features. *Eur Respir J*, 27 (1), 143-50.
- [98] Kim, H. G., Tashkin, D. P., Clements, P. J., et al. 2010. A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. *Clin Exp Rheumatol*, 28 (5 Suppl 62), S26-35.
- [99] Kim, H. J., Brown, M. S., Elashoff, R., et al. 2011. Quantitative texture-based assessment of one-year changes in fibrotic reticular patterns on HRCT in scleroderma lung disease treated with oral cyclophosphamide. *Eur Radiol*, 21 (12), 2455-65.
- [100] King, T. E., Jr., Behr, J., Brown, K. K., et al. 2008. BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 177 (1), 75-81.
- [101] Kolb, M., Richeldi, L., Behr, J., et al. 2017. Nintedanib in patients with idiopathic pulmonary fibrosis and preserved lung volume. *Thorax*, 72 (4), 340-346.
- [102] Kreuter, M., Costabel, U., Richeldi, L., et al. 2018. Statin Therapy and Outcomes in Trials of Nintedanib in Idiopathic Pulmonary Fibrosis. *Respiration*, 95 (5), 317-326.
- [103] Kupferberg, D. H., Kaplan, R. M., Slymen, D. J., et al. 2005. Minimal clinically important difference for the UCSD Shortness of Breath Questionnaire. *J Cardiopulm Rehabil*, 25 (6), 370-7.
- [104] Latsi, P. I., du Bois, R. M., Nicholson, A. G., et al. 2003. Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. *Am J Respir Crit Care Med*, 168 (5), 531-7.
- [105] Lettieri, C. J., Nathan, S. D., Browning, R. F., et al. 2006. The distance-saturation product predicts mortality in idiopathic pulmonary fibrosis. *Respir Med*, 100 (10), 1734-41.
- [106] Ley, B., Collard, H. R. & King, T. E., Jr. 2011. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*, 183 (4), 431-40.
- [107] Mangiapane, S. & Velasco Garrido, M. 2009. Use of Surrogate end points in HTA. *GMS Health Technol Assess*, 5 Doc12.
- [108] Matthys, H. & Seeger, W. 2008. Klinische Pneumologie. 4 84-86.
- [109] Miller, M. R., Crapo, R., Hankinson, J., et al. 2005. General considerations for lung function testing. *Eur Respir J*, 26 (1), 153-61.
- [110] Moher, D., Hopewell, S., Schulz, K. F., et al. 2010. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340 c869.
- [111] Nathan, S. D., du Bois, R. M., Albera, C., et al. 2015. Validation of test performance characteristics and minimal clinically important difference of the 6-minute walk test in patients with idiopathic pulmonary fibrosis. *Respir Med*, 109 (7), 914-22.
- [112] Osoba, D., Rodrigues, G., Myles, J., et al. 1998. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*, 16 (1), 139-44.
- [113] Parambil, J. G., Myers, J. L. & Ryu, J. H. 2005. Histopathologic Features and Outcome of Patients With Acute Exacerbation of Idiopathic Pulmonary Fibrosis Undergoing Surgical Lung Biopsy. *Chest*, 128 3310-3315.
- [114] Paterniti, M. O., Bi, Y., Rekić, D., et al. 2017. Acute Exacerbation and Decline in Forced Vital Capacity Are Associated with Increased Mortality in Idiopathic Pulmonary Fibrosis. *Ann Am Thorac Soc*, 14 (9), 1395-1402.
- [115] PharmNet.bund. 2013a. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on Forced Vital Capacity decline, in patients with Idiopathic Pulmonary Fibrosis (IPF) (2010-024252-29).

- Verfügbar: <https://portal.dimdi.de/clinical-trials/servlet/FlowController/DisplayDocuments?uid=12&docId=19&changebranch=true> [Aufgerufen am 17.01.2019].
- [116] PharmNet.bund. 2013b. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on annual Forced Vital Capacity decline, in patients with Idiopathic Pulmonary Fibrosis (IPF) (2010-024251-87). Verfügbar: <https://portal.dimdi.de/clinical-trials/servlet/FlowController/DisplayDocuments?uid=12&docId=20&changebranch=true> [Aufgerufen am 17.01.2019].
- [117] Pickard, A. S., Wilke, C., Jung, E., et al. 2008. Use of a preference-based measure of health (EQ-5D) in COPD and asthma. *Respir Med*, 102 (4), 519-36.
- [118] Rabin, R. & de Charro, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*, 33 (5), 337-43.
- [119] Raghu, G., Collard, H. R., Egan, J. J., et al. 2011. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med*, 183 (6), 788-824.
- [120] Raghu, G., Rochwerg, B., Zhang, Y., et al. 2015. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med*, 192 (2), e3-19.
- [121] Raghu, G., Wells, A. U., Nicholson, A. G., et al. 2017. Effect of nintedanib in subgroups of idiopathic pulmonary fibrosis by diagnostic criteria. *American Journal of Respiratory and Critical Care Medicine*, 195 (1), 78-85.
- [122] Reichmann, W. M., Yu, Y. F., Macaulay, D., et al. 2015. Change in forced vital capacity and associated subsequent outcomes in patients with newly diagnosed idiopathic pulmonary fibrosis. *BMC Pulm Med*, 15 167.
- [123] Richeldi, L., Ryerson, C. J., Lee, J. S., et al. 2012. Relative versus absolute change in forced vital capacity in idiopathic pulmonary fibrosis. *Thorax*, 67 (5), 407-11.
- [124] Richeldi, L., du Bois, R. M., Raghu, G., et al. 2014. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med*, 370 (22), 2071-82.
- [125] Richeldi, L., Cottin, V., du Bois, R. M., et al. 2016. Nintedanib in patients with idiopathic pulmonary fibrosis: Combined evidence from the TOMORROW and INPULSIS trials. *Respiratory Medicine*, 113 74-79.
- [126] Rinciog, C., Watkins, M., Chang, S., et al. 2017. A Cost-Effectiveness Analysis of Nintedanib in Idiopathic Pulmonary Fibrosis in the UK. *Pharmacoeconomics*, 35 (4), 479-491.
- [127] Schulz, K. F., Altman, D. G., Moher, D., et al. 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340 c332.
- [128] Song, J. W., Hong, S. B., Lim, C. M., et al. 2011. Acute exacerbation of idiopathic pulmonary fibrosis: incidence, risk factors and outcome. *Eur Respir J*, 37 (2), 356-63.
- [129] Swigris, J. J., Brown, K. K., Behr, J., et al. 2010a. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med*, 104 (2), 296-304.
- [130] Swigris, J. J., Wamboldt, F. S., Behr, J., et al. 2010b. The 6 minute walk in idiopathic pulmonary fibrosis: longitudinal changes and minimum important difference. *Thorax*, 65 (2), 173-7.
- [131] Swigris, J. J., Han, M., Vij, R., et al. 2012. The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med*, 106 (10), 1447-55.

- [132] Swigris, J. J., Esser, D., Conoscenti, C. S., et al. 2014. The psychometric properties of the St George's Respiratory Questionnaire (SGRQ) in patients with idiopathic pulmonary fibrosis: a literature review. *Health Qual Life Outcomes*, 12 124.
- [133] Swigris, J. J., Esser, D., Wilson, H., et al. 2017. Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis. *Eur Respir J*, 49 (1).
- [134] Swigris, J. J., Wilson, H., Esser, D., et al. 2018. Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis: insights from the INPULSIS trials. *BMJ Open Respir Res*, 5 (1), e000278.
- [135] Taniguchi, H., Xu, Z., Azuma, A., et al. 2016. Subgroup analysis of Asian patients in the INPULSIS trials of nintedanib in idiopathic pulmonary fibrosis. *Respirology*, 21 (8), 1425-1430.
- [136] Velikova, G., Keding, A., Harley, C., et al. 2010. Patients report improvements in continuity of care when quality of life assessments are used routinely in oncology practice: secondary outcomes of a randomised controlled trial. *Eur J Cancer*, 46 (13), 2381-8.
- [137] Wong, S. S., Wilczynski, N. L. & Haynes, R. B. 2006. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc*, 94 (4), 451-5.
- [138] World Health Organization - International Clinical Trials Registry Platform. 2013. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on Forced Vital Capacity decline , in patients with Idiopathic Pulmonary Fibrosis (IPF) (EUCTR2010-024252-29-FI). Verfügbar: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-024252-29-FI> [Aufgerufen am 17.01.2019].
- [139] World Health Organization - International Clinical Trials Registry Platform. 2014. A 52 weeks, double blind, randomized, placebo-controlled trial evaluating the effect of oral BIBF 1120, 150 mg twice daily, on annual Forced Vital Capacity decline , in patients with Idiopathic Pulmonary Fibrosis (IPF) (EUCTR2010-024251-87-DE) Verfügbar: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-024251-87-DE> [Aufgerufen am 17.01.2019].
- [140] World Health Organization - International Clinical Trials Registry Platform. 2018a. A clinical trial to study the safety and efficacy of BIBF 1120 at high dose in patients with lung fibrosis because of unknown cause (CTRI/2011/12/002245). Verfügbar: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2011/12/002245> [Aufgerufen am 17.01.2019].
- [141] World Health Organization - International Clinical Trials Registry Platform. 2018b. A clinical trial to study the safety and efficacy of BIBF 1120 at high dose in patients with lung fibrosis because of unknown cause (CTRI/2011/12/002248). Verfügbar: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2011/12/002248> [Aufgerufen am 17.01.2019].
- [142] World Health Organization - International Clinical Trials Registry Platform. 2018c. A Six Month Double Blind Randomized Placebo Controlled Trial Followed by Each Arm Being Converted to Oral Nintedanib 150 mg Twice Daily Comparing the Effect on High Resolution Computerized Tomography Quantitative Lung Fibrosis Score, Lung Function, Six Minute Walk Test Distance and St. George's Respiratory Questionnaire After Six Months of Treatment in Patients With Idiopathic Pulmonary Fibrosis With Continued Evaluations Over a Period of up to Eighteen Months (NCT01979952).

Verfügbar: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01979952>
[Aufgerufen am 17.01.2019].

- [143] Wuyts, W. A., Kolb, M., Stowasser, S., et al. 2016. First Data on Efficacy and Safety of Nintedanib in Patients with Idiopathic Pulmonary Fibrosis and Forced Vital Capacity of ≤ 50 % of Predicted Value. *Lung*, 194 (5), 739-743.
- [144] Yorke, J., Jones, P. W. & Swigris, J. J. 2010. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. *Thorax*, 65 (10), 921-6.
- [145] Zappala, C. J., Latsi, P. I., Nicholson, A. G., et al. 2010. Marginal decline in forced vital capacity is associated with a poor outcome in idiopathic pulmonary fibrosis. *Eur Respir J*, 35 (4), 830-5.

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 ²³] – 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

²³ 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.

14	or/12-13	773621
15	and/7,11,14	719

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

Datenbankname	Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	17.01.2019	
Zeitsegment	Stand: Dezember 2018	
Suchfilter	Es wurde kein Filter verwendet.	
#	Suchbegriffe	Ergebnis
1	exp Idiopathic Pulmonary Fibrosis/	182
2	exp Idiopathic Interstitial Pneumonias/	190
3	(IPF* or pulmon* fibros* or idiopathi* pulmon* fibros* or Idiopath* diffuse* interstit* pulmon* fibros* or fibros* alveolit* or cryptogen* fibros* alveolit* or diffus* fibrosi* alveolit* or idiopathi* interstiti* pneumoni* or usual* interstiti* pneumoni* or UIP* or interstiti* pneumoni* or IIP* or diffus* parenchym* lung* disease* or DPLD*).mp.	1585
4	1 or 2 or 3	1588
5	(nintedanib* or intedanib* or BIBF-1120* or BIBF1120*).mp.	325
6	(Ofev* or Ovev* or Ofef* or Ovef*).mp.	6
7	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	0
8	5 or 6 or 7	330
9	4 and 8	155

Datenbankname	Embase Classic + Embase	
Suchoberfläche	Ovid	
Datum der Suche	17.01.2019	
Zeitsegment	1947 bis 16.01.2019	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 (Wong et al., 2006) – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	exp fibrosing alveolitis/	21510
2	exp interstitial lung disease/	76610
3	exp interstitial pneumonia/	15183
4	exp lung fibrosis/	77595
5	(IPF* or pulmon* fibros* or idiopathi* pulmon* fibros* or Idiopath* diffuse* interstit* pulmon* fibros* or fibros* alveolit* or cryptogen* fibros* alveolit* or diffus* fibrosi* alveolit* or idiopathi* interstiti* pneumoni* or usual* interstiti* pneumoni* or UIP* or interstiti* pneumoni* or IIP* or diffus* parenchym* lung* disease* or DPLD*).mp.	60206
6	1 or 2 or 3 or 4 or 5	129554
7	exp nintedanib/	2101
8	(Ofev* or Ovev* or Ofef* or Ovef*).mp.	208
9	(nintedanib* or intedanib* or BIBF-1120* or BIBF1120*).mp.	2402
10	(928326-83-4 or 656247-17-5 or 656247-18-6).rn.	1852
11	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	0
12	7 or 8 or 9 or 10 or 11	2552
13	random*.tw. or placebo*.mp. or double-blind*.tw.	1641015
14	6 and 12 and 13	367

Datenbankname		Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R)
Suchoberfläche		Ovid
Datum der Suche		17.01.2019
Zeitsegment		1946 bis 16.01.2019
Suchfilter		Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity (Wong et al., 2006) ²⁴
#	Suchbegriffe	Ergebnis
1	exp Idiopathic Pulmonary Fibrosis/	2976
2	exp Pulmonary Fibrosis/	20675
3	exp Lung Diseases, Interstitial/	55843
4	exp Idiopathic Interstitial Pneumonias/	4197
5	(IPF* or pulmon* fibros* or idiopathi* pulmon* fibros* or Idiopath* diffuse* interstiti* pulmon* fibros* or fibros* alveolit* or cryptogen* fibros* alveolit* or diffus* fibrosi* alveolit* or idiopathi* interstiti* pneumoni* or usual* interstiti* pneumoni* or UIP* or interstiti* pneumoni* or IIP* or diffus* parenchym* lung* disease* or DPLD*).mp.	35989
6	1 or 2 or 3 or 4 or 5	82145
7	(nintedanib* or intedanib* or BIBF-1120* or BIBF1120*).mp.	657
8	(Ofev* or Ovev* or Ofef* or Ovef*).mp.	45
9	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	1
10	7 or 8 or 9	692
11	randomi#ed controlled trial.pt. or randomi#ed.mp. or placebo.mp.	858299
12	6 and 10 and 11	92

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-A3: Suche nach nicht-randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend

²⁴ Der RCT-Filter nach Wong et al. 2006 (Wong et al., 2006) wurde in Suchschritt 11 modifiziert („randomi#ed“ statt „randomized“), um sowohl Treffer mit der englischen als auch der amerikanischen Schreibweise zu identifizieren.

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](http://www.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

Studienregister/ Studienergebnisdatenbank	clinicaltrials.gov
Internetadresse	https://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	17.01.2019
Suchstrategie	(IPF OR idiopathic pulmonary fibrosis OR cryptogenic fibrosing alveolitis OR Idiopathic diffuse interstitial pulmonary fibrosis OR diffuse fibrosing alveolitis OR usual interstitial pneumonitis OR idiopathic interstitial pneumonia OR UIP) [Condition or disease] AND (nintedanib OR intedanib OR BIBF 1120 OR BIBF1120 OR Ofev OR Ovev OR Ofef OR Ovef OR 928326-83-4 OR 656247-17-5 OR 656247-18-6) [Intervention/treatment]
Trefferzahl	29

Studienregister/ Studienergebnisdatenbank	ICTRP
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	17.01.2019
Suchstrategie	nintedanib OR BIBF-1120 OR BIBF1120 OR Ofev OR 928326-83-4 OR 656247-17-5 OR 656247-18-6
Trefferzahl	176

Studienregister/ Studienergebnisdatenbank	EUCTR
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	17.01.2019
Suchstrategie	nintedanib OR (BIBF 1120) OR BIBF1120 OR Ofev OR 928326-83-4 OR 656247-17-5 OR 656247-18-6
Trefferzahl	51

Studienregister/ Studienergebnisdatenbank	PharmNet.Bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html
Datum der Suche	17.01.2019
Suchstrategie	?nintedanib? [Active Substance] ODER ?intedanib? [Active Substance] ODER ?BIBF 1120? [Active Substance] ODER ?BIBF1120? [Active Substance] ODER ?Ofev? [Active Substance] ODER ?928326-83-4? [Cas-number] ODER ?656247-17-5? [Cas-number] ODER ?656247-18-6? [Cas-number]
Trefferzahl	25

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-B3: Suche nach nicht-randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend

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

Nr.	Im Volltext ausgeschlossene Publikationen	Ausschlussgrund
1)	Anonymous. 2016. Nintedanib. Australian Prescriber, 39(2): 62-63	A7 Abweichende Publikationstypen
2)	Atkins, C. P.; Loke, Y. K.; Wilson, A. M.. 2014. Outcomes in idiopathic pulmonary fibrosis: A meta-analysis from placebo controlled trials. Respiratory Medicine, 108(2): 376-387	A2 Abweichende Intervention
3)	Bouvry, D.; Jouneau, S.. 2015. Idiopathic pulmonary fibrosis: Nintedanib, pirfenidone, beyond the randomized clinical trials. [French]. Revue des Maladies Respiratoires Actualites, 7(3): 256-262	A7 Abweichende Publikationstypen
4)	Canestaro, W. J.; Forrester, S. H.; Raghu, G.; Ho, L.; Devine, B. E.. 2016. Drug treatment of idiopathic pulmonary fibrosis systematic review and network meta-analysis. Chest, 149(3): 756-766	A7 Abweichende Publikationstypen
5)	Chan, P.; Bax, L.; Chen, C.; Zhang, N.; Huang, S. P.; Soares, H.; Rosen, G.; AbuTarif, M.. 2017. Model-based meta-analysis on the efficacy of pharmacological treatments for idiopathic pulmonary fibrosis. CPT: Pharmacometrics and Systems Pharmacology, 6(10): 695-704	A7 Abweichende Publikationstypen
6)	Crooks, M. G.; Hart, S. P.. 2014. A new era of drug therapy for idiopathic pulmonary fibrosis. The Lancet Respiratory Medicine, 2(12): 964-6	A7 Abweichende Publikationstypen
7)	Diamantopoulos, A.; Maher, T. M.; Schoof, N.; Esser, D.; LeReun, C.. 2018. Influence of Idiopathic Pulmonary Fibrosis Progression on Healthcare Resource Use. PharmacoEconomics Open, 25: 25	A5 Abweichender Studientyp
8)	Einecke, D.. 2015. [Two drugs inhibit disease progression. Idiopathic lung fibrosis treatable for the first time]. MMW Fortschritte der Medizin, 157(2): 24-5	A7 Abweichende Publikationstypen
9)	Fleetwood, K.; McCool, R.; Glanville, J.; Edwards, S. C.; Gsteiger, S.; Daigl, M.; Fisher, M.. 2017. Systematic Review and Network Meta-analysis of Idiopathic Pulmonary Fibrosis Treatments. Journal of managed care & specialty pharmacy, 23(3b Supplement): S5-S16	A7 Abweichende Publikationstypen
10)	Goh, N.; Raghu, G.; Wells, A.; Nicholson, A.; Richeldi, L.; Flaherty, K. et al.. 2016. Consistent effect of nintedanib on decline in FVC in patients across subgroups based on high-resolution computed tomography diagnostic criteria: Results from the INPULSIS trials in idiopathic pulmonary fibrosis. Respiriology, 21(148)	A7 Abweichende Publikationstypen

Nr.	Im Volltext ausgeschlossene Publikationen	Ausschlussgrund
11)	Jeldres, A.Labarca, G.. 2017. Is nintedanib effective for idiopathic pulmonary fibrosis?. Medwave, 17(Suppl2): e6918	A7 Abweichende Publikationstypen
12)	Loveman, E.; Copley, V. R.; Colquitt, J. L.; Scott, D. A.; Clegg, A. J.; Jones, J.; O'Reilly, K. M. A.; Singh, S.; Bausewein, C.Wells, A.. 2014. The effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: Systematic review, network meta-analysis and health economic evaluation. BMC Pharmacology and Toxicology, 15 (1) (no pagination)(63)	A7 Abweichende Publikationstypen
13)	Loveman, E.; Copley, V. R.; Scott, D. A.; Colquitt, J. L.; Clegg, A. J.O'Reilly, K. M. A.. 2015. Comparing new treatments for idiopathic pulmonary fibrosis - a network meta-analysis. BMC Pulmonary Medicine, 15 (1) (no pagination)(37)	A7 Abweichende Publikationstypen
14)	Paterniti, M. O.; Bi, Y.; Rekec, D.; Wang, Y.; Karimi-Shah, B. A.Chowdhury, B. A.. 2017. Acute exacerbation and decline in forced vital capacity are associated with increased mortality in idiopathic pulmonary fibrosis. Annals of the American Thoracic Society, 14(9): 1395-1402	A2 Abweichende Intervention
15)	Richeldi, L.; Costabel, U.; Selman, M.; Kim, D. S.; Hansell, D. M.; Nicholson, A. G.; Brown, K. K.; Flaherty, K. R.; Noble, P. W.; Raghu, G.; Brun, M.; Gupta, A.; Juhel, N.; Kluglich, M.Du Bois, R. M.. 2011. Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. New England Journal of Medicine, 365(12): 1079-1087	A2 Abweichende Intervention
16)	Richeldi, L.; Kreuter, M.; Selman, M.; Crestani, B.; Kirsten, A. M.; Wuyts, W. A.; Xu, Z.; Bernois, K.; Stowasser, S.; Quaresma, M.Costabel, U.. 2017. Long-term treatment of patients with idiopathic pulmonary fibrosis with nintedanib: results from the TOMORROW trial and its open-label extension. Thorax,	A2 Abweichende Intervention
17)	Rochweg, B.; Neupane, B.; Zhang, Y.; Garcia, C. C.; Raghu, G.; Richeldi, L.; Brozek, J.; Beyene, J.Schunemann, H.. 2016. Treatment of idiopathic pulmonary fibrosis: A network meta-analysis. BMC Medicine, 14 (1) (no pagination)(18)	A7 Abweichende Publikationstypen
18)	Rogliani, P.; Calzetta, L.; Cavalli, F.; Matera, M. G.Cazzola, M.. 2016. Pirfenidone, nintedanib and N-acetylcysteine for the treatment of idiopathic pulmonary fibrosis: A systematic review and meta-analysis. Pulmonary Pharmacology and Therapeutics, 40: 95-103	A7 Abweichende Publikationstypen
19)	Schmid, U.; Doege, C.; Dallinger, C.Freiwald, M.. 2018. Population pharmacokinetics of nintedanib in patients with idiopathic pulmonary fibrosis. Pulmonary Pharmacology & Therapeutics, 48: 136-143	A4 Abweichende Endpunkte
20)	Swigris, J. J.; Esser, D.; Wilson, H.; Conoscenti, C. S.; Schmidt, H.; Stansen, W.; Leidy, N. K.Brown, K. K.. 2017. Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis. European Respiratory Journal, 49 (1) (no pagination)(1601788)	A2 Abweichende Intervention
21)	Swigris, J. J.; Wilson, H.; Esser, D.; Conoscenti, C. S.; Stansen, W.; Kline Leidy, N.Brown, K. K.. 2018. Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis: insights from the INPULSIS trials. BMJ open respiratory research, 5(1): e000278	A5 Abweichender Studientyp

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-C3: Suche nach nicht-randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend

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

Tabelle 4-98: Übersicht ausgeschlossene Studienregistereinträge

Register	Trefferzahl Anhang 4-B1	Ausgeschlossene Registereinträge Anhang 4-D1	Eingeschlossene Registereinträge Tabelle 4-10
clinicaltrials.gov	29	26 (Position 1-26)	3
EU-CTR	51	49 (Position 27-75)	2
ICTRP	176	171 (Position 76-246)	5
PharmNet.Bund	25	23 (Position 247-269)	2
Summe	$\Sigma=281$	$\Sigma=269$	$\Sigma=12$

Tabelle 4-99: Ausgeschlossene Studienregistereinträge

Nr.	Studien-ID	Titel	Ausschlussgrund
Clinicaltrials.gov			
1)	NCT00514683	Safety And Efficacy of BIBF 1120 in Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2015 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT00514683	A2 Abweichende Intervention
2)	NCT01136174	Safety and PK Study of BIBF 1120 in Japanese Patients With IPF. ClinicalTrials.gov. 2015 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01136174	A6 zu kurze Studiendauer
3)	NCT01170065	Roll Over Study From 1199.30 BIBF 1120 in Idiopathic Pulmonary Fibrosis (IPF). ClinicalTrials.gov. 2017 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01170065	A5 Abweichender Studientyp
4)	NCT01417156	Safety and PK Study of BIBF 1120 in Japanese Patients With IPF: Follow up Study From 1199.31(NCT01136174). ClinicalTrials.gov. 2017 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01417156	A5 Abweichender Studientyp
5)	NCT01619085	Extension Trial of the Long Term Safety of BIBF 1120 in Patients With Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01619085	A5 Abweichender Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
6)	NCT01695408	Investigating Significant Health Trends in Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01695408	A5 Abweichender Studientyp
7)	NCT02171156	Expanded Access Program of Nintedanib in Patients With Idiopathic Pulmonary Fibrosis (EAP). ClinicalTrials.gov. 2017 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02171156	A5 Abweichender Studientyp
8)	NCT02230982	Expanded Access Program of Nintedanib in Patients With Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2016 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02230982	A5 Abweichender Studientyp
9)	NCT02551068	High Oxygen Delivery to Preserve Exercise Capacity in Idiopathic Pulmonary Fibrosis Patients Treated With Nintedanib. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02551068	A3 Abweichender Komparator
10)	NCT02579603	Safety, Tolerability and PK of Nintedanib in Combination With Pirfenidone in IPF. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02579603	A3 Abweichender Komparator
11)	NCT02597933	A Trial to Compare Nintedanib With Placebo for Patients With Scleroderma Related Lung Fibrosis. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02597933	A1 Falsche Population
12)	NCT02598193	Safety and Tolerability Study of Pirfenidone in Combination With Nintedanib in Participants With Idiopathic Pulmonary Fibrosis (IPF). ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02598193	A5 Abweichender Studientyp
13)	NCT02606877	A Study to Compare the Amount of Nintedanib and Pirfenidone in the Blood When Nintedanib and Pirfenidone Are Given Separately or in Combination. ClinicalTrials.gov. 2017 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02606877	A5 Abweichender Studientyp
14)	NCT02607722	All-Case Surveillance of Ofev in Patients With IPF in Japan. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02607722	A5 Abweichender Studientyp
15)	NCT02788474	Effect of Nintedanib on Biomarkers of Extracellular Matrix Turnover in Patients With Idiopathic Pulmonary Fibrosis and Limited Forced Vital Capacity Impairment. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02788474	A6 zu kurze Studiendauer
16)	NCT02802345	Efficacy and Safety of Nintedanib Co-administered With Sildenafil in Idiopathic Pulmonary Fibrosis Patients With Advanced Lung Function Impairment. ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT02802345	A3 Abweichender Komparator
17)	NCT03047031	Post Marketing Surveillance of Nintedanib in Indian Patients With Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03047031	A5 Abweichender Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
18)	NCT03104322	Patient-reported Monitoring of Symptoms and Spirometry Via the patientMpower Platform in Idiopathic Pulmonary Fibrosis. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03104322	A2 Abweichende Intervention
19)	NCT03242759	Non-Interventional Study (NIS) Collecting Experiences For IPF in Taiwan. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03242759	A5 Abweichender Studientyp
20)	NCT03281200	Chart Review Study to Describe the Clinical Profile of Idiopathic Pulmonary Fibrosis (IPF) Patients Treated With Nintedanib (OFEV®) in Real-world Practice in Spain. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03281200	A5 Abweichender Studientyp
21)	NCT03313180	A Trial to Evaluate the Safety of Long Term Treatment With Nintedanib in Patients With Scleroderma Related Lung Fibrosis.. ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03313180	A1 Falsche Population
22)	NCT03422068	This Study Tests Different Doses of BI 1015550 in Patients With Idiopathic Pulmonary Fibrosis (IPF). The Study Tests How BI 1015550 is Taken up by the Body and How Well it is Tolerated. The Study Also Looks at BI 1015550 When it is Taken in Addition to Nintedanib and Pirfenidone. ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03422068	A6 zu kurze Studiendauer
23)	NCT03502902	The Safety, Tolerability and Pharmacokinetic Study of HEC68498 in Healthy Male and Female Subjects. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03502902	A1 Falsche Population
24)	NCT03562416	Continuation of Nintedanib After Single Lung Transplantation in IPF Subjects. ClinicalTrials.gov. 2018 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03562416	A7 Keine Studienergebnisse
25)	NCT03710824	Investigating Trends in Quality of Life in Patients With Idiopathic Pulmonary Fibrosis (IPF) Under Treatment With Nintedanib. ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03710824	A5 Abweichender Studientyp
26)	NCT03717012	Study of Pulmonary Rehabilitation in Patients With Idiopathic Pulmonary Fibrosis (IPF). ClinicalTrials.gov. 2019 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://ClinicalTrials.gov/show/NCT03717012	A3 Abweichender Komparator
EU-CTR			
27)	2005-000317-35	A double-blind, randomized Phase II study to determine efficacy and safety of oral treatment with BIBF 1120 ES 250 mg twice daily versus 150 mg twice daily in patients suffering from advanced non-s.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2005-000317-35	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
28)	2005-002427-14	A Randomised Placebo-Controlled Phase II Study of Continuous Maintenance Treatment with BIBF 1120 Following Chemotherapy in Patients with Relapsed Ovarian Cancer. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2005-002427-14	A1 Falsche Population
29)	2005-005249-21	A multi-centre 3-arm randomized phase II trial of BIBF 1120 versus BIBW 2992 versus sequential administration of BIBF 1120 and BIBW 2992 in patients with hormone-resistant prostate cancer. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2005-005249-21	A1 Falsche Population
30)	2006-002875-42	A 12 month, double blind, randomized, placebo-controlled trial evaluating the effect of BIBF 1120 administered at oral doses of 50 mg qd, 50 mg bid, 100 mg bid and 150 mg bid on Forced Vital Capaci.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-002875-42	A2 Abweichende Intervention
31)	2006-004528-35	A multicenter, randomized, open label phase II trial evaluating the efficacy and safety of mFOLFOX7 plus weekly alternating sequential oral administration of BIBF 1120 250 mg twice daily and BIBW 2.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2006-004528-35	A1 Falsche Population
32)	2007-004803-36	Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy i.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-004803-36	A1 Falsche Population
33)	2008-002072-10	Multicenter, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard pemetrexed therapy compared to placebo plus standard pemetrexed therapy.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-002072-10	A1 Falsche Population
34)	2008-005364-14	A phase I-II study of BIBF 1120 and Folfox compared to Bevacizumab and Folfox in first line metastatic colorectal cancer patients.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-005364-14	A1 Falsche Population
35)	2008-006831-10	Multicenter, randomised, double-blind Phase III trial to investigate the efficacy and safety of BIBF 1120 in combination with carboplatin and paclitaxel compared to placebo plus carboplatin and pac.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-006831-10	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
36)	2009-009516-44	A randomised, open label, parallel group Phase II study comparing the efficacy and tolerability of BIBF 1120 versus sunitinib in previously untreated patients with Renal Cell Cancer.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-009516-44	A1 Falsche Population
37)	2009-011925-14	A multicenter, open label, phase I / II study to evaluate safety, pharmacokinetics and efficacy of BIBF 1120 in comparison with oral sorafenib for advanced hepatocellular carcinoma patients.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-011925-14	A1 Falsche Population
38)	2009-013788-21	Ensayo clínico de Fase II, abierto, y de extensión para evaluar la tolerabilidad a largo plazo, la seguridad y la eficacia de BIBF 1120 administrado por vía oral en pacientes con Fibrosis Pulmonar EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-013788-21	A5 Abweichender Studientyp
39)	2010-019707-32	LUME-Lung 3. A Phase I/II study of continuous oral treatment with BIBF 1120 added to standard gemcitabine/cisplatin therapy in first line NSCLC patients with squamous cell histology.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-019707-32	A1 Falsche Population
40)	2010-021407-24	Phase II study of BIBF 1120 in recurrent Glioblastoma multiforme. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-021407-24	A1 Falsche Population
41)	2010-023050-37	A multicenter, randomized, phase II trial: BIBF 1120 vs. placebo in patients receiving oxaliplatin plus fluorouracil and leucovorin (mFOLFOX6) for advanced, chemorefractory metastatic colorectal ca.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-023050-37	A1 Falsche Population
42)	2011-000921-61	A phase I/II, randomized, open-label, multi-centre study of BIBF1120 + reirradiation (R-RT) versus reirradiation in the treatment of patients with first or second progression of glioblastoma. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-000921-61	A1 Falsche Population
43)	2011-001086-41	A randomized, placebo-controlled, multi-center phase I/II trial to assess the safety and efficacy of nintedanib (BIBF 1120) added to low-dose cytarabine in elderly patients with AML unfit for an in.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-001086-41	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
44)	2011-001649-34	Ensayo clínico de fase I/II, aleatorizado, de paclitaxel neoadyuvante frente a imprinación con BIBF 1120 seguida por BIBF 1120 más paclitaxel en cáncer de mama HER-2 negativo con estudios correlati.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-001649-34	A1 Falsche Population
45)	2011-002766-21	An open-label extension trial of the long term safety of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis (IPF). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-002766-21	A5 Abweichender Studientyp
46)	2011-004250-26	A Phase I/II, Open-Label, Safety, Pharmacokinetic, and Preliminary Efficacy Study of Oral Rucaparib in Patients with gBRCA Mutation Ovarian Cancer, or Other Solid Tumor. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-004250-26	A1 Falsche Population
47)	2011-005814-12	Phase II, randomised, placebo controlled, multicentre, feasibility study of low dose (metronomic) cyclophosphamide with or without nintedanib (BIBF 1120) in advanced ovarian cancer (METRO-BIBF). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-005814-12	A1 Falsche Population
48)	2011-006288-23	Randomized double blind placebo-controlled phase II trial of Vargatef® in addition to first line chemotherapy with interval debulking surgery in patients with adenocarcinoma of the ovary, the fallo.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-006288-23	A1 Falsche Population
49)	2012-000095-42	A double-blind, randomised, placebo controlled Phase III study of nintedanib plus best supportive care (BSC) versus placebo plus BSC in patients with colorectal cancer refractory to standard therap.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-000095-42	A1 Falsche Population
50)	2012-004076-19	ENGOT-cx1/BGOG-cx1: “Randomized double-blind Phase II study comparing 3-weekly carboplatin (AUC 5) + paclitaxel 175 mg/m2 with or without concomitant and maintenance nintedanib (NINTEDANIB) in adv.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004076-19	A1 Falsche Population
51)	2012-004295-19	A phase II study exploring the safety and efficacy of nintedanib (BIBF1120) as second line therapy for patients with either differentiated or medullary thyroid cancer progressing after first line t.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004295-19	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
52)	2012-004895-21	Phase II randomised placebo controlled NEOadjuvant chemotherapy study of Nintedanib with Gemcitabine and Cisplatin in locally advanced muscle invasive BLADder cancEr. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-004895-21	A1 Falsche Population
53)	2012-005201-48	LUME-Meso: Double blind, randomised, multicentre, phase II/III study of nintedanib in combination with pemetrexed / cisplatin followed by continuing nintedanib monotherapy versus placebo in combina.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-005201-48	A1 Falsche Population
54)	2013-002109-73	A Randomised Phase II Study of Nintedanib (BIBF1120) Compared to Chemotherapy in Patients with Recurrent Clear Cell Carcinoma of the Ovary or Endometrium. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-002109-73	A1 Falsche Population
55)	2013-004324-11	Phase II study with oral fibroblast growth factor-1 inhibitor BIBF1120 as second line treatment in lung carcinoma patients harboring fibroblast growth factor receptor-1 gene amplification (NVALT-15.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-004324-11	A1 Falsche Population
56)	2013-004458-34	A Phase I/II, Multicenter, Randomized, Double-Blind, Placebo-Controlled Trial Evaluating the Efficacy and Safety of Nintedanib/Vargatef in Combination With Paclitaxel Chemotherapy for Treatment of EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-004458-34	A1 Falsche Population
57)	2014-003891-22	Nintedanib (BIBF 1120) plus docetaxel in NSCLC patients progressing after first-line CTX: angiogenic biomarker identification, phase II trial. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-003891-22	A1 Falsche Population
58)	2014-004058-32	The Efficacy and Mechanism Evaluation of Treating Idiopathic Pulmonary Fibrosis with the Addition of Co-trimoxazole (EMETIPAC). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-004058-32	A2 Abweichende Intervention
59)	2014-004782-24	A Phase 2 Trial to Evaluate the Efficacy of PRM-151 in Subjects with Idiopathic Pulmonary Fibrosis (IPF). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-004782-24	A7 Keine Studienergebnisse

Nr.	Studien-ID	Titel	Ausschlussgrund
60)	2014-005016-42	“Multicenter, prospective, open label study , in patients with non small cell lung cancer (NSCLC), stage IIIB / IV or recurrent, after the failure of first line chemotherapy in which it’s possible EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005016-42	A1 Falsche Population
61)	2014-005341-44	Molecular-biological tumor profiling for drug treatment selection in patients with advanced and refractory carcinoma. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005341-44	A1 Falsche Population
62)	2015-000392-28	A double blind, randomised, placebo-controlled trial evaluating efficacy and safety of oral nintedanib treatment for at least 52 weeks in patients with Systemic Sclerosis associated Interstitial Lu.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-000392-28	A1 Falsche Population
63)	2015-000640-42	A twelve week, open-label, randomised, parallel-group study evaluating safety, tolerability and pharmacokinetics (PK) of oral nintedanib in combination with oral pirfenidone, compared to treatment EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-000640-42	A3 Abweichender Komparator
64)	2015-002619-14	INSTAGE™: A 24-week, double-blind, randomized, parallel-group study evaluating the efficacy and safety of oral nintedanib coadministered with oral sildenafil, compared to treatment with ninteda.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-002619-14	A3 Abweichender Komparator
65)	2015-003148-38	A 12-week, double blind, randomised, placebo controlled, parallel group trial followed by a single active arm phase of 40 weeks evaluating the effect of oral nintedanib 150 mg twice daily on change.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003148-38	A6 zu kurze Studiendauer
66)	2015-003280-11	AN EXPLORATORY MULTICENTER, OPEN-LABEL, SINGLE ARM STUDY OF THE SAFETY AND TOLERABILITY OF PIRFENIDONE (ESBRIET®) IN COMBINATION WITH NINTEDANIB (OFEV®) IN PATIENTS WITH IDIOPATHIC PULMONARY FIBROSIS. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003280-11	A5 Abweichender Studientyp
67)	2015-003360-37	A double blind, randomized, placebo-controlled trial evaluating the efficacy and safety of nintedanib over 52 weeks in patients with Progressive Fibrosing Interstitial Lung Disease (PF-ILD). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003360-37	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
68)	2015-004919-20	A pilot study of nintedanib for lymphangioleiomyomatosis (LAM). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-004919-20	A1 Falsche Population
69)	2016-000193-38	ENGOT-EN1/FANDANGO: A randomized phase II trial of first-line combination chemotherapy with nintedanib / placebo for patients with advanced or recurrent endometrial cancer.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000193-38	A1 Falsche Population
70)	2016-000521-38	Nintedanib as maintenance treatment of malignant pleural mesothelioma (NEMO): a double-blind randomized phase II study of the EORTC Lung Cancer Group. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000521-38	A1 Falsche Population
71)	2016-000877-19	A phase ½ dose-escalation study to evaluate the safety, tolerability, and pharmacokinetics and preliminary efficacy of DTX-SPL8783 (a docetaxel (DTX)-dendrimer conjugate) as monotherapy in patien.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000877-19	A1 Falsche Population
72)	2016-002093-12	A Phase II multicenter study comparing the efficacy of the oral angionenesis inhibitor nintedanib with the intravenous cytotoxis compound ifosfamide for treatment of patients with advanced metastat.... EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-002093-12	A1 Falsche Population
73)	2016-003403-66	An open-label extension trial to assess the long term safety of nintedanib in patients with ‘Systemic Sclerosis associated Interstitial Lung Disease’ (SSc-ILD). EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-003403-66	A1 Falsche Population
74)	2017-003158-18	A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of BG00011 in Patients With Idiopathic Pulmonary Fibrosis. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003158-18	A2 Abweichende Intervention
75)	2017-004302-18	A Phase II randomized, double-blind, placebo-controlled, 26-week study to evaluate the efficacy, safety and tolerability of GLPG1205 in subjects with idiopathic pulmonary fibrosis.. EU-CTR. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-004302-18	A2 Abweichende Intervention
ICTRP			

Nr.	Studien-ID	Titel	Ausschlussgrund
76)	CTRI/2013/01/003290	A long term clinical study to determine safety of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis (IPF). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2013/01/003290	A5 Nicht randomisierte Studie
77)	CTRI/2013/02/003343	LUME-Lung 1: BIBF 1120 Plus Docetaxel as Compared to Placebo Plus Docetaxel in 2nd Line Non Small Cell Lung Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2013/02/003343	A1 Falsche Population
78)	CTRI/2016/02/006617	A Trial to Compare Nintedanib With Placebo for Patients With Scleroderma Related Lung Fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2016/02/006617	A1 Falsche Population
79)	CTRI/2017/01/007727	Study of Nintedanib and Sildenafil in advanced IPF patients. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2017/01/007727	A3 Abweichender Komparator
80)	CTRI/2017/02/007857	An active surveillance to monitor the real world safety of Nintedanib in Indian NSCLC patients. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2017/02/007857	A1 Falsche Population
81)	CTRI/2017/02/007867	An active surveillance to monitor the real world safety of Nintedanib in Indian Idiopathic Pulmonary Fibrosis (a type of lung disease in which tissue deep in our lungs becomes thick and stiff, or scarred, over the time) patients. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2017/02/007867	A5 Nicht randomisierte Studie
82)	CTRI/2018/03/012529	A study to review the safety of long term treatment with nintedanib in patients with scleroderma related lung fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2018/03/012529	A1 Falsche Population
83)	DRKS00003676	A Multicenter, Open Label, Phase I /Randomised Phase II Study to Evaluate Safety, Pharmacokinetics and Efficacy of BIBF 1120 in Comparison With Oral Sorafenib for Advanced Hepatocellular Carcinoma Patients. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00003676	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
84)	DRKS00003756	Multicentre, Randomised, Double-blind, Phase III Trial to Investigate the Efficacy and Safety of Oral BIBF 1120 Plus Standard Docetaxel Therapy Compared to Placebo Plus Standard Docetaxel Therapy in Patients With Stage IIIB/IV or Recurrent Non Small Cell Lung Cancer After Failure of First Line Chemotherapy. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00003756	A1 Falsche Population
85)	DRKS00003942	Multicenter, Randomised, Double-blind Phase III Trial to Investigate the Efficacy and Safety of BIBF 1120 in Combination With Carboplatin and Paclitaxel Compared to Placebo Plus Carboplatin and Paclitaxel in Patients With Advanced Ovarian Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00003942	A1 Falsche Population
86)	DRKS00008383	A Double-blind, Randomised, Placebo Controlled Phase III Study of Nintedanib Plus Best Supportive Care (BSC) Versus Placebo Plus BSC in Patients With Colorectal Cancer Refractory to Standard Therapies. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00008383	A1 Falsche Population
87)	DRKS00010759	A Phase I/II, Randomized, Open-label, Multi-centre Study of BIBF1120 + Reirradiation (R-RT) Versus Reirradiation in the Treatment of Patients With First or Second Progression of Glioblastoma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00010759	A1 Falsche Population
88)	EUCTR2005-000317-35-DE	A double-blind, randomized Phase II study to determine efficacy and safety of oral treatment with BIBF 1120 ES 250 mg twice daily versus 150 mg twice daily in patients suffering from advanced non-small-cell lung cancer. - -. ICTRP. 2012. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2005-000317-35-DE	A1 Falsche Population
89)	EUCTR2005-002427-14-GB	A Randomised Placebo-Controlled Phase II Study of Continuous Maintenance Treatment with BIBF 1120 Following Chemotherapy in Patients with Relapsed Ovarian Cancer - -. ICTRP. 2012. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2005-002427-14-GB	A1 Falsche Population
90)	EUCTR2005-005249-21-GB	A multi-centre 3-arm randomized phase II trial of BIBF 1120 versus BIBW 2992 versus sequential administration of BIBF 1120 and BIBW 2992 in patients with hormone-resistant prostate cancer. ICTRP. 2012. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2005-005249-21-GB	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
91)	EUCTR2006-002875-42-DE	A 12 month, double blind, randomized, placebo-controlled trial evaluating the effect of BIBF 1120 administered at oral doses of 50 mg qd, 50 mg bid, 100 mg bid and 150 mg bid on Forced Vital Capacity decline during one year, in patients with Idiopathic Pulmonary Fibrosis, with optional active treatment extension until last patient out. - Effect of orally administered BIBF 1120 on FVC decline in IPF. ICTRP. 2013. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2006-002875-42-DE	A2 Abweichende Intervention
92)	EUCTR2006-004528-35-DE	A multicenter, randomized, open label phase II trial evaluating the efficacy and safety of mFOLFOX7 plus weekly alternating sequential oral administration of BIBF 1120 250 mg twice daily and BIBW 2992 50 mg once daily (BB) versus mFOLFOX7 alone as first-line therapy in patients with metastatic colorectal cancer. - Combination of BIBF 1120, BIBW 2992 and mFOLFOX7 as first-line therapy for colorectal cancer. ICTRP. 2012. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2006-004528-35-DE	A1 Falsche Population
93)	EUCTR2007-004803-36-DE	LUME-Lung 1: BIBF 1120 plus docetaxel as compared to placebo plus docetaxel in 2nd line non small cell lung cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2007-004803-36-DE	A1 Falsche Population
94)	EUCTR2008-002072-10-HU	BIBF 1120 plus pemetrexed compared to placebo plus pemetrexed in 2nd line nonsquamous NSCLC. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-002072-10-HU	A1 Falsche Population
95)	EUCTR2008-005364-14-BE	A phase I-II study of BIBF 1120 and Folfox compared to Bevacizumab and Folfox in first line metastatic colorectal cancer patients. - BIBF 1120 vs Bevacizumab in CRC. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-005364-14-BE	A1 Falsche Population
96)	EUCTR2008-006831-10-DE	BIBF 1120 or placebo in combination with paclitaxel and carboplatin in first line treatment of ovarian cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-006831-10-DE	A1 Falsche Population
97)	EUCTR2009-009516-44-GB	Compare safety and efficacy of BIBF 1120 versus sunitinib. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-009516-44-GB	A1 Falsche Population
98)	EUCTR2009-011925-14-GB	Phase I/II comparison of efficacy and safety of BBF 1120 and sorafenib in patients with advanced hepatocellular carcinoma. ICTRP. 2014. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-011925-14-GB	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
99)	EUCTR2009-013788-21-BE	A phase II open label, roll over study of the long term tolerability, safety and efficacy of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-013788-21-BE	A5 Nicht randomisierte Studie
100)	EUCTR2009-013788-21-ES	Ensayo clínico de Fase II, abierto, y de extensión para evaluar la tolerabilidad a largo plazo, la seguridad y la eficacia de BIBF 1120 administrado por vía oral en pacientes con Fibrosis Pulmonar Idiopática A phase II open label, roll over study of the long term tolerability, safety and efficacy of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-013788-21-ES	A5 Nicht randomisierte Studie
101)	EUCTR2010-019707-32-GB	A Phase I/II study of continuous oral treatment with BIBF 1120 added to standard gemcitabine/cisplatin therapy in first line NSCLC patients with squamous cell histology. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-019707-32-GB	A1 Falsche Population
102)	EUCTR2010-021407-24-DK	Phase II study of BIBF 1120 in recurrent Glioblastoma multiforme. ICTRP. 2014. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-021407-24-DK	A1 Falsche Population
103)	EUCTR2010-023050-37-DE	A multicenter, randomized, phase II trial: BIBF 1120 vs. placebo in patients receiving oxaliplatin plus fluorouracil and leucovorin (mFOLFOX6) for advanced, chemorefractory metastatic colorectal cancer (mCRC) - TRICC-C. ICTRP. 2012. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-023050-37-DE	A1 Falsche Population
104)	EUCTR2011-000921-61-DE	Randomisierte, unverblindete, multizentrische klinische Prüfung der Phase I/II zum Vergleich einer Re-Bestrahlung plus einer Gabe von BIBF1120 gegenüber einer alleinigen Re-Bestrahlung zur Behandlung von Patienten mit erster oder zweiter Progression eines Glioblastoms. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-000921-61-DE	A1 Falsche Population
105)	EUCTR2011-001086-41-DE	A trial to test the effect of nintedanib together with chemotherapy in patients with acute myeloid leukemia. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-001086-41-DE	A1 Falsche Population
106)	EUCTR2011-001649-34-ES	Ensayo Clínico para pacientes con cáncer de mama que compara BIBF 1120 frente BIBF 1120 administrado junto a paclitaxel. ICTRP. 2014. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-001649-34-ES	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
107)	EUCTR2011-002766-21-GR	Long term safety study of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis (IPF). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-002766-21-GR	A5 Nicht randomisierte Studie
108)	EUCTR2011-006288-23-FR	Evaluation of efficacy and safety of using Nintedanib (Vargatef®) as part of the neoadjuvant and adjuvant treatment surrounding interval debulking surgery in patients with advanced ovarian cancer. ICTRP. 2013. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-006288-23-FR	A1 Falsche Population
109)	EUCTR2012-004076-19-BE	ENGOT-cx1/BGOG-cx1: “Randomized double-blind Phase II study comparing 3-weekly carboplatin + paclitaxel with or without concomitant and maintenance nintedanib (NINTEDANIB) in advanced or recurrent cervical carcinoma.”. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-004076-19-BE	A1 Falsche Population
110)	EUCTR2012-004895-21-GB	A study to investigate a new drug called Nintedanib for patients who have bladder cancer that has spread to the deeper walls of the bladder. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-004895-21-GB	A1 Falsche Population
111)	EUCTR2012-005201-48-DE	Nintedanib (BIBF 1120) in Mesothelioma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-005201-48-DE	A1 Falsche Population
112)	EUCTR2013-002109-73-GB	NiCCC - A Randomised Phase II Study Of Nintedanib (BIBF1120) Compared To Chemotherapy in Patients With Recurrent Clear Cell Carcinoma Of The Ovary Or Endometrium. ICTRP. 2014. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-002109-73-GB	A1 Falsche Population
113)	EUCTR2013-004324-11-NL	Study with BIBF1120 in lung cancer patients with an abnormal fibroblast growth factor-1 receptor (NVALT-15 study). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-004324-11-NL	A1 Falsche Population
114)	EUCTR2013-004458-34-DE	Clinical Trial to investigate the efficacy and safety of nintedanib/vargatef when combined with paclitaxel (chemotherapy) for treatment of patients suffering from metastatic melanoma with BRAF wildtyp. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-004458-34-DE	A1 Falsche Population
115)	EUCTR2014-003891-22-AT	Nintedanib (BIBF 1120) plus docetaxel in non-small cell lung cancer (NSCLC) patients progressing after first-line chemotherapy (CTX): angiogenic biomarker identification, phase II trial. ICTRP.	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-003891-22-AT	
116)	EUCTR2014-005016-42-IT	“Multicenter, prospective, open label study , in patients with non small cell lung cancer (NSCLC), stage IIIB / IV or recurrent, after the failure of first line chemotherapy in which it’s possible a treatment that evaluates the effectiveness and the safety of oral BIBF 1120 in combination with standard docetaxel therapy, which can be administered (at the discretion of the referring physician) with tri-weekly or weekly schedule. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-005016-42-IT	A1 Falsche Population
117)	EUCTR2015-000392-28-NL	A trial to compare nintedanib with placebo for patients with scleroderma related lung fibrosis. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-000392-28-NL	A1 Falsche Population
118)	EUCTR2015-000640-42-FR	Safety, tolerability and PK of nintedanib in combination with pirfenidone in IPF. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-000640-42-FR	A3 Abweichender Komparator
119)	EUCTR2015-002619-14-DE	Efficacy and safety of nintedanib when co-administered with sildenafil in IPF patients with advanced lung function impairment. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-002619-14-DE	A3 Abweichender Komparator
120)	EUCTR2015-003148-38-ES	Effect of nintedanib on biomarkers. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003148-38-ES	A6 zu kurze Studiendauer
121)	EUCTR2015-003280-11-DE	A Study of the Safety and Tolerability of Pirfenidone (Esbriet) in Combination with Nintedanib (Ofev) in Patients with Idiopathic Pulmonary Fibrosis. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003280-11-DE	A5 Nicht randomisierte Studie
122)	EUCTR2015-003360-37-ES	Efficacy and safety of nintedanib in patients with PF-ILD. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003360-37-ES	A1 Falsche Population
123)	EUCTR2015-004919-20-IT	A pilot study of nintedanib for lymphangioleiomyomatosis (LAM). ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004919-20-IT	A1 Falsche Population
124)	EUCTR2016-000193-38-SE	ENGOT-EN1/FANDANGO: A randomized double-blind phase II trial of combination chemotherapy with nintedanib or placebo for patients with endometrial cancer. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter:	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-000193-38-SE	
125)	EUCTR2016-000877-19-GB	A phase ½ study in which, DTX-SPL8783, the trial medication will be administered (alone or in combination with nintedanib) in order to assess the safety, tolerability, pharmacokinetics (distribution of DTX-SPL8783 in the body) and efficacy in patients with advanced malignant tumours. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-000877-19-GB	A1 Falsche Population
126)	ISRCTN50772895	A study of Nintedanib compared to chemotherapy in patients with recurrent Clear Cell Carcinoma of the ovary or endometrium. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ISRCTN50772895	A1 Falsche Population
127)	ISRCTN56349930	NEOadjuvant chemotherapy study of Nintedanib with Gemcitabine and Cisplatin in locally advanced muscle invasive BLADder cancEr. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ISRCTN56349930	A1 Falsche Population
128)	JPRN-UMIN000017114	Multicenter Phase 1b/2 Trial of Nintedanib with TAS-102 in Patients with Metastatic Colorectal Cancer (mCRC) who had the progression or intolerant to standard therapies. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000017114	A1 Falsche Population
129)	JPRN-UMIN000019436	Randomized Phase 2 study of Nintedanib and Pirfenidone versus Nintedanib following a clinically meaningful decline in forced vital capacity in patients with idiopathic pulmonary fibrosis administering pirfenidone. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000019436	A3 Abweichender Komparator
130)	JPRN-UMIN000019751	Prospective study of predictor for efficiency of treatment with Nintedanib for idiopathic pulmonary fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000019751	A5 Nicht randomisierte Studie
131)	JPRN-UMIN000020682	The comparison of the efficacy and safety of pirfenidone and nintedanib in patients with idiopathic pulmonary fibrosis. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000020682	A5 Nicht randomisierte Studie
132)	JPRN-UMIN000020722	Prospective study of efficiency of fibrosis score in computed tomography as predictor after treatment with Nintedanib for idiopathic pulmonary fibrosis. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter:	A5 Nicht randomisierte Studie

Nr.	Studien-ID	Titel	Ausschlussgrund
		http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000020722	
133)	JPRN-UMIN000021036	Effects of nintedanib in idiopathic pulmonary fibrosis - analysis on long term prognosis and predictive factors -. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000021036	A5 Nicht randomisierte Studie
134)	JPRN-UMIN000021591	The pilot study of Nintedanib and Carboplatin plus Paclitaxel for advanced non-small cell lung cancer with idiopathic interstitial pneumonia. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000021591	A5 Nicht randomisierte Studie
135)	JPRN-UMIN000022232	Prospective observational study to predict sensitive marker after administration of nintedanib in patients with idiopathic pulmonary fibrosis. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000022232	A5 Nicht randomisierte Studie
136)	JPRN-UMIN000026376	Long-term Effect of Pulmonary Rehabilitation under Nintedanib treatment in Idiopathic Pulmonary Fibrosis. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000026376	A3 Abweichender Komparator
137)	JPRN-UMIN000026799	A randomized phase II study of carboplatin plus nab-paclitaxel with or without nintedanib for advanced non-small-cell lung cancer with idiopathic pulmonary fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000026799	A6 zu kurze Studiendauer
138)	NCT00514683	Safety And Efficacy of BIBF 1120 in Idiopathic Pulmonary Fibrosis. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00514683	A2 Abweichende Intervention
139)	NCT00706628	A Multi-centre 3-arm Randomised Phase II Trial of BIBF 1120 Versus BIBW 2992 Versus Sequential Administration of BIBF 1120 and BIBW 2992 in Patients With Hormone-resistant Prostate Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00706628	A1 Falsche Population
140)	NCT00710762	A Randomised Placebo-Controlled Phase II Study of Continuous Maintenance Treatment With BIBF 1120 Following Chemotherapy in Patients With Relapsed Ovarian Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00710762	A1 Falsche Population
141)	NCT00715403	A Phase I/II Open Label Extension Study of BIBF 1120 Administered Orally Once or Twice Daily to Establish Safety, Pharmacokinetics and Efficacy in Patients With Advanced Solid Tumours and Clinical Benefit From Previous Therapy With BIBF 1120. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00715403	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
142)	NCT00730821	Phase I of BIBW 2992/BIBF 1120 Combination Therapy in Solid Tumors. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00730821	A1 Falsche Population
143)	NCT00801294	A Phase II Trial of Weekly Alternating Sequential Administration of BIBF 1120 and BIBW 2992 in Patients With Advanced Colorectal Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00801294	A1 Falsche Population
144)	NCT00876460	BIBF 1120 + Docetaxel (Japan) in Patients With Advanced Non-small-cell Lung Cancer, Phase I. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00876460	A1 Falsche Population
145)	NCT00979576	BIBF 1120 in Combination With Pemetrexed in Advanced Non Small Cell Lung Cancer (NSCLC). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00979576	A1 Falsche Population
146)	NCT00987935	A Phase I / II Trial of Nintedanib in Asian Hepatocellular Carcinoma Patients. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00987935	A1 Falsche Population
147)	NCT00998296	Phase I Dose Escalation Study of Concomitant BIBF 1120 and BIBW 2992 in Patients With Advanced Solid Tumours. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00998296	A1 Falsche Population
148)	NCT01022853	Combination of BI6727 (Volasertib) and BIBF1120 in Solid Tumors. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01022853	A1 Falsche Population
149)	NCT01024920	Compare Safety and Efficacy of BIBF 1120 Versus Sunitinib. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01024920	A1 Falsche Population
150)	NCT01136174	Safety and PK Study of BIBF 1120 in Japanese Patients With IPF. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01136174	A6 zu kurze Studiendauer
151)	NCT01225887	Nintedanib in Treating Patients With Recurrent or Persistent Endometrial Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01225887	A1 Falsche Population
152)	NCT01251484	BIBF 1120 in Recurrent Glioblastoma Multiforme. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01251484	A1 Falsche Population
153)	NCT01314105	BIBF 1120 + Carboplatin/Pegylated Liposomal Doxorubicin (PLD) in Patients With Advanced Ovarian Cancer, Fallopian Tube Carcinoma or Primary Peritoneal Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01314105	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
154)	NCT01329549	Dose Escalation Study of BIBF 1120 in Combination With Carboplatin and PLD in Relapsed Ovarian Cancer (OC). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01329549	A1 Falsche Population
155)	NCT01346540	A Phase I/II Study of Continuous Oral Treatment With BIBF 1120 Added to Standard Gemcitabine/Cisplatin Therapy in First Line NSCLC Patients With Squamous Cell Histology. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01346540	A1 Falsche Population
156)	NCT01349296	BIBF 1120 and RAD001 in Solid Tumors - Phase I. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01349296	A1 Falsche Population
157)	NCT01380782	BIBF 1120 for Recurrent High-Grade Gliomas. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01380782	A1 Falsche Population
158)	NCT01417156	Safety and PK Study of BIBF 1120 in Japanese Patients With IPF: Follow up Study From 1199.31(NCT01136174). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01417156	A5 Nicht randomisierte Studie
159)	NCT01441297	BIBF 1120 as Second Line Treatment for Small Cell Lung Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01441297	A1 Falsche Population
160)	NCT01484080	Neoadjuvant Paclitaxel Versus BIBF 1120 Priming Followed by BIBF 1120 Plus Paclitaxel in Early HER-2 Negative Breast Cancer With Proteomic and Dynamic Imaging Correlates. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01484080	A1 Falsche Population
161)	NCT01485874	Doxorubicin + BIBF 1120 in Patients for Ovarian Cancer. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01485874	A1 Falsche Population
162)	NCT01548924	Determination of Dose of Antiangiogenic Multitargeted DOVITINIB (TKI258) Plus Paclitaxel in Patients With Solid Tumors. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01548924	A1 Falsche Population
163)	NCT01583322	Vargatef in Addition to First Line Chemotherapy With Interval Debulking Surgery in Patients With Ovarian Cancer. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01583322	A1 Falsche Population
164)	NCT01594125	Dose Escalation Study of Nintedanib (BIBF 1120) in Japanese Patients With Hepatocellular Carcinoma. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01594125	A1 Falsche Population
165)	NCT01610869	Low Dose Cyclophosphamide +/- Nintedanib in Advanced Ovarian Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01610869	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
166)	NCT01619085	Extension Trial of the Long Term Safety of BIBF 1120 in Patients With Idiopathic Pulmonary Fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01619085	A5 Nicht randomisierte Studie
167)	NCT01658462	Phase II Study of Docetaxel +/- Nintedanib in Breast Cancer. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01658462	A1 Falsche Population
168)	NCT01666600	NOA-12: BIBF1120 and R-RT in Glioblastoma. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01666600	A1 Falsche Population
169)	NCT01679613	Investigation of Drug-drug Interaction of Nintedanib and Ketoconazole in Healthy Male Volunteers. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01679613	A1 Falsche Population
170)	NCT01683682	An Dose Escalation Trial of Oral BIBF 1120 in Combination With Intravenous Carboplatin and Vinorelbine in Elderly Patients With Advanced Non-Small Cell Lung Cancer - Stage IV (VENUS-2). ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01683682	A1 Falsche Population
171)	NCT01684111	An Dose Escalation Trial of Oral BIBF 1120 in Combination With Intravenous Vinorelbine in Elderly Patients With Advanced Non Small Lung Cell Cancer - Stage IV (VENUS-1). ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01684111	A1 Falsche Population
172)	NCT01770392	Relative Bioavailability of a Single Dose of Nintedanib Given Alone and in Combination With Multiple Doses of Rifampicin. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01770392	A1 Falsche Population
173)	NCT01788982	Nintedanib(BIBF1120) in Thyroid Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01788982	A1 Falsche Population
174)	NCT01948141	Nintedanib in Treating Patients With Advanced Non-Small Cell Lung Cancer Who Have Failed Up to Two Previous Chemotherapy Regimens. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01948141	A1 Falsche Population
175)	NCT01951846	To Determine the Maximum Tolerated Dose (MTD) of BIBF 1120 in Patients With Solid Tumours. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01951846	A1 Falsche Population
176)	NCT02009579	ENGOT-cx1/BGOG-cx1: 3 Weekly Carboplatin/Paclitaxel With or Without Nintedanib in Cervix Cancer. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02009579	A1 Falsche Population
177)	NCT02149108	Nintedanib (BIBF 1120) vs Placebo in Refractory Metastatic Colorectal Cancer (LUME-Colon 1). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02149108	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
178)	NCT02152059	A Study to Evaluate the Good and Bad Effects of BIBF1120 in Small Cell Lung Cancer Patients Who Have Previously Benefited From First-line Platinum-based Chemotherapy. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02152059	A1 Falsche Population
179)	NCT02171156	Expanded Access Program of Nintedanib in Patients With Idiopathic Pulmonary Fibrosis (EAP). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02171156	A5 Nicht randomisierte Studie
180)	NCT02182050	Efficacy and Safety of Oral Treatment With BIBF 1120 ES in Advanced Non-small-cell Lung Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182050	A1 Falsche Population
181)	NCT02182063	A Study With BIBF 1120 in Patients With Hormone Refractory Prostate Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182063	A1 Falsche Population
182)	NCT02182076	Pharmacokinetics/Bioavailability of BIBF 1120 Administered to Healthy Male Volunteers. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182076	A1 Falsche Population
183)	NCT02182102	Dose Escalation of BIBF 1120 Combined With Pemetrexed in Advanced Non Small Cell Lung Cancer. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182102	A1 Falsche Population
184)	NCT02182128	A Dose-escalation Study of BIBF 1120 in Japanese Patients With Advanced Solid Tumours. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182128	A1 Falsche Population
185)	NCT02182141	An Dose Escalation Study of BIBF 1120 Administered in Patients With Relapsed or Refractory Multiple Myeloma. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182141	A1 Falsche Population
186)	NCT02182154	Metabolism and Pharmacokinetics of [14C]-BIBF 1120 in Healthy Male Volunteers. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182154	A1 Falsche Population
187)	NCT02182193	Safety and Relative Bioavailability of BIBF 1120 Soft Gelatine Capsules Charge 1, BIBF 1120 Soft Gelatine Capsules Charge 2 and BIBF 1120 Drinking Solution in Healthy Male Volunteers. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182193	A1 Falsche Population
188)	NCT02182206	An Dose Escalation Study of Treatment With BIBF 1120 in Patients With Advanced Solid Tumours. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182206	A1 Falsche Population
189)	NCT02182219	Oral Treatment With BIBF 1120 Together With Docetaxel and Prednisone in Patients With Hormone Refractory Prostate Cancer.	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182219	
190)	NCT02182232	A Dose Escalation Study of BIBF 1120 Together With Paclitaxel and Carboplatin in Patients With Advanced Stage Non-small-cell Lung Cancer. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182232	A1 Falsche Population
191)	NCT02182245	Dose Escalation Study of BIBF 1120 in Patients With Advanced Gynaecological Malignancies. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182245	A1 Falsche Population
192)	NCT02182258	Safety and Tolerability Study of BIBF 1120 as Intravenous Infusion and Absolute Bioavailability of BIBF 1120 as Soft Gelatine Capsule in Healthy Subjects. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02182258	A1 Falsche Population
193)	NCT02191865	Nintedanib in Volunteers With Hepatic Impairment Compared With Healthy Volunteers. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02191865	A1 Falsche Population
194)	NCT02225405	Cisplatin, Docetaxel, and Nintedanib Before Surgery in Treating Patients With Previously Untreated Stage IB-III A Non-small Cell Lung Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02225405	A3 Abweichender Komparator
195)	NCT02230982	Expanded Access Program of Nintedanib in Patients With Idiopathic Pulmonary Fibrosis. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02230982	A5 Nicht randomisierte Studie
196)	NCT02231164	LUME-Columbus: Nintedanib Plus Docetaxel in Advanced Non-small Cell Lung Cancer With Translational Research. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02231164	A1 Falsche Population
197)	NCT02234596	Nintedanib in Patients With Advanced Esophagogastric Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02234596	A1 Falsche Population
198)	NCT02278978	BIBF1120 in Patients With Advanced FGFR3 Mutated, Overexpressed, or Wild Type Urothelial Carcinoma. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02278978	A1 Falsche Population
199)	NCT02299141	Nintedanib in Molecularly Selected Patients With Advanced Non-Small Cell Lung Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02299141	A1 Falsche Population
200)	NCT02300298	Nintedanib Plus Docetaxel in Japanese Patients With Adenocarcinoma Subtype Non-small Cell Lung Cancer After Failure of First Line Chemotherapy. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02300298	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
201)	NCT02308553	Efficacy and Safety of Nintedanib Combined With Paclitaxel Chemotherapy for Patients With BRAF wt Metastatic Melanoma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02308553	A1 Falsche Population
202)	NCT02389764	Nintedanib For HER2-Negative Metastatic Inflammatory Breast Cancer (MIBC). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02389764	A1 Falsche Population
203)	NCT02393755	Nintedanib and Capecitabine in Treating Patients With Refractory Metastatic Colorectal Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02393755	A1 Falsche Population
204)	NCT02399215	Nintedanib in Treating Patients With Locally Advanced or Metastatic Neuroendocrine Tumors. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02399215	A1 Falsche Population
205)	NCT02452463	Nintedanib Compared With Placebo in Treating Against Radiation-Induced Pneumonitis in Patients With Non-small Cell Lung Cancer That Cannot Be Removed by Surgery and Are Undergoing Chemoradiation Therapy. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02452463	A1 Falsche Population
206)	NCT02496585	Study to Evaluate the Efficacy and Safety of Nintedanib (BIBF 1120) + Prednisone Taper in Patients With Radiation Pneumonitis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02496585	A1 Falsche Population
207)	NCT02531737	Nintedanib (Vargatef®) Plus Docetaxel in Second Line of Treatment in Patients With Lung Cancer. ICTRP. 2016. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02531737	A1 Falsche Population
208)	NCT02551068	High Oxygen Delivery to Preserve Exercise Capacity in Idiopathic Pulmonary Fibrosis Patients Treated With Nintedanib. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02551068	A3 Abweichender Komparator
209)	NCT02558387	Trial of BIBF1120 (Nintedanib) in Patients With Recurrent or Metastatic Salivary Gland Cancer of the Head and Neck. ICTRP. 2015. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02558387	A1 Falsche Population
210)	NCT02568449	Nintedanib in Treating Patients With Malignant Pleural Mesothelioma That Is Recurrent. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02568449	A1 Falsche Population
211)	NCT02572752	Relative Bioavailability of 2 Oral Formulations of Nintedanib. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02572752	A1 Falsche Population
212)	NCT02606877	A Study to Compare the Amount of Nintedanib and Pirfenidone in the Blood When Nintedanib and Pirfenidone Are Given Separately or in Combination. ICTRP. 2018. [Zugriffsdatum: 17.01.2019].	A5 Nicht randomisierte Studie

Nr.	Studien-ID	Titel	Ausschlussgrund
		Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02606877	
213)	NCT02607722	All-Case Surveillance of Ofev in Patients With IPF in Japan. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02607722	A5 Nicht randomisierte Studie
214)	NCT02619162	Nintedanib+Letrozole in Postmenopausal Women With Breast Cancer: Clinical Trial Safety and Pharmacodynamics. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02619162	A1 Falsche Population
215)	NCT02665143	A Randomized Trial of a Combination of Nintedanib/Placebo in Combination With Induction Chemotherapy for Patients With Refractory or First Relapse Acute Myeloid Leukemia. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02665143	A1 Falsche Population
216)	NCT02667704	Influence of Bosentan on the Pharmacokinetics of Nintedanib. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02667704	A1 Falsche Population
217)	NCT02668393	Nintedanib and Weekly Docetaxel in Lung Adenocarcinoma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02668393	A1 Falsche Population
218)	NCT02730416	Combination Chemotherapy With Nintedanib / Placebo in Endometrial Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02730416	A1 Falsche Population
219)	NCT02751385	Investigation of the Effect of Nintedanib on the Pharmacokinetics of a Combination of Ethinylestradiol and Levonorgestrel in Patients With Non-small Cell Lung Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02751385	A1 Falsche Population
220)	NCT02780700	Nintedanib Alone or in Combination With Capecitabine in Refractory Metastatic Colorectal Cancer [LUME-Colon 2]. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02780700	A1 Falsche Population
221)	NCT02788474	Effect of Nintedanib on Biomarkers of Extracellular Matrix Turnover in Patients With Idiopathic Pulmonary Fibrosis and Limited Forced Vital Capacity Impairment. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02788474	A6 zu kurze Studiendauer
222)	NCT02802345	Efficacy and Safety of Nintedanib When Co-administered With Sildenafil in Idiopathic Pulmonary Fibrosis Patients With Advanced Lung Function Impairment. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02802345	A3 Abweichender Komparator
223)	NCT02808247	Ph II Nintedanib vs. Ifosfamide in Soft Tissue Sarcoma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02808247	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
224)	NCT02835833	Study of Nintedanib Plus Bevacizumab in Advanced Solid Tumors. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02835833	A1 Falsche Population
225)	NCT02856425	Trial Of Pembrolizumab And Nintedanib. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02856425	A1 Falsche Population
226)	NCT02856867	Nintedanib Plus mFOLFOX6 for Previously Untreated Metastatic Esophagogastric Adenocarcinoma (MEGAN). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02856867	A1 Falsche Population
227)	NCT02863055	Nintedanib as Switch Maintenance Treatment of Pleural Malignant Mesothelioma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02863055	A1 Falsche Population
228)	NCT02902484	Study of Nintedanib and Chemotherapy for Advanced Pancreatic Cancer. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02902484	A1 Falsche Population
229)	NCT02999178	Efficacy and Safety of Nintedanib in Patients With Progressive Fibrosing Interstitial Lung Disease (PF-ILD). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02999178	A1 Falsche Population
230)	NCT03017885	Post Marketing Surveillance of Nintedanib in Indian Patients With Non-small Cell Lung Cancer (NSCLC) After First-line Therapy. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03017885	A1 Falsche Population
231)	NCT03047031	Post Marketing Surveillance of Nintedanib in Indian Patients With Idiopathic Pulmonary Fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03047031	A5 Nicht randomisierte Studie
232)	NCT03062943	A Study of Nintedanib for LymphAngioleiomyomatosis (LAM). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03062943	A1 Falsche Population
233)	NCT03242759	Non-Interventional Study (NIS) Collecting Experiences For IPF in Taiwan. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03242759	A5 Nicht randomisierte Studie
234)	NCT03281200	Chart Review Study to Describe the Clinical Profile of Idiopathic Pulmonary Fibrosis (IPF) Patients Treated With Nintedanib (OFEV®) in Real-world Practice in Spain. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03281200	A5 Nicht randomisierte Studie
235)	NCT03283007	Nintedanib in Lung Transplant Recipients With Bronchiolitis Obliterans Syndrome Grade 1-2. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03283007	A1 Falsche Population
236)	NCT03287947	LCI-GI-APX-NIN-001: Nintedanib in Metastatic Appendiceal Carcinoma. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03287947	
237)	NCT03292250	Korean Cancer Study Group: Translational biomarker Driven UMBrella Project for Head and Neck (TRIUMPH), Esophageal Squamous Cell Carcinoma- Part 1 (HNSCC). ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03292250	A1 Falsche Population
238)	NCT03313180	A Trial to Evaluate the Safety of Long Term Treatment With Nintedanib in Patients With Scleroderma Related Lung Fibrosis. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03313180	A1 Falsche Population
239)	NCT03361319	Combination Nab-paclitaxel (N-P) and Nintedanib or N-P and Placebo in Relapsed NSCLC Adenocarcinoma. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03361319	A1 Falsche Population
240)	NCT03363815	A Study to Evaluate the Effect of Multiple Doses of CC-90001 on the Pharmacokinetics of Omeprazole, Midazolam, Warfarin, Rosuvastatin, Metformin, Digoxin, and Nintedanib in Healthy Adult Subjects. ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03363815	A1 Falsche Population
241)	NCT03422068	This Study Tests Different Doses of BI 1015550 in Patients With Idiopathic Pulmonary Fibrosis (IPF). The Study Tests How BI 1015550 is Taken up by the Body and How Well it is Tolerated. The Study Also Looks at BI 1015550 When it is Taken in Addition to Nintedanib and Pirfenidone. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03422068	A6 zu kurze Studiendauer
242)	NCT03513484	Nintedanib and Azacitidine in Treating Participants With HOX Gene Overexpression Relapsed or Refractory Acute Myeloid Leukemia. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03513484	A1 Falsche Population
243)	NCT03675581	A Study to Test Whether Nintedanib Influences the Components of Birth-control Pills in Women With Systemic Sclerosis Associated Interstitial Lung Disease (SSc-ILD). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03675581	A1 Falsche Population
244)	NCT03710824	Investigating Trends in Quality of Life in Patients With Idiopathic Pulmonary Fibrosis (IPF) Under Treatment With Nintedanib. ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03710824	A5 Nicht randomisierte Studie
245)	NCT03717012	Study of Pulmonary Rehabilitation in Patients With Idiopathic Pulmonary Fibrosis (IPF). ICTRP. 2018. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03717012	A3 Abweichender Komparator
246)	NTR4588	Treatment with BIBF1120 capsules lung carcinoma patients with an abnormal fibroblast growth factor-1 receptor (NVALT15).	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		ICTRP. 2017. [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NTR4588	
PharmNet.Bund			
247)	2006-002875-42	A 12 month, double blind, randomized, placebo-controlled trial evaluating the effect of BIBF 1120 administered at oral doses of 50 mg qd, 50 mg bid, 100 mg bid and 150 mg bid on Forced Vital Capacity decline during one year, in patients with Idiopathic Pulmonary Fibrosis, with optional active treatment extension until last patient out.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A2 Abweichende Intervention
248)	2006-004528-35	A multicenter, randomized, open label phase II trial evaluating the efficacy and safety of mFOLFOX7 plus weekly alternating sequential oral administration of BIBF 1120 250 mg twice daily and BIBW 2992 50 mg once daily (BB) versus mFOLFOX7 alone as first-line therapy in patients with metastatic colorectal cancer.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
249)	2007-004803-36	Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of first line chemotherapy. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
250)	2008-002072-10	Multicenter, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard pemetrexed therapy compared to placebo plus standard pemetrexed therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of firstline chemotherapy. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
251)	2008-005364-14	A phase I-II study of BIBF 1120 and Folfex compared to Bevacizumab and Folfex in first line metastatic colorectal cancer patients.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
252)	2009-011925-14	A multicenter, open label, phase I / II study to evaluate safety, pharmacokinetics and efficacy of BIBF 1120 in comparison with oral sorafenib for advanced hepatocellular carcinoma patients.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
253)	2009-013788-21	A phase II open label, roll over study of the long term tolerability, safety and efficacy of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter:	A5 Abweichender Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
		http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	
254)	2010-019707-32	LUME-Lung 3. A Phase I/II study of continuous oral treatment with BIBF 1120 added to standard gemcitabine/cisplatin therapy in first line NSCLC patients with squamous cell histology.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
255)	2010-023050-37	A multicenter, randomized, phase II trial: BIBF 1120 vs. placebo in patients receiving oxaliplatin plus fluorouracil and leucovorin (mFOLFOX6) for advanced, chemorefractory metastatic colorectal cancer (mCRC). PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
256)	2011-000921-61	A phase I/II, randomized, open-label, multi-centre study of BIBF1120 +reirradiation (R-RT) versus reirradiation in the treatment of patients with first or second progression of glioblastoma. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
257)	2011-001086-41	A randomized, placebo-controlled, multi-center phase I/II trial to assess the safety and efficacy of nintedanib (BIBF 1120) added to low-dose cytarabine in elderly patients with AML unfit for an intensive induction therapy. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
258)	2011-002766-21	An open-label extension trial of the long term safety of oral BIBF 1120 in patients with Idiopathic Pulmonary Fibrosis (IPF). PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A5 Abweichender Studientyp
259)	2012-000095-42	A double-blind, randomised, placebo controlled Phase III study of nintedanib plus best supportive care (BSC) versus placebo plus BSC in patients with colorectal cancer refractory to standard therapies.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
260)	2012-004076-19	ENGOT-cx1/BGOG-cx1: “Randomized double-blind Phase II study comparing 3-weekly carboplatin (AUC 5) + paclitaxel 175 mg/m2 with or without concomitant and maintenance nintedanib (NINTEDANIB) in advanced or recurrent cervical carcinoma.”. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
261)	2012-005201-48	LUME-Meso: Double blind, randomised, multicentre, phase II/III study of nintedanib in combination with pemetrexed / cisplatin followed by continuing nintedanib monotherapy versus placebo in combination with pemetrexed / cisplatin followed by continuing placebo monotherapy for the treatment of patients with unresectable malignant pleural mesothelioma. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter:	A1 Falsche Population

Nr.	Studien-ID	Titel	Ausschlussgrund
		http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	
262)	2013-004458-34	A Phase I/II, Multicenter, Randomized, Double-Blind, Placebo-Controlled Trial Evaluating the Efficacy and Safety of Nintedanib/Vargatef in Combination With Paclitaxel Chemotherapy for Treatment of Patients with BRAF Wildtype Metastatic Melanoma - Multizentrische, randomisierte, doppelt verblindete, placebo-kontrollierte Phase I/II-Studie zur Prüfung der Wirksamkeit und Sicherheit von Nintedanib/Vargatef in Kombination mit Paclitaxel-Chemotherapie zur Behandlung von Patienten mit BRAF Wildtyp metastasiertem Melanom. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
263)	2015-000392-28	A double blind, randomised, placebo-controlled trial evaluating efficacy and safety of oral nintedanib treatment for at least 52 weeks in patients with Systemic Sclerosis associated Interstitial Lung Disease (SSc-ILD). PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
264)	2015-000640-42	A twelve week, open-label, randomised, parallel-group study evaluating safety, tolerability and pharmacokinetics (PK) of oral nintedanib in combination with oral pirfenidone, compared to treatment with nintedanib alone, in patients with idiopathic pulmonary fibrosis (IPF). PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A3 Abweichender Komparator
265)	2015-002619-14	A 24-week, double-blind, randomized, parallel-group study evaluating the efficacy and safety of oral nintedanib coadministered with oral sildenafil, compared to treatment with nintedanib alone, in patients with idiopathic pulmonary fibrosis (IPF) and advanced lung function impairment. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A3 Abweichender Komparator
266)	2015-003148-38	A 12-week, double blind, randomised, placebo controlled, parallel group trial followed by a single active arm phase of 40 weeks evaluating the effect of oral nintedanib 150 mg twice daily on change in biomarkers of extracellular matrix (ECM) turnover in patients with idiopathic pulmonary fibrosis (IPF) and limited forced vital capacity (FVC) impairment.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A6 zu kurze Studiendauer
267)	2015-003280-11	AN EXPLORATORY MULTICENTER, OPEN-LABEL, SINGLE ARM STUDY OF THE SAFETY AND TOLERABILITY OF PIRFENIDONE (ESBRIET®) IN COMBINATION WITH NINTEDANIB (OFEV®) IN PATIENTS WITH IDIOPATHIC PULMONARY FIBROSIS. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A5 Abweichender Studientyp

Nr.	Studien-ID	Titel	Ausschlussgrund
268)	2016-000193-38	ENGOT-EN1/FANDANGO: A randomized phase II trial of first-line combination chemotherapy with nintedanib / placebo for patients with advanced or recurrent endometrial cancer.. PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population
269)	2016-003403-66	An open-label extension trial to assess the long term safety of nintedanib in patients with 'Systemic Sclerosis associated Interstitial Lung Disease' (SSc-ILD). PharmNet.Bund. 0000 [Zugriffsdatum: 17.01.2019]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	A1 Falsche Population

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend

Anhang 4-D3: Suche nach nicht-randomisierten vergleichenden Studien

Nicht zutreffend

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend

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 inklusive eines Flow-Charts für den Patientenfluss (Tabelle 4-100, Tabelle 4-101).

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 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-100 (Anhang): Studiendesign und -methodik für Studie INPULSIS 1

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2b	Genaue Ziele, Fragestellung und Hypothesen	Zielsetzung Untersuchung der Wirksamkeit sowie des Nutzen-Risiko-Verhältnisses von oral verabreichtem Nintedanib 150 mg bid im Vergleich zu Placebo bei IPF. Als primärer Endpunkt wird die Verringerung der Lungenfunktionsabnahme, gemessen als jährliche Abnahme der FVC, ermittelt. Hypothesen Bewertung der Überlegenheit von Nintedanib 150 mg bid im Vergleich zu Placebo in Bezug auf die jährliche Abnahme der FVC bei Patienten mit IPF. FVC-Abnahme in 52 Wochen: Nullhypothese (H₀): Keine Unterschiede der jährlichen Abnahme der FVC vom Ausgangswert bis Woche 52 zwischen den Gruppen Nintedanib 150 mg bid und Placebo. Alternativhypothese (H_a): Die jährliche Abnahme der FVC vom Ausgangswert bis Woche 52 ist bei Patienten unter Nintedanib 150 mg bid geringer als unter Placebo.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Allokation: randomisiert Zuteilungsverhältnis: 3:2 Randomisierungseinheit: Studienteilnehmer Verblindung: doppelblind Studienhorizont: 52 Wochen + 28 Tage Nachbeobachtung Design: parallel, zweiarstig, placebokontrolliert Studienorganisation: multizentrisch Phase: III
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Amendment 1 (11/2011) Möglichkeit eines Re-Tests bei abnormalen

Item ^a	Charakteristikum	Studieninformation
		Laborparametern zu Visite 1, bei Verdachtsfällen auf Messfehler oder fehlendem bzw. angezweifelttem Bezug auf den Patientenzustand • Studienausschluss, wenn Patienten den Studienablauf nicht verstehen bzw. die Fragebögen nicht selbstständig ausfüllen können • Ergänzung der „schwerwiegenden unerwünschten Ereignisse“ entsprechend den neuen BI-Standards, um deren ordnungsgemäße Dokumentation zu gewährleisten • Einschlusskriterium 4 wurde abgeändert: „Thorax-HRCT innerhalb von 52 Wochen vor Visite 1“ anstelle von „Thorax HRCT innerhalb von 52 Wochen vor Visite 2“ Amendment 2 (09/2012) • Ergänzung explorativer Biomarker-Analysen, um den Effekt von Nintedanib auf Biomarker in Verbindung mit IPF-Pathologie und prognostischen Erkrankungsmarkern zu untersuchen, und Ergänzung pharmakogenomischer Analysen • Schlechte Compliance wurde als Protokollverletzung definiert
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: • Schriftliche Einwilligungserklärung vor Studienbeginn • Alter ≥ 40 Jahre zu Visite 1 • Diagnose IPF nach ATS-, ERS-, JRS- und ALAT-Richtlinien (innerhalb von 5 Jahren vor Ausgangswert) • Thorax HRCT innerhalb von 52 Wochen vor Visite 1) • HRCT-Muster und, wenn verfügbar, Lungenbiopsie-Muster übereinstimmend mit IPF-Diagnose bewertet von zentralen Reviewern • DL_{CO}: 30-79% (zu Visite 2; angepasst an Hämoglobin) • FVC% pred. $\geq 50\%$ (zu Visite 2) Ausschlusskriterien: • Laborparameter (zu Visite 1) • AST und ALT $> 1,5$-fach ULN • Bilirubin $> 1,5$-fach ULN • Lungenfunktion (zu Visite 2) • Atemwegsobstruktion (FEV₁/FVC $< 0,7$) • Lungentransplantation im Verlauf der Studie wahrscheinlich (nach Prüfarzt-Einschätzung) • Andere Krankheiten • Herzerkrankung - Myokardinfarkt innerhalb 6 Monate vor Visite 2 - Instabile Angina innerhalb 1 Monat vor Visite 2 • Blutungsrisiko - Bekannte Prädisposition für Blutungen

Item ^a	Charakteristikum	Studieninformation
		- Fibrinolyse mit im Studienprotokoll festgelegten Interventionen - Hämorrhagisches ZNS-Ereignis innerhalb von 52 Wochen vor Visite 2 - weitere Kriterien innerhalb von 3 Monaten vor Visite 2 - Hämoptysis oder Hämaturie - Aktive gastrointestinale Blutung oder Geschwüre - Größere Verletzung oder Operation - Koagulations Parameter: INR >2, Verlängerung der Prothrombinzeit und PTT (sek) um >50% ULN (Visite 1) • Thromboserisiko - Bekannte erbliche Prädisposition für Thrombose - Thrombotisches Ereignis innerhalb von 52 Wochen (inkl. Schlaganfall und transiente ischämische Attacke) vor Visite 2 • Bekannte Hypersensitivität gegenüber Studienmedikation oder deren Bestandteilen • Andere Krankheiten, die das Testprozedere oder die Studienteilnahme beeinflussen oder den Patienten bei Studienteilnahme in Gefahr bringen (Ermessen des Prüfarztes) • Lebenserwartung aufgrund anderer Krankheiten als IPF <2,5 Jahre (Ermessen des Prüfarztes) • Allgemeine Ausschlusskriterien • Frühere Behandlung mit Nintedanib länger als 4 Wochen) • Andere Interventionen (Forschungsstudie) innerhalb von 8 Wochen vor Visite 1 • Einnahme von N-Acetylcystein (NAC), Prednison >15 mg/Tag oder vergleichbare Arzneimittel innerhalb von 2 Wochen vor Visite 1 • Einnahme von Pirfenidon, Azathioprin, Cyclophosphamid, Cyclosporin A innerhalb von 8 Wochen vor Visite 1 • Geplante größere Operationen während der Studienperiode • Schwangere oder stillende Frauen oder Frauen im gebärfähigen Alter, die keine zwei adäquaten Verhütungsmethoden für mind. 1 Monat vor Behandlungsbeginn und bis 3 Monate nach Behandlungsende verwenden • Sexuell aktive Männer, die sich nicht verpflichten, Kondome während und 3 Monate nach Ende der Behandlung zu nutzen (außer wenn Partner nicht im gebärfähigem Alter sind)

Item ^a	Charakteristikum	Studieninformation
		• Aktiver Alkohol- oder Drogenmissbrauch • Patienten, die das Studienprozedere nicht verstehen oder einhalten (inkl. Ausfüllen von Fragebögen ohne Hilfe)
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie, 98 Studienzentren in 13 Ländern: Australien, Belgien, China, Tschechische Republik, Frankreich, Deutschland, Indien, Irland, Israel, Italien, Japan, UK, USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Alle Gruppen Visite 1: Screening Visite 2: Therapiestart und nachfolgende Einnahme des Arzneimittels über 52 Wochen Nintedanib (oral) • 2-mal täglich 150 mg im Intervall von 12 Stunden über 52 Wochen • Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich 100 mg im Intervall von 12 Stunden über 52 Wochen bzw. in besonderen Fällen Unterbrechung der Therapie für maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Placebo (oral) • 2-mal täglich im Intervall von 12 Stunden über 52 Wochen • Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich Placebo zu 100 mg im Intervall von 12 Stunden über 52 Wochen bzw. in besonderen Fällen Unterbrechung der Therapie von maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Alle Gruppen Follow-up: 28 Tage nach Therapieende
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	Primäres Zielkriterium Untersuchung der Wirksamkeit sowie des Nutzen/Risiko-Verhältnisses von Nintedanib 150 mg bid im Vergleich zu Placebo bei IPF Primäre Endpunkte • Jährliche FVC-Abnahme über 52 Wochen (ml) Key sekundäre Endpunkte • Veränderung des SGRQ-Gesamtscores vom Ausgangswert bis Woche 52 (in Punkten) • Zeit bis zur ersten <i>akuten Exazerbation</i> über 52 Wochen (in Tagen) Weitere sekundäre Endpunkte

Item ^a	Charakteristikum	Studieninformation
		• Weitere Analysen der FVC • Absolute und relative Veränderung der FVC (ml) und FVC % des Sollwertes über 52 Wochen • Absolute kategorielle Veränderung der FVC % des Sollwertes bis zu 52 Wochen: Abfall >5%, Anstieg >5% und Veränderungen ≤5% • Absolute kategorielle Veränderung der FVC % des Sollwertes vom Ausgangswert bis Woche 52: Abfall >10%, Anstieg >10% und Veränderung ≤10% • Anteil FVC-Responder mit Abfall <5% vom FVC-Sollwert nach 52 Wochen • Anteil FVC-Responder mit Abfall <10% vom FVC-Sollwert nach 52 Wochen • Patientenberichteter Endpunkt (PRO) gemessen vom Ausgangswert bis Woche 52: • Anteil SGRQ-(≤4 Punkte)-Responder (definiert als absolute Veränderung im SGRQ-Gesamtscore um ≤4 Punkte) • Veränderung SGRQ-Domänen in Punkten • Veränderung SGRQ-I-Gesamtscore • Veränderung SOBQ • Veränderung Hustenimpact und Hustensymptome (CASA-Q[CD])-Score • Anteil PGI-C-Responder • Veränderung EQ-5D-VAS • Akute Exazerbation: • Risiko einer akuten Exazerbation • Überlebenszeitanalysen gemessen über 52 Wochen: • Zeit bis zum Tod • Zeit bis zum respiratorisch bedingten Tod • Zeit bis zum Tod unter Behandlung • Zeit bis zum Tod oder Lungentransplantation • Zeit bis zum Tod oder Lungentransplantation oder Qualifizierung für Lungentransplantation (Kriterien: FVC <45% des Sollwertes oder DL_{CO} <30% des Sollwertes oder SpO₂<88% in Ruhe) • Veränderung SpO₂ in Ruhe • Veränderung DL_{CO} in Ruhe Sicherheit • Unerwünschte Ereignisse • Vitalzeichen • Körperliche Untersuchung • Gewicht • Klinische Labortests • 12-Kanal-EKG (japanische Studienzentren) Andere Endpunkte • Prä-Dosis Plasmakonzentration (C_{pre,ss,N}) von

Item ^a	Charakteristikum	Studieninformation
		Nintedanib, BIBF 1202 und BIBF 1202 Glukuronid Visite 4 (4 Wochen) und Visite 7 (24 Wochen) - Pharmakogenomische Auswertung - Biomarker-Auswertung
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	- Gepowert auf primären Endpunkt: Jährliche FVC-Abnahme über 52 Wochen (ml) - Erwarteter Effekt: Jährliche Differenz von 100 ml zwischen Nintedanib- und Placebo-Gruppe - Erwartete Abweichung: 300 ml (auf Basis der Phase II-Studienergebnisse) - Erwarteter Anteil nicht auswertbarer Studienteilnehmer: 2% - Zuteilungsverhältnis von 3:2 führt zu einer Fallzahl von: 291 zu 194 (n=485) in Nintedanib zu Placebo - Verwendeter Test: t-Test - Teststärke: 90% (α): 0,025, einseitig - H₀-Hypothese: Es gibt keinen Unterschied zwischen den Behandlungsgruppen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Zwischenanalyse - keine Kriterien für Patientenausschluss - direkt nach Visite 2 (vor Arzneimittel-Ersteinnahme) - AST, ALT >1,5-fach ULN - Bilirubin >1,5-fach ULN - INR >2 - Prothrombinzeit oder PTT >50% ULN - Sonst keine weiteren Regeln für Patientenausschluss vorgesehen Kriterien für patientenindividuellen Therapieausschluss - Leberfunktion - Leberenzyme - ALT oder AST $\geq$8-fach ULN - ALT oder AST $\geq$3-fach ULN für mehr als 2 Wochen - Anzeichen akuter Leberschaden - ALT oder AST $\geq$3-fach ULN und Gesamt-Bilirubin $\geq$1,5-fach ULN oder INR >1,5 oder Anzeichen von Fatigue, Übelkeit, Erbrechen, Schmerz oder Druckempfindlichkeit im rechten oberen abdominalen Quadranten, Fieber, Ausschlag, und/oder Eosinophile >5%

Item ^a	Charakteristikum	Studieninformation
		- Andere Kriterien - Wiederauftreten einer Diarrhö (>4 Stuhlgänge/Tag) für 8 oder mehr aufeinander folgende Tage trotz Dosisreduktion und Begleitbehandlung - Gleichzeitige Einnahme von gemäß den Ausschlusskriterien verbotenen Arzneimitteln - Schwangerschaft - Widerruf der Einwilligung Kriterien für dauerhaften Abbruch der Studienmedikation - Größere Operation (abdominal und intestinal) - Patienten mit vollständiger Antikoagulation - Größere thromboembolische Ereignisse (Schlaganfall, tiefe Venenthrombose, Lungenembolie, Myokardinfarkt) - Erhöhtes Blutungsrisiko (hämorrhagisches ZNS-Ereignis, Hämoptysis oder Hämaturie, aktive gastrointestinale Blutung oder Geschwüre) Kriterien für Studienausschluss - Widerruf der Einwilligung Kriterien für Studienabbruch durch Sponsor - Nichteinhaltung der Aufnahmekriterien - Auftreten von Wirkung/Sicherheit-Informationen, die möglicherweise die Fortsetzung des Tests beeinflusst haben - Verstoß GCP, CTP oder Vertragsbruch seitens Prüfzentrum oder Prüfer
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung über Interaktives Telefon/Internet-Response-System (IXRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	- Blockrandomisierung: Blockgröße 5 - Nicht stratifizierte Randomisierung
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	- Zuteilung eines nummerierten Behandlungskits erfolgte mittels IVXS gemäß Randomisierungsliste des Sponsors. - Zugewiesene Arzneimittelnummer wurde im eCRF dokumentiert. - Zuteilung der Arzneimittel durch ausgewiesenes Personal - Das Studienpersonal erhielt keine Information über die Zuteilung.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die	Erstellung der Randomisierungsliste durch den Sponsor und Abruf der Zuteilung durch den Prüfarzt über IVXS

Item ^a	Charakteristikum	Studieninformation
	Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Aufnahme der Probanden in die Studie durch Prüfarzt und weiteres Studienpersonal
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Endpunkte beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) verblindet b) verblindet c) verblindet Verblindung wurde gewährleistet durch - Zuteilung der Behandlung durch IVXS - Identifikation der Studienmedikation über Identifikations-Codes - Identisches Aussehen der Verpackung und Beschriftung sowie der Kapseln von Verum und Placebo
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	- Identische Verpackung und Beschriftung der Studienmedikation - Verschiedene Farben der 150 mg- und 100 mg-Kapseln, jedoch für Nintedanib und Placebo identisch - Identische Einnahmezeitpunkte von Verum und Placebo
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Primäre Zielkriterien - Jährliche Abnahme FVC in ml über 52 Wochen mittels Random-Coefficient-Modell Sekundäre Zielkriterien - Veränderung des SGRQ-Gesamtscore bis Woche 52 mittels MMRM - Zeit der ersten <i>akuten Exazerbation</i> über 52 Wochen (proportionales Cox-Hazardmodell, Log-rank-Test, Kaplan-Meier). Für alle weiteren sekundäre Zielkriterien - Veränderung FVC bis Woche 52 mittels MMRM - Absolute und relative Veränderung FVC bis Woche 52 mittels MMRM - Alle weiteren kontinuierlichen Endpunkte (PRO) mittels MMRM - Alle Responderanalysen (95%-KI nach Wilson, logistische Regression) - Alle weiteren Überlebensanalysen (proportionales Cox-Hazardmodell, Log-rank-Test, Kaplan-Meier)
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	- Japanisch oder Nicht-Japanisch Zusätzlich für die gepoolte Auswertung der INPULSIS Studien: - Geschlecht (männlich/weiblich) - Alter (<65, ≥65 Jahre)

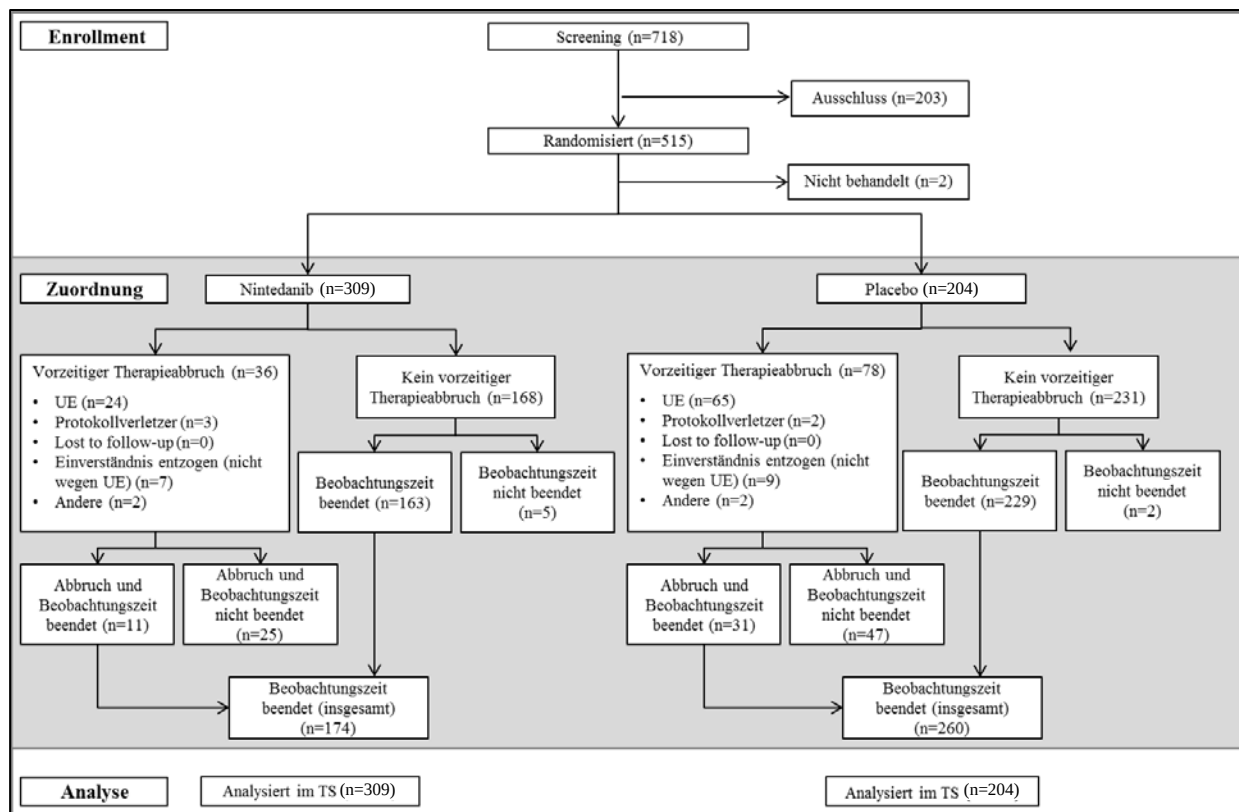
Item ^a	Charakteristikum	Studieninformation
		- Ethnie (kaukasisch, asiatisch) - Raucherstatus (Nichtraucher, Ex-Raucher/Raucher) - FVC % pred. ($\leq 70\%$, $> 70\%$) - SGRQ-Gesamtscore (≤ 40, > 40 Punkte) - Systemische Kortikosteroid-Therapie (ja, nein) - Verwendung Bronchodilatator (ja, nein) - Dosisintensität ($\leq 90\%$, $> 90\%$) (nur für den primären Endpunkt)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	s. Flow-Chart
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	a) n=515 Nintedanib 150 mg bid: n=309 Placebo: n=206 b) n=513 Nintedanib 150 mg bid: n=309 Placebo: n=204 c) n=513 Nintedanib 150 mg bid: n=309 Placebo: n=204
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Vorzeitiger Abbruch der Studienmedikation <u>Nintedanib: n=78 (25,2%)</u> - n=65 (21,0%) unerwünschte Ereignisse - n=9 (2,9%) unerwartete Verschlechterung der IPF - n=2 (0,6%) unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=54 (17,5%) andere unerwünschte Ereignisse - n=2 (0,6%) Protokollverletzer - n=0 Lost to follow-up - n=9 (2,9%) Verweigerung der Medikation - n=2 (0,6%) andere <u>Placebo: n=36 (17,6%)</u> - n=24 (11,8%) unerwünschte Ereignisse - n=9 (4,4%) unerwartete Verschlechterung der IPF - n=2 (1,0%) unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=13 (6,4%) andere unerwünschte Ereignisse - n=3 (1,5%) Protokollverletzer - n=0 Lost to follow-up - n=7 (3,4%) Verweigerung der Medikation - n=2 (1,0%) andere
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden /	Studienaufnahme: 05/2011 bis 10/2013 Follow-up: 28 Tage

Item ^a	Charakteristikum	Studieninformation
	Patienten und der Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	Beendigung der Studie gemäß Studienprotokoll

a: nach CONSORT 2010

ALAT: Latin American Thoracic Association; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; ATS: American Thoracic Society; BI: Boehringer Ingelheim; CASA-Q: Cough and Sputum Assessment Questionnaire; CTP: Clinical Trial Protocol (Studienprotokoll); DL_{CO}: Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); eCRF: Electronic Case Report Form; EKG: Elektrokardiogramm; EQ-5D: European Quality of Life-5 Dimensions; ERS: European Respiratory Society; FEV₁: Forced Expiratory Volume in one Second (Forcierte expiratorische Einsekunden-Kapazität); FVC: Forced Vital Capacity (Forcierte Vitalkapazität); GCP: Good Clinical Practice; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); INR: International Normalized Ratio; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); IXRS: Interaktives Telefon/Internet-Response-System; JRS: Japanese Respiratory Society; KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; NAC: N-Acetylcystein; PGI-C: Patient's Global Impression of Change; PRO: Patient Reported Outcome (Patientenberichteter Endpunkt); PTT: Partielle Thromboplastinzeit; SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SpO₂: Sauerstoffsättigung; ULN: Upper Limits of Normal (Oberer Normgrenzwert); VAS: Visual Analog Scale; ZNS: Zentrales Nervensystem

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a)

Abbildung 39: Patientenfluss INPULSIS 1 Studie

Tabelle 4-101 (Anhang): Studiendesign und -methodik für Studie INPULSIS 2

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2b	Genaue Ziele, Fragestellung und Hypothesen	Zielsetzung Untersuchung der Wirksamkeit sowie des Nutzen-Risiko-Verhältnisses von oral verabreichtem Nintedanib 150 mg bid im Vergleich zu Placebo bei IPF. Als primärer Endpunkt wird die Verringerung der Lungenfunktionsabnahme, gemessen als jährliche Abnahme der FVC, ermittelt. Hypothesen Bewertung der Überlegenheit von Nintedanib 150 mg bid im Vergleich zu Placebo in Bezug auf die jährliche Abnahme der FVC bei Patienten mit IPF. FVC-Abnahme in 52 Wochen: • Nullhypothese (H₀): Keine Unterschiede der jährlichen Abnahme der FVC vom Ausgangswert bis Woche 52 zwischen den Gruppen Nintedanib 150 mg bid und Placebo. • Alternativhypothese (H_a): Die jährliche Abnahme der FVC vom Ausgangswert bis Woche 52 ist bei Patienten unter Nintedanib 150 mg bid geringer als unter Placebo.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Allokation: randomisiert Zuteilungsverhältnis: 3:2 Randomisierungseinheit: Studienteilnehmer Verblindung: doppelblind Studienhorizont: 52 Wochen + 28 Tage Nachbeobachtung Design: parallel, zweiarstig, placebokontrolliert Studienorganisation: multizentrisch Phase: III
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Amendment 1 (11/2011) • Möglichkeit eines Re-Tests bei abnormalen Laborparametern zu Visite 1, bei Verdachtsfällen auf Messfehler oder fehlendem bzw. angezweifelterm Bezug auf den Patientenzustand • Studienausschluss, wenn Patienten den Studienablauf nicht verstehen bzw. die Fragebögen nicht selbstständig ausfüllen können. • Ergänzung der „schwerwiegenden unerwünschten Ereignisse“ entsprechend den neuen BI-Standards, um deren ordnungsgemäße Dokumentation zu gewährleisten

Item ^a	Charakteristikum	Studieninformation
		- Einschlusskriterium 4 wurde abgeändert: „Thorax-HRCT innerhalb von 52 Wochen vor Visite 1“ anstelle von „Thorax HRCT innerhalb von 52 Wochen vor Visite 2“ Amendment 2 (09/2012) - Ergänzung explorativer Biomarker-Analysen, um den Effekt von Nintedanib auf Biomarker in Verbindung mit IPF-Pathologie und prognostischen Erkrankungsmarkern zu untersuchen, und Ergänzung pharmakogenomischer Analysen - Schlechte Compliance wurde als Protokollverletzung definiert
4	Probanden / Patienten	.
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: - Schriftliche Einwilligungserklärung vor Studienbeginn - Alter ≥ 40 Jahre zu Visite 1 - Diagnose IPF nach ATS-, ERS-, JRS- und ALAT-Richtlinien (innerhalb von 5 Jahren vor Ausgangswert) - Thorax HRCT innerhalb von 52 Wochen vor Visite 1) - HRCT-Muster und, wenn verfügbar, Lungenbiopsie-Muster übereinstimmend mit IPF-Diagnose bewertet von zentralen Reviewern - DL_{CO}: 30-79% (zu Visite 2; angepasst an Hämoglobin) - FVC $\geq 50\%$ des Sollwertes (zu Visite 2) Ausschlusskriterien: - Laborparameter (zu Visite 1) - AST und ALT $> 1,5$-fach ULN - Bilirubin $> 1,5$-fach ULN - Lungenfunktion (zu Visite 2) - Atemwegsobstruktion ($FEV_1/FVC < 0,7$) - Lungentransplantation im Verlauf der Studie wahrscheinlich (nach Prüfarzt-Einschätzung) - Andere Krankheiten - Herzerkrankung - Myokardinfarkt innerhalb 6 Monaten vor Visite 2 - Instabile Angina innerhalb 1 Monat vor Visite 2 - Blutungsrisiko - Bekannte Prädisposition für Blutungen - Fibrinolyse mit im Studienprotokoll festgelegten Interventionen - Hämorrhagisches ZNS-Ereignis innerhalb von 52 Wochen vor Visite 2 - weitere Kriterien innerhalb von 3 Monaten vor Visite 2 - Hämoptysis oder Hämaturie - Aktive gastrointestinale Blutung oder Geschwüre - Größere Verletzung oder Operation

Item ^a	Charakteristikum	Studieninformation
		- Koagulations Parameter: INR >2, Verlängerung der Prothrombinzeit und PTT (sek) um >50% ULN (Visite 1) • Thromboserisiko - Bekannte erbliche Prädisposition für Thrombose - Thrombotisches Ereignis innerhalb von 52 Wochen (inkl. Schlaganfall und transiente ischämische Attacke) vor Visite 2 • Bekannte Hypersensitivität gegenüber Studienmedikation oder deren Bestandteilen • Andere Krankheiten, die das Testprozedere oder die Studienteilnahme beeinflussen oder den Patienten bei Studienteilnahme in Gefahr bringen (Ermessen des Prüfarztes) • Lebenserwartung aufgrund anderer Krankheiten außer IPF <2,5 Jahre (Ermessen des Prüfarztes) • Allgemeine Ausschlusskriterien • Frühere Behandlung mit Nintedanib länger als 4 Wochen) • Andere Interventionen (Forschungsstudie) innerhalb von 8 Wochen vor Visite 1 • Einnahme von N-Acetylcystein (NAC), Prednison >15 mg/Tag oder vergleichbare Arzneimittel innerhalb von 2 Wochen vor Visite 1 • Einnahme von Pirfenidon, Azathioprin, Cyclophosphamid, Cyclosporin A innerhalb von 8 Wochen vor Visite 1 • Geplante größere Operationen während der Studienperiode • Schwangere oder stillende Frauen oder Frauen im gebärfähigen Alter, die keine zwei adäquaten Verhütungsmethoden für mind. 1 Monat vor Behandlungsbeginn und bis 3 Monate nach Behandlungsende verwenden. • Sexuell aktive Männer, die sich nicht verpflichten, Kondome während und 3 Monate nach Ende der Behandlung zu nutzen (außer wenn Partner nicht im gebärfähigem Alter sind). • Aktiver Alkohol- oder Drogenmissbrauch Patienten, die das Studienprozedere nicht verstehen oder einhalten (inkl. Ausfüllen von Fragebögen ohne Hilfe).
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie, 107 Studienzentren in 17 Ländern: Kanada, Chile, China, Finnland, Frankreich, Deutschland, Griechenland, Indien, Japan, Korea, Mexiko, Niederlande, Portugal, Russland, Spanien, Türkei, USA
5	Interventionen	Alle Gruppen

Item ^a	Charakteristikum	Studieninformation
	Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Visite 1: Screening Visite 2: Therapiestart und nachfolgende Einnahme des Arzneimittels über 52 Wochen Nintedanib (oral) • 2-mal täglich 150 mg im Intervall von 12 Stunden über 52 Wochen • Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich 100 mg im Intervall von 12 Stunden über 52 Wochen bzw. in besonderen Fällen Unterbrechung der Therapie von maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Placebo (oral) • 2-mal täglich im Intervall von 12 Stunden über 52 Wochen • Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich Placebo zu 100 mg im Intervall von 12 Stunden über 52 Wochen bzw. in besonderen Fällen Unterbrechung der Therapie von maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Alle Gruppen Follow-up: 28 Tage nach Therapieende
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	Primäres Zielkriterium Untersuchung der Wirksamkeit sowie des Nutzen/Risiko-Verhältnisses von Nintedanib 150 mg bid im Vergleich zu Placebo bei IPF Primäre Endpunkte • Jährliche FVC- Abnahme über 52 Wochen (ml) Key sekundäre Endpunkte • Veränderung des SGRQ-Gesamtscores vom Ausgangswert bis Woche 52 (in Punkten) • Zeit bis zur ersten <i>akuten Exazerbation</i> über 52 Wochen (in Tagen) Weitere sekundäre Endpunkte • Weitere Analysen der FVC • Absolute und relative Veränderung der FVC (ml) und FVC (% des Sollwertes) über 52 Wochen • Absolute kategorielle Veränderung der FVC (% des Sollwertes) bis zu 52 Wochen: Abfall >5%, Anstieg >5% und Veränderungen ≤5% • Absolute kategorielle Veränderung der FVC (% des Sollwertes) vom Ausgangswert bis Woche 52: Abfall >10%, Anstieg >10% und Veränderung ≤10%

Item ^a	Charakteristikum	Studieninformation
		• Anteil FVC-Responder mit Abfall <5% vom FVC-Sollwert nach 52 Wochen • Anteil FVC-Responder mit Abfall <10% vom FVC-Sollwert nach 52 Wochen • Patientenberichteter Endpunkt (PRO) gemessen vom Ausgangswert bis Woche 52: • Anteil SGRQ-(≤4 Punkte)-Responder (definiert als absolute Veränderung im SGRQ-Gesamtscore um ≤4 Punkte) • Veränderung SGRQ-Domänen in Punkten • Veränderung SGRQ-I-Gesamtscore • Veränderung SOBQ • Veränderung Hustenimpact und Hustensymptome (CASA-Q[CD])-Score • Anteil PGI-C-Responder • Veränderung EQ-5D-VAS • <i>Akute Exazerbation:</i> • Risiko einer akuten Exazerbation • Überlebenszeitanalysen gemessen über 52 Wochen: • Zeit bis zum Tod • Zeit bis zum respiratorisch bedingten Tod • Zeit bis zum Tod unter Behandlung • Zeit bis zum Tod oder Lungentransplantation • Zeit bis zum Tod oder Lungentransplantation oder Qualifizierung für Lungentransplantation (Kriterien: FVC <45% des Sollwertes oder DL_{CO} <30% des Sollwertes oder SpO₂<88% in Ruhe • Veränderung SpO₂ in Ruhe • Veränderung DL_{CO} in Ruhe Sicherheit • Unerwünschte Ereignisse • Vitalzeichen • Körperliche Untersuchung • Gewicht • Klinische Labortests • 12-Kanal-EKG (japanische Studienzentren) Andere Endpunkte • Prä-Dosis Plasmakonzentration (C_{pre,ss,N}) von Nintedanib, BIBF 1202 und BIBF 1202 Glukuronid Visite 4 (4 Wochen) und Visite 7 (24 Wochen) • Pharmakogenomische Auswertung • Biomarker-Auswertung
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	keine
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	• Gepowert auf primären Endpunkt: Jährliche FVC-Abnahme über 52 Wochen (ml)

Item ^a	Charakteristikum	Studieninformation
		• Erwarteter Effekt: Jährliche Differenz von 100 ml zwischen Nintedanib- und Placebo-Gruppe • Erwartete Abweichung: 300 ml (auf Basis der Phase II-Studienergebnisse) • Erwarteter Anteil nicht auswertbarer Studienteilnehmer: 2% • Zuteilungsverhältnis von 3:2 führt zu einer Fallzahl von: 291 zu 194 (n=485) in Nintedanib zu Placebo • Verwendeter Test: t-Test • Teststärke: 90% • (α): 0,025, einseitig • H_0-Hypothese: Es gibt keinen Unterschied zwischen den Behandlungsgruppen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Zwischenanalyse • keine Kriterien für Patientenausschluss • direkt nach Visite 2 (vor Arzneimittel-Ersteinnahme) • AST, ALT >1,5-fach ULN • Bilirubin >1,5-fach ULN • INR >2 • Prothrombinzeit oder PTT >50% ULN • Sonst keine weiteren Regeln für Patientenausschluss vorgesehen Kriterien für Therapieausschluss • Leberfunktion • Leberenzyme - ALT oder AST $\geq$ 8-fach ULN - ALT oder AST $\geq$ 3-fach ULN für mehr als 2 Wochen • Anzeichen akuter Leberschaden • ALT oder AST $\geq$ 3-fach ULN und Gesamt-Bilirubin $\geq$ 1,5-fach ULN oder INR >1,5 oder Anzeichen von Fatigue, Übelkeit, Erbrechen, Schmerz oder Druckempfindlichkeit im rechten oberen abdominalen Quadranten, Fieber, Ausschlag, und/oder Eosinophile >5% • Andere Kriterien • Wiederauftreten einer Diarrhö (>4 Stuhlgänge/Tag) für 8 oder mehr aufeinander folgende Tage trotz Dosisreduktion und Begleitbehandlung • Gleichzeitige Einnahme von gemäß der Ausschlusskriterien verbotenen Arzneimitteln • Schwangerschaft • Widerruf der Einwilligung

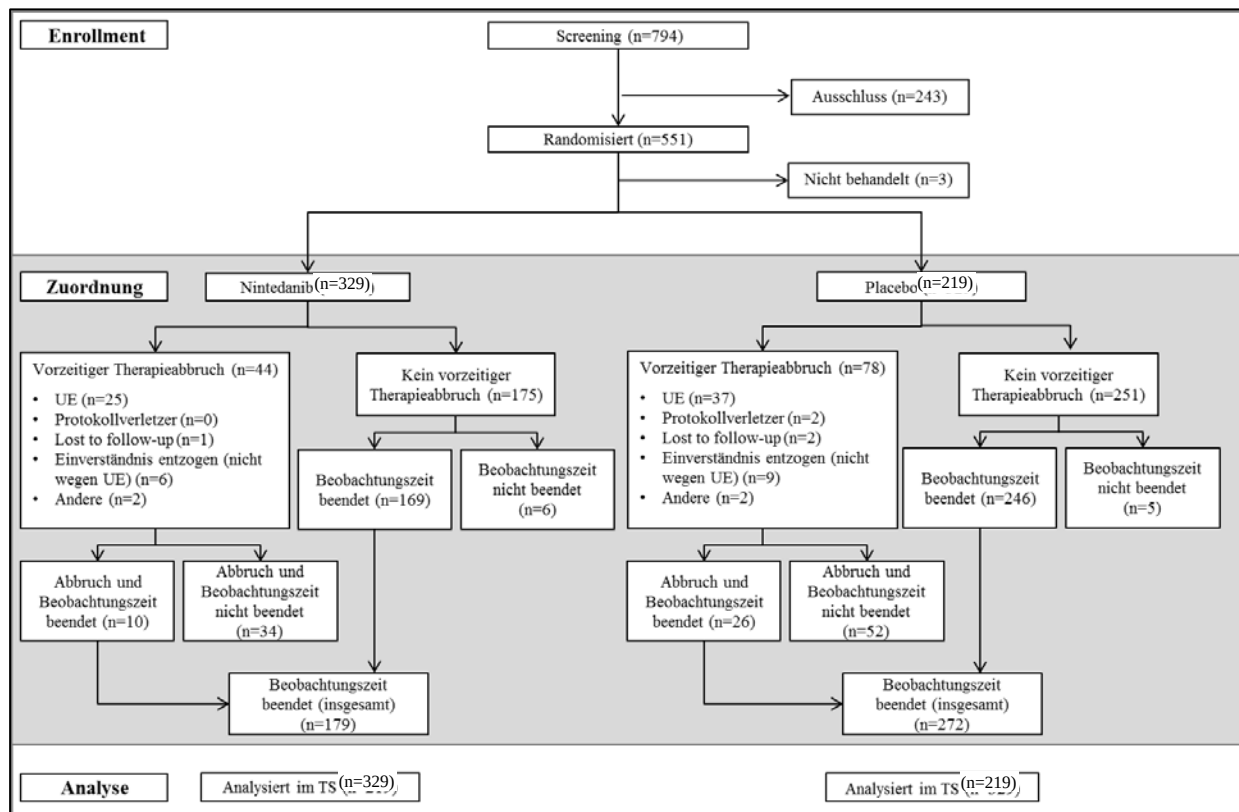
Item ^a	Charakteristikum	Studieninformation
		Kriterien für dauerhaften Abbruch der Studienmedikation • Größere Operation (abdominal und intestinal) • Patienten mit vollständiger Antikoagulation • Größere thromboembolische Ereignisse (Schlaganfall, tiefe Venenthrombose, Lungenembolie, Myokardinfarkt) • Erhöhtes Blutungsrisiko (hämorrhagisches ZNS-Ereignis, Hämoptysis oder Hämaturie, aktive gastrointestinale Blutung oder Geschwüre) Kriterien für Studienausschluss • Widerruf der Einwilligung Kriterien für Studienabbruch durch Sponsor • Nichteinhaltung der Aufnahmekriterien • Auftreten von Wirkung/Sicherheit Informationen, die möglicherweise die Fortsetzung des Tests beeinflusst haben • Verstoß GCP, CTP oder Vertragsbruch seitens Prüfzentrum oder Prüfer
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung über Interaktives Telefon/Internet-Response-System (IXRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	• Blockrandomisierung: Blockgröße 5 • Nicht stratifizierte Randomisierung

Item ^a	Charakteristikum	Studieninformation
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	• Zuteilung eines nummerierten Behandlungskits erfolgte mittels IVXS gemäß Randomisierungsliste des Sponsors. • Zugewiesene Arzneimittelnummer wurde im eCRF dokumentiert. • Zuteilung der Arzneimittel durch ausgewiesenes Personal • Das Studienpersonal erhielt keine Information über die Zuteilung.
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?	• Erstellung der Randomisierungsliste durch den Sponsor und Abruf der Zuteilung durch den Prüfarzt über IVXS • Aufnahme der Probanden in die Studie durch Prüfarzt und weiteres Studienpersonal
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Endpunkte beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) verblindet b) verblindet c) verblindet Verblindung wurde gewährleistet durch • Zuteilung der Behandlung durch IVXS • Identifikation der Studienmedikation über Identifikations-Codes • Identisches Aussehen der Verpackung und Beschriftung sowie der Kapseln von Verum und Placebo
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	• Identische Verpackung und Beschriftung der Studienmedikation • Verschiedene Farben der 150 mg- und 100 mg-Kapseln, jedoch für Nintedanib und Placebo identisch • Identische Einnahmezeitpunkte von Verum und Placebo
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Primäre Zielkriterien • Jährliche Abnahme FVC in ml über 52 Wochen mittels Random-Coefficient-Modell Sekundäre Zielkriterien • Veränderung des SGRQ-Gesamtscores bis Woche 52 mittels MMRM • Zeit der ersten <i>akuten Exazerbation</i> über 52 Wochen (proportionales Cox-Hazardmodell, Log-rank-Test, Kaplan-Meier). Für alle weiteren sekundäre Zielkriterien • Veränderung FVC bis Woche 52 mittels MMRM

Item ^a	Charakteristikum	Studieninformation
		• Absolute und relative Veränderung FVC bis Woche 52 mittels MMRM • Alle weiteren kontinuierlichen Endpunkte (PRO) mittels MMRM • Alle Responderanalysen (95%-KI nach Wilson, logistische Regression). • Alle weiteren Überlebensanalysen (proportionales Cox-Hazardmodell, Log-rank-Test, Kaplan-Meier)
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	• Japanisch oder Nicht-Japanisch Zusätzlich für die gepoolte Auswertung der INPULSIS Studien: • Geschlecht (männlich/weiblich) • Alter (<65, ≥65 Jahre) • Ethnie (kaukasisch, asiatisch) • Raucherstatus (Nichtraucher, Ex-Raucher/Raucher) • FVC% des Sollwertes (≤70%, >70%) • SGRQ-Gesamtscore (≤40, >40 Punkte) • Systemische Kortikosteroid-Therapie (ja, nein) • Verwendung Bronchodilatator (ja, nein) • Dosisintensität (≤90%, >90%) (nur für den primären Endpunkt)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	s. Flow-Chart
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	a) n=551 Nintedanib 150 mg bid: n=331 Placebo: n=220 b) n=548 Nintedanib 150 mg bid: n=329 Placebo: n=219 c) n=548 Nintedanib 150 mg bid: n=329 Placebo: n=219
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Vorzeitiger Abbruch der Studienmedikation <u>Nintedanib: n=78 (23,7%)</u> • n=62 (18,8%) unerwünschte Ereignisse - n=10 (3,0%) unerwartete Verschlechterung der Krankheit - n=3 (0,9%) unerwartete Verschlechterung einer bereits vorhandenen Krankheit - n=49 (14,9%) andere unerwünschte Ereignisse • n=2 (0,6%) Protokollverletzer • n=0 Lost to follow-up • n=11 (3,3%) Verweigerung der Medikation

Item ^a	Charakteristikum	Studieninformation
		- n=3 (0,9%) andere <u>Placebo: n=44 (20,1%)</u> - n=35 (16,0) unerwünschte Ereignisse - n=18 (8,2%) unerwartete Verschlechterung der Krankheit - n=2 (0,9%) unerwartete Verschlechterung einer bereits vorhandenen Krankheit - n=15 (6,8%) andere unerwünschte Ereignisse - n=1 (0,5%) Protokollverletzer - n=1 (0,5%) Lost to follow-up - n=6 (2,7%) Verweigerung der Medikation - n=1 (0,5%) andere
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	- Studienaufnahme: 05/2011 bis 10/2013 - Follow-up: 28 Tage
14b	Informationen, warum die Studie endete oder beendet wurde	Beendigung der Studie gemäß Studienprotokoll
a: nach CONSORT 2010 ALAT: Latin American Thoracic Association; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; ATS: American Thoracic Society; BI: Boehringer Ingelheim; CASA-Q: Cough and Sputum Assessment Questionnaire; CTP: Clinical Trial Protocol (Studienprotokoll); DL_{CO}: Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); eCRF: Electronic Case Report Form; EKG: Elektrokardiogramm; EQ-5D: European Quality of Life-5 Dimensions; ERS: European Respiratory Society; FEV₁: Forced Expiratory Volume in one Second (Forcierte expiratorische Einsekunden-Kapazität); FVC: Forced Vital Capacity (Forcierte Vitalkapazität); GCP: Good Clinical Practice; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); INR: International Normalized Ratio; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); IXRS: Interaktives Telefon/Internet-Response-System; JRS: Japanese Respiratory Society; KI: Konfidenzintervall; MMRM: Mixed Effect Model for Repeated Measures; NAC: N-Acetylcystein; PGI-C: Patient's Global Impression of Change; PRO: Patient Reported Outcome (Patientenberichteter Endpunkt); PTT: Partielle Thromboplastinzeit; SGRQ: St. George's Respiratory Questionnaire; SGRQ-I: St. George's Respiratory Questionnaire IPF-specific version; SOBQ: Shortness of Breath Questionnaire; SpO₂: Sauerstoffsättigung; ULN: Upper Limits of Normal (Oberer Normgrenzwert); VAS: Visual Analog Scale; ZNS: Zentrales Nervensystem		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b)

Abbildung 40: Patientenfluss INPULSIS 2 Studie

Tabelle 4-102 (Anhang): Studiendesign und -methodik für Studie 1199.187

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2b	Genaue Ziele, Fragestellung und Hypothesen	Zielsetzung Nach der Implementierung des globalen Protokoll Amendments Nr.1 war das primäre Ziel der Studie die Untersuchung der Wirksamkeit von oral verabreichtem Nintedanib 150 mg bid nach 6 Monaten im Vergleich zu Placebo bei IPF-Patienten. Primärer Endpunkt der Studie war die relative Änderung des HRCT-QLF Scores zu Monat 6. Hypothesen Nach der Implementierung des globalen Protokoll Amendments Nr. 1 änderte sich das Design der Studie in ein exploratives. Es wurden keine formalen Null- oder Alternativhypothesen aufgestellt oder getestet.
Methoden		
3	Studiendesign	

Item ^a	Charakteristikum	Studieninformation
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Allokation: randomisiert Zuteilungsverhältnis: 1:1 Randomisierungseinheit: Studienteilnehmer Verblindung: doppelblind Studienhorizont: 6 Monate verblindet + 18 Monate Open-label Nintedanib Design: parallel, zweiarstig, placebokontrolliert Studienorganisation: multizentrisch Phase: IIIb
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Kanada – lokales Amendment 1 (09/2013) - Entfernen der arteriellen Blutgasanalyse im kompletten Protokoll sowie den assoziierten exploratorischen Endpunkten Globales Amendment 1 (10/2014) - Änderung des Studiendesigns zu einer 6-monatigen randomisierten, doppelblinden, placebokontrollierten Periode, gefolgt von einer Konvertierung beider Arme zu Nintedanib 150 mg bid (unverblindet) für bis zu 18 Monate - Überarbeiteter primärer Endpunkt: Änderung des HRCT-QLF Scores (bezogen auf den Ausgangswert) nach 6 Monaten statt nach 52 Wochen - Wegfall sekundärer Endpunkte - Hinzunahme zusätzlicher sekundärer Endpunkte: relative Änderung des Ausgangswert HRCT-QLF Scores nach 52 Wochen; absolute, relative und kategoriale Änderung der Ausgangswert FVC nach 24 Wochen; 6 MWT-Distanzänderungen vom Ausgangswert bis Woche 24. Andere sekundäre Endpunkte wurden so überarbeitet, dass die Bewertung nach 24 anstelle von 52 Wochen stattfand. - Umbenennen von Exploratorischen Endpunkten in Weitere Endpunkte und Modifikation aus Konsistenzgründen mit dem überarbeiteten Studiendesign - Studienergebnisse umfassten Punktschätzer und 95%-Konfidenzintervalle der Behandlungsunterschiede. Kein Testen formaler Null- oder Alternativhypothesen. - Zutreffen von Einschlusskriterium #5 (FVC$\geq$50%) bei Visite 1 und 2 - Beibehalten des Ausschlusskriteriums #9 als Patienten Sicherheit - Entfernen des Ausschlusskriteriums #13 bezüglich sexuell aktiver männlicher Patienten, die sich nicht

Item ^a	Charakteristikum	Studieninformation
		dazu verpflichteten, Kondome während und bis zu 3 Monaten nach Ende der Behandlung zu nutzen. • Entfernen von arteriellen Blutgas Messungen • Hinzufügen zusätzlicher Studienvisiten nach 6 Monaten nach Konvertierung beider Arme zu Nintedanib 150 mg bid, um Patientensicherheit zu kontrollieren. • Reduzierung geplanter Patientenzahlen und Studienzentren. Erhöhung der Anzahl potenziell teilnehmender Länder. Kanada – lokales Amendment 2 (02/2015) • Modifizierte Behandlungsdauer über 18 Monate hinaus, um IPF-Patienten, die mindestens eine Behandlungsdauer von 18 Monaten abgeschlossen haben, eine Fortsetzung der Nintedanib-Behandlung anzubieten. Das CTP wurde überarbeitet, um zu definieren, dass die Behandlung beendet wird, sobald ein Kriterium für die Absetzung getroffen oder sobald Nintedanib zugelassen wurde und in jeder Provinz, in der es Studienpatienten gab, Nintedanib erstattet wird. • Zusätzliche Studienvisiten wurden zum Flow-Chart zugefügt, um die verlängerte Behandlungsdauer einzubauen
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: • Schriftliche Einwilligungserklärung vor Studienbeginn • Alter ≥ 40 Jahre zur Visite 1 • Diagnose IPF nach ATS/ERS/JRS/ALAT-Richtlinien (2011) innerhalb von 5 Jahren vor Ausgangswert und erneute Bestätigung nach den Richtlinien (2011), falls die Diagnose >2 und bis zu 5 Jahren vor Visite 1 liegt. Die Diagnose wurde durch ein Thorax HRCT innerhalb von 24 Monaten nach Visite 1 bestätigt. • DL_{CO}: 30-79% (angepasst an Hämoglobin) • FVC% pred. $\geq 50\%$ (zu Visite 1 und 2) Ausschlusskriterien: • AST und ALT $>1,5$-fach ULN • Bilirubin $>1,5$-fach ULN • Blutungsrisiko: - Fibrinolyse, Antikoagulation, Thrombozytenaggregationshemmer - Hämorrhagisches ZNS-Ereignis innerhalb von 52 Wochen - weitere Kriterien innerhalb von 3 Monaten:

Item ^a	Charakteristikum	Studieninformation
		- o Hämoptysis oder Hämaturie - o Aktive gastrointestinale Blutung oder Geschwüre - o Größere Verletzung oder Operation - Koagulations Parameter: INR >2, Verlängerung der Prothrombinzeit und PTT (sek) um >150% ULN • Geplante größere Operationen innerhalb der nächsten 3 Monate, was eine Lungentransplantation, eine größere abdominale oder größere intestinale Operation einschließt. • Thromboserisiko - Bekannte erbliche Prädisposition für Thrombose - Thrombotisches Ereignis innerhalb von 52 Wochen (inkl. Schlaganfall und transiente ischämische Attacke). • Aktuelle oder geplante andere Interventionen (Forschungsstudie) während der Studie • Frühere Behandlung mit Nintedanib innerhalb einer klinischen Studie innerhalb der letzten 3 Monate und Abbruch von Nintedanib aufgrund eines unerwünschten Ereignisses • Bekannte Hypersensitivität gegenüber Studienmedikation oder deren Bestandteile • Andere Erkrankungen oder Zustände, welche nach Ermessen des Prüfarztes den Patienten durch Teilnahme an der Studie in Gefahr versetzte oder die Fähigkeit des Patienten an der Studie teilzunehmen limitierte. Patienten wurden ausgeschlossen, wenn sie mehr als 12 l/min Sauerstoff benötigten, nicht ambulant waren oder eine Gehhilfe oder einen Stock während des 6 MWT benötigten. Patienten, die den 6 MWT nicht abschließen konnten, wurden aus der Studie ausgeschlossen. • Aktiver Alkohol- oder Drogenmissbrauch. • Schwangere oder stillende Frauen oder Frauen im gebärfähigen Alter, die keine zwei adäquaten Verhütungsmethoden für mind. 1 Monat vor Behandlungsbeginn und bis 3 Monate nach Behandlungsende verwendeten.
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie, 26 Studienzentren in 3 Ländern: USA, Kanada, Türkei
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Alle Gruppen Visite 1: Screening Visite 2: Therapiestart und nachfolgende Einnahme des Arzneimittels über 6 Monate Nintedanib (oral) • 2-mal täglich 150 mg im Intervall von 12 Stunden über 6 Monate

Item ^a	Charakteristikum	Studieninformation
		- Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich 100 mg im Intervall von 12 Stunden über 6 Monate bzw. in besonderen Fällen Unterbrechung der Therapie für maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Placebo (oral) 2-mal täglich im Intervall von 12 Stunden über 24 Wochen - Bei unerwünschten Ereignissen Reduktion der Dosis auf 2-mal täglich Placebo zu 100 mg im Intervall von 12 Stunden über 24 Wochen bzw. in besonderen Fällen Unterbrechung der Therapie von maximal 4 Wochen bei therapieassoziierten unerwünschten Ereignissen und bis zu 8 Wochen bei Exazerbationen Alle Gruppen Visite 7 (24 Wochen): orale Einnahme von Nintedanib 2-mal täglich im Intervall von 12 Stunden über weitere 24-52 Wochen Follow-up: 28 Tage nach Therapieende
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	Primäres Zielkriterium Untersuchung der Wirksamkeit von oral verabreichtem Nintedanib 150 mg bid nach 6 Monaten im Vergleich zu Placebo bei IPF-Patienten Primäre Endpunkte - Relative Änderung des HRCT-QLF Scores vom Ausgangswert bis Woche 24 Sekundäre Endpunkte - Effekt des 24 Wochen späteren Behandlungsbeginns: relative Änderung des HRCT-QLF Scores vom Ausgangswert bis Woche 52 - Absolute, relative und kategorielle Änderung der FVC vom Ausgangswert bis Woche 24 - Änderung des SGRQ-Gesamtscores vom Ausgangswert bis Woche 24 - Änderung der gelaufenen Distanz beim 6 MWT vom Ausgangswert bis Woche 24 - Änderung des UCSD-SOBQ vom Ausgangswert bis Woche 24 - Folgende Time-to-event-Analysen: - Gesamtmortalität bis Woche 24 - Respiratorisch bedingte Hospitalisierungen bis Woche 24 - Respiratorisch bedingte Mortalität bis Woche 24

Item ^a	Charakteristikum	Studieninformation
		- Akute Exazerbationen bis Woche 24 Weitere sekundäre Endpunkte • Absolute, relative und kategoriale Änderung der FVC vom Ausgangswert bis Woche 52 und bis zum Ende der Behandlung • Änderung des SGRQ-Gesamtscores vom Ausgangswert bis Woche 52 und bis zum Ende der Behandlung • Änderung der gelaufenen Distanz beim 6 MWT vom Ausgangswert bis Woche 52 und bis zum Ende der Behandlung • 6 MWT-Ergebnisse vom Ausgangswert bis Woche 24, bis Woche 52 und bis zum Ende der Behandlung - Änderung in Nadir-Entsättigung vom Ausgangswert - Änderung in Delta-Entsättigung (Ruhewert SaO₂ – Nadir SaO₂) vom Ausgangswert - Produkt aus Sättigung und Strecke - Entsättigung bis 88% oder weniger - 4% oder größere Entsättigung - Borg-Score vor und nach zurückgelegter Gehstrecke - Geschwindigkeit (d/t) während der ersten 3 Minuten des 6 MWT - Geschwindigkeit (d/t) während der zweiten 3 Minuten des 6 MWT - Geschwindigkeit (d/t) x Nadir-Sättigung • Relative Änderung des HRCT-QLF Scores vom Ausgangswert bis zum Ende der Behandlung • Änderung des UCSD-SOBQ vom Ausgangswert bis Woche 24, Woche 52 und bis zum Ende der Behandlung Sicherheit • Vitalzeichen • Körperliche Untersuchung • Gewicht • Klinische Labortests • Unerwünschte Ereignisse • 12-Kanal-EKG
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Durch das globale Amendment 1 wurde das Design der Studie in ein exploratives Design geändert und die kontrollierte Beobachtungsdauer wurde verkürzt.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	• Ursprünglich geplante 52-wöchige Studie: Zuteilungsverhältnis von 1:1 mit einer Fallzahl von 83 Patienten pro Behandlungsarm, um mit einer Power von 90% einen Behandlungseffekt von 50% (Verhältnis des Ratios der geometrischen

Item ^a	Charakteristikum	Studieninformation
		Mittel der QLF-Scores (auf einer logarithmierten Skala) zu entdecken, falls er tatsächlich vorliegt. - Nach Überführung in ein exploratives Design wurde die Zahl der Studienteilnehmer reduziert. Durchgeführte 24-wöchige Studie: aufgrund der Änderung des Studiendesigns war es nicht mehr möglich die Studie zwischen den Behandlungsgruppen angemessen zu powern. - Durch fehlende Power für eine geschätzte Effektgröße von 1,25 nach 24 Wochen, liefert diese Studie nur Punktschätzer und 95%-Konfidenzintervalle für die Behandlungsunterschiede. Statistische Hypothesentests werden nicht durchgeführt.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Zwischenanalyse - keine Kriterien für Patientenausschluss - vor Visite 2 (vor Arzneimittel-Ersteinnahme) - AST, ALT >1,5-fach ULN - Bilirubin >1,5-fach ULN - Sonst keine weiteren Regeln für Patientenausschluss vorgesehen Kriterien für patientenindividuellen Therapieausschluss - Leberfunktion - Leberenzyme - ALT oder AST ≥8-fach ULN - ALT oder AST ≥3-fach ULN für mehr als 2 Wochen - Anzeichen akuter Leberschaden: ALT oder AST ≥3-fach ULN und - Anzeichen von Fatigue, Übelkeit, Erbrechen, Schmerz oder Druckempfindlichkeit im rechten oberen abdominalen Quadranten, Fieber, Ausschlag und/oder Eosinophile >5% - Gesamt-Bilirubin >1,5-fach ULN - INR >1,5 Andere Kriterien - Schwangerschaft - Widerruf der Einwilligung - Gleichzeitige Einnahme von gemäß der Ausschlusskriterien verbotenen Arzneimitteln - Wiederauftreten einer Diarrhö - Unterbrechung der Behandlung (spätestens 8 Wochen nach Unterbrechung aufgrund von unerwünschten Ereignissen musste die Medikation fortgesetzt werden)

Item ^a	Charakteristikum	Studieninformation
		Kriterien für dauerhaften Abbruch der Studienmedikation • Größere Operation (abdominal und intestinal) • Patienten mit vollständiger Antikoagulation • Größere thromboembolische Ereignisse (Schlaganfall, tiefe Venenthrombose, Lungenembolie, Myokardinfarkt) • Erhöhtes Blutungsrisiko (hämorrhagisches ZNS-Ereignis, Hämoptysis oder Hämaturie, aktive gastrointestinale Blutung oder Geschwüre) Kriterien für Studienausschluss • Widerruf der Einwilligung Kriterien für Studienabbruch durch Sponsor • Nichteinhaltung der Aufnahmekriterien • Auftreten von Wirkungs/Sicherheits-Informationen, die möglicherweise die Fortsetzung des Tests beeinflusst haben • Verstoß GCP, CTP oder Vertragsbruch seitens Prüfzentrum oder Prüfer
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Randomisierung und Zuteilung auf die Behandlungsgruppen über Interactive Response Technology (IRT) im Verhältnis 1:1
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Eine Blockgröße war im Studienbericht nicht angegeben; die Randomisierung erfolgte nicht stratifiziert.

Item ^a	Charakteristikum	Studieninformation
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	- Randomisierung sowie Zuteilung der Studienmedikation unter Verwendung eines unabhängigen IRT (Interactive Response Technology System) - Zugang zur Randomisierungsliste während der Studie wurde kontrolliert und dokumentiert. - Dokumentierte Verteilung der Arzneimittel durch autorisiertes Personal - Studienmedikation wurde über eine Codenummer identifiziert. - Randomisierungscode wurde an einem sicheren Ort bis nach Datenbanksperre verwahrt. - Das Studienpersonal (u. a. Prüfarzt, Personal bei Sponsor und Dritter) erhielt keine Information über die Zuteilung und war bis nach Datenbanksperre verblindet.
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?	IRT
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Endpunkte beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) verblindet b) verblindet c) verblindet Verblindung wurde gewährleistet durch - Zuteilung der Behandlung durch IRT - Identifikation der Studienmedikation über eine Codenummer - Identisches Aussehen der Verpackung und Beschriftung sowie der Kapseln von Verum und Placebo
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	- Identische Verpackung und Beschriftung der Studienmedikation - Verschiedene Farben der 150 mg- und 100 mg-Kapseln, jedoch für Nintedanib und Placebo identisch - Identische Einnahmezeitpunkte von Verum und Placebo
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Primäre Zielkriterien Relative Änderung des HRCT-QLF Scores vom Ausgangswert bis Woche 24, was anhand nicht-transformierter Daten analysiert wurde. Der Logarithmus des QLF-Scores (log[QLF Score +1]) wurde nur für Sensitivitätsanalysen verwendet.

Item ^a	Charakteristikum	Studieninformation
		Ergebnisse wurden anhand von ANCOVA analysiert (adjudiziert nach Ausgangswert HRCT-QLF). Sekundäre Zielkriterien - Relative Änderung des HRCT-QLF Scores vom Ausgangswert bis Woche 52 (Effekt der 6 Monate später beginnenden Behandlung) mittels Sensitivitätsanalyse - Absolute und relative Änderung der FVC vom Ausgangswert bis Woche 24 mittels MMRM - Änderung des SGRQ-Gesamtscores vom Ausgangswert bis Woche 24 mittels MMRM - Änderung der gelaufenen Distanz des 6 MWT vom Ausgangswert bis Woche 24 mittels MMRM - Änderung des UCSD-SOBQ vom Ausgangswert bis Woche 24 mittels MMRM - Gesamtmortalität, respiratorisch bedingte Hospitalisierung, respiratorisch bedingte Mortalität und akute IPF Exazerbationen bis Woche 24 mittels einer Analyse der Überlebensrate
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Keine Subgruppenanalysen
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	s. Flow-Chart
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	a) n=113 Nintedanib 150 mg bid: n=56 Placebo: n=57 b) n=113 Nintedanib 150 mg bid: n=56 Placebo: n=57 c) n=87 Nintedanib 150 mg bid: n=42 Placebo: n=45
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Vorzeitiger Abbruch der Studienmedikation <u>Nintedanib: n=13 (23,2%)</u> - n=7 (12,5%) unerwünschte Ereignisse - n=1 (1,8%) unerwartete Verschlechterung der IPF - n=1 (1,8%) unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=5 (8,9%) andere unerwünschte Ereignisse - n=0 Protokollverletzer - n=0 Lost to follow-up - n=2 (3,6%) Verweigerung der Medikation

Item ^a	Charakteristikum	Studieninformation
		- n=4 (7,1%) andere <u>Placebo: n=14 (24,6%)</u> - n=6 (10,5%) unerwünschte Ereignisse - n=3 (5,3%) unerwartete Verschlechterung der IPF - n=0 unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=3 (5,3%) andere unerwünschte Ereignisse - n=1 (1,8%) Protokollverletzer - n=0 Lost to follow-up - n=0 Verweigerung der Medikation - n=7 (12,3%) andere Vorzeitiger Abbruch der Studie Nintedanib: n=4 (7,1%) - n=2 (3,6%) unerwünschte Ereignisse - n=1 (1,8%) unerwartete Verschlechterung der IPF - n=0 unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=1 (1,8%) andere unerwünschte Ereignisse - n=0 Protokollverletzer - n=0 Lost to follow-up - n=1 (1,8%) Widerruf der Einverständniserklärung - n=1 (1,8%) andere Placebo: n=6 (10,5%) - n=5 (8,8%) unerwünschte Ereignisse - n=3 (5,3%) unerwartete Verschlechterung der IPF - n=0 unerwartete Verschlechterung einer anderen bereits vorhandenen Krankheit - n=2 (3,5%) andere unerwünschte Ereignisse - n=0 Protokollverletzer - n=0 Lost-to-follow-up - n=1 (1,8%) Widerruf der Einverständniserklärung - n=0 andere
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienaufnahme: 12/2013 bis 10/2016 Follow-up: 28 Tage
14b	Informationen, warum die Studie endete oder beendet wurde	Beendigung der Studie gemäß Globales Amendment 1 (10/2014)

Item ^a	Charakteristikum	Studieninformation
a: nach CONSORT 2010 6 MWT: 6-Minute-Walk-Test (6-Minuten-Gehtest); ALAT: Latin American Thoracic Association; ALT: Alanin-Aminotransferase; ANCOVA: Kovarianzanalyse; AST: Aspartat-Aminotransferase; ATS: American Thoracic Society; CTP: Clinical Trial Protocol (Studienprotokoll); DL_{CO}: Diffusing Capacity of the Lung for Carbon Monoxide (Kohlenmonoxid-Diffusionskapazität); EKG: Elektrokardiogramm; ERS: European Respiratory Society; FVC: Forced Vital Capacity (Forcierte Vitalkapazität); GCP: Good Clinical Practice; HRCT: High Resolution Computed Tomography (Hochauflösende Computer-Tomografie); INR: International Normalized Ratio; IPF: Idiopathic Pulmonary Fibrosis (Idiopathische Lungenfibrose); IRT: Interactive Response Technology System; JRS: Japanese Respiratory Society; MMRM: Mixed Effect Model for Repeated Measures; PTT: Partielle Thromboplastinzeit; QLF: Quantitative Lung Fibrosis Score (Quantitativer Lungenfibrosescore); SGRQ: St. George's Respiratory Questionnaire; UCSD-SOBQ: University of California San Diego Shortness of Breath Questionnaire; ULN: Upper Limits of Normal (Oberer Normgrenzwert); ZNS: Zentrales Nervensystem		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

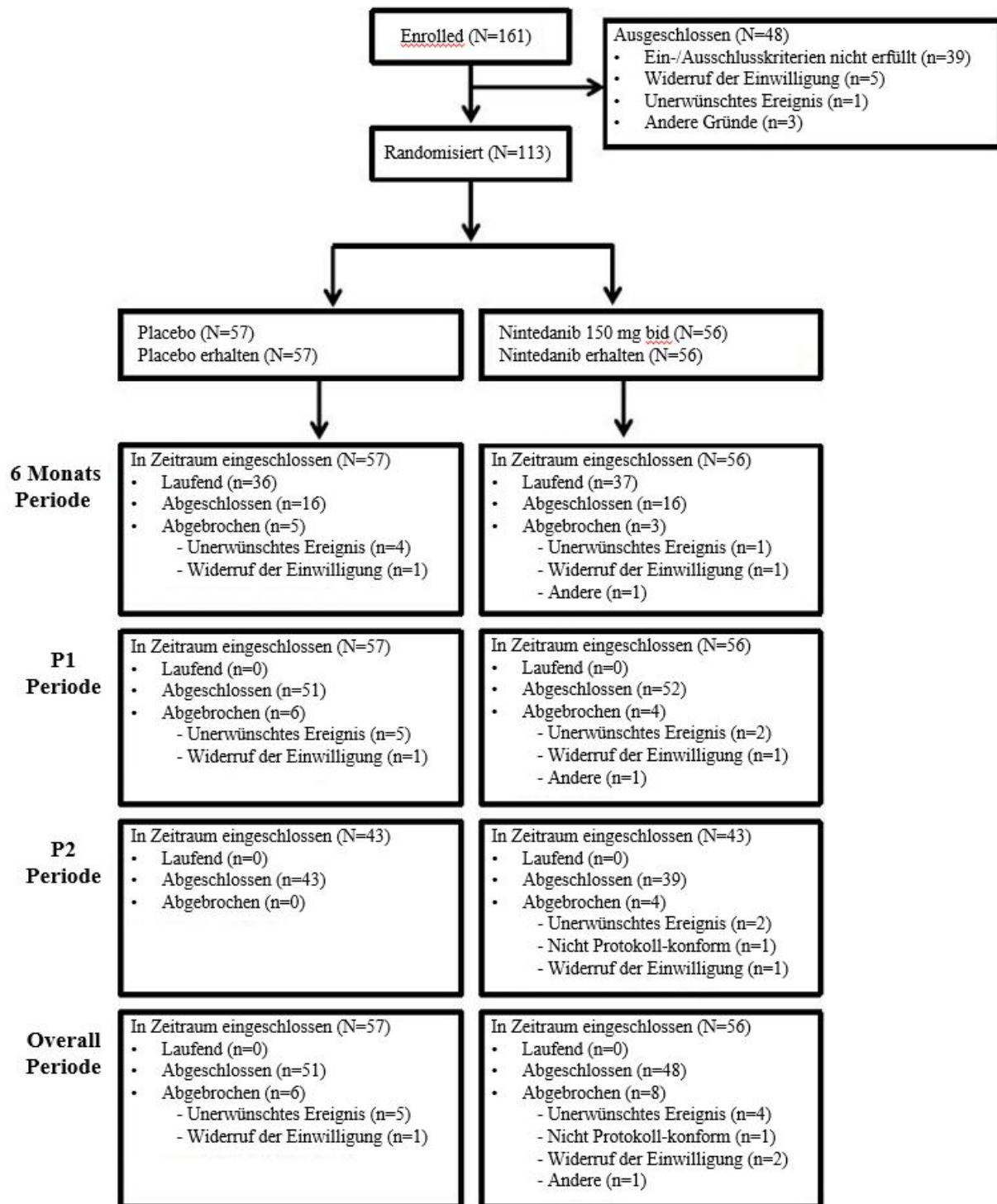


Abbildung 41: Patientenfluss Studie 1199.187

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. Quellspezifische 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 4-103 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INPULSIS 1

Studie: INPULSIS 1

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Boehringer Ingelheim. Clinical Trial Report 1199-0032	INPULSIS 1 Studienbericht
Boehringer Ingelheim. Clinical Trial Protocol 1199-0032	INPULSIS 1 CTP
Boehringer Ingelheim Trial Statistical Analysis Plan 1199-0032	INPULSIS 1 TSAP
Richeldi 2014b Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis	Richeldi 2014

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:

Computergenerierte Zuteilung zu einer Behandlungsgruppe.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

3. **Verblindung von Patienten und behandelnden Personen****Patient:**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass der Patient keine Rückschlüsse auf die Behandlung ziehen konnte.

behandelnde bzw. weiterbehandelnde Personen:☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass die behandelnde bzw. weiterbehandelnde Person keine Rückschlüsse auf die Behandlung ziehen konnte.

4. **Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

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:

Es handelt sich um eine doppelblinde, randomisierte, placebokontrollierte, multizentrische Studie. Randomisierung und Verblindung wurden adäquat durchgeführt und über den gesamten Studienzeitraum gewährleistet. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Studienebene ist somit als niedrig einzustufen.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**Endpunkt: Zeit bis zum Tod, unabhängig von der Ursache****1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zum respiratorisch bedingten Tod**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich. Die Beurteilung der Todesursache wurde sowohl durch den Prüfarzt als auch durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung der Todesursache ausgegangen werden.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Jährliche FVC-Abnahme (ml)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: FVC-Responder (Schwellenwert <5% und <10% des Sollwertes)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Schwellenwerte 5% und 10% des Sollwertes sind anerkannt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten akuten Exazerbation**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt, einem Facharzt mit Erfahrung in der Behandlung der IPF.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten adjudizierten akuten Exazerbation**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Zusätzlich wurden die Exazerbationen durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung des Ereignisses Exazerbation ausgegangen werden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-(≤ -4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-I-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-I-(≤ -4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SOBQ-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung CASA-Q Husten–Symptome und Husten-Impact**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: PGI-C-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung EQ-5D-VAS**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Unerwünschte Ereignisse (Gesamtzahl UE, SUE, UE zum Therapieabbruch oder Tod führend, UE von besonderen Interesse)

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Auswertungskriterien waren im Prüfplan a priori definiert. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Tabelle 4-104 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INPULSIS 2

Studie: INPULSIS 2**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

Genaue Benennung der Quelle	Kürzel
Boehringer Ingelheim. Clinical Trial Report 1199-0034	INPULSIS 2 Studienbericht
Boehringer Ingelheim. Clinical Trial Protocol 1199-0034	INPULSIS 2 CTP
Boehringer Ingelheim Trial Statistical Analysis Plan 1199-0034	INPULSIS 2 TSAP
Richeldi 2014. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis	Richeldi 2014

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:

Computergenerierte Zuteilung zu einer Behandlungsgruppe.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

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:

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 Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass der Patient keine Rückschlüsse auf die Behandlung ziehen konnte.

behandelnde bzw. weiterbehandelnde Personen:

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass die behandelnde bzw. weiterbehandelnde Person keine Rückschlüsse auf die Behandlung ziehen konnte.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

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:

Es handelt sich um eine doppelblinde, randomisierte, placebokontrollierte, multizentrische Studie. Randomisierung und Verblindung wurden adäquat durchgeführt und über den gesamten Studienzeitraum gewährleistet. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Studienebene ist somit als niedrig einzustufen.

B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**Endpunkt: Zeit bis zum Tod, unabhängig von der Ursache****1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zum respiratorisch bedingten Tod**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich. Die Beurteilung der Todesursache wurde sowohl durch den Prüfarzt als auch durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung der Todesursache ausgegangen werden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Jährliche FVC-Abnahme (ml)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: FVC-Responder (Schwellenwert 5% und 10% des Sollwertes)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Schwellenwerte 5% und 10% des Sollwertes sind anerkannt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten akuten Exazerbation**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt, einem Facharzt mit Erfahrung in der Behandlung der IPF.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten adjudizierten akuten Exazerbation**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Zusätzlich wurden die Exazerbationen durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung des Ereignisses Exazerbation ausgegangen werden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-(≤ -4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-I-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-I-($\leq$ -4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SOBQ-Gesamtscore**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung CASA-Q Husten–Symptome und Husten-Impact**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: PGI-C-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung EQ-5D-VAS**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Unerwünschte Ereignisse (Gesamtzahl UE, SUE, UE zum Therapieabbruch oder Tod führend, UE von besonderem Interesse)

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Auswertungskriterien waren im Prüfplan a priori definiert. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Tabelle 4-105 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 1199.187

Studie: 1199.187

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Boehringer Ingelheim. Clinical Trial Report 1199-0187	1199.0187 Studienbericht
Boehringer Ingelheim. Clinical Trial Protocol 1199-0187	1199.0187 CTP
Boehringer Ingelheim Trial Statistical Analysis Plan 1199-0187	1199.0187 TSAP

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:

Interactive Response Technology (IRT)

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

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

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:

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 Gruppenzuteilung blieb während der gesamten für die Nutzenbewertung relevanten Phase der Studie (kontrollierte Phase) verdeckt, sodass der Patient keine Rückschlüsse auf die Behandlung ziehen konnte

behandelnde bzw. weiterbehandelnde Personen:

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Gruppenzuteilung blieb während der gesamten Studie verdeckt, sodass der Patient keine Rückschlüsse auf die Behandlung ziehen konnte

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ ja ☐ unklar ☐ nein

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

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:

Es handelt sich um eine doppelblinde, randomisierte, placebokontrollierte, multizentrische Studie. Randomisierung und Verblindung wurden adäquat durchgeführt und über den gesamten kontrollierten Studienzeitraum gewährleistet. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Studienebene ist somit als niedrig einzustufen.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Zeit bis zum Tod, unabhängig von der Ursache

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zum respiratorisch bedingten Tod,**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die Feststellung eines Todesfalles ist zudem objektiv möglich. Die Beurteilung der Todesursache wurde sowohl durch den Prüfarzt als auch durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung der Todesursache ausgegangen werden.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Jährliche FVC-Abnahme (ml)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: FVC-Responder (Schwellenwert 5% und 10% des Sollwertes)**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Die FVC ist zudem objektiv bestimmbar.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Schwellenwerte 5% und 10% des Sollwertes sind anerkannt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten akuten Exazerbation**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt, einem Facharzt mit Erfahrung in der Behandlung der IPF.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Zeit bis zur ersten adjudizierten akuten Exazerbation**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind durch den Prüfarzt. Zusätzlich wurden die Exazerbationen durch ein unabhängiges Bewertungskomitee anhand der Patientenakten beurteilt. Dadurch kann von einer objektiven Einschätzung des Ereignisses Exazerbation ausgegangen werden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung des HRCT-QLF zu Woche 24**1. Verblindung der Endpunkterheber**

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☐ ja ☐ unklar ☒ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Für die Veränderung des HRCT-QLF zu Woche 24 ergibt sich aufgrund des hohen Anteils an Studienteilnehmern mit fehlenden Daten ein hohes Verzerrungspotenzial.

Endpunkt: Veränderung des 6 MWT zu Woche 24**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-(≤4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SGRQ-I-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: SGRQ-I-(≤4 Punkte)-Responder**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Als Cut-off-Wert für die Response wurde die für COPD validierte und für IPF ebenfalls akzeptierte MCID von vier Punkten zugrunde gelegt. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Veränderung SOBQ-Gesamtscore**1. Verblindung der Endpunkterheber**☒ ja ☐ unklar ☐ neinAngaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

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

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Rücklaufquoten waren in den beiden Studienarmen vergleichbar. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

Endpunkt: Unerwünschte Ereignisse (Gesamtzahl UE, SUE, UE zum Therapieabbruch oder Tod führend, UE von besonderem Interesse)

1. Verblindung der Endpunkterheber

☒ ja ☐ unklar ☐ nein

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Datenerhebung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

☒ ja ☐ unklar ☐ nein

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ ja ☐ unklar ☐ nein

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

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:

Die Verblindung der Endpunkterheber war über den gesamten kontrollierten Studienzeitraum gewährleistet. Das ITT-Prinzip wurde konsequent eingehalten. Weder eine ergebnisgesteuerte Berichterstattung noch sonstige Aspekte, die das Verzerrungspotenzial beeinflussen könnten, lagen vor. Die Auswertungskriterien waren im Prüfplan a priori definiert. Das Verzerrungspotenzial auf Endpunktebene ist somit als niedrig einzustufen.

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:
