

Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Nintedanib (Vargatef[®])

Boehringer Ingelheim Pharma GmbH & Co. KG

Modul 4 A

*Kombinationstherapie mit Docetaxel bei erwachsenen
Patienten mit lokal fortgeschrittenem, metastasiertem
oder lokal rezidiertem nicht-kleinzelligen
Lungenkarzinom (NSCLC) mit Adenokarzinom-
Histologie nach Erstlinien-Chemotherapie*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 19.12.2014

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	9
Abkürzungsverzeichnis	10
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik.....	24
4.2.1 Fragestellung.....	24
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	26
4.2.3 Informationsbeschaffung.....	30
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	30
4.2.3.2 Bibliografische Literaturrecherche.....	31
4.2.3.3 Suche in Studienregistern.....	32
4.2.3.4 Selektion relevanter Studien.....	33
4.2.4 Bewertung der Aussagekraft der Nachweise.....	34
4.2.5 Informationssynthese und -analyse.....	36
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	36
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	37
4.2.5.3 Meta-Analysen.....	49
4.2.5.4 Sensitivitätsanalysen.....	50
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	51
4.2.5.6 Indirekte Vergleiche.....	53
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	55
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	58
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	58
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	58
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	60
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	62
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	64
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	65
4.3.1.2.1 Studiendesign und Studienpopulationen.....	65
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	74
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	75
4.3.1.3.1 Endpunkte – RCT.....	76
4.3.1.3.1.1 Gesamtüberleben – RCT.....	77
4.3.1.3.1.2 Progressionsfreies Überleben – RCT.....	80
4.3.1.3.1.3 Objektive Ansprechrate – RCT.....	84
4.3.1.3.1.4 Krankheitskontrollrate – RCT.....	86

4.3.1.3.1.5	Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 – RCT	88
4.3.1.3.1.6	Gesundheitsbezogene Lebensqualität – RCT	98
4.3.1.3.1.7	Unerwünschte Ereignisse – RCT.....	103
4.3.1.3.1.8	Schwerwiegende unerwünschte Ereignisse – RCT	106
4.3.1.3.1.9	Unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 – RCT.....	108
4.3.1.3.1.10	Unerwünschte Ereignisse, die zum Therapieabbruch führten – RCT	109
4.3.1.3.1.11	Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ – RCT.....	111
4.3.1.3.2	Subgruppenanalysen – RCT.....	114
4.3.1.3.2.1	Subgruppenanalysen zu OS und PFS – RCT.....	115
4.3.1.3.2.2	Subgruppenanalysen zu ORR und DCR – RCT	117
4.3.1.3.2.3	Subgruppenanalysen zur Symptomatik – RCT	119
4.3.1.3.2.4	Subgruppenanalysen zur gesundheitsbezogenen Lebensqualität – RCT	136
4.3.1.3.2.5	Subgruppenanalysen zu unerwünschten Ereignissen – RCT	143
4.3.1.3.2.6	Subgruppenanalysen zu Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ – RCT.....	147
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	151
4.3.2	Weitere Unterlagen.....	155
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	155
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	155
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	156
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	156
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	156
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	159
4.3.2.2	Nicht randomisierte vergleichende Studien.....	159
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	159
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	160
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	160
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	160
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	162
4.3.2.3	Weitere Untersuchungen.....	162
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	162
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	163
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	163
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	163
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	164
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	164
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	164
4.4.1	Beurteilung der Aussagekraft der Nachweise	165

4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	166
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	176
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	176
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	176
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	177
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	177
4.5.4	Verwendung von Surrogatendpunkten.....	177
4.6	Liste der eingeschlossenen Studien.....	178
4.7	Referenzliste.....	180
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		188
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		194
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		197
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		199
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		204
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		215

Tabellenverzeichnis

	Seite
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Nintedanib	15
Tabelle 4-2: Ergebnisse zum medizinischen Zusatznutzen.....	21
Tabelle 4-3: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels.....	28
Tabelle 4-4: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	59
Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	60
Tabelle 4-6: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	63
Tabelle 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	65
Tabelle 4-8: Charakterisierung der eingeschlossenen Studie LUME-Lung 1 – RCT mit dem zu bewertenden Arzneimittel.....	67
Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	70
Tabelle 4-10: Charakterisierung der Studienpopulationen in der LUME-Lung 1 Studie (Patienten mit Adenokarzinom, RS) – RCT.....	73
Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	75
Tabelle 4-12: Matrix der Endpunkte in den eingeschlossenen RCT.....	75
Tabelle 4-13: Operationalisierung von Gesamtüberleben.....	77
Tabelle 4-14: Bewertung des Verzerrungspotenzials für Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel.....	78
Tabelle 4-15: Ergebnisse zum Gesamtüberleben aus RCT	79
Tabelle 4-16: Operationalisierung progressionsfreies Überleben	81
Tabelle 4-17: Bewertung des Verzerrungspotenzials für progressionsfreies Überleben in RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-18: Ergebnisse für progressionsfreies Überleben aus RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-19: Operationalisierung der objektiven Ansprechrate	84
Tabelle 4-20: Bewertung des Verzerrungspotenzials für die objektive Ansprechrate in RCT mit dem zu bewertenden Arzneimittel	84
Tabelle 4-21: Ergebnisse für die objektive Ansprechrate aus RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4-22: Operationalisierung der Krankheitskontrollrate	86

Tabelle 4-23: Bewertung des Verzerrungspotenzials für die Krankheitskontrollrate in RCT mit dem zu bewertenden Arzneimittel	86
Tabelle 4-24: Ergebnisse für die Krankheitskontrollrate aus RCT mit dem zu bewertenden Arzneimittel.....	87
Tabelle 4-25: Operationalisierung der Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13.....	88
Tabelle 4-26: Bewertung des Verzerrungspotenzials für Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 in RCT mit dem zu bewertenden Arzneimittel	89
Tabelle 4-27: Ausgangswerte der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	91
Tabelle 4-28: Ausgangswerte der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	92
Tabelle 4-29: Ergebnisse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel.....	93
Tabelle 4-30: Ergebnisse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalyse, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel.....	94
Tabelle 4-31: Ergebnisse zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel.....	96
Tabelle 4-32: Ergebnisse zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel.....	97
Tabelle 4-33: Operationalisierung der gesundheitsbezogenen Lebensqualität	98
Tabelle 4-34: Bewertung des Verzerrungspotenzials für gesundheitsbezogene Lebensqualität in RCT mit dem zu bewertenden Arzneimittel.....	99
Tabelle 4-35: Ausgangswerte der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-36: Ausgangswerte zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-37: Ergebnisse zur Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalyse, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel.....	101
Tabelle 4-38: Ergebnisse zur Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	102

Tabelle 4-39: Ergebnisse zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	103
Tabelle 4-40: Operationalisierung der unerwünschten Ereignisse.....	104
Tabelle 4-41: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel	104
Tabelle 4-42: Ergebnisse für die Zeit bis zum ersten Auftreten eines unerwünschten Ereignisses aus RCT mit dem zu bewertenden Arzneimittel.....	105
Tabelle 4-43: Operationalisierung von schwerwiegenden unerwünschten Ereignissen	106
Tabelle 4-44: Bewertung des Verzerrungspotenzials für schwerwiegende unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel.....	106
Tabelle 4-45: Ergebnisse für die Zeit bis zum ersten Auftreten eines schwerwiegenden unerwünschten Ereignisses aus RCT mit dem zu bewertenden Arzneimittel.....	107
Tabelle 4-46: Operationalisierung von unerwünschten Ereignissen mit CTCAE-Grad ≥ 3 ...	108
Tabelle 4-47: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 in RCT mit dem zu bewertenden Arzneimittel.....	108
Tabelle 4-48: Ergebnisse für die Zeit bis zum ersten Auftreten eines unerwünschten Ereignisses mit CTCAE-Grad ≥ 3 aus RCT mit dem zu bewertenden Arzneimittel.....	109
Tabelle 4-49: Operationalisierung von UE, die zum Therapieabbruch führten.....	110
Tabelle 4-50: Bewertung des Verzerrungspotenzials für UE, die zum Therapieabbruch führten, in RCT mit dem zu bewertenden Arzneimittel.....	110
Tabelle 4-51: Ergebnisse für die Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, aus RCT mit dem zu bewertenden Arzneimittel	111
Tabelle 4-52: Operationalisierung von Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$	112
Tabelle 4-53: Bewertung des Verzerrungspotenzials für Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ in RCT mit dem zu bewertenden Arzneimittel.....	112
Tabelle 4-54: Ergebnisse für die Zeit bis zum ersten Auftreten einer Nebenwirkung mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel	113
Tabelle 4-55: Anzahl der beobachteten und der zu erwartenden zufällig signifikanten Ergebnisse von Interaktionstests	115
Tabelle 4-56: p-Werte der Interaktionstests für die Endpunkte OS und PFS aus der Studie LUME-Lung 1.....	116
Tabelle 4-57: Einzelergebnisse der untersuchten Subgruppen für die Endpunkte OS und PFS aus der Studie LUME-Lung 1	117
Tabelle 4-58: p-Werte der Interaktionstests für die Endpunkte ORR und DCR aus der Studie LUME-Lung 1.....	117
Tabelle 4-59: Einzelergebnisse der untersuchten Subgruppen für die Endpunkte ORR und DCR aus der Studie LUME-Lung 1	119

Tabelle 4-60: p-Werte der Interaktionstests für die Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1	120
Tabelle 4-61: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1	122
Tabelle 4-62: p-Werte der Interaktionstests für die Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalysen) aus der LUME-Lung 1 Studie	124
Tabelle 4-63: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalysen) aus der Studie LUME-Lung 1	126
Tabelle 4-64: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 aus der Studie LUME-Lung 1	128
Tabelle 4-65: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 aus der Studie LUME-Lung 1	130
Tabelle 4-66: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 aus der Studie LUME-Lung 1	132
Tabelle 4-67: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 aus der Studie LUME-Lung 1	135
Tabelle 4-68: p-Werte der Interaktionstests für die Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1	137
Tabelle 4-69: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1	139
Tabelle 4-70: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 aus der Studie LUME-Lung 1	140
Tabelle 4-71: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 aus der Studie LUME-Lung 1	141
Tabelle 4-72: p-Werte der Interaktionstests für die gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS aus der Studie LUME-Lung 1	142
Tabelle 4-73: Einzelergebnisse der untersuchten Subgruppen zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS aus der Studie LUME-Lung 1	143

Tabelle 4-74: p-Werte der Interaktionstests für die Zeit bis zum ersten Auftreten unerwünschter Ereignisse aus der Studie LUME-Lung 1	144
Tabelle 4-75: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zum ersten Auftreten unerwünschter Ereignisse aus der Studie LUME-Lung 1.....	146
Tabelle 4-76: p-Werte der Interaktionstests für die Zeit bis zum ersten Auftreten der Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der Studie LUME-Lung 1	148
Tabelle 4-77: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zum ersten Auftreten der Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der Studie LUME-Lung 1.....	150
Tabelle 4-78: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	156
Tabelle 4-79: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	157
Tabelle 4-80: Operationalisierung von <Endpunkt xxx>.....	157
Tabelle 4-81: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	158
Tabelle 4-82: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	158
Tabelle 4-83: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	160
Tabelle 4-84: Operationalisierung von <Endpunkt xxx>.....	161
Tabelle 4-85: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	161
Tabelle 4-86: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	163
Tabelle 4-87: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	176
Tabelle 4-88: Liste der für die Nutzenbewertung berücksichtigten Studien und Untersuchungen.....	179
Tabelle 4-89 (Anhang): Studiendesign und -methodik für Studie LUME-Lung 1	205
Tabelle 4-90 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie LUME-Lung 1.....	216

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach Studien mit der dem zu bewertenden Arzneimittel Nintedanib	62
Abbildung 4-2: Kaplan-Meier-Kurven zum Gesamtüberleben (LUME-Lung 1 Studie, 2. Datenschnitt)	80
Abbildung 4-3: Kaplan-Meier-Kurven zum progressionsfreien Überleben (2. Datenschnitt, LUME-Lung 1 Studie).....	83
Abbildung 4-4: Kaplan-Meier-Kurven zum progressionsfreien Überleben (1. Datenschnitt, LUME-Lung 1 Studie).....	83

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALK	Anaplastic Lymphoma Kinase
ALT	Alanin-Aminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
AST	Aspartat-Aminotransferase
BI	Boehringer Ingelheim
b.i.d.	lat. <i>bis in die</i> (zweimal täglich)
CONSORT	Consolidated Standards of Reporting Trials
CR	Complete Response (komplettes Ansprechen)
CRF	Case Report Form
CT	Computertomographie
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease Control Rate (Krankheitskontrollrate)
DMC	Data Monitoring Committee
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal Growth Factor Receptor
EMA	European Medicines Agency (Europäische Arzneimittelbehörde)
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D	EuroQoL-5 Dimensions (Patientenfragebogen zur Lebensqualität)
EQ-VAS	EuroQoL-Visual Analogue Scale (visuelle Analogskala)
EU-CTR	EU Clinical Trials Register
EuroQoL	European Quality of Life
FGFR	Fibroblast Growth Factor Receptor
G-BA	Gemeinsamer Bundesausschuss
HR	Hazard Ratio
HRQoL	Health Related Quality of Life (gesundheitsbezogene Lebensqualität)
ICH-GCP	International Conference on Harmonisation-Good Clinical Practice
ICTRP	International Clinical Trials Registry Platform
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRR	Independent Radiologic Review (unabhängige, radiologische Beurteilung)
ITT	Intention-to-Treat

IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
KI	Konfidenzintervall
LDH	Lactatdehydrogenase
MRI	Magnetic Resonance Imaging (Magnetresonanztomographie)
NSCLC	Non-Small Cell Lung Cancer (nicht-kleinzelliges Lungenkarzinom)
NYHA	New York Heart Association
OR	Odds Ratio
ORR	Objective Response Rate (objektive Ansprechrage)
OS	Overall Survival (Gesamtüberleben)
PDGFR	Platelet Derived Growth Factor Receptor
PFS	Progression Free Survival (progressionsfreies Überleben)
PR	Partial Response (partielleres Ansprechen)
PS	Performance Status (Allgemeinzustand)
PT	Preferred Term
QLG	Quality of Life Group
QLQ-C30	Quality of Life Questionnaire Core-30
QLQ-LC13	Quality of Life Questionnaire Lung Cancer-13
RCT	Randomized Controlled Trial (randomisierte kontrollierte Studie)
RD	Risikodifferenz
RECIST	Response Evaluation Criteria In Solid Tumors
RR	Relatives Risiko
RS	Randomized Set
RTK	Rezeptor-Tyrosinkinase
SD	Stable Disease (stabile Erkrankung)
SGB	Sozialgesetzbuch
SSC	Special Search Category
STD	Standard Deviation (Standardabweichung)
SUE	Schwerwiegendes unerwünschtes Ereignis
TKI	Tyrosinkinase-Inhibitor
TS	Treated Set
TSAP	Trial Statistical Analysis Plan (statistischer Analyseplan)
UE	Unerwünschtes Ereignis

UICC	Union for International Cancer Control
VEGFR	Vascular Endothelial Growth Factor Receptor
VerfO	Verfahrensordnung des Gemeinsamen Bundesausschusses
WHO	World Health Organization
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 Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Die Neubildung von Blutgefäßen (Angiogenese) ist von zentraler Bedeutung bei pathologischen Prozessen, wie z. B. dem Wachstum und der Ausbreitung solider Tumore (Folkman, 1995). Nintedanib (Vargatef[®]) blockiert zielgerichtet drei wichtige Rezeptor-Tyrosinkinasen (RTK) der Angiogenese: VEGFR (Vascular Endothelial Growth Factor Receptor), PDGFR (Platelet Derived Growth Factor Receptor) sowie FGFR (Fibroblast Growth Factor Receptor) und ermöglicht die Inhibition mehrerer wichtiger Signalwege der Angiogenese.

Ziel dieses Dossiers ist die Bewertung des Zusatznutzens von Nintedanib in Kombination mit Docetaxel im Vergleich zur Docetaxel-Monotherapie bei erwachsenen Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem nicht-kleinzelligem Lungenkarzinom (Non-Small Cell Lung Cancer [NSCLC]) mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie. Der Vergleich gegen eine Monotherapie mit Docetaxel entspricht für das vorliegende Anwendungsgebiet den Empfehlungen des Gemeinsamen Bundesausschusses (G-BA) zur zweckmäßigen Vergleichstherapie (ZVT).

Die Nutzenbewertung wurde bezüglich patientenrelevanter Endpunkte aus den Endpunktkategorien Mortalität (Gesamtüberleben [Overall Survival (OS)]), Morbidität (progressionsfreies Überleben [Progression Free Survival (PFS)]), objektive Ansprechraten [Objective Response Rate (ORR)], Krankheitskontrollrate [Disease Control Rate (DCR)], generelle patientenberichtete Symptome einer Krebserkrankung und spezifische Symptome bei Lungenkrebs), gesundheitsbezogene Lebensqualität (Health Related Quality of Life [HRQoL]) sowie Nebenwirkungen vorgenommen.

Datenquellen

Die Daten zum Nachweis des medizinischen Nutzens und des medizinischen Zusatznutzens von Nintedanib+Docetaxel basieren auf der pivotalen, randomisiert kontrollierten, doppelblinden Phase-III-Studie (Randomized Controlled Trial [RCT]) LUME-Lung 1.

Ein-/Ausschlusskriterien für Studien

Die für die Nutzenbewertung relevanten Studien wurden nach folgenden Kriterien eingegrenzt:

Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Nintedanib

Einschlusskriterien		
E1	Population	Erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie.
E2	Intervention	Die zu bewertende Intervention ist die Kombination aus Nintedanib+Docetaxel gemäß Fachinformation.
E3	Vergleichstherapie	Docetaxel gemäß Fachinformation*.
E4	Endpunkte	Mindestens einer der patientenrelevanten Endpunkte wird berichtet: <ul style="list-style-type: none"> • Mortalität <ul style="list-style-type: none"> ○ OS • Morbidität/Symptomatik <ul style="list-style-type: none"> ○ PFS ○ ORR ○ DCR ○ Tumorbedingte und therapieassoziierte Symptomatik • HRQoL • Nebenwirkungen
E5	Studientyp	RCT
E6	Publikationstyp	Vollpublikation, Studienbericht oder ausführlicher Ergebnisbericht aus einem Studienregister verfügbar in deutscher oder englischer Sprache.
E7	Studiendauer	Keine Einschränkung
Ausschlusskriterien		
A1	Außerhalb der Zulassung	Anwendung außerhalb des von den zuständigen Zulassungsbehörden genehmigten Anwendungsgebiets
A2	Präklinische Studien	Präklinische <i>in vitro</i> und <i>in vivo</i> Studien
* Der von BI gezeigte Vergleich gegen Docetaxel entspricht den Empfehlungen des G-BA zur ZVT. Die bibliografische Literaturrecherche sowie die Suche in Studienregistern wurden nicht auf eine bestimmte Vergleichstherapie eingeschränkt.		

Neben den Ausschlusskriterien führte das Nichterfüllen eines in Tabelle 4-1 aufgeführten Einschlusskriteriums zum Ausschluss der Studie.

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

Das Design und die Methodik der eingeschlossenen RCT (LUME-Lung 1 Studie) wurde anhand der Consolidated Standards of Reporting Trials (CONSORT)-Kriterien dargestellt. Die Studienpopulation wurde anhand demografischer und krankheitsspezifischer Faktoren beschrieben. Verzerrungsaspekte wurden endpunktübergreifend für die gesamte eingeschlossene Studie und endpunktspezifisch bewertet. Die Bewertung folgte den Vorgaben der Dossiervorlage. Die Patientenrelevanz und die Validität der Endpunkte bzw. Messinstrumente wurden überprüft und diskutiert.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Der medizinische Nutzen von Nintedanib in Kombination mit Docetaxel (Nintedanib+Docetaxel) wurde mit der Zulassung durch die Europäische Arzneimittelbehörde (European Medicines Agency [EMA]) bestätigt. Nachfolgend wird der medizinische Zusatznutzen von Nintedanib in Kombination mit Docetaxel im Vergleich zur Docetaxel-Monotherapie bei erwachsenen Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie beschrieben. Bei der Wahl der ZVT konnte Boehringer Ingelheim (BI) aus den gleichermaßen zweckmäßigen Vergleichstherapien Docetaxel oder Pemetrexed (Alimta[®]) bzw. Docetaxel, Pemetrexed, Gefitinib (Iressa[®]) oder Erlotinib (Tarceva[®]) für Patienten mit aktivierenden Epidermal Growth Factor Receptor (EGFR)-Mutationen bzw. Docetaxel, Pemetrexed oder Crizotinib für Patienten mit aktivierenden Anaplastic Lymphoma Kinase (ALK)-Mutationen auswählen (Gemeinsamer Bundesausschuss, 2014d; Gemeinsamer Bundesausschuss, 2014a). BI folgte bezüglich der ZVT den Empfehlungen des G-BA: Da die Zulassungsstudie LUME-Lung 1 in Übereinstimmung hiermit einen direkten Vergleich von Nintedanib+Docetaxel zu Placebo+Docetaxel ermöglicht, zeigt BI den Zusatznutzen dazu. Mit der LUME-Lung 1 Studie liegt eine doppelblinde RCT vor, so dass die zur Bewertung des Zusatznutzens herangezogene Evidenz der höchsten Evidenzstufe entspricht (Evidenzstufe Ib).

Endpunkte

Gesamtüberleben

Für Patienten mit Adenokarzinom war das Gesamtüberleben (OS) unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0359$). Das Sterberisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 17%, verglichen mit Patienten, die mit Placebo+Docetaxel behandelt wurden (Hazard Ratio [HR] [95%-Konfidenzintervall (KI)]: 0,83 [0,70; 0,99]). Dies entsprach einer Verlängerung des medianen OS um 2,3 Monate von 10,3 auf 12,6 Monate.

Progressionsfreies Überleben

Progressionsfreies Überleben (PFS) war der primäre Endpunkt der LUME-Lung 1 Studie. Für Patienten mit Adenokarzinom war das PFS unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0485$). Das Progressionsrisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 16%, verglichen mit Patienten, die mit Placebo+Docetaxel behandelt wurden (HR [95%-KI]: 0,84 [0,71; 1,00]). Dies entsprach einer Verlängerung des medianen PFS um 1,4 Monate von 2,8 auf 4,2 Monate. Der PFS-Vorteil zugunsten der Kombinationstherapie war bereits zum Zeitpunkt der primären PFS-Analyse nachweisbar (medianes PFS: 4,0 Monate im Nintedanib+Docetaxel-Arm vs. 2,8 Monate im Placebo+Docetaxel-Arm; HR [95%-KI]: 0,77 [0,62; 0,96]; $p=0,0193$).

Objektives Ansprechen

Hinsichtlich der ORR lag kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen vor ($p=0,4823$). Das Relative Risiko (RR) [95%-KI] belief sich auf 1,30 [0,62; 2,74], das Odds Ratio (OR) [95%-KI] auf 1,32 [0,61; 2,86] und die Risikodifferenz (RD) [95%-KI] auf 1,09 [-1,95; 4,13].

Krankheitskontrollrate

Die DCR war für Patienten mit der Kombinationstherapie Nintedanib+Docetaxel statistisch signifikant besser im Vergleich zur Placebo+Docetaxel-Therapie ($p<0,0001$). Im Nintedanib+Docetaxel-Behandlungsarm erreichten 60,2% der Patienten eine Krankheitskontrolle, im Placebo+Docetaxel-Arm nur 44,0% (RR [95%-KI]: 1,37 [1,18; 1,59] / OR [95%-KI]: 1,93 [1,41; 2,63] / RD [95%-KI]: 16,20 [8,67; 23,73]).

Symptomatik

Die Symptomatik wurde anhand der Fragebögen European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 (Quality of Life Questionnaire Core-30) und EORTC QLQ-LC13 (Quality of Life Questionnaire Lung Cancer-13) erhoben. Die Ergebnisse werden in dem Dossier anhand zweier Analysen dargestellt, die klinisch relevante Veränderungen für unterschiedliche patientenrelevante Nutzenaspekte abbilden:

1. Responderanalyse des Anteils an Patienten, die durchschnittlich eine Verbesserung des jeweiligen Scores um mindestens 10 Punkte erreicht haben (Osoba et al., 1998)
2. Zeit bis zur Verschlechterung des Scores der jeweiligen Symptome um mindestens 10 Punkte (Osoba et al., 1998)

Nachfolgend wird nur auf Symptome mit statistisch signifikanten Ergebnissen zwischen den Vergleichsarmen eingegangen. Hierzu zählen Schmerzen, Diarrhö sowie Übelkeit und Erbrechen. Für alle anderen Symptome lagen keine statistisch signifikanten Unterschiede vor.

Schwerwiegende Symptome

Schmerzen

Bei der Responderanalyse zur Verbesserung der Symptomatik Schmerz, erhoben mit dem Fragebogen EORTC QLQ-C30, ergab sich ein statistisch signifikanter Vorteil zugunsten der Kombinationstherapie Nintedanib+Docetaxel ($p=0,005$). Das RR [95%-KI] lag bei 1,56 [1,13; 2,14], das OR [95%-KI] bei 1,76 [1,18; 2,62] und die RD [95%-KI] bei 9,52 [2,84; 16,20]. Bei Betrachtung der Zeit bis zur Verschlechterung des Symptoms Schmerzen (EORTC QLQ-C30) lag kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen vor.

Die Zeit bis zur Verschlechterung der Schmerzen in Arm/Schulter (EORTC QLQ-LC13) war unter der Therapie mit Nintedanib+Docetaxel signifikant länger als unter der Therapie mit Placebo+Docetaxel (HR [95%-KI]: 0,80 [0,63; 1,00]; $p=0,0456$). Bezüglich der

Responderanalyse zur Verbesserung der Symptomatik von Schmerzen in Arm/Schulter, gemessen mit EORTC QLQ-LC13, zeigte sich kein statistisch signifikanter Effekt.

Weitere Symptome

Diarrhö

Das Symptom Diarrhö wird über den Fragebogen EORTC QLQ-C30 erfasst. Für die Verbesserung der Symptomatik ergab sich kein statistisch signifikanter Effekt. Bei Betrachtung der Zeit bis zur Verschlechterung der Symptomatik zeigte sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (HR [95%-KI]: 1,90 [1,54; 2,34]; $p < 0,0001$). Empfehlungen für geeignete Maßnahmen und proaktives Therapiemanagement finden sich in der Fach- und Gebrauchsinformation für Nintedanib wieder (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Übelkeit und Erbrechen

Das Symptom Übelkeit und Erbrechen wurde über den Fragebogen EORTC QLQ-C30 gemessen. Für die Verbesserung der Symptomatik ergab sich kein statistisch signifikanter Effekt. Bezüglich der Zeit bis zur Verschlechterung der Symptomatik zeigte sich für Übelkeit und Erbrechen ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (HR [95%-KI]: 1,23 [1,00; 1,50]; $p = 0,0467$). Empfehlungen für geeignete Maßnahmen und proaktives Therapiemanagement finden sich in der Fach- und Gebrauchsinformation für Nintedanib wieder (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Gesundheitsbezogene Lebensqualität

Die HRQoL wurde anhand der Patientenfragebögen EORTC QLQ-C30 und EQ-5D (EuroQoL [European Quality of Life]-5 Dimensions) erhoben. Analog zur Darstellung der Symptomatik werden für den EORTC QLQ-C30 die klinisch relevanten Veränderungen durch die Responderanalysen sowie die Zeit bis zur Verschlechterung dargestellt.

1. Responderanalyse des Anteils an Patienten, die durchschnittlich eine Verbesserung des jeweiligen Scores um mindestens 10 Punkte erreicht haben (Osoba et al., 1998)
2. Zeit bis zur Verschlechterung des Scores der jeweiligen Symptome um mindestens 10 Punkte (Osoba et al., 1998)

Die Analysen des EQ-5D Fragebogens umfassen die Mittelwerte bis zur medianen Follow-up-Zeit sowie die entsprechende Mittelwertdifferenz der beiden Behandlungsarme. Die Ergebnisse wurden im Dossier auf Basis des EQ-5D UK Utility Scores sowie der EQ-VAS (EuroQoL-Visual Analogue Scale [visuelle Analogskala]) dargestellt.

Nachfolgend wird nur auf die statistisch signifikanten Ergebnisse zwischen den Vergleichsarmen eingegangen. Diese umfassen die Daten zur Verbesserung des globalen Gesundheitsstatus sowie der kognitiven Funktion, die beide mittels des Fragebogens EORTC QLQ-C30 erhoben wurden. Alle anderen Skalen des EORTC QLQ-C30 sowie die Daten, die

anhand des EQ-5D Fragebogens erhoben wurden, wiesen keine statistisch signifikanten Unterschiede auf.

Globaler Gesundheitsstatus

Es zeigte sich bei der Verbesserung der HRQoL für den globalen Gesundheitsstatus ein statistisch signifikanter Vorteil für die Kombinationstherapie Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel ($p=0,026$). Das RR [95%-KI] lag bei 1,56 [1,05; 2,33], das OR [95%-KI] bei 1,69 [1,06; 2,69] und die RD [95%-KI] bei 6,64 [0,82; 12,45]. Für die Operationalisierung Zeit bis zur Verschlechterung lag kein statistisch signifikanter Unterschied vor.

Kognitive Funktion

Es zeigte sich bei der Verbesserung der HRQoL bezüglich der kognitiven Funktion ein statistisch signifikanter Vorteil für Nintedanib in Kombination mit Docetaxel gegenüber der Vergleichstherapie mit Placebo+Docetaxel ($p=0,009$). Es ergaben sich ein RR [95%-KI] von 1,64 [1,12; 2,41], ein OR [95%-KI] von 1,81 [1,15; 2,84] und eine RD [95%-KI] von 8,00 [2,00; 14,00]. Die Zeit bis zur Verschlechterung unterschied sich in den beiden Behandlungsgruppen nicht signifikant.

Nebenwirkungen

Im Rahmen dieser Nutzenbewertung wurden unerwünschten Ereignisse (UE), die schwerwiegenden unerwünschten Ereignisse (SUE), die UE vom Common Terminology Criteria for Adverse Events (CTCAE)-Grad ≥ 3 nach Version 3.0 (National Cancer Institute, 2006) sowie die UE, die zum Therapieabbruch führten, berücksichtigt. Zusätzlich wurden jeweils die Nebenwirkungen einzeln betrachtet, welche UE mit einem CTCAE-Grad ≥ 3 umfassen, die in $\geq 5\%$ der Patienten eines Behandlungsarms auftraten (siehe Abschnitt 4.2.5.2). Insgesamt unterschieden sich die Zeiten bis zum ersten Auftreten der patientenrelevanten Nebenwirkungen der mit Nintedanib+Docetaxel behandelten Patienten nicht signifikant von denen der mit Placebo+Docetaxel behandelten Patienten:

- Alle UE: HR [95%-KI]: 1,10 [0,94; 1,29]; $p=0,2067$
- Alle SUE: HR [95%-KI]: 1,01 [0,78; 1,32]; $p=0,9316$
- Alle UE mit CTCAE-Grad ≥ 3 : HR [95%-KI]: 1,10 [0,92; 1,32]; $p=0,2655$
- UE, die zum Therapieabbruch führten: HR [95%-KI]: 1,08 [0,76; 1,54]; $p=0,6557$
- Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$: Keine signifikanten Unterschiede in den einzelnen UEs (siehe Tabelle 4-54)

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Tabelle 4-2 fasst die signifikanten Ergebnisse zu den einzelnen patientenrelevanten Endpunkten der LUME-Lung 1 Studie zusammen und leitet für diese Endpunkte das Ausmaß und die Wahrscheinlichkeit des Zusatznutzens für erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie, ab.

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

Tabelle 4-2: Ergebnisse zum medizinischen Zusatznutzen

Endpunkt	Effektschätzer [95%-KI] / Nintedanib+Docetaxel vs. Placebo+Docetaxel / p-Wert	Beurteilung nach Arzneimittel-Nutzenbewertungsverordnung (AM-Nutzen V)	Ausmaß des Zusatznutzens einschließlich dessen Wahrscheinlichkeit
Mortalität			
Gesamtüberleben (OS)	HR: 0,83 [0,70; 0,99]; p=0,0359	Bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne einer moderaten Verlängerung der Lebensdauer	Beträchtlicher Zusatznutzen; Wahrscheinlichkeit: Hinweis
Morbidität			
Progressionsfreies Überleben (PFS)	HR: 0,84 [0,71; 1,00]; p=0,0485	Bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapielevanten Nutzens im Sinne einer Verlängerung der Zeit ohne Tumorprogression/Tod	Geringer Zusatznutzen; Wahrscheinlichkeit: Hinweis
Krankheitskontrollrate (DCR)	RR: 1,37 [1,18; 1,59]; p<0,0001	Bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne einer bedeutsamen Stabilisierung der Erkrankung	Beträchtlicher Zusatznutzen; Wahrscheinlichkeit: Hinweis
Symptomatik (schwerwiegend)			
Schmerzen (allgemein), Symptomverbesserung (EORTC QLQ-C30)	RR (EORTC QLQ-C30): 1,56 [1,13; 2,14]; p=0,005	Bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne einer Abschwächung eines schwerwiegenden Symptoms	Beträchtlicher Zusatznutzen; Wahrscheinlichkeit: Hinweis
Schmerzen (Arm/Schulter), Zeit bis zur Verschlechterung (EORTC QLQ-LC13)	HR (EORTC QLQ-LC13): 0,80 [0,63; 1,00]; p=0,0456	Bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapielevanten Nutzens im Sinne einer Verringerung eines schwerwiegenden Symptoms	Geringer Zusatznutzen; Wahrscheinlichkeit: Hinweis

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

Endpunkt	Effektschätzer [95%-KI] / Nintedanib+Docetaxel vs. Placebo+Docetaxel / p-Wert	Beurteilung nach Arzneimittel- Nutzenbewertungsverordnung (AM-Nutzen V)	Ausmaß des Zusatznutzens einschließlich dessen Wahrscheinlichkeit
Symptomatik (weitere)			
Diarrhö, Zeit bis zur Verschlechterung (EORTC QLQ-C30)	HR: 1,90 [1,54; 2,34]; p<0,0001	Signifikant kürzere Zeit bis zur Verschlechterung	Nachteil; Wahrscheinlichkeit: Hinweis
Übelkeit und Erbrechen, Zeit bis zur Verschlechterung (EORTC QLQ-C30)	HR: 1,23 [1,00; 1,50]; p=0,0467	Signifikant kürzere Zeit bis zur Verschlechterung	Nachteil; Wahrscheinlichkeit: Hinweis
Gesundheitsbezogene Lebensqualität (HRQoL)			
Globaler Gesundheitsstatus, Verbesserung (EORTC QLQ-C30)	RR: 1,56 [1,05; 2,33]; p=0,026	Bisher nicht erreichte deutliche Verbesserung der Lebensqualität	Beträchtlicher Zusatznutzen; Wahrscheinlichkeit: Hinweis
Kognitive Funktion, Verbesserung (EORTC QLQ-C30)	RR: 1,64 [1,12; 2,41]; p=0,009	Bisher nicht erreichte deutliche Verbesserung der Lebensqualität	Beträchtlicher Zusatznutzen; Wahrscheinlichkeit: Hinweis
Ein HR <1 bzw. ein RR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.			

Gesamtfazit

Aus Tabelle 4-2 geht hervor, dass die Behandlung mit Nintedanib+Docetaxel für erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie eine wirksame und im Allgemeinen gut verträgliche Therapieoption darstellt. Mit einer Verlängerung des medianen Gesamtüberlebens um 2,3 Monate auf 12,6 Monate ist mit Nintedanib in Kombination mit Docetaxel als Therapieoption für Adenokarzinom-Patienten ein relevanter Fortschritt in der Zweitlinien-Therapie des fortgeschrittenen NSCLC erreicht worden. Die Daten der LUME-Lung 1 Studie zum wichtigsten sekundären Endpunkt OS konnten erstmals seit fast 10 Jahren einen signifikanten Überlebensvorteil im vorliegenden Anwendungsgebiet zeigen.

Aus dieser klinisch relevanten Verlängerung des Gesamtüberlebens, der Verbesserung der aus Patientensicht wichtigen Endpunkte PFS, DCR und Lebensqualität bei gut kontrollierbaren Nebenwirkungen resultiert ein **beträchtlicher Zusatznutzen** für Nintedanib.

4.2 Methodik

Zusammenfassung des Abschnitts 4.2

Dieser Abschnitt beschreibt die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und Zusatznutzens von Nintedanib. Ziel dieses Dossiers ist die Bewertung des Zusatznutzens von Nintedanib in Kombination mit Docetaxel im Vergleich zur Docetaxel-Monotherapie. Die Docetaxel-Monotherapie ist als mögliche ZVT vom G-BA für das vorliegende Anwendungsgebiet empfohlen.

Die Daten zum Nachweis des medizinischen Nutzens und des medizinischen Zusatznutzens von Nintedanib+Docetaxel basieren auf der pivotalen, randomisiert kontrollierten, doppelblinden Phase-III-Studie (Randomized Controlled Trial) LUME-Lung 1. Das Design und die Methodik der eingeschlossenen RCT (LUME-Lung 1 Studie) wurden anhand der Consolidated Standards of Reporting Trials (CONSORT)-Kriterien dargestellt. Die Studienpopulation wurde anhand demografischer und krankheitsspezifischer Faktoren beschrieben. Verzerrungsaspekte wurden endpunktübergreifend für die gesamte eingeschlossene Studie und endpunktspezifisch bewertet. Die Bewertung folgte den Vorgaben der Dossiervorlage. Die Patientenrelevanz und die Validität der Endpunkte bzw. Messinstrumente wurden überprüft und diskutiert.

Eine bibliografische Literaturrecherche wurde entsprechend der Vorgaben in den Datenbanken EMBASE, EMBASE Daily Alerts, MEDLINE, MEDLINE In-Process und Cochrane Central Register of Controlled Trials durchgeführt. Die Studienregistersuche entsprach den Vorgaben der Dossiervorlage.

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.

Nintedanib in Kombination mit Docetaxel ist vorgesehen für die Anwendung bei erwachsenen Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem nicht-kleinzelligem Lungenkarzinom (NSCLC) mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Basis der europäischen Zulassung war die pivotale randomisierte, kontrollierte und doppelblinde Phase-III-Studie LUME-Lung 1 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a).

Ziel dieses Dossiers ist die Bewertung des Zusatznutzens von Nintedanib in Kombination mit Docetaxel im Vergleich zur Docetaxel-Monotherapie. Die Docetaxel-Monotherapie ist als mögliche ZVT vom G-BA für das vorliegende Anwendungsgebiet empfohlen.

Der G-BA legte anhand der Kriterien gemäß 5. Kapitel § 6 Verfahrensordnung (VerfO) (Beratungsanforderung 2013-B-119 Nintedanib) die Chemotherapien Docetaxel oder Pemetrexed als ZVT für Nintedanib in Kombination mit Docetaxel im zugelassenen Anwendungsgebiet fest (siehe Modul 3 Abschnitt 3.1). Für Patienten mit aktivierenden EGFR-Mutationen benannte der G-BA zusätzlich eine Behandlung mit den Proteinkinase-Inhibitoren Gefitinib oder Erlotinib als ZVT und für Patienten mit aktivierenden ALK-Mutationen zusätzlich eine Behandlung mit Crizotinib (Gemeinsamer Bundesausschuss, 2014d; Gemeinsamer Bundesausschuss, 2014a). Die Geschäftsstelle des G-BA wies BI darauf hin, dass der pharmazeutische Unternehmer den Nachweis des Zusatznutzens im Dossier gegen einen oder mehrere der genannten Vergleichstherapien (Docetaxel oder Pemetrexed bzw. Docetaxel, Pemetrexed, Gefitinib oder Erlotinib für Patienten mit aktivierenden EGFR-Mutationen bzw. Docetaxel, Pemetrexed oder Crizotinib für Patienten mit aktivierenden ALK-Mutationen) darlegen kann (Gemeinsamer Bundesausschuss, 2014d; Gemeinsamer Bundesausschuss, 2014a). BI folgte den Empfehlungen des G-BA und vergleicht sich gegen Docetaxel. Da eine vergleichende Studie zu Docetaxel vorliegt (LUME-Lung 1), wird in diesem Dossier die beste verfügbare Evidenz dargestellt und der medizinische Zusatznutzen gegenüber dieser ZVT abgeleitet.

Die Bewertung erfolgt bezüglich patientenrelevanter Endpunkte im Hinblick auf Mortalität, Morbidität/Symptomatik (tumorbedingt und therapieassoziiert), Lebensqualität und Nebenwirkungen auf Basis von RCT ohne Begrenzung der Studiendauer.

Die einzelnen Komponenten dieser Fragestellungen (Patientenpopulation, Intervention, Vergleichstherapie, Endpunkte, Studientypen) werden in Abschnitt 4.2.2 näher spezifiziert.

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.

Diejenigen Studien, die grundsätzlich für die Beantwortung der Fragestellung gemäß Abschnitt 4.2.1 geeignet sind, wurden anhand von Ein- und Ausschlusskriterien charakterisiert. Maßgeblich für die Wahl der Kriterien für den Einschluss von Studien in die Nutzenbewertung ist grundsätzlich der Zulassungsstatus des zu bewertenden Arzneimittels (Gemeinsamer Bundesausschuss, 2014b).

Population

Gemäß dem zugelassenen Anwendungsgebiet wurden für die Nutzenbewertung von Nintedanib folgende Studien berücksichtigt:

Studien, die erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie, einschlossen.

Intervention

Die zu bewertende Intervention ist Nintedanib in Kombination mit Docetaxel. In die Nutzenbewertung sind Studien einzuschließen, in denen Nintedanib+Docetaxel gemäß Angaben in der Fachinformation gegeben wurde. Diese sieht eine zweimal tägliche Einnahme von 200 mg Nintedanib vor. Die beiden Nintedanib-Einnahmen sollen im Abstand von ungefähr 12 Stunden erfolgen, jeweils an Tag 2 bis 21 eines Standard-Docetaxel-Behandlungszyklus (toleranzadaptierte Dosisreduktion nach Schema auf zweimal täglich 150 mg oder zweimal täglich 100 mg möglich) (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b). Der Standardbehandlungszyklus mit Docetaxel sieht eine Infusion alle drei Wochen mit einer empfohlenen Dosierung von 75mg/m² vor (Aventis Pharma S.A., 2014).

Vergleichstherapie

Da eine vergleichende Studie zu Docetaxel vorliegt (LUME-Lung 1), wird in diesem Dossier die beste verfügbare Evidenz dargestellt und der medizinische Zusatznutzen gegenüber dieser ZVT abgeleitet.

Die von BI gewählte Suchstrategie folgt den Empfehlungen des G-BA (Gemeinsamer Bundesausschuss, 2014d), die bibliografische Literaturrecherche sowie die Suche in Studienregistern für den Nachweis des Zusatznutzens nicht allein auf die ausgewählte(n) Vergleichstherapie(n) einzuschränken. Im Dossier sind alle verfügbaren Studienberichte zu direkten Vergleichsstudien mit den Optionen zur zweckmäßigen Vergleichstherapie enthalten.

Sowohl das zu bewertende Arzneimittel als auch die herangezogenen Vergleichstherapie(n) müssen in den Studien zulassungskonform verabreicht worden sein.

Endpunkte

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

Für die vorliegende Nutzenbewertung wurden Studien berücksichtigt, die Ergebnisse zu mindestens einem der folgenden patientenrelevanten Endpunkte berichten:

- Mortalität
 - Gesamtüberleben (OS)
- Morbidität/Symptomatik
 - Progressionsfreies Überleben (PFS)
 - Objektive Ansprechrate (ORR)
 - Krankheitskontrollrate (DCR)
 - Tumorbedingte und therapieassoziierte Symptomatik
- Gesundheitsbezogene Lebensqualität (HRQoL)
- Nebenwirkungen

Eine ausführliche Diskussion und Begründung der Patientenrelevanz dieser Endpunkte, sowie der Validität der für die Symptomatik und HRQoL verwendeten Messinstrumente und Fragebögen, findet sich im Abschnitt 4.2.5.2.

Studientyp

Für die gegebene Fragestellung liegt mit der Zulassungsstudie LUME-Lung 1 zu Nintedanib+Docetaxel eine RCT vor.

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.

Publikationstyp

Für die Nutzenbewertung wurden Studienberichte und Vollpublikationen berücksichtigt, die die CONSORT-Kriterien erfüllen. Darüber hinaus konnten auch Studien berücksichtigt werden, für die ein ausführlicher Ergebnisbericht aus einem Studienregister vorlag.

Mehrfachpublikationen ohne relevante Zusatzinformation wurden von der Nutzenbewertung ausgeschlossen. Ebenfalls ausgeschlossen wurden Studien, zu denen Ergebnisse lediglich als Kongressbeiträge oder Poster vorlagen. Diese Publikationstypen wurden als nicht ausreichend für eine Bewertung der Studie angesehen.

Studiendauer

Hinsichtlich der Studiendauer wurde keine Einschränkung vorgenommen.

Außerhalb der Zulassung

Falls die Anwendung des Wirkstoffs in den zu betrachtenden Studien außerhalb des zugelassenen Anwendungsgebiets erfolgte, wurden diese ausgeschlossen.

Präklinische Studien

Präklinische *in vitro* und *in vivo* Studien wurden nicht betrachtet und ausgeschlossen.

Eine Übersicht über die Ein- und Ausschlusskriterien und deren Begründung findet sich in Tabelle 4-3.

Tabelle 4-3: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels

Einschlusskriterien		
E1	Population	<p>Erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie</p> <p><u>Begründung:</u> Die im Einschlusskriterium E1 definierte Population entspricht der Patientenpopulation, für die Nintedanib zugelassen ist. Eine weitere Differenzierung der Population war nicht notwendig, da die Festlegung der ZVT durch den G-BA für das gesamte Anwendungsgebiet zutrifft.</p>

Einschlusskriterien		
E2	Intervention	<p>Das zu bewertende Arzneimittel ist Nintedanib. Nintedanib wird in Kombination mit Docetaxel angewendet. In die Nutzenbewertung sind Studien einzuschließen, in denen Nintedanib gemäß Angaben in der Fachinformation gegeben wurde. Diese sieht eine zweimal tägliche Einnahme von 200 mg Nintedanib vor. Die beiden Einnahmen von Nintedanib sollen im Abstand von ungefähr 12 Stunden erfolgen, jeweils an Tag 2 bis 21 eines Standard-Docetaxel-Behandlungszyklus (bei Auftreten von Nebenwirkungen waren Dosisreduktionen nach Schema auf zweimal täglich 150 mg oder zweimal täglich 100 mg möglich). Der Standardbehandlungszyklus mit Docetaxel sieht eine Infusion alle drei Wochen mit einer empfohlenen Dosierung von 75mg/m² vor.</p> <p><u>Begründung:</u> Intervention gemäß Zulassungsstatus von Nintedanib</p>
E3	Vergleichstherapie	<p>Docetaxel gemäß Fachinformation.</p> <p><u>Begründung:</u> BI folgte den Empfehlungen des G-BA bezüglich der ZVT (siehe Modul 3 Abschnitt 3.1) und vergleicht sich gegen Docetaxel, da eine vergleichende Studie zu Docetaxel vorliegt (LUME-Lung 1). Es wird in diesem Dossier die beste verfügbare Evidenz dargestellt und der medizinische Zusatznutzen gegenüber Docetaxel abgeleitet.</p> <p>Dosierung und Anwendungsmodalitäten erfolgten gemäß Zulassung: Für die Behandlung nach einer vorausgegangenen, platinhaltigen Chemotherapie wird die Gabe von 75 mg/m² Docetaxel als Monotherapie empfohlen (einstündige Infusion, alle 3 Wochen).</p>
E4	Endpunkte	<p>Mindestens einer der patientenrelevanten Endpunkte wird berichtet:</p> <ul style="list-style-type: none"> • Mortalität <ul style="list-style-type: none"> ○ OS • Morbidität / Symptomatik <ul style="list-style-type: none"> ○ PFS ○ ORR ○ DCR ○ Tumorbedingte und therapieassoziierte Symptomatik • HRQoL • Nebenwirkungen <p><u>Begründung:</u> Zur Darstellung des Zusatznutzens soll nach § 35 Abs. 1b Satz 5 Sozialgesetzbuch (SGB) V, § 5 Abs. 2 Satz 3 AM-NutzenV und 5. Kapitel, 2. Abschnitt § 5 Abs. 2 Satz 3 Verfahrensordnung (VerfO) des G-BA (Bundesministerium der Justiz, 1988; Bundesministerium der Justiz, 2010; Gemeinsamer Bundesausschuss, 2014b) auf patientenrelevante Zielgrößen insbesondere im Hinblick auf</p> <ul style="list-style-type: none"> - Mortalität - Morbidität - Lebensqualität <p>Bezug genommen werden.</p>

Einschlusskriterien		
E5	Studientyp	RCT <u>Begründung:</u> „RCT und systematischen Übersichten von RCT wird im Rahmen von Therapiestudien der höchste Evidenzgrad zugeordnet“ (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013b) (Evidenzklassifizierung siehe 2. Kapitel, 3. Abschnitt § 11 Abs. 3 VerfO [Gemeinsamer Bundesausschuss, 2014b]).
E6	Publikationstyp	Vollpublikation, Studienbericht oder ausführlicher Ergebnisbericht aus einem Studienregister verfügbar in deutscher oder englischer Sprache. <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.
E7	Studiendauer	Keine Einschränkung
Ausschlusskriterien		
A1	Außerhalb der Zulassung	Anwendung des Wirkstoffs außerhalb des zugelassenen Anwendungsgebiets <u>Begründung:</u> „Für Arzneimittel mit neuen Wirkstoffen [...] wird ein Zusatznutzen für das jeweilige zugelassene Anwendungsgebiet nachgewiesen im Vergleich zu der nach § 6 bestimmten zweckmäßigen Vergleichstherapie auf der Grundlage von Unterlagen zum Nutzen des Arzneimittels in den zugelassenen Anwendungsgebieten.“ (Bundesministerium der Justiz, 2010)
A2	Präklinische Studien	Präklinische <i>in vitro</i> und <i>in vivo</i> Studien <u>Begründung:</u> „Der Nachweis einer therapeutischen Verbesserung erfolgt aufgrund der Fachinformationen und durch Bewertung von klinischen Studien nach den internationalen Standards der evidenzbasierten Medizin.“ (Bundesministerium der Justiz, 2010)

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 war es, RCT mit dem zu bewertenden Arzneimittel Nintedanib im Anwendungsgebiet zu identifizieren, die einen Vergleich von Nintedanib+Docetaxel zu einer Vergleichstherapie in Übereinstimmung mit den Empfehlungen des G-BA zur ZVT erlauben.

Die Suche erfolgte in folgenden Datenbanken:

- MEDLINE, MEDLINE In-Process

- EMBASE, EMBASE Daily Alerts
- Cochrane Central Register of Controlled Trials

Der Informationsstand bezieht sich auf den 16.10.2014. 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.

Die Studien wurden nach den in Tabelle 4-3 prädefinierten Ein- und Ausschlusskriterien selektiert. Die Selektion nach den einzelnen relevanten Endpunkten fand erst auf Volltextbasis statt. Es wurden nur Volltexte in deutscher oder englischer Sprache berücksichtigt.

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 Unternehmen 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.

Um sicherzustellen, dass alle verfügbaren Studiendaten mit dem zu bewertenden Arzneimittel Nintedanib für die frühe Nutzenbewertung berücksichtigt werden, wurde eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien in folgenden medizinischen Studienregistern durchgeführt:

- Clinicaltrials.gov
- International Clinical Trials Registry Platform (ICTRP) Search Portal, Suchportal der World Health Organization (WHO)
- EU Clinical Trials Register (EU-CTR)
- PharmNet.Bund

Der Informationsstand bezieht sich auf den 16.10.2014. Die einzelnen Suchstrategien für jede Datenbankabfrage sind in Anhang 4-B dokumentiert.

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 bibliographische Literaturrecherche (Abschnitt 4.2.3.2) und Studienregisterabfrage (Abschnitt 4.2.3.3) identifizierten Studien zu Nintedanib wurden gemäß den in Abschnitt 4.2.2 prädefinierten Kriterien zum Ein- und Ausschluss relevanter Studien (Tabelle 4-3) von zwei unabhängigen Personen selektiert. Auf Basis aller verfügbaren Informationen (Titel, Abstract und Volltext bzw. Studienregistereintrag und eventuell darin hinterlegte Dokumente) wurden Studien bei Erfüllung aller Kriterien als relevant eingestuft, oder bei Nichterfüllung eines oder mehrerer Kriterien begründet ausgeschlossen. Diskrepanzen im Selektionsprozess wurden durch Diskussion aufgelöst.

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 in die Nutzenbewertung eingeschlossenen Studien wurden hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung erfolgte gemäß den Vorgaben der Dossievorlage anhand der zur Verfügung stehenden Informationen aus dem Studienbericht und aus Vollpublikationen.

Die Bewertung der Aussagekraft der eingeschlossenen Studien erfolgte zunächst endpunktübergreifend, 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 (siehe Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene:

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

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

Gemäß G-BA Verfo 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.

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. Abweichungen davon wurden begründet.

Die Einstufung „hoch“ für das Verzerrungspotenzial eines Endpunktergebnisses führte nicht zum Ausschluss der Daten. Die Klassifizierung diente zur Einschätzung der Aussagekraft der Nachweise. 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 in 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 selektierten RCT entsprechend den 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 Zielgrößen 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 Patienten 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 Patienten gemacht. Studienabbrüche oder Ausschlüsse von Patienten 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 herangezogen.

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 Studienpopulation (Patienten mit einem NSCLC der Tumorphistologie Adenokarzinom) wurde durch die folgenden demografischen Daten sowie krankheitsspezifischen Charakteristika beschrieben:

Demografie

- Alter (als stetige und kategoriale Variable [<65 Jahre, ≥ 65 Jahre])
- Geschlecht (männlich, weiblich)

- Ethnie⁴ (asiatisch, nicht-asiatisch)
- Geografische Region (Asien, Europa, Südafrika)
- Rauchverhalten (Nieraucher, aktiver Raucher, Ex-Raucher)

Krankheitsspezifische Charakteristika

- Erkrankungsstadium bei Erstdiagnose (<IIIB/IV, IIIB, IV)
- Krankheitsstatus zur Baseline (lokal rezidiert oder metastasiert)
- Allgemeinzustand (ECOG [Eastern Cooperative Oncology Group]-PS [Performance Status] 0, 1⁵)
- Hirnmetastasen zur Baseline (Ja, Nein)
- Vorbehandlung mit Bevacizumab (Ja, Nein)
- Zeit seit Beginn der Erstlinien-Therapie (<9 Monate, ≥9 Monate)

Endpunkte

Im vorliegenden Dossier wurden die Ergebnisse der folgenden patientenrelevanten Endpunkte dargestellt:

- Mortalität
 - Gesamtüberleben (OS)
- Morbidität/Symptomatik
 - Progressionsfreies Überleben (PFS)
 - Objektive Ansprechrage (ORR)
 - Krankheitskontrollrate (DCR)
 - Tumorbedingte Symptomatik und therapieassoziierte Symptomatik gemessen anhand von validierten Messinstrumenten zur Erfassung genereller Symptome einer Krebserkrankung bzw. spezifischer Symptome bei Lungenkrebs (Husten, Dyspnoe, Schmerz [allgemein, in der Brust, in Arm und Schulter, an anderen Stellen], Mundschmerzen, Schluckbeschwerden, periphere Neuropathie, Haarausfall, Bluthusten, Fatigue, Übelkeit und Erbrechen, Schlaflosigkeit, Appetitverlust, Verstopfung, Diarrhö)

⁴ Im Studienbericht der LUME-Lung 1 Studie als „Race class“ definiert.

⁵ In der Gruppe der ECOG-PS 1-Patienten wurde fälschlicherweise im Nintedanib+Docetaxel-Arm ein Patient mit ECOG-PS 2 eingeschlossen.

- Gesundheitsbezogene Lebensqualität (HRQoL), gemessen anhand validierter Messinstrumente
(z. B. globaler Gesundheitsstatus, körperliche, kognitive, emotionale, soziale Funktion und Rollenfunktion)
- Nebenwirkungen
 - Unerwünschtes Ereignis (UE)
 - Schwerwiegendes unerwünschtes Ereignis (SUE)
 - UE der CTCAE-Grade ≥ 3
 - UE, die zum Therapieabbruch führten
 - Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$

Aus Sicht von BI stellen alle hier aufgeführten Zielgrößen patientenrelevante Endpunkte dar. Für eine hohe Validität dieser patientenrelevanten Endpunkte spricht das durchweg niedrige Verzerrungspotential (siehe 4.3.1.3). Die Endpunkte der Kategorien Morbidität (PFS, ORR, DCR) wurden zudem zentral und unabhängig beurteilt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b). Die Patientenrelevanz sowie die Operationalisierung der hier aufgeführten Endpunkte im Anwendungsgebiet des lokal fortgeschrittenen, metastasierten oder lokal rezidierten NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie, werden nachfolgend diskutiert und begründet.

Mortalität

Gesamtüberleben

Das Ziel einer palliativen Krebstherapie sollte immer die Verlängerung des Überlebens und die Verbesserung der Lebensqualität der Patienten sein (Peppercorn et al., 2011). Die Verlängerung des OS stellt einen direkten klinischen Nutzen für den Patienten dar (Pazdur, 2008). Neben der Symptomatik der Erkrankung, den unerwünschten Wirkungen der Therapie sowie der Lebensqualität ist OS ein vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (Rapid Report „Aussagekraft von Surrogatendpunkten in der Onkologie“) anerkannter patientenrelevanter Endpunkt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011).

In der LUME-Lung 1 Studie wurde OS definiert als die Zeit von der Randomisierung bis zum Tod, unabhängig von der Todesursache. Der Endpunkt OS ermöglicht damit eine direkte Beurteilung der Verlängerung des Überlebens und stellt einen patientenrelevanten Endpunkt dar.

Morbidität

Progressionsfreies Überleben

Bridges et al. zeigen in einer Patientenpräferenzstudie mit NSCLC-Patienten mittels einer wahlbasierten Conjoint-Analyse, dass aus Patientensicht eine Verlängerung des PFS als direkter Vorteil wahrgenommen wird (Bridges et al., 2012). In einem zweiten umfassenderen

Discrete Choice Experiment mit 211 NSCLC-Patienten im Stadium IV (Union for International Cancer Control [UICC] 7. Auflage) in Deutschland bestätigen Mühlbacher et al. diese Ergebnisse. Es zeigte sich, dass bei der Betrachtung von 7 Attributen aus Patientensicht die zwei wichtigsten Entscheidungskriterien für die Wahl einer Therapie die Zeit ohne Tumorprogress sowie die Behandlung von tumorassoziierten Symptomen sind (Mühlbacher und Bethge, 2014)⁶.

In einer deutschen Studie hatten Krebspatienten am meisten Angst vor der Progression des Tumors (Herschbach et al., 2004). Unabhängig davon, wie der Progress einer Erkrankung letztendlich gemessen wird, spürt der Patient die direkten Folgen. Lungenkrebspatienten müssen neben unspezifischen psychosozialen Belastungen (Konfrontation mit einer lebensbedrohlichen Diagnose; häufig radikale, langwierige und nebenwirkungsreiche Behandlung mit ungewissem Ausgang und gravierenden körperlichen wie seelischen Folgen; gegebenenfalls Krankheitsprogredienz) eine Reihe von krankheitsspezifischen Belastungen (ungünstige Prognose; Krankheitsprogress häufig nach kurzem therapiefreiem Intervall etc.) bewältigen; hinzu kommen schwerwiegende Symptome (Goeckenjan et al., 2010). Gelingt den Patienten die Bewältigung dieser häufig massiven Belastungen nicht, kommt es unter Umständen zu gravierenden psychischen Störungen (z. B. Depressionen und/oder Angst- und Panikstörungen) (Goeckenjan et al., 2010).

Die EMA erkennt das PFS neben dem OS als einen akzeptablen, eigenständigen und patientenrelevanten primären Endpunkt an (European Medicines Agency, 2005; European Medicines Agency, 2008; European Medicines Agency, 2012). Studien mit PFS als Endpunkt haben den Vorteil, dass sie im Vergleich zu OS-Studien eine geringere Anzahl an Studienpatienten und eine kürzere Nachbeobachtungszeit benötigen (Food and Drug Administration, 2007). Auch können PFS-Daten bereits früher erhoben werden als OS-Daten, da PFS das Tumorwachstum erfasst (Food and Drug Administration, 2007).

PFS wurde in der LUME-Lung 1 Studie definiert als Zeit von der Randomisierung bis zur Progression der Erkrankung oder bis zum Tod. Gemäß World Health Organization (WHO) ist die Krankheitsprogression eines von vier Kriterien, um die Veränderung eines Tumorleidens zu beschreiben (Miller et al., 1981). Diese Kriterien wurden von der RECIST (Response Evaluation Criteria In Solid Tumors)-Arbeitsgruppe entsprechend für die soliden Tumore definiert. Die Erhebung des PFS in der Studie LUME-Lung 1 erfolgte mithilfe bildgebender Verfahren gemäß der RECIST-Kriterien Version 1.0 (Therasse et al., 2000) basierend auf einer zentralen und unabhängigen Beurteilung.

Objektive Ansprechrare (ORR)/Krankheitskontrollrate (DCR)

Die objektive Ansprechrare und die Krankheitskontrollrate sind beide ein Maß der Kontrolle des Tumorwachstums. Tumoransprechen und Tumorstabilisierung sind eng mit der

⁶ Beide Studien widerlegten die Nullhypothese, dass die unerwünschten Effekte eine vergleichbare Relevanz wie tumorassoziierte Symptome oder PFS haben. Obwohl die Forschungsfrage in beiden Studien OS nicht berücksichtigte, ist von einer Umkehr der Ränge nicht auszugehen.

Progression der Erkrankung verbunden. Eine unmittelbare Kontrolle der Progression ist vor allem notwendig bei Tumoren, die lebenswichtige Organe bedrohen, und viele Patienten und ihre Behandler sehen daher eine schnelle und effektive Tumorkontrolle sowie das Verzögern der Progression des Tumors als Vorteil an, unabhängig von der Verlängerung des Gesamtüberlebens (Zhuang et al., 2009).

Die ORR ist als Endpunkt nur dann sinnvoll, wenn sie auf breit akzeptierten und einfach anzuwendenden Standardkriterien basiert (Eisenhauer et al., 2009). Daher sollte laut Committee for Medicinal Products for Human Use (CHMP) die ORR nach internationalen Standards definiert werden (z. B. RECIST-, Volumetric RECIST- oder WHO-Kriterien) (European Medicines Agency, 2012). In der LUME-Lung 1 Studie wurden die standardisierten und etablierten RECIST-Kriterien (Version 1.0) herangezogen und die ORR folgendermaßen operationalisiert:

- ORR: Anteil der Patienten, die gemäß zentraler und unabhängiger Beurteilung ein bestätigtes komplettes Ansprechen (Complete Response [CR]) oder ein bestätigtes partielles Ansprechen (Partial Response [PR]) erreicht haben. Die Bestätigung des CR oder PR musste mindestens 28 Tage nach dem ersten dokumentierten CR oder PR erfolgen.

Nach den RECIST-Kriterien Version 1.0 versteht man unter CR den vollständigen Rückgang aller Läsionen sowie die Normalisierung des Tumormarker-Levels (Therasse et al., 2000). Unter PR versteht man eine Reduktion des Tumors um mindestens 30% im Durchmesser im Vergleich zum Ausgangswert (Therasse et al., 2000).

Auch für die DCR wurden bei der LUME-Lung 1 Studie die RECIST-Kriterien Version 1.0 verwendet und folgendermaßen operationalisiert:

- DCR: Anteil der Patienten, die gemäß zentraler und unabhängiger Beurteilung ein bestätigtes CR, ein bestätigtes PR oder eine stabile Erkrankung (Stable Disease [SD]; keine signifikante Änderung der Tumorgroße) erreicht haben. Die Bestätigung des CR oder PR musste mindestens 28 Tage nach dem ersten dokumentierten CR oder PR erfolgen.

Neben dem CR und PR wird bei der DCR das Erreichen einer SD berücksichtigt. Hierunter versteht man nach RECIST-Kriterien (Version 1.0) den Zustand, bei dem weder das Schrumpfen des Tumors einem partiellen Ansprechen, noch das Wachstum des Tumors einer progressiven Erkrankung zugeordnet werden kann (Therasse et al., 2000). Generell erreichen nur wenige Patienten mit fortgeschrittenem NSCLC nach einer Standardchemotherapie ein Schrumpfen des Tumors, die meisten erfahren eine Stabilisierung der Erkrankung oder einen Progress der Erkrankung. Dementsprechend erwies sich bei Patienten mit fortgeschrittenem NSCLC die DCR als ein stärkerer Prädiktor für das Überleben als die ORR (Lara et al., 2008).

Messinstrumente der Symptomatik und gesundheitsbezogenen Lebensqualität

Die Symptomatik der Erkrankung sowie die HRQoL der Patienten können in klinischen Studien anhand von validierten Fragebögen zur Lebensqualität erfasst werden. Für die Erfassung der Symptomatik kamen in der LUME-Lung 1 Studie die Fragebögen QLQ-C30 und QLQ-LC13 der EORTC-Studiengruppe für Lebensqualität (EORTC Quality of Life Group [QLG]) zum Einsatz, und für die Erfassung der HRQoL der Fragebogen EORTC QLQ-C30 sowie der Fragebogen EQ-5D der internationalen Forschungsgruppe EuroQoL. Um die Antworten der Patienten nicht zu beeinflussen, wurden alle Fragebögen in LUME-Lung 1 von den Patienten vor dem Arztkontakt und bevor sie neue Informationen zu ihrem Krankheitsstatus erhielten, ausgefüllt (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a). So wurde insbesondere sichergestellt, dass der Patient nicht durch externe Faktoren, wie z. B. durch Kenntnis eines radiologischen Befunds, in seinen Angaben zur Lebensqualität beeinflusst wurde.

Fragebögen EORTC QLQ-C30 und EORTC QLQ-LC13

Die EORTC QLG hat ein modulares System entwickelt, basierend auf 1) dem Kernfragebogen EORTC QLQ-C30, der neben der gesundheitsbezogenen Lebensqualität generelle Symptome einer Krebserkrankung abfragt, sowie 2) krankheits- und behandlungsspezifischen Fragebögen (Aaronson et al., 1993; Bergman et al., 1994). Der für Lungenkrebs spezifisch entwickelte Fragebogen ist der EORTC QLQ-LC13 Zusatzfragebogen (Bergman et al., 1994).

Der EORTC QLQ-C30 besteht aus 30 Fragen. Er enthält sowohl Multi-Item Skalen als auch Einzel-Items:

- Eine Skala für die allgemeine Gesundheit und Lebensqualität (2 Items)
- Fünf Funktionsskalen: Körperliche Funktion (5 Items), kognitive Funktion (2 Items), emotionale Funktion (4 Items), soziale Funktion (2 Items), Rollenfunktion (2 Items)
- Drei Symptom-Skalen: Fatigue (3 Items), Schmerzen (2 Items), Übelkeit und Erbrechen (2 Items)
- Sechs Einzel-Items: Dyspnoe, Schlaflosigkeit, Appetitverlust, Verstopfung, Diarrhö und finanzielle Schwierigkeiten⁷ (Aaronson et al., 1993)

Der Patient gibt seine Einschätzung mittels einer kategorialen 4-Punkte-Skala von 1 (überhaupt nicht) bis 4 (sehr) an; mit Ausnahme der beiden Fragen zum globalen Gesundheitsstatus. Hier wird eine 7-Punkte Skala verwendet von 1 (sehr schlecht) bis 7 (exzellent). Diese Werte werden durch lineare Transformation in eine Skala von 0-100 übertragen. Hohe Werte bei der allgemeinen Gesundheit und Lebensqualität und den

⁷ Der Endpunkt finanzielle Schwierigkeiten ist für den deutschen Versorgungskontext nicht relevant, da in Deutschland die Mehrheit der Patienten eine Krankenversicherung besitzt, die die Kosten der Behandlung übernimmt. Er wird daher im Dossier nicht weiter adressiert.

Funktionsskalen deuten auf eine bessere Lebensqualität hin. Bezüglich einzelner Symptome hingegen sprechen hohe Werte für eine starke Ausprägung dieser Symptome (Aaronson et al., 1993). Die Validität und Reliabilität des EORTC QLQ-C30 wurde für Lungenkrebs (Aaronson et al., 1993) und andere Tumorarten bestätigt (Hjermstad et al., 1995; Groenvold et al., 1997). Generell ist der EORTC QLQ-C30 einer der meist verwendeten Fragebögen in randomisierten Studien der Onkologie (Velikova et al., 2012). Er wurde in zahlreichen klinischen Studien in der Indikation NSCLC (Braun et al., 2011; Teckle et al., 2011) sowie anderen onkologischen Indikationen verwendet (Maisey et al., 2002; Efficace et al., 2004; Fang et al., 2004).

Der EORTC QLQ-LC13 enthält 13 Fragen. Diese umfassen eine Multi-Item Skala für Dyspnoe und eine Reihe von Einzel-Items für Schmerz (in der Brust, in Arm und Schulter, an anderen Stellen), Husten, wunder Mund, Dysphagie, periphere Neuropathie, Haarausfall und Bluthusten. Die Auswertung des EORTC QLQ-LC13 erfolgt analog zum EORTC QLQ-C30 (Bergman et al., 1994). Der EORTC QLQ-LC13 wurde speziell für Patienten mit Lungenkrebs entwickelt und validiert (Bergman et al., 1994) und findet häufig in NSCLC-Studien Verwendung (Efficace et al., 2006; Larsson et al., 2012).

Innerhalb der EORTC QLQ-C30 Scores wird eine Differenz von mindestens 10 Punkten (auf der 0-100 Skala) als klinisch relevant auf Patientenebene angesehen (Osoba et al., 1998; Cocks et al., 2011). Diese Skalierung ist auch für den QLQ-LC13 Score anerkannt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013a). Die Auswertung der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 erfolgte daher anhand des Anteils der Patienten, die während des Studienverlaufs im Mittel eine Verbesserung von mindestens 10 Punkten gegenüber dem Ausgangswert über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben (Responderanalysen). Ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen ist somit gleichzeitig als klinisch relevant zu werten. Außerdem werden Analysen für die Zeit bis zur Verschlechterung präsentiert, die auf patientenindividuellen Schwellenwerten für klinisch relevante Verschlechterungen (Verschlechterung um mindestens 10 Punkte gegenüber dem Ausgangswert) basieren. Auch hier sind somit entsprechende Unterschiede zwischen den Behandlungsgruppen als klinisch relevant zu werten.

Diese beiden Auswertungen bilden unterschiedliche patientenrelevante Nutzenaspekte ab, die jeweils eine eigenständige Berechtigung haben. Die Responderanalysen zeigen die Verbesserung initial bereits vorhandener Symptome an. Dies ist aus Patientensicht ein primäres Therapieziel. Die Zeit bis zur Verschlechterung beschreibt die Verzögerung des Neuauftretens bzw. der Verschlechterung von Symptomen und ist eher mit dem Aufhalten des Fortschreitens der Krebserkrankung verbunden. Beide Analysemethoden sind vom IQWiG im Rahmen der frühen Nutzenbewertung von Crizotinib (Xalkori[®]) und Afatinib (Giotrif[®]) anerkannt worden (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013a; Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2014).

Fragebogen EQ-5D

Der EQ-5D wurde von der internationalen Forschungsgruppe EuroQoL entwickelt, um andere gesundheitsbezogene Messungen zu unterstützen, diese aber nicht zu ersetzen (Rabin und de Charro, 2001). Der Fragebogen setzt sich zusammen aus einem deskriptiven Teil und einer visuellen, analogen Skala (EQ-VAS) (EuroQoL Group, 2013):

- Der deskriptive Teil (EQ-5D self classifier) misst fünf Dimensionen mittels fünf Indikatoren: Beweglichkeit/Mobilität, Selbstversorgung, allgemeine Tätigkeiten, Schmerzen/körperliche Beschwerden und Angst/Niedergeschlagenheit. In der 3L-Version des Fragebogens, die auch bei der Studie LUME-Lung 1 zum Einsatz kam, ist jede dieser fünf Dimensionen in drei Level (keine Probleme, einige Probleme oder extreme Probleme) unterteilt. Die Patienten geben Auskunft über ihren Gesundheitszustand, indem sie in jeder der fünf Dimensionen die Kategorie ankreuzen, die ihre Situation am Tag der Erhebung des Fragebogens am besten beschreibt. Die EuroQoL weist jedoch explizit darauf hin, dass die für die Einzelindikatoren vergebenen Ausprägungen 1-3 nicht auf einer Kardinalskala basieren und daher nicht als metrische Variable interpretierbar sind. Daher werden im vorliegenden Dossier keine Ergebnisse für die einzelnen Dimensionen des Fragebogens vorgelegt. Stattdessen werden die üblichen Index-Werte (Utility Scores) verwendet, die sich durch die Aggregation aller fünf Dimensionen ergeben. Durch Kombination der drei Antwortmöglichkeiten über die fünf Dimensionen lassen sich insgesamt 243 Gesundheitszustände abbilden⁸. Jeder der 243 möglichen Gesundheitszustände kann einem Utility Score zugeordnet werden, der auf den Präferenzen der Allgemeinbevölkerung basiert. Die Berechnung des Utility Scores bei der LUME-Lung 1 Studie erfolgte anhand der britischen (EQ-5D UK Utility) und belgischen (EQ-5D Belgium Utility) Präferenzwerte. Da die Ergebnisse sehr ähnlich waren, werden in diesem Dossier die gängigeren UK Utilities präsentiert.
- Anhand der EQ-VAS können die Patienten ihren Gesundheitsstatus mittels einer Skalierung von 0 (schlechtester denkbarer Gesundheitszustand) bis 100 (bester denkbarer Gesundheitszustand) bewerten.

Der generische EQ-5D Fragebogen ist ein verlässliches und validiertes Messinstrument (Brazier et al., 1993; de Boer et al., 2004) mit guter Evidenz (Rabin und de Charro, 2001) und ist auch für Lungenkrebs validiert (Teckle et al., 2011; Schuette et al., 2012).

Symptomatik der Erkrankung

Die Symptomatik der Erkrankung wird vom IQWiG (Rapid Report „Aussagekraft von Surrogatendpunkten in der Onkologie“) als patientenrelevanter Endpunkt anerkannt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011). Der G-BA bestätigte diese

⁸ Der Status 11223 bedeutet z. B., dass der Patient keine Probleme bezüglich Beweglichkeit/Mobilität und Selbstversorgung, einige Probleme hinsichtlich allgemeiner Tätigkeiten und Schmerzen/körperlicher Beschwerden und extreme Probleme bezüglich Angst/Niedergeschlagenheit verspürt.

Einschätzung im Rahmen von Beschlüssen über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen (Gemeinsamer Bundesausschuss, 2012; Gemeinsamer Bundesausschuss, 2013b; Gemeinsamer Bundesausschuss, 2013a; Gemeinsamer Bundesausschuss, 2014c). Wie bereits zuvor erwähnt, kann die Symptomatik der Erkrankung anhand von validierten Fragebögen zur Lebensqualität erfasst werden. Zu diesen zählen z. B. der EORTC QLQ-LC13 oder die beiden nachfolgenden Teile des EORTC QLQ-C30-Fragebogens (Aaronson et al., 1993; Bergman et al., 1994):

- Drei Symptom-Skalen: Fatigue (3 Items), Schmerzen (2 Items), Übelkeit und Erbrechen (2 Items)
- Sechs Einzel-Items: Dyspnoe, Schlaflosigkeit, Appetitverlust, Verstopfung, Diarrhö und finanzielle Schwierigkeiten⁹ (Aaronson et al., 1993)

In der LUME-Lung 1 Studie wurde die Krankheitssymptomatik anhand dieser beiden Fragebögen (EORTC QLQ-LC13 und EORTC QLQ-C30) ermittelt. Die in den Fragebögen vom Patienten einzuschätzenden Symptome sind differenziert zu betrachten. Sie umfassen sowohl tumorbedingte als auch therapieassoziierte Symptome. Einzelne Symptome, wie Fatigue, können sowohl tumorbedingt als auch therapieassoziiert sein:

Tumorbedingte Symptome

Schmerzen, Dyspnoe, Fatigue, Bluthusten, Schlaflosigkeit, Husten

Therapieassoziierte Symptome

Appetitverlust, Diarrhö, Übelkeit und Erbrechen, Verstopfung, Fatigue, Haarausfall, Mundschmerzen, periphere Neuropathie, Schluckbeschwerden

Die Vorgehensweise bei der Auswertung dieser beiden Fragebögen wurde im Detail im oberen Abschnitt „Messinstrumente der Symptomatik und gesundheitsbezogenen Lebensqualität“ beschrieben. Hohe Werte sprechen für eine starke Ausprägung des Symptoms (Aaronson et al., 1993).

Bei Patienten mit fortgeschrittenem Lungenkrebs stellen die krankheitsspezifischen Symptome eine große Belastung dar, was sich wiederum negativ auf die Funktionalität und die Lebensqualität der Tumorerkrankten auswirken kann (McCannon und Temel, 2012). Dem Erhalt der Lebensqualität durch die Kontrolle von Symptomen kommt eine besondere Bedeutung zu (McCannon und Temel, 2012). Aufgrund dessen und der im Folgenden dargestellten Argumente werden folgende Symptome des NSCLC im Dossier als schwerwiegend eingestuft: Dyspnoe, Husten, Bluthusten, Schmerzen und Fatigue.

⁹ Der Endpunkt finanzielle Schwierigkeiten ist für den deutschen Versorgungskontext nicht relevant, da in Deutschland eine Krankenversicherungspflicht besteht, wodurch die Kosten einer Behandlung übernommen werden. Er wird daher im Dossier nicht weiter adressiert.

Dyspnoe

Dyspnoe (Atemnot bzw. die erschwerte Atmung) stellt ein komplexes, multidimensionales und subjektives Symptom dar. Physikalische, psychologische, emotionale und funktionelle Faktoren beeinflussen die Einschränkung der Atmung. Diese Faktoren sowie die subjektive Wahrnehmung des Patienten bestimmen den Schweregrad der Dyspnoe und die Beeinträchtigung der Lebensqualität (Goeckenjan et al., 2010).

Die Dyspnoe hat einen negativen Effekt auf die Lebensqualität, besonders in den physischen und psychischen Domänen (McCannon und Temel, 2012). In einer Studie mit 120 Patienten mit Lungenkarzinom aller Krankheitsstadien, in der 87% der Patienten von Dyspnoe betroffen waren, berichteten Patienten mit stark ausgeprägter Dyspnoe von einer signifikant geringeren Lebensqualität und nahmen außerdem Schmerzen und Angstzustände verstärkt wahr (Smith et al., 2001). Von 157 ambulant behandelten Patienten mit fortgeschrittenem Lungenkrebs berichteten 55% über Dyspnoe. Die Dyspnoe wirkte sich nicht nur negativ auf körperliche Aktivitäten wie Gehen oder Arbeiten aus, sondern auch auf psychische Aspekte wie die Stimmung und die Lebenslust (Tanaka et al., 2002). Außerdem zeigte eine Studie mit 105 Patienten mit fortgeschrittenem Lungenkarzinom, dass die Dyspnoe über die Zeit zunimmt, während die Lebensqualität sinkt (Henoch et al., 2007).

Husten und Bluthusten

Husten ist das am häufigsten beobachtete tumorbedingte Symptom, gefolgt von Dyspnoe, Brustschmerzen und Bluthusten (Hämoptyse) bzw. Auswurf (Expektoration) eines blutig tingierten Sputums (Goeckenjan et al., 2010). Wie bei der Dyspnoe ist die Ätiologie von Husten oftmals von mehreren Faktoren abhängig. Husten kann auf direkte (extrinsische oder intrinsische Obstruktionen der Atemwege) und indirekte Effekte (Pleuraerguss oder Pneumonitis aufgrund der Strahlen- oder Chemotherapie) des Tumors zurückgeführt werden, aber auch auf chronische Bedingungen wie z. B. Asthma, chronisch-obstruktive Lungenerkrankung (Chronic Obstructive Pulmonary Disease [COPD]) oder überlagerte, akute Krankheiten wie eine Pneumonie (McCannon und Temel, 2012). Husten ist mit Fatigue, Brustschmerzen, Kopfschmerzen, Übelkeit, Erbrechen und Inkontinenz assoziiert und hat außerdem einen negativen Effekt auf die Lebensqualität, vor allem auf psychosoziale Aspekte (French et al., 1998; Irwin et al., 2002). Obwohl Bluthusten viel seltener auftritt als Dyspnoe und Husten, kann Bluthusten beängstigend für die Patienten und deren Familien sein (McCannon und Temel, 2012).

Schmerzen

Bei pneumologischen Krankheitsbildern (z. B. NSCLC) zählen Schmerzen neben Dyspnoe und Husten zu den Leitsymptomen (Heigener et al., 2011). Schmerz ist ein körperliches Symptom, dessen Charakter und Ausmaß nur der Patient selbst beurteilen kann. Das Schmerzempfinden ist wesentlich durch die psychische Wahrnehmung und Verarbeitung geprägt. Bei Tumorpatienten sind Schmerzen überwiegend tumorbedingt (60%-90%), seltener therapiebedingt (10%-25%), tumorassoziiert (5%-20%) oder tumorunabhängig (3%-10%) (Goeckenjan et al., 2010). Der Schmerz tritt am häufigsten in der Brust (25%-67%) sowie dem Lendenwirbelbereich (10%-40%) auf (Krech et al., 1992; Mercadante et al., 1994; Grond

et al., 1996; Potter und Higginson, 2004). Schmerzen haben einen signifikant negativen Einfluss auf die von NSCLC-Patienten berichtete gesundheitsbezogene Lebensqualität (Iyer et al., 2013) und werden von Lungenkrebs-Patienten häufig als schwerwiegend wahrgenommen (Krech et al., 1992).

Fatigue

Fatigue ist eines der verbreitetsten und stark einschränkenden Symptome bei Krebspatienten (Hofman et al., 2007) und bisher nur wenig untersucht (Gutstein, 2001). Fatigue hat einen dramatischen Einfluss auf das Leben der Patienten und kann durch die Krankheit verursacht werden, durch deren Behandlung sowie durch tumorbedingten physiologischen und psychologischen Stress (Gutstein, 2001). Fatigue hat einen starken und direkt negativen Einfluss auf alle Lebensqualitätsaspekte (Meyerowitz et al., 1983; Padilla und Grant, 1985), insbesondere auf das körperliche Wohlbefinden der betroffenen Patienten (Devlen et al., 1987a; Devlen et al., 1987b; Rhodes et al., 1988). Die Definition der krebsvermittelten Fatigue variiert, sie beinhaltet ein subjektives Gefühl von Müdigkeit oder Erschöpfung, das durch den Krebs selbst oder die Behandlung verursacht wird, durch Ruhepausen nicht gelindert werden kann und dadurch den Alltag negativ beeinflusst (Portenoy, 2000). Krebsvermittelte Fatigue wird bei 50%-90% der Patienten berichtet und hat einen gravierenden Einfluss auf die Lebensqualität und die funktionale Belastbarkeit der Patienten (Campos et al., 2011). Die Fatigue wird aus der Sicht von NSCLC-Patienten als relevantes bzw. bedeutendes Behandlungsrisiko gesehen (Bridges et al., 2012; Mühlbacher und Bethge, 2014). Bei einer Umfrage bezüglich Fatigue, basierend auf einem Fragebogen, fühlten sich die Krebspatienten (6% Lungenkrebspatienten) durch Fatigue signifikant stärker beeinträchtigt als durch andere Symptome. Bei der Wahl zwischen Fatigue, Schmerz oder Übelkeit und Erbrechen empfanden diese Krebspatienten die Fatigue mit Abstand als das stärkste Problem, gefolgt von Schmerzen (Stone et al., 2000).

Weitere Symptome

Mit dem Fragebogen QLQ-C30 der EORTC QLG werden neben den zuvor beschriebenen Symptomen Schmerz, Dyspnoe und Fatigue die Symptome Übelkeit und Erbrechen, Appetitverlust, Schlafstörungen, Verstopfung und Diarrhö erfasst (Aaronson et al., 1993). Über den lungenkrebspezifischen Fragebogen EORTC QLQ-LC13 werden neben den bereits beschriebenen Symptomen Husten, Bluthusten, Dyspnoe und Schmerz (Brust, Arm und Schultern sowie an anderen Stellen) auch Nebenwirkungen konventioneller Chemo- und Strahlentherapien (Haarausfall, periphere Neuropathie, Mundschmerzen, Schluckbeschwerden) gemessen (Bergman et al., 1994).

Patientenpräferenz

Die Patientenpräferenzstudie von Bridges et al. zeigt, dass Patienten mit NSCLC eine Verlängerung des PFS zusammen mit Verbesserungen der Symptomschwere (Schmerzen, Husten und Dyspnoe) als am wichtigsten erachten (Bridges et al., 2012). Diese Ergebnisse bestätigte ein umfassenderes Discrete Choice Experiment: Aus Patientensicht waren bei der Betrachtung von 7 Attributen die zwei wichtigsten Entscheidungskriterien für die Wahl einer

Therapie die Zeit ohne Tumorprogress sowie die Behandlung von tumorbedingten Symptomen (Dyspnoe, Husten, Schmerz) (Mühlbacher und Bethge, 2014)¹⁰.

Zum Ende des Lebens hin tritt die Dyspnoe vermehrt auf (Currow et al., 2010). Ebenso nimmt die Häufigkeit von Husten, Brustschmerzen und Bluthusten im Verlauf der Erkrankung zu (Lutz et al., 2001). In der Patientenpräferenzstudie von Osoba et al. listeten die Patienten mit NSCLC (n=99) Übelkeit und Erbrechen (63%), Schmerzen (56%), emotionale Funktion (48%), Dyspnoe (43%) sowie die Rollenfunktion (43%) als die fünf wichtigsten Effekte, die es zu vermeiden gilt (Osoba et al., 2006). In der Gruppe der Patienten im späten Stadium der Erkrankung (n=74) stand die Dyspnoe an erster Stelle (60%) (Osoba et al., 2006). Bei einer Studie, bei der die Befragten die Schwere der Symptome (mild, moderat, schwer) selbst einstufen durften, waren die häufigsten und schwerwiegendsten Symptome Schmerzen, Dyspnoe und Appetitverlust (Krech et al., 1992).

Gesundheitsbezogene Lebensqualität

Die Beurteilung der Lebensqualität kann ein wichtiger Endpunkte für den klinischen Nutzen sein, insbesondere bei überwiegend symptomatischen Erkrankungen, wie dem NSCLC (Food and Drug Administration, 2011). Die Bestimmung der Lebensqualität ist mittlerweile als wichtiger Ergebnisparameter für die Tumorthherapie anerkannt (Plunkett et al., 2003). Neben dem OS, der Symptomatik der Erkrankung und den unerwünschten Wirkungen der Therapie ist die Lebensqualität ein vom IQWiG (Rapid Report „Aussagekraft von Surrogatendpunkten in der Onkologie“) anerkannter patientenrelevanter Endpunkt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011).

Die Messung der HRQoL erfolgte in der LUME-Lung 1 Studie mittels des Fragebogens EORTC QLQ-C30 sowie des Fragebogens EQ-5D. Die Bestandteile sowie die Auswertung des von der EORTC QLG entwickelten Fragebogens EORTC QLQ-C30 sowie des von der internationalen Forschungsgruppe EuroQoL entwickelten Fragebogens EQ-5D wurden im Detail im oberen Abschnitt „Messinstrumente der Symptomatik und gesundheitsbezogenen Lebensqualität“ erläutert.

Nebenwirkungen

Die Erfassung von UE sollte bei einer Studie generell durchgeführt werden (European Medicines Agency, 1996). Die Daten der UE dienen unter anderem zum Nachweis des Nutzens einer Therapie. Die Verringerung von Nebenwirkungen zählt zu den Aspekten der therapiebedingten Morbidität (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013b) und ist ein vom IQWiG anerkannter Endpunkt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011).

¹⁰ Beide Studien widerlegten die Nullhypothese, dass die unerwünschten Effekte eine vergleichbare Relevanz wie tumorassoziierte Symptome oder PFS haben. Obwohl die Forschungsfrage in beiden Studien OS nicht berücksichtigte, ist von einer Umkehr der Ränge nicht auszugehen.

Im Rahmen dieser Nutzenbewertung wurden neben der Gesamtrate der UE¹¹ auch SUE¹², UE vom CTCAE-Grad ≥ 3 nach Version 3.0¹³ (National Cancer Institute, 2006) sowie UE, die zum Therapieabbruch führten¹⁴, berücksichtigt. Zusätzlich wurden die jeweils einzelnen UE mit einem CTCAE-Grad ≥ 3 betrachtet, die in $\geq 5\%$ der Patienten eines Behandlungsarms auftraten¹⁵. Aufgrund der Zugehörigkeit der Ereignisse febrile Neutropenie, Neutropenie und Verminderung der neutrophilen Granulozyten zum medizinischen Konzept Neutropenie wird zusätzlich die Special Search Category (SSC) „Neutropenie“ dargestellt, die alle entsprechenden Ereignisse umfasst. Die Laborparameter „weiße Blutkörperchen“ und „Alanin-Aminotransferasen (ALT)“ werden im vorliegenden Dossier nicht separat dargestellt, da durch frühzeitiges Monitoring ein adäquates Nebenwirkungsmanagement möglich ist (siehe Modul 3 Abschnitt 3.4.4; [Boehringer Ingelheim Pharma GmbH & Co. KG, 2014a]).

Zu den im Dossier betrachteten Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ zählen somit: Diarrhö, Dyspnoe, febrile Neutropenie, Neutropenie (PT [Preferred Term] sowie SSC) und Verminderung der neutrophilen Granulozyten.

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 Effekten¹⁶ 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

¹¹ In der LUME-Lung 1 Studie operationalisiert als Zeit von der Randomisierung bis zum ersten Auftreten eines UE. Es gingen alle Ereignisse unter Behandlung ein, wobei Ereignisse innerhalb von 28 Tagen nach Ende der Behandlung auch zu den Ereignissen unter Behandlung gezählt wurden.

¹² In der LUME-Lung 1 Studie operationalisiert als Zeit von der Randomisierung bis zum ersten Auftreten eines SUE

¹³ In der LUME-Lung 1 Studie operationalisiert als Zeit von der Randomisierung bis zum ersten Auftreten eines UE mit CTCAE-Grad ≥ 3

¹⁴ In der LUME-Lung 1 Studie operationalisiert als die Zeit von der Randomisierung bis ersten Auftreten eines UE, das zum Therapieabbruch führte

¹⁵ Dieses Vorgehen wurde analog zur IQWiG-Bewertung von Afatinib (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2014) umgesetzt.

¹⁶ DerSimonian R, Laird N. *Meta-analysis in clinical trials. Control Clin Trials* 1986;7(3):177-188.

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 verwendet¹⁷ 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 Heterogenität¹⁸ 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.

Für die vorliegende Bewertung des Zusatznutzens von Nintedanib in Kombination mit Docetaxel liegt eine RCT vor. Somit sind Meta-Analysen nicht vorgesehen. Auf eine Beschreibung meta-analytischer Methodik wird daher verzichtet.

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.

¹⁷ 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.

¹⁸ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. *Measuring inconsistency in meta-analyses*. *BMJ* 2003;327(7414):557-560.

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.

Sensitivitätsanalysen dienen zur Überprüfung der Robustheit der Ergebnisse sowie der Schlussfolgerung gegenüber der Variation methodischer Faktoren. Sensitivitätsanalysen können die Robustheit der Resultate der Einzelstudien im Hinblick auf methodische Variation, wie z. B. das Ersetzen fehlender Werte, untersuchen. Dabei ist man auf die Verfügbarkeit solcher Analysen aus den Einzelstudien angewiesen.

Die in der eingeschlossenen Studie durchgeführten und im Studienbericht referenzierten Sensitivitätsanalysen wurden dargestellt. Hierbei handelt es sich um Sensitivitätsanalysen für den Endpunkt OS.

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. Diese wurden entweder aus bereits vorliegenden Ergebnissen aus Regressionsmodellen mit Interaktionstermen oder aus den Ergebnissen von Subgruppenanalysen berechnet. Subgruppenanalysen wurden basierend auf individuellen Patientendaten durchgeführt. Für kategoriale Variablen wurden die Interaktionstests mittels Cochran Q-Test berechnet. Bei einer zur Berechnung des RR unzureichenden Zellbesetzung von null Ereignissen, wurde eine Korrektur um 0,5 vorgenommen, damit eine Berechnung des RR möglich war.

Ein p-Wert $\geq 0,05$ und $< 0,2$ im Interaktionstest wurde als Hinweis auf unterschiedliche Effekte zwischen den Subgruppen gewertet. Die Subgruppenergebnisse wurden in einer Tabelle präsentiert. 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 (multiples Testen, Power, Präspezifizierung, qualitative versus quantitative Interaktion) diskutiert. Falls kein Hinweis auf eine signifikante Interaktion zwischen Behandlung und dem subgruppenbildenden Faktor bestand, wurde keine nach Subgruppen getrennte Analyse präsentiert. Bei Vorliegen eines signifikanten Gesamtschätzers und signifikanter Interaktion wurden die entsprechenden Ergebnisse diskutiert (z. B. hinsichtlich biologischer Plausibilität oder methodischer Aspekte wie Stichprobengröße und dem Problem des multiplen Testens).

Geplante Subgruppenanalysen waren in der LUME-Lung 1 Studie auf die Endpunkte PFS und OS beschränkt.

Für die vorliegende Nutzenbewertung wurden folgende Subgruppen für alle patientenrelevanten Endpunkte betrachtet, um zu beurteilen, ob der Therapieeffekt über die Subgruppen hinweg konstant ist oder ob Effektmodifikationen vorliegen:

- Alter (< 65 Jahre, ≥ 65 Jahre)
- Geschlecht (männlich, weiblich)
- Geografische Region (Asien, Europa, Südafrika)
- ECOG-PS zur Baseline (ECOG 0, 1)

- Hirnmetastasen zur Baseline (Ja, Nein)
- Vorbehandlung mit Bevacizumab (Ja, Nein)
- Raucherstatus (Nieraucher, aktiver Raucher/Ex-Raucher)
- Zeit seit Beginn der Erstlinien-Therapie (<9 Monate, ≥9 Monate)

Dabei handelt es sich um die laut VerFO vorgeschriebenen Subgruppenanalysen bezüglich Alter, Geschlecht, Schweregrad und geografischer Region (Länder und Zentren werden abgebildet durch die Regionen) sowie die a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen. Die Subgruppenanalysen für Ethnie (asiatisch, nicht-asiatisch) werden nicht zusätzlich dargestellt, da diese inhaltlich bereits über die Subgruppenanalysen für geografische Region (Asien, Europa, Südafrika) abgedeckt sind.

Da die finale Analyse für den wichtigsten sekundären Endpunkt OS einer hierarchischen Teststrategie folgte, wurde zudem die Zeit seit Beginn der Erstlinien-Therapie in den hier vorliegenden Analysen betrachtet (siehe 4.3.1.2.1).

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich¹⁹. 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“²⁰, „Multiple-Treatment-Meta-Analysen“²¹ oder auch „Netzwerk-Meta-Analysen“²², 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

¹⁹ 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.

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

²¹ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

²² Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17(3): 279-301.

Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist²³.

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 untersuchen²⁴.

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

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

Ein indirekter Vergleich war nicht notwendig, da eine direkt vergleichende Studie vorlag. Auf eine Beschreibung der Methodik indirekter Vergleiche wird daher verzichtet.

²³ B. Schöttker, D. Lüthmann, D. Boukhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.

²⁴ 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.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

Zusammenfassung des Abschnitts 4.3

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Der medizinische Nutzen von Nintedanib in Kombination mit Docetaxel wurde mit der Zulassung durch die Europäische Arzneimittelbehörde bestätigt. Nachfolgend wird der medizinische Zusatznutzen von Nintedanib+Docetaxel gegenüber Docetaxel bei erwachsenen Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie, beschrieben.

Endpunkte

Gesamtüberleben

Für Patienten mit Adenokarzinom war das Gesamtüberleben (OS) unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0359$). Das Sterberisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 17%, verglichen mit Patienten, die mit Placebo+Docetaxel behandelt wurden (HR [Hazard Ratio] [95%-KI (Konfidenzintervall)]: 0,83 [0,70; 0,99]). Dies entsprach einer Verlängerung des medianen OS um 2,3 Monate von 10,3 auf 12,6 Monate.

Progressionsfreies Überleben

Progressionsfreies Überleben (PFS) war der primäre Endpunkt der LUME-Lung 1 Studie. Für Patienten mit Adenokarzinom war das PFS unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0485$). Das Progressionsrisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 16%, verglichen mit Patienten, die mit Placebo+Docetaxel behandelt wurden (HR [95%-KI]: 0,84 [0,71; 1,00]). Dies entsprach einer Verlängerung des medianen PFS um 1,4 Monate von 2,8 auf 4,2 Monate. Der PFS-Vorteil zugunsten der Kombinationstherapie war bereits zum Zeitpunkt der primären PFS-Analyse nachweisbar (medianes PFS: 4,0 Monate im Nintedanib+Docetaxel-Arm vs. 2,8 Monate im Placebo+Docetaxel-Arm; HR [95%-KI]: 0,77 [0,62; 0,96]; $p=0,0193$).

Objektives Ansprechen

Hinsichtlich der ORR lag kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen vor ($p=0,4823$). Das RR (Relatives Risiko) [95%-KI] belief sich auf 1,30 [0,62; 2,74], das OR (Odds Ratio) [95%-KI] auf 1,32 [0,61; 2,86] und die RD (Risikodifferenz) [95%-KI] auf 1,09 [-1,95; 4,13].

Krankheitskontrollrate

Die DCR war für Patienten mit der Kombinationstherapie Nintedanib+Docetaxel statistisch signifikant besser im Vergleich zur Placebo+Docetaxel-Therapie ($p<0,0001$). Im Nintedanib+Docetaxel-Behandlungsarm erreichten 60,2% der Patienten eine

Krankheitskontrolle, im Placebo+Docetaxel-Arm nur 44,0% (RR [95%-KI]: 1,37 [1,18; 1,59] / OR [95%-KI]: 1,93 [1,41; 2,63] / RD [95%-KI]: 16,20 [8,67; 23,73]).

Symptomatik

Nachfolgend wird nur auf Symptome mit statistisch signifikanten Ergebnissen zwischen den Vergleichsarmen eingegangen. Hierzu zählen Schmerzen, Diarrhö sowie Übelkeit und Erbrechen. Für alle anderen Symptome lagen keine statistisch signifikanten Unterschiede vor.

Schwerwiegende Symptome

Schmerzen

Bei der Responderanalyse zur Verbesserung der Symptomatik, erhoben mit dem Fragebogen EORTC QLQ-C30, ergab sich ein statistisch signifikanter Vorteil zugunsten der Kombinationstherapie Nintedanib+Docetaxel ($p=0,005$). Das RR [95%-KI] lag bei 1,56 [1,13; 2,14], das OR [95%-KI] bei 1,76 [1,18; 2,62] und die RD [95%-KI] bei 9,52 [2,84; 16,20]. Bei Betrachtung der Zeit bis zur Verschlechterung des Symptoms Schmerzen (EORTC QLQ-C30) lag kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen vor.

Die Zeit bis zur Verschlechterung der Schmerzen in Arm/Schulter (EORTC QLQ-LC13) war unter der Therapie mit Nintedanib+Docetaxel signifikant länger als unter der Therapie mit Placebo+Docetaxel (HR [95%-KI]: 0,80 [0,63; 1,00]; $p=0,0456$). Bezüglich der Responderanalyse zur Verbesserung der Symptomatik von Schmerzen in Arm/Schulter, gemessen mit EORTC QLQ-LC13, zeigte sich kein statistisch signifikanter Effekt.

Weitere Symptome

Diarrhö

Das Symptom Diarrhö wird über den Fragebogen EORTC QLQ-C30 erfasst. Für die Verbesserung der Symptomatik ergab sich kein statistisch signifikanter Effekt. Bei Betrachtung der Zeit bis zur Verschlechterung der Symptomatik zeigte sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (HR [95%-KI]: 1,90 [1,54; 2,34]; $p<0,0001$).

Übelkeit und Erbrechen

Das Symptom Übelkeit und Erbrechen wurde über den Fragebogen EORTC QLQ-C30 gemessen. Für die Verbesserung der Symptomatik ergab sich kein statistisch signifikanter Effekt. Bezüglich der Zeit bis zur Verschlechterung der Symptomatik zeigte sich für Übelkeit und Erbrechen ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (HR [95%-KI]: 1,23 [1,00; 1,50]; $p=0,0467$). Empfehlungen für geeignete Maßnahmen und proaktives Therapiemanagement finden sich in der Fach- und Gebrauchsinformation für Nintedanib wieder (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Gesundheitsbezogene Lebensqualität

Nachfolgend wird nur auf die statistisch signifikanten Ergebnisse zwischen den

Vergleichsarmen eingegangen. Diese umfassen die Daten zur Verbesserung des globalen Gesundheitsstatus sowie der kognitiven Funktion, die beide mittels des Fragebogens EORTC QLQ-C30 erhoben wurden. Alle anderen Skalen des EORTC QLQ-C30 sowie die Daten, die anhand des EQ-5D Fragebogens erhoben wurden, wiesen keine statistisch signifikanten Unterschiede auf.

Globaler Gesundheitsstatus

Es zeigte sich bei der Verbesserung der HRQoL für den globalen Gesundheitsstatus ein statistisch signifikanter Vorteil für die Kombinationstherapie Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel ($p=0,026$). Das RR [95%-KI] lag bei 1,56 [1,05; 2,33], das OR [95%-KI] bei 1,69 [1,06; 2,69] und die RD [95%-KI] bei 6,64 [0,82; 12,45]. Für die Operationalisierung Zeit bis zur Verschlechterung lag kein statistisch signifikanter Unterschied vor.

Kognitive Funktion

Es zeigte sich bei der Verbesserung der HRQoL bezüglich der kognitiven Funktion ein statistisch signifikanter Vorteil für Nintedanib in Kombination mit Docetaxel gegenüber der Vergleichstherapie mit Placebo+Docetaxel ($p=0,009$). Es ergaben sich ein RR [95%-KI] von 1,64 [1,12; 2,41], ein OR [95%-KI] von 1,81 [1,15; 2,84] und eine RD [95%-KI] von 8,00 [2,00; 14,00]. Die Zeit bis zur Verschlechterung unterschied sich in den beiden Behandlungsgruppen nicht signifikant.

Nebenwirkungen

Im Rahmen dieser Nutzenbewertung wurden unerwünschten Ereignisse (UE), die schwerwiegenden unerwünschten Ereignisse (SUE), die UE vom Common Terminology Criteria for Adverse Events (CTCAE)-Grad ≥ 3 nach Version 3.0 (National Cancer Institute, 2006) sowie die UE, die zum Therapieabbruch führten, berücksichtigt. Zusätzlich wurden jeweils die Nebenwirkungen einzeln betrachtet, welche UE mit einem CTCAE-Grad ≥ 3 umfassen, die in $\geq 5\%$ der Patienten eines Behandlungsarms auftraten (siehe Abschnitt 4.2.5.2). Insgesamt unterschieden sich die Zeiten bis zum ersten Auftreten der patientenrelevanten Nebenwirkungen der mit Nintedanib+Docetaxel behandelten Patienten nicht signifikant von denen der mit Placebo+Docetaxel behandelten Patienten:

- Alle UE: HR [95%-KI]: 1,10 [0,94; 1,29]; $p=0,2067$
- Alle SUE: HR [95%-KI]: 1,01 [0,78; 1,32]; $p=0,9316$
- Alle UE mit CTCAE-Grad ≥ 3 : HR [95%-KI]: 1,10 [0,92; 1,32]; $p=0,2655$
- UE, die zum Therapieabbruch führten: HR [95%-KI]: 1,08 [0,76; 1,54]; $p=0,6557$
- Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$: Keine signifikanten Unterschiede in den einzelnen UEs (siehe Tabelle 4-54)

Fazit

Die Behandlung mit Nintedanib+Docetaxel für erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie stellt eine wirksame und im Allgemeinen gut verträgliche Therapieoption dar. Mit einer Verlängerung des medianen Gesamtüberlebens um 2,3 Monate auf 12,6 Monate ist mit Nintedanib in Kombination mit Docetaxel als Therapieoption für Adenokarzinom-Patienten ein relevanter Fortschritt in der Zweitlinien-Therapie des fortgeschrittenen NSCLC erreicht worden. Die Daten der LUME-Lung 1 Studie zum wichtigsten sekundären Endpunkt OS konnten erstmals seit fast 10 Jahren einen signifikanten Überlebensvorteil im vorliegenden Anwendungsgebiet zeigen.

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

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
LUME-Lung 1	ja	ja	Laufende Studie, Rekrutierung abgeschlossen	Beginn: 23. Dezember 2008 1. Datenschnitt: 02. November 2010 (Primäranalyse PFS) 2. Datenschnitt: 15. Februar 2013 (finale OS-Analyse)	Nintedanib+Docetaxel vs. Placebo+Docetaxel
LUME-Lung 2	ja - nur supportiv	ja	Laufende Studie, Rekrutierung abgeschlossen	Beginn: 23. Dezember 2008	Nintedanib+Pemetrexed vs. Placebo+ Pemetrexed
NCT02231164	nein	ja	Laufende Studie	voraussichtlich ~42 Monate Beginn: Oktober 2014	Nintedanib+Docetaxel vs. Placebo+Docetaxel

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

Der Informationsstand bezieht sich auf den 16.10.2014.

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

Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
LUME-Lung 2	Kombination Nintedanib+Pemetrexed nicht zugelassen
NCT02231164	Laufende Studie, keine (Zwischen-)Ergebnisse verfügbar

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

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

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

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

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

Wie in Abschnitt 4.2.3.2 beschrieben, wurde eine systematische bibliografische Literaturrecherche nach RCT, die einen direkten Vergleich mit Nintedanib+Docetaxel und einer Vergleichstherapie in Übereinstimmung mit den G-BA Empfehlungen zur ZVT (Gemeinsamer Bundesausschuss, 2014d) ermöglichen, durchgeführt.

Die Recherche wurde in den Datenbanken MEDLINE, MEDLINE In-Process, EMBASE, EMBASE Daily Alerts und Cochrane Central Register of Controlled Trials mithilfe der Plattform Ovid durchgeführt. Die Suchstrategien in der jeweiligen Datenbank sind in Anhang 4-A abgebildet. Die Suchen in den Datenbanken fanden am 16.10.2014 statt.

Die Suchergebnisse wurden anhand der in Abschnitt 4.2.2, Tabelle 4-3 prädefinierten Ein- und Ausschlusskriterien von zwei Personen unabhängig voneinander hinsichtlich ihrer

Relevanz selektiert. Publikationen, die beide Personen als potenziell relevant erachteten, wurden anhand ihres Volltextes auf Relevanz geprüft. Zitate, die nur eine Person als potenziell relevant einstufte, wurden von beiden Personen nochmals gesichtet und anschließend nach Diskussion entweder als irrelevant ausgeschlossen oder ebenfalls anhand ihres Volltextes auf Relevanz geprüft.

Für die Suche nach RCT mit Nintedanib im Anwendungsgebiet ergab sich eine Trefferzahl von insgesamt 107 Zitaten. Nach automatisiertem Ausschluss der Duplikate (n=12) wurden die verbleibenden 95 Publikationen gemäß den Ein- und Ausschlusskriterien selektiert.

Basierend auf den Informationen aus Titel und Abstract wurden 86 Publikationen als nicht relevant eingestuft und aus der weiteren Betrachtung ausgeschlossen. Für die neun verbliebenen Publikationen, deren Titel oder Abstract die Einschlusskriterien erfüllten, wurde im Anschluss der Volltext hinsichtlich dieser Kriterien überprüft.

Im Ergebnis entsprachen acht Publikationen nicht den gesetzten Anforderungen und wurden begründet ausgeschlossen (Anhang 4-C). Somit ergab die bibliografische Literaturrecherche nach RCT mit Nintedanib+Docetaxel eine relevante Publikation (Reck et al., 2014a). Diese Publikation bezieht sich auf eine klinische Studie (Tabelle 4-7).

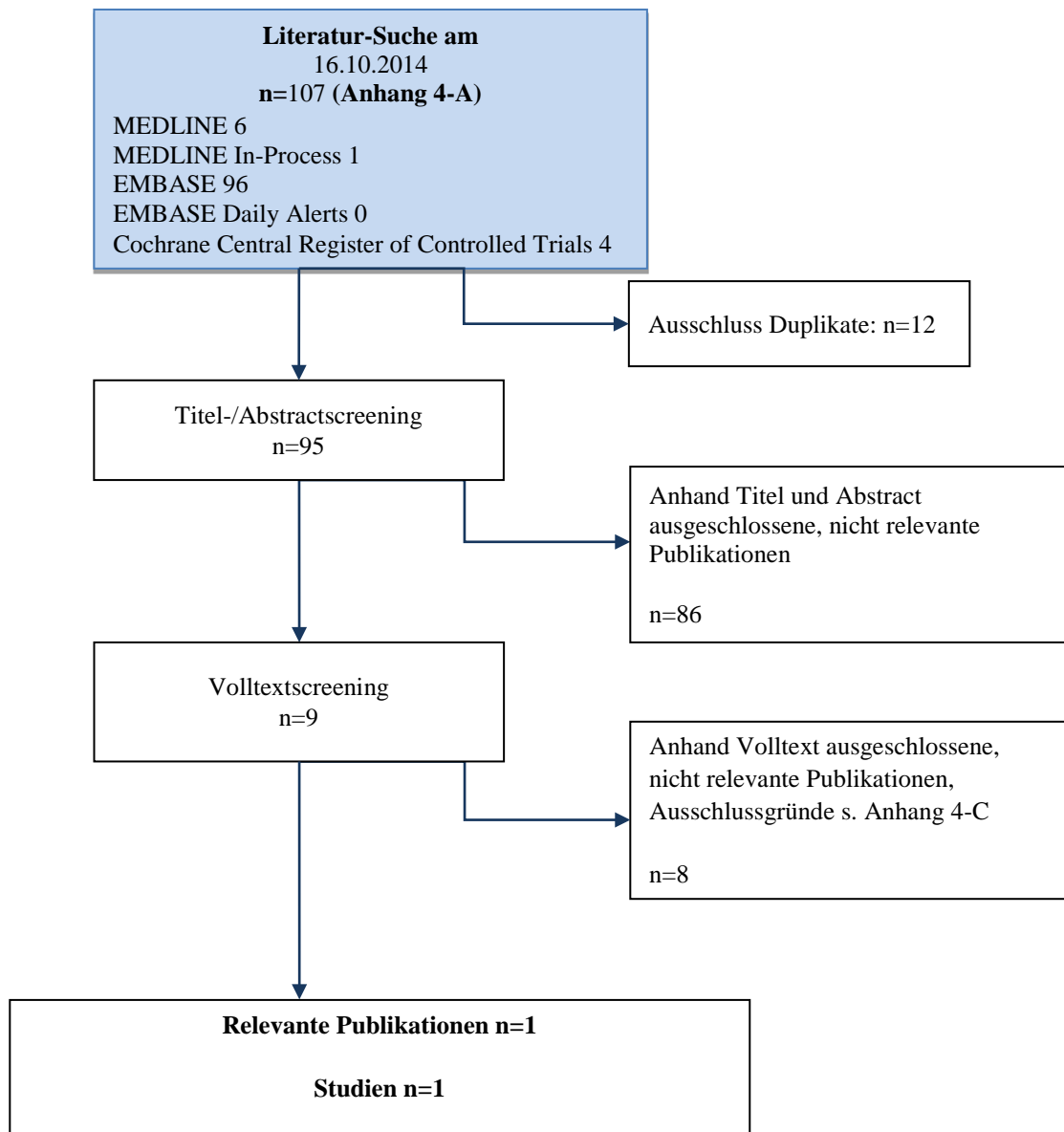


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach Studien mit der dem zu bewertenden Arzneimittel Nintedanib

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 4-4) 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 4-6: 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)
LUME- Lung 1	ClinicalTrials.gov (NCT00805194) (ClinicalTrials.gov, 2014b) WHO-ICTRP (EUCTR2007-004803-36- DE) (World Health Organization - International Clinical Trial Registry Platform, 2014a) EU-CTR (2007-004803-36) (EU Clinical Trials Register, 2014) PharmNet.bund (2007- 004803-36 - 4034498) (PharmNet.bund, 2014)	ja	ja: (Reck et al., 2014a)	Laufende Studie, Rekrutierung abgeschlossen. Finale OS Analyse abgeschlossen.
NCT022311 64	ClinicalTrials.gov (NCT02231164) (ClinicalTrials.gov, 2014a) WHO-ICTRP (NCT02231164) (World Health Organization - International Clinical Trial Registry Platform, 2014b)	ja	nein	Laufende Studie
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-6 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Der Informationsstand bezieht sich auf den 16.10.2014.

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 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
-	-	-	-	-	-	-
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
LUME-Lung 1	ja	ja	nein	ja: Studienbericht Primäre PFS-Analyse (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b) Studienbericht Finale OS-Analyse (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a)	ja: (ClinicalTrials.gov, 2014b; EU Clinical Trials Register, 2014; PharmNet.bund, 2014; World Health Organization - International Clinical Trial Registry Platform, 2014a)	ja: (Reck et al., 2014a)
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 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.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

Die Charakterisierung (Studiendesign, Studienpopulation, Studienverlauf) der eingeschlossenen Studie LUME-Lung 1 wird in Tabelle 4-8 dargestellt. Daran schließt sich die Charakterisierung der Interventionen (Tabelle 4-9) an.

Die Studienpopulation der LUME-Lung 1 Studie umfasste 658 randomisierte Patienten mit Adenokarzinom, die der Zulassungspopulation entsprechen. Die Charakterisierung in Tabelle 4-10 bezieht sich daher ausschließlich auf diese Zulassungspopulation. Die 658 Adenokarzinom-Patienten der LUME-Lung 1 Studie verteilen sich im Randomized Set (RS) auf 322 Patienten im Nintedanib+Docetaxel-Arm und 336 Patienten im Placebo+Docetaxel-Arm.

Tabelle 4-8: Charakterisierung der eingeschlossenen Studie LUME-Lung 1 – 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
LUME- Lung 1	RCT, doppelblind, parallel, placebokontrolliert, Phase-III-Studie, multizentrisch, multinational	Erwachsene NSCLC-Patienten im Stadium IIIB/IV oder Patienten mit rezidiertem NSCLC nach einer Erstlinien- Chemotherapie	3-wöchentlicher Behandlungszyklus <u>Nintedanib+Docetaxel:</u> Intravenöse Infusion von Docetaxel (75 mg/m ²) an Tag 1; Nintedanib (200 mg b.i.d.) an Tag 2 bis Tag 21 n=655 (Tumorhistologie Adenokarzinom: n=322) <u>Placebo+Docetaxel:</u> Intravenöse Infusion von Docetaxel (75 mg/m ²) an Tag 1; Placebo (b.i.d.) an Tag 2 bis Tag 21 n=659 (Tumorhistologie Adenokarzinom: n=336)	Voraussichtlich ~72 Monate <u>Run-in:</u> Keine Run-in Phase; die Behandlung startete nach dem Screening mit Tag 1 <u>Behandlung:</u> Nintedanib+Docetaxel Behandlung bis Krankheitsprogress oder inakzeptables UE Placebo+Docetaxel Behandlung bis Krankheitsprogress oder inakzeptables UE <u>Nachbeobachtung:</u> Follow-up-Visiten alle 6 bis 8 Wochen nach Behandlungsende bis zum Tod bzw. bis „lost- to-follow-up“. Alle Folgetherapien mussten dokumentiert werden.	<u>Ort:</u> Asien, Europa, Südafrika <u>Zeitraum:</u> Laufende Studie, Rekrutierung abgeschlossen Beginn: 23. Dezember 2008 1. Datenschnitt: 02. November 2010 (Primäranalyse PFS) 2. Datenschnitt: 15. Februar 2013 (finale OS- Analyse)	<u>Primärer Endpunkt:</u> PFS <u>Patientenrelevante sekundäre Endpunkte:</u> OS, ORR, DCR, Symptomatik (EORTC QLQ-C30, EORTC QLQ-LC13), HRQoL (EORTC QLQ-C30, EQ-5D), UE

b.i.d.: lat. *bis in die* (zweimal täglich); DCR: Disease Control Rate (Krankheitskontrollrate); EORTC: European Organization for Research and Treatment of Cancer; HRQoL: Health Related Quality of Life (gesundheitsbezogene Lebensqualität); n: Anzahl der Patienten; NSCLC: Non-Small Cell Lung Cancer (nicht-kleinzelliges Lungenkarzinom); ORR: Objective Response Rate (objektive Ansprechrate); OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); QLQ-C30: Quality of Life Questionnaire Core-30; QLQ-LC 13: Quality of Life Questionnaire Lung Cancer-13; RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie); UE: Unerwünschtes Ereignis.

Bei der LUME-Lung 1 Studie handelt es sich um eine zweiarmige, randomisierte, doppelblinde, placebokontrollierte Phase-III-Studie im Parallelgruppen-Design zum Vergleich der Wirksamkeit und Sicherheit von Nintedanib+Docetaxel gegenüber Placebo+Docetaxel bei NSCLC-Patienten im Stadium IIIB/IV oder Patienten mit rezidiviertem NSCLC nach einer Erstlinien-Chemotherapie (Tabelle 4-8). Die multinationale, multizentrische Studie wurde an 211 Zentren in 27 Ländern (Asien, Europa, Südafrika) durchgeführt.

Eingeschlossen wurden erwachsene NSCLC-Patienten im Stadium IIIB/IV oder Patienten mit rezidiviertem NSCLC nach einer Erstlinien-Chemotherapie. Weitere Einschlusskriterien waren ein ECOG-PS 0-1, eine Lebenserwartung von mindestens 3 Monaten sowie eine vorliegende Einverständniserklärung nach International Conference on Harmonisation-Good Clinical Practice (ICH-GCP). Ausgeschlossen wurden Patienten mit mehr als einer vorherigen palliativen Chemotherapie zur Behandlung des fortgeschrittenen und/oder metastasierten oder rezidivierten NSCLC. Alle weiteren Ausschlusskriterien sind im Anhang 4-E aufgelistet.

Die Randomisierung erfolgte in Blöcken und stratifiziert nach ECOG-PS (0 vs. 1), Histologie (Plattenepithelkarzinom vs. Nicht-Plattenepithelkarzinom), Hirnmetastasen (Ja vs. Nein) sowie nach vorheriger Bevacizumab-Behandlung (Ja vs. Nein).

Der primäre Endpunkt der LUME-Lung 1 Studie war das PFS nach den RECIST-Kriterien (Version 1.0), basierend auf einer zentralen und unabhängigen Beurteilung. Die sekundären Endpunkte waren: OS, ORR, DCR, Symptomatik (EORTC QLQ-C30, EORTC QLQ-LC13), HRQoL (EORTC QLQ-C30, EQ-5D) sowie UE.

Es wurden zwei Hauptanalysen durchgeführt: Die erste Hauptanalyse erfolgte auf Basis des 1. Datenschnittes vom 02. November 2010 (Primäranalyse PFS). Zu diesem Zeitpunkt wurde eine Zwischenauswertung des wichtigsten sekundären Endpunkts OS durchgeführt. Die zweite Hauptanalyse erfolgte auf Basis des 2. Datenschnittes vom 15. Februar 2013 (finale OS-Analyse). Zu diesem Zeitpunkt wurde auch eine Follow-up Analyse des PFS durchgeführt. In der finalen OS-Analyse wurde das Alpha-Niveau adjustiert, um der Zwischenauswertung für OS zum Zeitpunkt der primären PFS-Analyse Rechnung zu tragen. Die finale OS-Analyse folgte zudem einer hierarchischen Teststrategie: (1) Testung in der Population der Adenokarzinom-Patienten mit Zeit seit Beginn der Erstlinien-Therapie <9 Monate, (2) Testung in der Population aller Adenokarzinom-Patienten, (3) Testung in der Population aller Patienten der LUME-Lung 1 Studie. Die Grundlage für diese Teststrategie wurde aus den Daten der LUME-Lung 2 Studie abgeleitet (Kaiser et al., 2013).

Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+ Docetaxel	Placebo+ Docetaxel	Weitere Behandlungscharakteristika
LUME-Lung 1	Nintedanib (200 mg, b.i.d.) + Docetaxel (75 mg/m ²)	Placebo (b.i.d.) + Docetaxel (75 mg/m ²)	<p>Die Patienten erhielten entweder Nintedanib 200 mg oral zweimal täglich oder oral gegebenes Placebo zweimal täglich an den Tagen 2-21 jedes 21-tägigen Behandlungszyklus.</p> <p>75 mg/m² Docetaxel wurde als einstündige Infusion am 1. Tag jedes 21-tägigen Behandlungszyklus verabreicht.</p> <p>Um sowohl Häufigkeit und Schweregrad von Flüssigkeitsretentionen als auch den Schweregrad von Überempfindlichkeitsreaktionen zu reduzieren, erhielten alle Patienten eine Prämedikation mit einem oralen Kortikosteroid. Hierzu konnte z. B. Dexamethason über 3 Tage, beginnend einen Tag vor der Gabe von Docetaxel, in einer Dosierung von 16 mg pro Tag (z. B. 8 mg zweimal täglich), verabreicht werden.</p> <p>Die Behandlungszyklen wurden so lange wiederholt, bis eine Krankheitsprogression auftrat, die Therapie nicht mehr vertragen wurde, keiner der prädefinierten Therapieabbruchgründe erfüllt war, oder entweder der Patient oder der Prüfarzt einen Abbruch der Therapie mit der Studienmedikation forderten.</p> <p>Beim Auftreten folgender präspezifizierter medikamentenbedingter UE konnte die Dosis von Nintedanib bzw. Placebo in maximal zwei Schritten reduziert werden, z. B. von 200 mg zweimal täglich auf 150 mg zweimal täglich im ersten Schritt und von 150 mg zweimal täglich auf 100 mg zweimal täglich im zweiten Schritt:</p> <ul style="list-style-type: none"> • Erhöhung von AST- und/oder ALT-Werten vom CTCAE-Grad 2 in Verbindung mit einem Anstieg des Gesamtbilirubins vom CTCAE-Grad ≥ 1 <u>oder</u> Erhöhung von AST- und/oder ALT-Werten vom CTCAE-Grad ≥ 3 • Diarrhö vom CTCAE-Grad 2 an mehr als 7 aufeinanderfolgenden Tagen trotz Behandlung der Diarrhö • Diarrhö vom CTCAE-Grad ≥ 3 trotz Behandlung der Diarrhö • Erbrechen vom CTCAE-Grad ≥ 2 und/oder Übelkeit vom CTCAE-Grad ≥ 3 trotz einer antiemetischen Behandlung • Nicht-hämatologische Nebenwirkungen vom CTCAE-Grad ≥ 3 <p>Für Docetaxel war, entsprechend der Fachinformation, bei bestimmten präspezifizierten medikamentenbedingten UE eine Dosisreduktion, z. B. von 75 auf 60 mg/m², erlaubt.</p> <p>Nach einer vorangegangenen Dosisreduktion konnte keine Dosiserhöhung der Studienmedikation mehr erfolgen.</p> <p>Musste die Therapie mit Docetaxel abgebrochen werden, konnten die Patienten mit einer Monotherapie mit Nintedanib bzw. Placebo weiterbehandelt werden, sofern sie mindestens 4 Zyklen der Kombinationstherapie ohne Progression der Erkrankung durchlaufen hatten und keiner der prädefinierten Therapieabbruchgründe erfüllt war. Musste die Behandlung von Nintedanib bzw. Placebo abgebrochen werden, konnten die Patienten, die weiterhin für eine Docetaxel-Therapie geeignet waren, Docetaxel als Monotherapie erhalten. Für die Monotherapien galten eigene prädefinierte Dosisreduktionsschemata.</p>

ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; b.i.d.: *bis in die* (zweimal täglich); CTCAE: Common Terminology Criteria for Adverse Events; UE: Unerwünschtes Ereignis.

Die Intervention in der LUME-Lung 1 Studie bestand aus der Kombinationstherapie Nintedanib+Docetaxel sowie einer Placebo+Docetaxel-Vergleichstherapie im Kontrollarm. Die Standarddosierung von Nintedanib entsprach 200 mg (b.i.d. [lat. *bis in die* {zweimal täglich}]), die gemäß einem Dosis-Anpassungsplan zunächst auf 150 mg (b.i.d.) und im Falle einer zweiten Dosisreduktion auf 100 mg (b.i.d.) reduziert werden konnte. Die Medikation im Kontrollarm der LUME-Lung 1 Studie war Placebo+Docetaxel. Docetaxel wurde als intravenöse Infusion von 75 mg/m² an Tag 1 des 21-tägigen Behandlungszyklus verabreicht. Für Docetaxel war, entsprechend der Fachinformation, eine Dosisreduktion auf 60 mg/m², erlaubt. Die Dosierung von Docetaxel im Placebo+Docetaxel Kontrollarm erfolgte analog zur Dosierung im Nintedanib+Docetaxel-Behandlungsarm.

Sowohl das Dosierungsschema für Nintedanib als auch das Dosierungsschema für Docetaxel in der LUME-Lung 1 Studie entsprachen den aktuell für Deutschland gültigen Fachinformationen (Aventis Pharma S.A., 2014; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Wie aus Tabelle 4-10 ersichtlich ist, waren in der LUME-Lung 1 Studie beide Behandlungsarme hinsichtlich Demografie sowie krankheitsspezifischer Patientencharakteristika ausgeglichen.

Der Anteil der Patienten unter 65 Jahren betrug 72,0% in der Nintedanib+Docetaxel-Gruppe bzw. 71,4% in der Placebo+Docetaxel-Gruppe. Die Patienten im Nintedanib+Docetaxel-Arm waren durchschnittlich 58,5 ($\pm 10,1$) Jahre alt und die Patienten im Placebo+Docetaxel-Arm 58,6 ($\pm 9,5$) Jahre alt (Mittelwert [\pm STD {Standard Deviation, Standardabweichung}]). 37,0% der Nintedanib+Docetaxel-Patienten waren weiblich. In der Gruppe der Placebo+Docetaxel-Patienten betrug der Anteil der weiblichen Patienten 38,1%. Für einen Vergleich mit der aktuellen Versorgungssituation in Deutschland wurden aktuelle Daten des Tumorregisters Lungenkarzinom (Datenstand Januar 2014) betrachtet. Hierbei zeigt sich, ein im Vergleich zur LUME-Lung 1 Studie vergleichbarer Altersdurchschnitt (Mittelwert \pm STD: 64,6 $\pm 10,3$ Jahre) und ein etwas geringerer Frauenanteil (29,8%) bei den Patienten, die in der Zweitlinie palliativ versorgt werden (Tessen, 2014).

Der überwiegende Teil der Studienpopulation war nicht-asiatisch (Nintedanib+Docetaxel: 79,5% vs. Placebo+Docetaxel: 76,5%) und wurde in europäischen Zentren eingeschlossen (Nintedanib+Docetaxel: 71,1% vs. Placebo+Docetaxel: 69,6%). Da die Empfehlungen der europäischen Leitlinie (Reck et al., 2014b) den Empfehlungen der deutschen Leitlinien (Goekenjan et al., 2010; Griesinger et al., 2012) entsprechen, ist eine Übertragbarkeit der Daten der LUME-Lung 1 Studienpopulation auf die deutsche Versorgungssituation gegeben.

Zusammenfassend ist daher eine Übertragbarkeit der Studiendaten gegeben, da die Patientenpopulation der Studie der deutschen Versorgungsrealität entspricht.

Tabelle 4-10: Charakterisierung der Studienpopulationen in der LUME-Lung 1 Studie (Patienten mit Adenokarzinom, RS) – RCT

	Nintedanib+Docetaxel (N=322)	Placebo+Docetaxel (N=336)
Alter in Jahren		
Mittelwert (STD)	58,5 (10,1)	58,6 (9,5)
Median (min-max)	60,0 (29-80)	59,0 (30-80)
Altersgruppen: n (%)		
<65 Jahre	232 (72,0)	240 (71,4)
≥65 Jahre	90 (28,0)	96 (28,6)
Geschlecht: n (%)		
Weiblich	119 (37,0)	128 (38,1)
Männlich	203 (63,0)	208 (61,9)
Ethnie: n (%)		
Nicht-asiatisch	256 (79,5)	257 (76,5)
Asiatisch ¹	66 (20,5)	79 (23,5)
Geografische Region: n (%)		
Europa ²	229 (71,1)	234 (69,6)
Asien ³	86 (26,7)	96 (28,6)
Südafrika ⁴	7 (2,2)	6 (1,8)
Rauchverhalten: n (%)		
Nieraucher	115 (35,7)	115 (34,2)
Ex-Raucher	151 (46,9)	162 (48,2)
Aktiver Raucher	56 (17,4)	59 (17,6)
ECOG-PS: n (%)		
0	96 (29,8)	99 (29,5)
1 ⁵	226 (70,2)	237 (70,5)
Erkrankungsstadium bei Erstdiagnose: n (%)		
UICC 6. Edition	154 (47,8)	158 (47,0)
Stadium IV	86 (26,7)	103 (30,7)
Stadium IIIB	34 (10,6)	22 (6,5)
Stadium <IIIB/IV	34 (10,6)	33 (9,8)
UICC 7. Edition	166 (51,6)	178 (53,0)
Stadium IV	129 (40,1)	134 (39,9)
Stadium IIIB	21 (6,5)	23 (6,8)
Stadium <IIIB/IV	16 (5,0)	21 (6,3)
Stadium fehlend	2 (0,6)	0 (0)

	Nintedanib+Docetaxel (N=322)	Placebo+Docetaxel (N=336)
Krankheitsstatus zur Baseline: n (%)		
Lokal rezidiert ⁶	22 (6,8)	16 (4,8)
Metastasiert ⁷	300 (93,2)	320 (95,2)
Hirnmetastasen⁸: n (%)		
Ja	26 (8,1)	23 (6,8)
Nein	296 (91,9)	313 (93,2)
Bevacizumab-Vorbehandlung: n (%)		
Ja	24 (7,5)	21 (6,3)
Nein	298 (92,5)	315 (93,8)
Zeit seit Beginn der Erstlinien- Therapie: n (%)		
<9 Monate	206 (64,0)	199 (59,2)
≥9 Monate	112 (34,8)	134 (39,9)
Fehlende Angabe	4 (1,2)	3 (0,9)
ECOG: Eastern Cooperative Oncology Group; n: Anzahl der Patienten; PS: Performance Status (Allgemeinzustand); STD: Standard Deviation (Standardabweichung); UICC: Union for International Cancer Control.		
¹ Die Kategorie „asiatisch“ umfasst auch die Kategorie „amerikanische Indianer/Ureinwohner Alaskas“; ² Belgien, Bulgarien, Dänemark, Deutschland, Finnland, Frankreich, Georgien, Griechenland, Israel, Italien, Kroatien, Litauen, Österreich, Polen, Portugal, Rumänien, Russland, Schweden, Schweiz, Slowakische Republik, Spanien, Tschechische Republik, Ukraine, Vereinigtes Königreich, Weißrussland; ³ China, Indien, Korea; ⁴ Südafrika; ⁵ Inklusiv eines Patienten im Nintedanib+Docetaxel-Arm, der zur Baseline ein ECOG-PS von 2 aufwies; ⁶ Basierend auf der Angabe „Lokal rezidiert ohne Metastasen zur Baseline“ im Case Report Form (CRF); ⁷ Basierend auf der Angabe „Metastasen vorhanden zur Baseline“ im CRF; ⁸ Hirnmetastasen zur Baseline entsprechend der Dokumentation im CRF		

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

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

Nicht zutreffend.

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 4-11: 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			
LUME-Lung 1	Ja	Ja	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei der Phase-III-Studie LUME-Lung 1 handelt es sich um eine multizentrische, randomisierte, doppelblinde, placebokontrollierte Vergleichsstudie. Die Randomisierungssequenz wurde adäquat (computergestützt) generiert. Die Gruppenzuteilung erfolgte verdeckt mithilfe eines interaktiven Sprach- oder Webdialogsystems (IVRS [Interactive Voice Response System] oder IWRS [Interactive Web Response System]). Sowohl die Patienten als auch die behandelnden Personen waren über die gesamte Studiendauer verblindet. Sonstige, das Verzerrungspotenzial beeinflussende, Faktoren wurden nicht identifiziert. Das Verzerrungspotenzial auf Studienebene wird für die LUME-Lung 1 Studie daher insgesamt als niedrig eingestuft.

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 4-12: Matrix der Endpunkte in den eingeschlossenen RCT

Studie	OS	PFS	ORR	DCR	Symptomatik	HRQoL	Nebenwirkungen
LUME-Lung 1	•	• ¹	•	•	•	•	•
<p>¹ LUME-Lung 1: Für das PFS wurde sowohl der 1. Datenschnitt vom 02. November 2010 (Primäre PFS-Analyse) als auch der 2. Datenschnitt vom 15. Februar 2013 (finale OS-Analyse) dargestellt. Für die Bewertung des Zusatznutzens wird der 2. Datenschnitt vom 15. Februar 2013 verwendet. Die Ergebnisse zu allen anderen Endpunkten basieren auf dem 2. Datenschnitt vom 15. Februar 2013.</p> <p>DCR: Disease Control Rate (Krankheitskontrollrate); HRQoL: Health Related Quality of Life (gesundheitsbezogene Lebensqualität); ORR: Objective Response Rate (objektive Ansprechrage); OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben).</p>							

Verwendete Daten

Die im Folgenden dargestellten Daten stellen die Ergebnisse der Studie LUME-Lung 1 für die Zulassungspopulation der Adenokarzinom-Patienten dar. Es werden die Ergebnisse für alle patientenrelevanten Endpunkte (siehe Abschnitt 4.2.5.2) präsentiert (Tabelle 4-12).

Grundsätzlich basieren alle im Folgenden präsentierten Daten auf dem 2. Datenschnitt vom 15. Februar 2013 und wurden dem Studienbericht der finalen OS-Analyse entnommen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a) oder zusätzlich für dieses Dossier berechnet (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014e).²⁵ Für die Analysen zur gesundheitsbezogenen Lebensqualität gemäß EQ-5D und EQ-VAS wurde der Health Economic Report als Quelle verwendet (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d; Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c). Für den primären Endpunkt PFS wird zusätzlich der 1. Datenschnitt vom 2. November 2010 berichtet, da dieser dem Zeitpunkt der primären PFS-Analyse entspricht. Die PFS-Daten der primären Analyse der Adenokarzinom-Patienten wurden dem Biomarker-Report entnommen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013c). Die Bewertung des Zusatznutzens sowie die Subgruppenanalysen zum PFS basieren ausschließlich auf dem späteren, und damit umfangreicheren, 2. Datenschnitt vom 15. Februar 2013.

4.3.1.3.1 Endpunkte – 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

²⁵ Die Zusatzanalysen zur LUME-Lung 1 Studie enthalten ebenfalls die Kaplan-Meier-Kurven zur Verschlechterung der Symptomatik bzw. der gesundheitsbezogenen Lebensqualität, gemessen anhand des Fragebogens EORTC QLQ-C30 und EORTC QLQ-LC13, sowie zu den Analysen der UE.

- 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.

Im Folgenden werden die Operationalisierung, das Verzerrungspotenzial und die Ergebnisse jedes einzelnen patientenrelevanten Endpunkts in einem separaten Abschnitt betrachtet.

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.

4.3.1.3.1.1 Gesamtüberleben – RCT

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

Tabelle 4-13: Operationalisierung von Gesamtüberleben

Studie	Operationalisierung
LUME-Lung 1	OS war definiert als die Zeit von der Randomisierung bis zum Tod, unabhängig von der Todesursache.
OS: Overall Survival (Gesamtüberleben).	

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 4-14: Bewertung des Verzerrungspotenzials für Gesamtüberleben 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das OS zählt aufgrund seiner eindeutigen und klaren Bestimmbarkeit zu den harten klinischen Endpunkten, da es verlässlich und unabhängig von subjektiven Einschätzungen durch Patienten oder andere Beurteilende erhoben werden kann. Alle randomisierten Patienten wurden bei der Analyse des OS berücksichtigt, die Zuordnung zu den Behandlungsgruppen in der Analyse erfolgte gemäß Randomisierung (Analyse Set: RS). Somit wurde das ITT-Prinzip adäquat umgesetzt. Die Analyse dieses Endpunkts erfolgte gemäß dem vorab erstellten TSAP (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013b). Sonstige, das Verzerrungspotenzial beeinflussende Punkte lagen nicht vor, und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Da außerdem das Verzerrungspotenzial auf Studienebene niedrig ist und die Folgetherapien in beiden Behandlungsgruppen ausgeglichen waren, ist für den Endpunkt OS von einem niedrigen Verzerrungspotenzial auszugehen.

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

Die Ergebnisse zum OS aus der Studie LUME-Lung 1 sind in Tabelle 4-15 dargestellt.

Tabelle 4-15: Ergebnisse zum Gesamtüberleben aus RCT

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median [95%-KI] (Monate)	N	n (%)	Median [95%-KI] (Monate)	HR [95%-KI]	p-Wert
LUME-Lung 1	322	259 (80,4)	12,6 [10,6;15,1]	336	276 (82,1)	10,3 [8,6;12,2]	0,83 [0,70; 0,99]	0,0359
LUME-Lung 1 ^a (Sensitivitätsanalyse)	322	259 (80,4)		336	276 (82,1)		0,83 [0,70; 0,98]	0,0295
LUME-Lung 1 ^b (Sensitivitätsanalyse)	322	259 (80,4)		336	276 (82,1)		0,81 [0,69; 0,97]	0,0186
LUME-Lung 1 ^c (Sensitivitätsanalyse)	322	259 (80,4)		336	276 (82,1)		0,80 [0,68; 0,95]	0,0123

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a In dieser Sensitivitätsanalyse wurden die Stratifizierungsvariablen als Kovariaten berücksichtigt

^b In dieser Sensitivitätsanalyse wurde zusätzlich zu den Stratifizierungsvariablen für die Summe der größten Durchmesser der Zielläsionen adjustiert

^c In dieser Sensitivitätsanalyse wurde zusätzlich zu den Stratifizierungsvariablen für die Zeit seit Beginn der Erstlinien-Therapie adjustiert

HR: Hazard Ratio aus (stratifizierter) Cox-Regression (Stratifizierungsvariablen: ECOG-PS zur Baseline, Bevacizumab-Vorbehandlung, Hirnmetastasen zur Baseline); KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl verstorbener Patienten; RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Zum Zeitpunkt der finalen OS-Analyse waren 82,1% der Patienten in der Placebo+Docetaxel-Gruppe und 80,4% der Patienten in der Nintedanib+Docetaxel-Gruppe verstorben (Tabelle 4-15). In der stratifizierten Analyse war das Gesamtüberleben der Patienten mit Adenokarzinom unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0359$). Das Sterberisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 17 Prozent, verglichen mit Patienten, die mit Placebo+Docetaxel behandelt wurden (HR [95%-KI]: 0,83 [0,70; 0,99]). Dies entsprach einer Verlängerung des medianen OS um 2,3 Monate von 10,3 auf 12,6 Monate.

Die Sensitivitätsanalyse, in der die drei Stratifikationsvariablen als Kovariaten ins Modell aufgenommen wurden, bestätigte den Überlebensvorteil von Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel (HR [95%-KI]: 0,83 [0,70; 0,98]; $p=0,0295$). Ähnlich fielen die Ergebnisse von zwei weiteren Sensitivitätsanalysen aus, in denen zusätzlich entweder für die Summe der größten Durchmesser der Zielläsionen (HR [95%-KI]: 0,81 [0,69; 0,97]; $p=0,0186$) oder für die Zeit seit Beginn der Erstlinien-Therapie (HR [95%-KI]: 0,80 [0,68; 0,95]; $p=0,0123$) adjustiert wurde.

Die im Folgenden dargestellten Kaplan-Meier-Kurven beschreiben den statistisch signifikanten Überlebensvorteil der Nintedanib+Docetaxel-Therapie gegenüber der Placebo+Docetaxel-Vergleichstherapie (Abbildung 4-2).

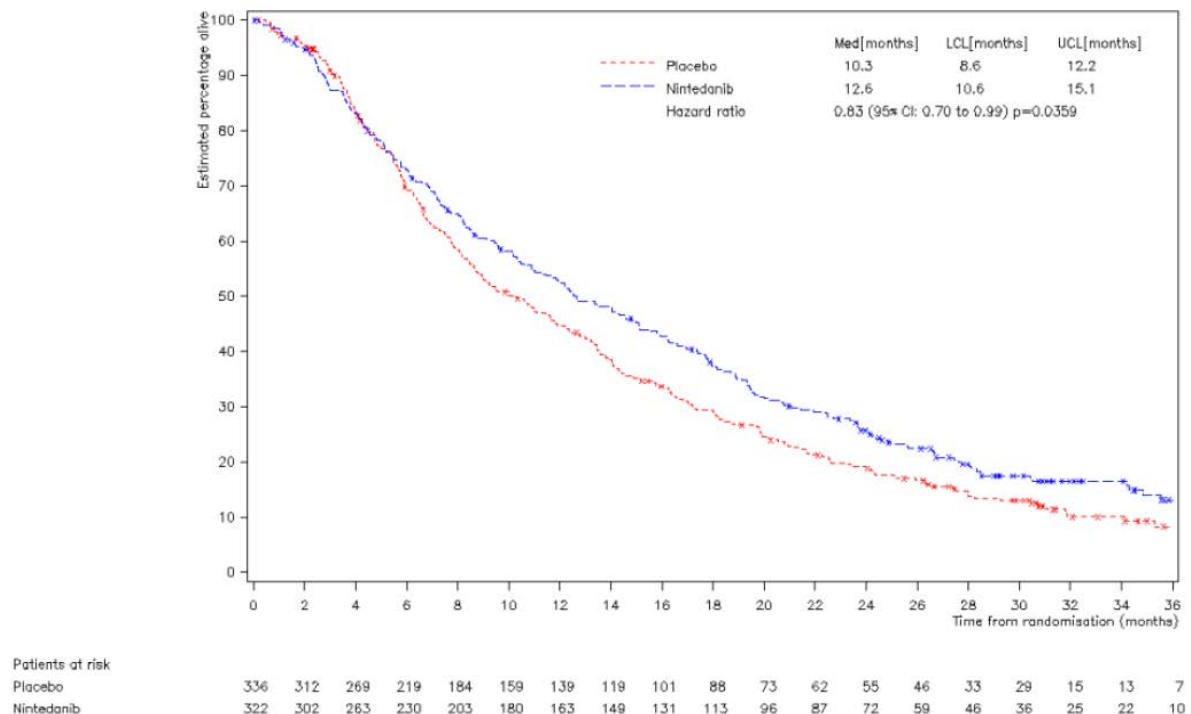


Abbildung 4-2: Kaplan-Meier-Kurven zum Gesamtüberleben (LUME-Lung 1 Studie, 2. Datenschnitt)

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.2 Progressionsfreies Überleben – RCT

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

Tabelle 4-16: Operationalisierung progressionsfreies Überleben

Studie	Operationalisierung
LUME-Lung 1	Das PFS war definiert als die Zeit von der Randomisierung bis zur Progression der Erkrankung oder bis zum Tod. Die Progression der Erkrankung wurde mithilfe bildgebender Verfahren anhand der RECIST-Kriterien Version 1.0 zentral und unabhängig beurteilt.
PFS: Progression Free Survival (progressionsfreies Überleben); RECIST: Response Evaluation Criteria In Solid Tumors.	

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 4-17: Bewertung des Verzerrungspotenzials für progressionsfreies Überleben 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

PFS war der primäre Endpunkt der LUME-Lung 1 Studie. PFS ist definiert als die Zeit (in Monaten) zwischen dem Datum der Randomisierung und dem Datum der Progression der Erkrankung oder dem Tod, falls ein Patient vorher verstarb. Die objektive Krankheitsprogression wurde verblindet durch eine unabhängige, radiologische Beurteilung (Independent Radiologic Review [IRR]) anhand der RECIST-Kriterien Version 1.0 festgestellt. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt, denn es wurden alle randomisierten Patienten (Analyse Set: RS) berücksichtigt und entsprechend der Zuordnung gemäß Randomisierung analysiert. Die Analyse dieses primären Endpunkts erfolgte gemäß dem vorab erstellten TSAP (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011). Sonstige, das Verzerrungspotenzial beeinflussende, Punkte existieren nicht, und es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung. Da außerdem das Verzerrungspotenzial auf Studienebene niedrig ist, ist für den Endpunkt PFS von einem niedrigen Verzerrungspotenzial auszugehen.

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

Die Ergebnisse zum PFS aus der Studie LUME-Lung 1 sind in Tabelle 4-18 dargestellt.

Tabelle 4-18: Ergebnisse für progressionsfreies Überleben aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median [95%-KI] (Monate)	N	n (%)	Median [95%-KI] (Monate)	HR [95%-KI]	p-Wert
LUME-Lung 1 (2. Datenschnitt)	322	255 (79,2)	4,2 [3,6;4,4]	336	267 (79,5)	2,8 [2,6;3,2]	0,84 [0,71; 1,00]	0,0485
LUME-Lung 1 (1. Datenschnitt)	277	152 (54,9)	4,0 [3,2;4,4]	285	180 (63,2)	2,8 [2,6;3,3]	0,77 [0,62; 0,96]	0,0193

Verwendete Datenschnitte: 2. Datenschnitt vom 15. Februar 2013 (Finale OS-Analyse) und 1. Datenschnitt vom 02. November 2010 (Primäre PFS-Analysen); Analyse Set: RS

HR: Hazard Ratio aus stratifizierter Cox-Regression (Stratifizierungsvariablen: ECOG-PS zur Baseline, Bevacizumab-Vorbehandlung, Hirnmetastasen zur Baseline); KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit PFS-Ereignis; RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Zum Zeitpunkt der finalen OS-Analyse war das PFS unter der Kombinationstherapie mit Nintedanib+Docetaxel gegenüber der Placebo+Docetaxel-Vergleichstherapie statistisch signifikant länger. Die Verlängerung im medianen PFS betrug 1,4 Monate (4,2 Monate im Nintedanib+Docetaxel-Arm versus 2,8 Monate im Placebo+Docetaxel-Arm). Im Beobachtungszeitraum trat bei 79,2% der Nintedanib+Docetaxel-Patienten eine Progression oder der Tod ein. Dies war bei 79,5% der Placebo+Docetaxel-Patienten der Fall (Tabelle 4-18). Das auf Basis einer stratifizierten Cox-Regression berechnete HR [95%-KI] von 0,84 [0,71; 1,00]; p=0,0485 zeigt den statistisch signifikanten Vorteil zugunsten der Kombinationstherapie aus Nintedanib+Docetaxel. Der Vorteil bezüglich des PFS für Nintedanib+Docetaxel gegenüber Placebo+Docetaxel war auch schon beim 1. Datenschnitt zum Zeitpunkt der primären PFS-Analyse vorhanden (HR [95%-KI]: 0,77 [0,62; 0,96]; p=0,0193).

Die im Folgenden dargestellten Kaplan-Meier-Kurven verdeutlichen den Vorteil der Nintedanib+Docetaxel-Therapie gegenüber der Placebo+Docetaxel-Vergleichstherapie bezüglich des PFS zum finalen Datenschnitt (Abbildung 4-3) sowie zum Zeitpunkt der primären Analyse (Abbildung 4-4).

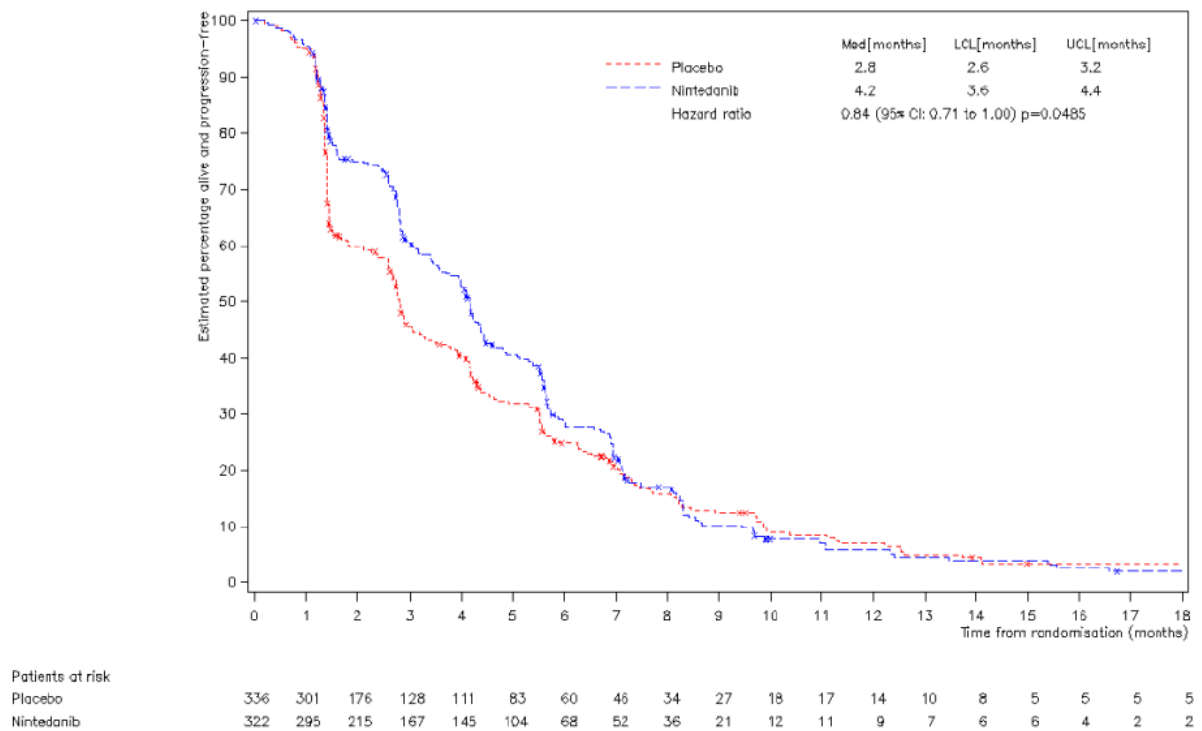


Abbildung 4-3: Kaplan-Meier-Kurven zum progressionsfreien Überleben (2. Datenschnitt, LUME-Lung 1 Studie)

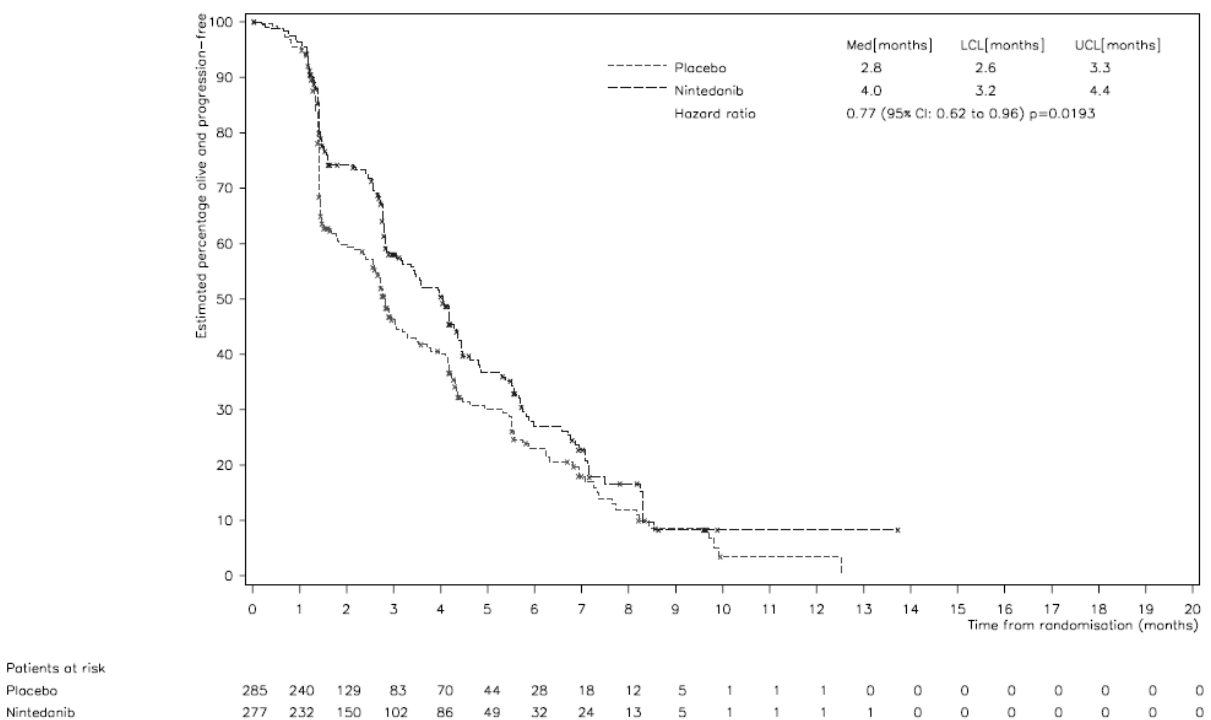


Abbildung 4-4: Kaplan-Meier-Kurven zum progressionsfreien Überleben (1. Datenschnitt, LUME-Lung 1 Studie)

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.3 Objektive Ansprechrate – RCT

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

Tabelle 4-19: Operationalisierung der objektiven Ansprechrate

Studie	Operationalisierung
LUME-Lung 1	Die ORR war definiert als der Anteil der Patienten, die gemäß zentraler und unabhängiger Beurteilung ein bestätigtes CR oder ein bestätigtes PR erreicht haben (gemäß RECIST-Kriterien 1.0). Die Bestätigung des CR oder PR musste mindestens 28 Tage nach dem ersten dokumentierten CR oder PR erfolgen.
CR: Complete Response (komplettes Ansprechen); ORR: Objective Response Rate (objektive Ansprechrate); PR: Partial Response (partielles Ansprechen); RECIST: Response Evaluation Criteria In Solid Tumors.	

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 4-20: Bewertung des Verzerrungspotenzials für die objektive Ansprechrate 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Endpunkterheber der ORR waren bzgl. der Behandlungsgruppe der Patienten verblindet. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt, denn es wurden alle randomisierten Patienten berücksichtigt und entsprechend der randomisiert zugewiesenen Behandlungsgruppe analysiert. Die Auswertung dieses Endpunkts erfolgte gemäß dem vorab erstellten TSAP (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011). Sonstige, das Verzerrungspotenzial beeinflussende, Punkte existieren nicht, und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Somit ist für den Endpunkt ORR von einem niedrigen Verzerrungspotenzial auszugehen.

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

Die Ergebnisse zur ORR aus der Studie LUME-Lung 1 sind in Tabelle 4-21 dargestellt.

Tabelle 4-21: Ergebnisse für die objektive Ansprechrate aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
LUME-Lung 1	322	15 (4,7)	336	12 (3,6)	1,30 [0,62; 2,74] 1,32 [0,61; 2,86] 1,09 [-1,95; 4,13]	0,4823

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a 95%-Konfidenzintervall basierend auf der Normal-Approximation; ^b p-Wert des Chi-Quadrat-Tests

KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten, die gemäß verblindeter und unabhängiger Beurteilung eine CR oder PR erreicht haben (RECIST 1.0); OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; RS: Randomized Set.

Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die Ergebnisse zur objektiven Ansprechrate zeigten keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie Nintedanib+Docetaxel und der Placebo+Docetaxel-Vergleichstherapie (p=0,4823). In der Nintedanib+Docetaxel-Gruppe wiesen 4,7% und in der Placebo+Docetaxel-Gruppe 3,6% der Patienten ein objektives Ansprechen auf. Die Ergebnisse zur ORR aus der Studie LUME-Lung 1 sind in Tabelle 4-21 dargestellt. Das RR [95%-KI] lag bei 1,30 [0,62; 2,74], das OR [95%-KI] bei 1,32 [0,61; 2,86] und die RD [95%-KI] bei 1,09 [-1,95; 4,13].

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.4 Krankheitskontrollrate – RCT

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

Tabelle 4-22: Operationalisierung der Krankheitskontrollrate

Studie	Operationalisierung
LUME-Lung 1	Die DCR war definiert als der Anteil der Patienten, die gemäß zentraler und unabhängiger Beurteilung ein bestätigtes CR, ein bestätigtes PR oder eine SD erreicht haben (gemäß RECIST-Kriterien 1.0). Die Bestätigung des CR oder PR musste mindestens 28 Tage nach dem ersten dokumentierten CR oder PR erfolgen.
CR: Complete Response (komplettes Ansprechen); DCR: Disease Control Rate; PR: Partial Response (partielleres Ansprechen); RECIST: Response Evaluation Criteria In Solid Tumors; SD: Stable Disease (stabile Erkrankung).	

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 4-23: Bewertung des Verzerrungspotenzials für die Krankheitskontrollrate 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Endpunkterheber der DCR waren ebenso wie beim ORR bzgl. der Behandlungsgruppe der Patienten verblindet. Bei der Auswertung wurde das ITT-Prinzip adäquat umgesetzt, denn es wurden alle randomisierten Patienten (Analyse Set: RS) berücksichtigt und entsprechend der randomisiert zugewiesenen Behandlungsgruppe analysiert. Die Auswertung dieses Endpunkts erfolgte gemäß dem vorab erstellten TSAP (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011). Sonstige, das Verzerrungspotenzial beeinflussende, Punkte existieren nicht, und es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Somit ist für den Endpunkt DCR von einem niedrigen Verzerrungspotenzial auszugehen.

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

Die Ergebnisse zur DCR aus der Studie LUME-Lung 1 sind in Tabelle 4-24 dargestellt.

Tabelle 4-24: Ergebnisse für die Krankheitskontrollrate aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
LUME-Lung 1	322	194 (60,2)	336	148 (44,0)	1,37 [1,18; 1,59] 1,93 [1,41; 2,63] 16,20 [8,67; 23,73]	<0,0001

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a 95%-Konfidenzintervall basierend auf der Normal-Approximation; ^b p-Wert des Chi-Quadrat-Tests

KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten, die gemäß verblindeter und unabhängiger Beurteilung eine CR, PR oder SD erreicht haben (RECIST 1.0); OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; RS: Randomized Set.

Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die DCR war für Patienten mit der Kombinationstherapie Nintedanib+Docetaxel statistisch signifikant besser im Vergleich zur Placebo+Docetaxel-Therapie (p<0,0001). Im Nintedanib+Docetaxel-Arm erreichten 60,2% eine Kontrolle der Erkrankung, im Placebo+Docetaxel-Behandlungsarm 44,0%. Die Ergebnisse zur DCR aus der Studie LUME-Lung 1 sind in Tabelle 4-24 dargestellt. Das RR [95%-KI] lag bei 1,37 [1,18; 1,59], das OR [95%-KI] bei 1,93 [1,41; 2,63] und die RD [95%-KI] betrug 16,20 [8,67; 23,73] zugunsten der Kombinationstherapie Nintedanib+Docetaxel.

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.5 Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 – RCT

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 der Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13

Studie	Operationalisierung
LUME-Lung 1	a) Jeweils für alle Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13: Anteil der Patienten, die während des Studienverlaufs im Mittel eine Verbesserung von mindestens 10 Punkten gegenüber dem Ausgangswert über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben. (Responderanalysen) b) Jeweils für alle Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13: Zeit bis zur Verschlechterung des Scores um mindestens 10 Punkte gegenüber dem Ausgangswert.
EORTC: European Organization for Research and Treatment of Cancer; QLQ-C30: Quality of Life Questionnaire Core-30; QLQ-LC13: Quality of Life Questionnaire Lung Cancer-13.	

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 Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Symptomatik wurde mittels des Patientenfragebogens EORTC QLQ-C30 sowie des EORTC QLQ-LC13 jeweils zu Beginn eines Behandlungszyklus sowie alle drei Wochen während der Follow-up Phase beurteilt. Hierbei handelt es sich um validierte, patientenrelevante und weltweit eingesetzte Instrumente, welche eine gute Korrelation zum klinischen Verlauf einer Erkrankung aufweisen (Aaronson et al., 1993; Osoba et al., 1998; Velikova et al., 2012).

Die Compliance-Raten für das Ausfüllen der Fragebögen EORTC QLQ-C30 und EORTC QLQ-LC13 waren während der Behandlungsperiode in beiden Behandlungsarmen hoch (>80% in beiden Behandlungsarmen) (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a). Für Patienten mit der Tumorphistologie Adenokarzinom lag die Rücklaufquote für beide Fragebögen ebenfalls über 80% und war zwischen beiden Behandlungsarmen ausgeglichen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d).

Die Patienten waren verblindet und es gab keine weiteren Aspekte, die das Verzerrungspotenzial beeinflussten. Deshalb wurde das endpunktspezifische Verzerrungspotenzial für die Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13, als niedrig bewertet.

Durch die Analyse der Fragebogendaten, einmal hinsichtlich des Erreichens einer Response und einmal hinsichtlich der Zeit bis zur Verschlechterung, können sowohl klinisch relevante Verbesserungen als auch Verschlechterungen der Symptomatik abgebildet werden.

Stellen Sie die Ergebnisse für den Endpunkt Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Zur Symptomatik, die anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 untersucht wurde, seien einige Bemerkungen vorangestellt. Die in den Fragebögen vom Patienten einzuschätzenden Symptome sind differenziert zu betrachten. Sie umfassen sowohl tumorbedingte als auch therapieassoziierte Symptome. Einzelne Symptome, wie Fatigue, können sowohl tumorbedingt als auch therapieassoziiert sein:

Tumorbedingte Symptome

Schmerzen, Dyspnoe, Fatigue, Bluthusten, Schlaflosigkeit, Husten

Therapieassoziierte Symptome

Appetitverlust, Diarrhö, Übelkeit und Erbrechen, Verstopfung, Fatigue, Haarausfall, Mundschmerzen, periphere Neuropathie, Schluckbeschwerden

Die Ergebnisse der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 werden in diesem Nutzendossier anhand von Responderanalysen und Analysen der Zeit bis zu einer klinisch relevanten Verschlechterung dargestellt. Diese Analysen basieren auf patientenindividuellen Schwellenwerten für klinisch relevante Verbesserungen bzw. Verschlechterungen (Osoba et al., 1998). Ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen ist somit gleichzeitig als klinisch relevant zu werten.

Die beiden Auswertungen bilden unterschiedliche patientenrelevante Nutzenaspekte ab, die jeweils eine eigenständige Berechtigung haben. Die Responderanalysen zeigen die Verbesserung initial bereits vorhandener Symptome an. Dies ist aus Patientensicht ein wichtiges Therapieziel. Die Zeit bis zur Verschlechterung beschreibt die Verzögerung des Neuaufretens bzw. der Verschlechterung von Symptomen und ist dadurch mit dem Aufhalten des Fortschreitens der Krebserkrankung verbunden und ebenso ein wichtiges Therapieziel.

Zur Beschreibung der Ausgangssituation der Patienten werden zunächst die Ausgangswerte der Symptomatik zur Baseline (Tabelle 4-27 und Tabelle 4-28) präsentiert. Daran schließt die Diskussion der Ergebnisse der Responderanalysen (Tabelle 4-29 und Tabelle 4-30) sowie der Zeit bis zur Verschlechterung (Tabelle 4-31 und Tabelle 4-32) an.

Tabelle 4-27: Ausgangswerte der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel			Placebo+Docetaxel		
	N	Mittelwert	STD	N	Mittelwert	STD
Appetitverlust	296	22,86	29,67	315	22,33	26,58
Diarrhö	295	5,88	14,40	311	4,72	13,36
Dyspnoe	296	31,76	27,95	315	32,80	28,83
Fatigue	297	36,70	24,66	315	34,89	22,13
Schlaflosigkeit	297	26,82	29,81	315	26,03	28,89
Schmerzen	297	28,11	28,27	316	27,58	26,15
Übelkeit und Erbrechen	295	9,15	19,17	316	7,75	16,28
Verstopfung	287	14,87	25,31	304	13,27	23,05

Werte zur Baseline; Analyse Set: RS
N: Anzahl Patienten in der Analyse; RS: Randomized Set; STD: Standardabweichung.

In Tabelle 4-27 sind die Ausgangswerte der Symptome, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30, dargestellt. Generell waren die Ausgangswerte in beiden Behandlungsgruppen sehr ausgeglichen. Die patientenberichteten Angaben zur Baseline zeigten, dass die Symptome Diarrhö sowie Übelkeit und Erbrechen zu Beginn der Studie eine sehr geringe Belastung für die Patienten in beiden Behandlungsarmen dargestellt haben. Bezüglich der Symptome Dyspnoe, Fatigue und Schmerzen berichteten sowohl die Nintedanib+Docetaxel-Patienten als auch die Placebo+Docetaxel-Patienten eine größere Belastung.

Tabelle 4-28: Ausgangswerte der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel			Placebo+Docetaxel		
	N	Mittelwert	STD	N	Mittelwert	STD
Bluthusten	294	2,61	9,78	313	2,45	9,49
Dyspnoe	290	29,12	20,38	308	26,91	20,40
Haarausfall	292	18,84	31,13	313	15,44	28,11
Husten	294	37,76	27,80	314	33,23	26,84
Mundschmerzen	293	5,12	13,81	315	4,87	13,73
Periphere Neuropathie	293	16,84	25,68	312	15,92	25,77
Schluckbeschwerden	293	5,01	14,25	314	6,05	15,84
Schmerzen (Arm/Schulter)	292	19,63	26,70	313	19,91	27,04
Schmerzen (Brust)	293	19,45	26,83	312	22,01	25,22
Schmerzen (andere)	275	23,15	29,89	300	20,22	26,40

Werte zur Baseline; Analyse Set: RS
N: Anzahl Patienten in der Analyse; RS: Randomized Set; STD: Standardabweichung.

In Tabelle 4-28 ist ersichtlich, dass die Patienten in beiden Behandlungsarmen ihre Symptombelastung, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13, zu Beginn der Studie ähnlich einschätzten. Bluthusten und Schluckbeschwerden waren nur eine geringfügige Belastung für die Patienten. Husten und Dyspnoe hingegen stellten zu Beginn der Studie bereits eine größere Belastung sowohl für Nintedanib+Docetaxel-Patienten als auch für Placebo+Docetaxel-Patienten dar.

Tabelle 4-29: Ergebnisse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
Appetitverlust	288	52 (18,1)	291	61 (21,0)	0,86 [0,62; 1,20] 0,83 [0,55; 1,25] -2,91 [-9,36; 3,54]	0,378
Diarrhö	289	21 (7,3)	289	26 (9,0)	0,81 [0,47; 1,40] 0,79 [0,44; 1,44] -1,73 [-6,18; 2,72]	0,447
Dyspnoe	289	67 (23,2)	292	84 (28,8)	0,81 [0,61; 1,06] 0,75 [0,51; 1,08] -5,58 [-12,70; 1,53]	0,125
Fatigue	289	57 (19,7)	292	55 (18,8)	1,05 [0,75; 1,46] 1,06 [0,70; 1,60] 0,89 [-5,53; 7,30]	0,786
Schlaflosigkeit	289	75 (26,0)	292	78 (26,7)	0,97 [0,74; 1,28] 0,96 [0,66; 1,39] -0,76 [-7,92; 6,40]	0,835
Schmerzen	289	77 (26,6)	292	50 (17,1)	1,56 [1,13; 2,14] 1,76 [1,18; 2,62] 9,52 [2,84; 16,20]	0,005
Übelkeit und Erbrechen	289	34 (11,8)	292	30 (10,3)	1,15 [0,72; 1,82] 1,16 [0,69; 1,96] 1,49 [-3,60; 6,58]	0,566
Verstopfung	285	60 (21,1)	282	44 (15,6)	1,35 [0,95; 1,92] 1,44 [0,94; 2,22] 5,45 [-0,90; 11,80]	0,094

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a 95%-Konfidenzintervall basierend auf der Normal-Approximation; ^b p-Wert des Chi-Quadrat-Tests

KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten, die während des Studienverlaufs im Mittel eine Verbesserung von mindestens 10 Punkten über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; RS: Randomized Set.

Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die Responderanalysen in Tabelle 4-29 beschreiben den Anteil von Patienten mit einer Verbesserung der Symptomatik anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30. Bezüglich der Schmerzen zeigte sich ein statistisch signifikanter Vorteil für die Kombination von Nintedanib+Docetaxel gegenüber der Therapie mit

Placebo+Docetaxel (p=0,005). Das RR [95%-KI] lag bei 1,56 [1,13; 2,14], das OR [95%-KI] bei 1,76 [1,18; 2,62] und die RD [95%-KI] betrug 9,52 [2,84; 16,20]. Für alle weiteren Symptome (Appetitverlust, Diarrhö, Dyspnoe, Fatigue, Schlaflosigkeit, Übelkeit und Erbrechen sowie Verstopfung) ergaben sich bezüglich des Anteils der Patienten mit einer Symptomverbesserung keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsarmen.

Tabelle 4-30: Ergebnisse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalyse, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
Bluthusten	286	10 (3,5)	289	10 (3,5)	1,01 [0,43; 2,39] 1,01 [0,41; 2,47] 0,04 [-2,96; 3,03]	0,981
Dyspnoe	284	49 (17,3)	289	53 (18,3)	0,94 [0,66; 1,34] 0,93 [0,60; 1,42] -1,09 [-7,35; 5,18]	0,734
Haarausfall	285	44 (15,4)	289	32 (11,1)	1,39 [0,91; 2,13] 1,47 [0,90; 2,39] 4,37 [-1,17; 9,91]	0,123
Husten	286	91 (31,8)	291	80 (27,5)	1,16 [0,90; 1,49] 1,23 [0,86; 1,76] 4,33 [-3,12; 11,77]	0,255
Mund- schmerzen	285	25 (8,8)	291	21 (7,2)	1,22 [0,70; 2,12] 1,24 [0,68; 2,26] 1,56 [-2,87; 5,99]	0,491
Periphere Neuropathie	284	45 (15,8)	289	37 (12,8)	1,24 [0,83; 1,85] 1,28 [0,80; 2,05] 3,04 [-2,69; 8,78]	0,298
Schluck- beschwerden	285	23 (8,1)	290	21 (7,2)	1,11 [0,63; 1,97] 1,12 [0,61; 2,08] 0,83 [-3,52; 5,18]	0,709
Schmerzen (Arm/Schulter)	285	72 (25,3)	290	63 (21,7)	1,16 [0,86; 1,56] 1,22 [0,83; 1,79] 3,54 [-3,39; 10,47]	0,317
Schmerzen (Brust)	285	58 (20,4)	290	57 (19,7)	1,04 [0,75; 1,44] 1,04 [0,69; 1,57] 0,70 [-5,84; 7,24]	0,835

Symptom	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
Schmerzen (andere)	270	67 (24,8)	278	56 (20,1)	1,23 [0,90; 1,68] 1,31 [0,87; 1,96] 4,67 [-2,31; 11,65]	0,190

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a 95%-Konfidenzintervall basierend auf der Normal-Approximation; ^b p-Wert des Chi-Quadrat-Tests

KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten, die während des Studienverlaufs im Mittel eine Verbesserung von mindestens 10 Punkten über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; RS: Randomized Set.

Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an;
p-Wert <0,05 fett geschrieben

Die Responderanalysen in Tabelle 4-30 beschreiben die Verbesserung der Symptomatik anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13. Für keines der im Fragebogen EORTC QLQ-LC13 erfassten Symptome (Bluthusten, Dyspnoe, Haarausfall, Husten, Mundschmerzen, Periphere Neuropathie, Schluckbeschwerden, Schmerzen – Arm/Schulter, Schmerzen – Brust, Schmerzen – andere) wurde ein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen ermittelt.

Tabelle 4-31: Ergebnisse zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Monate)	N	n (%)	Median (Monate)	HR [95%-KI]	p-Wert
Appetitverlust	322	194 (60,2)	2,1	336	180 (53,6)	3,0	1,15 [0,94; 1,40]	0,1778
Diarrhö	322	219 (68,0)	2,1	336	154 (45,8)	4,2	1,90 [1,54; 2,34]	<0,0001
Dyspnoe	322	177 (55,0)	3,6	336	166 (49,4)	3,6	1,05 [0,85; 1,30]	0,6327
Fatigue	322	230 (71,4)	1,4	336	211 (62,8)	2,1	1,15 [0,95; 1,38]	0,1414
Schlaflosigkeit	322	176 (54,7)	3,6	336	171 (50,9)	3,5	0,98 [0,80; 1,21]	0,8661
Schmerzen	322	206 (64,0)	2,8	336	196 (58,3)	2,8	0,94 [0,77; 1,14]	0,5265
Übelkeit und Erbrechen	322	197 (61,2)	2,8	336	173 (51,5)	3,8	1,23 [1,00; 1,50]	0,0467
Verstopfung	322	145 (45,0)	5,1	336	140 (41,7)	4,9	0,91 [0,72; 1,14]	0,4005

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit einer Verschlechterung des Scores um mindestens 10 Punkte; RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben

In Tabelle 4-31 wird die Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand des Fragebogens EORTC QLQ-C30, dargestellt. Für die Symptome Diarrhö (HR [95%-KI]: 1,90 [1,54; 2,34]; p<0,0001) sowie das Symptom Übelkeit und Erbrechen (HR [95%-KI]: 1,23 [1,00; 1,50]; p=0,0467) zeigt sich ein statistisch signifikanter Nachteil für Nintedanib+Docetaxel. Die Zeit bis zur Verschlechterung dieser Symptome war gegenüber der Vergleichstherapie Placebo+Docetaxel kürzer. Für alle weiteren Symptome des Fragebogens EORTC QLQ-C30 (Appetitverlust, Dyspnoe, Fatigue, Schlaflosigkeit, Schmerzen und Verstopfung) wurde kein statistisch signifikanter Unterschied zwischen beiden Behandlungsgruppen ermittelt.

Die Kaplan-Meier-Kurven zur Verschlechterung der Symptomatik, gemessen anhand des Fragebogens EORTC QLQ-C30, wurden erstellt und mit dem Dossier eingereicht (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014e).

Tabelle 4-32: Ergebnisse zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Symptom	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Monate)	N	n (%)	Median (Monate)	HR [95%-KI]	p-Wert
Bluthusten	322	98 (30,4)	14,0	336	100 (29,8)	9,0	0,90 [0,68; 1,19]	0,4545
Dyspnoe	322	222 (68,9)	1,8	336	220 (65,5)	2,1	1,03 [0,86; 1,25]	0,7135
Haarausfall	322	218 (67,7)	1,4	336	230 (68,5)	1,0	0,87 [0,72; 1,04]	0,1094
Husten	322	168 (52,2)	4,2	336	166 (49,4)	4,2	0,97 [0,78; 1,20]	0,7640
Mund- schmerzen	322	156 (48,4)	4,3	336	145 (43,2)	4,2	1,02 [0,81; 1,28]	0,8468
Periphere Neuropathie	322	176 (54,7)	3,5	336	161 (47,9)	3,7	1,02 [0,83; 1,27]	0,8314
Schluck- beschwerden	322	132 (41,0)	7,9	336	136 (40,5)	5,6	0,92 [0,72; 1,17]	0,4801
Schmerzen (Arm/Schulter)	322	143 (44,4)	5,8	336	160 (47,6)	4,2	0,80 [0,63; 1,00]	0,0456
Schmerzen (Brust)	322	170 (52,8)	4,2	336	157 (46,7)	4,2	1,03 [0,83; 1,28]	0,7751
Schmerzen (andere)	322	150 (46,6)	5,1	336	153 (45,5)	4,1	0,86 [0,69; 1,08]	0,1835

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit einer Verschlechterung des Scores um mindestens 10 Punkte; RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben

Die Zeit bis zur Verschlechterung der anhand des Fragebogens EORTC QLQ-LC13 erfassten Symptome wird in Tabelle 4-32 dargestellt. Anhand der Schmerzen (Arm/Schulter) zeigte sich ein statistisch signifikanter Vorteil für die Kombinationstherapie Nintedanib+Docetaxel gegenüber der Placebo+Docetaxel-Therapie. Das HR [95%-KI] betrug 0,80 [0,63; 1,00]; p=0,0456 zugunsten von Nintedanib+Docetaxel. Für alle weiteren Symptome des Fragebogens EORTC QLQ-LC13 ergaben sich bezüglich der Zeit bis zum ersten Auftreten der Symptome keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsarmen.

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.6 Gesundheitsbezogene Lebensqualität – RCT

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 der gesundheitsbezogenen Lebensqualität

Studie	Operationalisierung
LUME-Lung 1	<ul style="list-style-type: none"> a) Jeweils für die Skalen globaler Gesundheitsstatus, Rollenfunktion, körperliche, emotionale, kognitive und soziale Funktion aus dem Fragebogen EORTC QLQ-C30: Anteil an Patienten, die eine durchschnittliche Verbesserung des Scores um mindestens 10 Punkte gegenüber dem Ausgangswert über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben. (Responderanalyse) b) Jeweils für die Skalen globaler Gesundheitsstatus, Rollenfunktion, körperliche, emotionale, kognitive und soziale Funktion aus dem Fragebogen EORTC QLQ-C30: Zeit bis zur Verschlechterung des Scores um mindestens 10 Punkte gegenüber dem Ausgangswert. c) Für die Lebensqualitätsskala EQ-5D UK Utility sowie für die Lebensqualitätsskala EQ-VAS: Mittlerer Wert bis zur medianen Follow-up-Zeit sowie die Mittelwertdifferenzen.
<small>EORTC: European Organization for Research and Treatment of Cancer; EQ-5D: EuroQoL-5 Dimensions (Patientenfragebogen zur Lebensqualität); EQ-VAS: EuroQoL-Visual Analogue Scale (visuelle Analogskala); QLQ-C30: Quality of Life Questionnaire Core-30.</small>	

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 gesundheitsbezogene Lebensqualität 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die gesundheitsbezogene Lebensqualität wurde mittels der Patientenfragebögen EORTC QLQ-C30 und EQ-5D jeweils zu Beginn eines Behandlungszyklus sowie alle drei Wochen während der Follow-up Phase bewertet. Hierbei handelt es sich um validierte, patientenrelevante und weltweit eingesetzte Instrumente zur Bewertung der HRQoL, welche eine gute Korrelation zum klinischen Verlauf einer Erkrankung aufweisen (Aaronson et al., 1993; Osoba et al., 1998; Velikova et al., 2012).

Die Compliance-Raten über die Behandlungszeit für den EORTC QLQ-C30 und die EQ-5D-Befragung waren für die Patienten der Tumorphistologie Adenokarzinom durchweg hoch und zwischen beiden Behandlungsarmen ausgeglichen. Die Rücklaufquote für den Fragebogen EORTC QLQ-C30 lag für Nintedanib-Patienten und für Placebo-Patienten bei über 80%. Für die EQ-5D-Befragung waren die Compliance-Raten noch höher und betragen im Nintedanib-Arm sowie im Placebo-Arm über 90% (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d).

Aufgrund der Verblindung der Patienten und keiner weiteren, das Verzerrungspotenzial beeinflussenden, Aspekte wurde das endpunktspezifische Verzerrungspotenzial für die gesundheitsbezogene Lebensqualität als niedrig bewertet.

Die Auswertung erfolgte analog zu den Symptomskalen.

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

Einleitend werden zunächst die Ausgangswerte der gesundheitsbezogenen Lebensqualität, gemessen anhand des Fragebogens EORTC QLQ-C30 (Tabelle 4-35) sowie der Skalen EQ-5D und EQ-VAS (Tabelle 4-36) berichtet. Die Ergebnisse zur Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30, werden in Tabelle 4-37 präsentiert. Daran anschließend wird die Operationalisierung Zeit bis zur Verschlechterung der gesundheitsbezogenen

Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30, in Tabelle 4-38 dargestellt. Und schließlich zeigt Tabelle 4-39 die Auswertungen zu den aggregierten Lebensqualitätsskalen EQ-5D und EQ-VAS.

Tabelle 4-35: Ausgangswerte der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Domäne	Nintedanib+Docetaxel			Placebo+Docetaxel		
	N	Mittelwert	STD	N	Mittelwert	STD
Globaler Gesundheitsstatus	293	61,49	19,87	311	63,29	19,83
Emotionale Funktion	294	75,85	22,37	313	77,45	20,82
Kognitive Funktion	294	83,56	19,37	313	86,00	17,75
Körperliche Funktion	297	72,18	21,45	314	74,72	18,45
Rollenfunktion	297	72,11	28,03	315	74,18	27,16
Soziale Funktion	294	71,26	28,99	313	75,45	26,08

Werte zur Baseline; Analyse Set: RS
N: Anzahl Patienten in der Analyse; STD: Standardabweichung

Für die gesundheitsbezogene Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30, werden in Tabelle 4-35 die Ausgangswerte zu Beginn der Studie dargestellt. Hohe Werte auf der Skala von 0-100 zeigen hier eine hohe Lebensqualität an. In der Studie LUME-Lung 1 zeigten die Patienten in beiden Behandlungsarmen zu Studienbeginn eine vergleichbare Lebensqualität. Zudem wurde die Lebensqualität in allen Domänen ähnlich gut berichtet. Die kognitive Funktion wurde sowohl von Nintedanib+Docetaxel-Patienten als auch von Placebo+Docetaxel-Patienten besonders gut beschrieben, während der globale Gesundheitstatus jeweils die niedrigste Bewertung erhielt.

Tabelle 4-36: Ausgangswerte zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Domäne	Nintedanib+Docetaxel			Placebo+Docetaxel		
	N	Mittelwert	STD	N	Mittelwert	STD
EQ-5D UK Utility	293	0,697	0,259	315	0,726	0,235
EQ-VAS	295	67,7	18,0	315	68,9	18,6

EQ-5D: EuroQoL-5 Dimensions (Patientenfragebogen zur Lebensqualität); EQ-VAS: EuroQoL-Visual Analogue Scale (visuelle Analogskala); N: Anzahl der Patienten in der Analyse.

Die Ausgangswerte des EQ-5D UK Utility Scores (Tabelle 4-36) lagen bei 0,697 für Nintedanib+Docetaxel-Patienten und bei 0,726 für Placebo+Docetaxel-Patienten. Bezüglich der EQ-VAS Skala betragen die Ausgangswerte 67,7 im Nintedanib+Docetaxel-Arm und 68,9 im Placebo+Docetaxel-Arm. Die Ausgangssituation der Patienten in beiden Behandlungsarmen war somit bezüglich der gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS, vergleichbar.

Tabelle 4-37: Ergebnisse zur Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalyse, Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Domäne	Nintedanib+Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	N	n (%)	RR [95%-KI] ^a OR [95%-KI] ^a RD [95%-KI] ^a	p-Wert ^b
Globaler Gesundheitsstatus	288	53 (18,4)	289	34 (11,8)	1,56 [1,05; 2,33] 1,69 [1,06; 2,69] 6,64 [0,82; 12,45]	0,026
Emotionale Funktion	289	60 (20,8)	290	54 (18,6)	1,11 [0,80; 1,55] 1,15 [0,76; 1,73] 2,14 [-4,34; 8,62]	0,517
Kognitive Funktion	289	59 (20,4)	290	36 (12,4)	1,64 [1,12; 2,41] 1,81 [1,15; 2,84] 8,00 [2,00; 14,00]	0,009
Körperliche Funktion	289	39 (13,5)	291	31 (10,7)	1,27 [0,81; 1,97] 1,31 [0,79; 2,16] 2,84 [-2,46; 8,14]	0,294
Rollenfunktion	289	55 (19,0)	291	55 (18,9)	1,01 [0,72; 1,41] 1,01 [0,67; 1,53] 0,13 [-6,25; 6,51]	0,968
Soziale Funktion	289	69 (23,9)	290	55 (19,0)	1,26 [0,92; 1,72] 1,34 [0,90; 2,00] 4,91 [-1,76; 11,58]	0,150

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

^a 95%-Konfidenzintervall basierend auf der Normal-Approximation; ^b p-Wert des Chi-Quadrat-Tests

KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten, die während des Studienverlaufs im Mittel eine Verbesserung von mindestens 10 Punkten über alle Zeitpunkte erreicht haben, zu denen sie den Fragebogen ausgefüllt haben; RR: Relatives Risiko; OR: Odds Ratio; RD: Risikodifferenz; RS: Randomized Set.

Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die Ergebnisse der Responderanalyse für die gesundheitsbezogene Lebensqualität aus dem Fragebogen EORTC QLQ-C30 sind in Tabelle 4-37 dargestellt. Der Vergleich der Kombinationstherapie Nintedanib+Docetaxel mit der Placebo+Docetaxel-Therapie zeigte für den globalen Gesundheitsstatus eine statistisch signifikante Überlegenheit von Nintedanib+Docetaxel (p=0,026). Im Nintedanib+Docetaxel-Arm trat bei 18,4% und im

Placebo+Docetaxel-Arm bei 11,8% der Patienten eine Verbesserung des globalen Gesundheitsstatus ein. Das RR [95%-KI] belief sich auf 1,56 [1,05; 2,33], das OR [95%-KI] auf 1,69 [1,06; 2,69] und die RD [95%-KI] betrug 6,64 [0,82; 12,45].

Auch in Bezug auf die kognitive Funktion war der Anteil der Patienten mit einer Verbesserung im Nintedanib+Docetaxel-Arm höher als im Vergleichsarm (Nintedanib+Docetaxel: 20,4% vs. Placebo+Docetaxel: 12,4%). Dieser Unterschied war statistisch signifikant ($p=0,009$). Es ergab sich ein RR [95%-KI] von 1,64 [1,12; 2,41], ein OR [95%-KI] von 1,81 [1,15; 2,84] und eine RD [95%-KI] von 8,00 [2,00; 14,00]. Für alle weiteren Aspekte der gesundheitsbezogenen Lebensqualität wurden keine statistisch signifikanten Unterschiede ermittelt.

Tabelle 4-38: Ergebnisse zur Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Domäne	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Monate)	N	n (%)	Median (Monate)	HR [95%-KI]	p-Wert
Globaler Gesundheitsstatus	322	195 (60,6)	2,9	336	198 (58,9)	2,8	0,88 [0,72; 1,07]	0,1912
Emotionale Funktion	322	175 (54,3)	3,5	336	164 (48,8)	3,8	1,07 [0,86; 1,32]	0,5348
Kognitive Funktion	322	187 (58,1)	2,8	336	177 (52,7)	3,0	1,03 [0,84; 1,27]	0,7695
Körperliche Funktion	322	191 (59,3)	2,8	336	196 (58,3)	2,8	0,92 [0,75; 1,12]	0,3931
Rollenfunktion	322	217 (67,4)	2,1	336	202 (60,1)	2,1	1,04 [0,86; 1,26]	0,7076
Soziale Funktion	322	195 (60,6)	2,8	336	189 (56,3)	2,8	0,97 [0,80; 1,19]	0,7974

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit einer Verschlechterung des Scores um mindestens 10 Punkte; RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität wird in Tabelle 4-38 präsentiert. Für alle Domänen des Fragebogen EORTC QLQ-C30 zeigten sich keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie Nintedanib+Docetaxel sowie der Placebo+Docetaxel-Vergleichstherapie.

Tabelle 4-39: Ergebnisse zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Domäne	Nintedanib+Docetaxel	Placebo+Docetaxel	Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	Mittelwert	Mittelwert	Mittelwertdifferenz [95%-KI]	p-Wert
EQ-5D UK Utility	0,684	0,679	0,005 [-0,022; 0,032]	0,6983
EQ-VAS	66,7	66,7	-0,0 [-1,6; 1,5]	0,9630

Schätzer aus longitudinalem Modell adjustiert für die Stratifizierungsvariablen ECOG Status, Bevacizumab-Vorbehandlung und Hirnmetastasen; Mittelwert bis zur medianen Follow-up-Zeit, der mit Hilfe der „area under the curve“ der individuellen Behandlungsprofile unter Verwendung des adjustierten longitudinalen Modells geschätzt wurde.

EQ-5D: EuroQoL-5 Dimensions (Patientenfragebogen zur Lebensqualität); EQ-VAS: EuroQoL-Visual Analogue Scale (visuelle Analogskala); KI: Konfidenzintervall.

Eine Mittelwertdifferenz >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben

Die Ergebnisse zur gesundheitsbezogenen Lebensqualität auf Basis des EQ-5D UK Utility Scores sowie der EQ-VAS sind in Tabelle 4-39 dargestellt. Anhand dieser beiden Auswertungen zeigten sich keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie mit Nintedanib+Docetaxel sowie der Placebo+Docetaxel-Vergleichstherapie.

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.7 Unerwünschte Ereignisse – RCT

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

Tabelle 4-40: Operationalisierung der unerwünschten Ereignisse

Studie	Operationalisierung
LUME-Lung 1	Zeit bis zum ersten Auftreten eines UE: Definiert als die Zeit von der Randomisierung bis zu dessen ersten Auftreten.
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 4-41: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der TSAP der LUME-Lung 1 Studie sah für die UE von besonderem Interesse die Analyse der Zeit bis zum ersten Auftreten vor (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011). Die Auswertungen für alle weiteren UE erfolgten im Rahmen der LUME-Lung 1 Studie deskriptiv. Für das vorliegende Dossier wurden alle UE mit der Operationalisierung „Zeit bis zum ersten Auftreten“ analysiert und präsentiert. In die Analyse gingen alle Patienten ein, die mindestens eine Dosis der Studienmedikation erhalten haben (TS [Treated Set]). Diese Zahl der Patienten war in der Studie nur geringfügig kleiner als die Zahl der randomisierten Patienten. Dadurch wird nicht von einem erhöhten Verzerrungspotenzial diesbezüglich ausgegangen. In die Analyse der UE gingen alle Ereignisse unter Behandlung sowie Ereignisse, die innerhalb von 28 Tagen nach Ende der Behandlung auftraten, ein.

Bezogen auf das relative Risiko, das Odds Ratio und die Risikodifferenz basierend auf naiven Proportionen von Patienten mit einem unerwünschten Ereignis kann es zu Verzerrungen aufgrund verschieden langer Behandlungsdauern (und damit auch Beobachtungsdauern) in den Therapiearmen kommen. Die mittlere Behandlungsdauer für Patienten mit Nintedanib+Docetaxel-Therapie betrug 5,4 Monate, während die Patienten in der

Placebo+Docetaxel-Gruppe im Mittel 4,7 Monate behandelt wurden (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a). So ist die Wahrscheinlichkeit, dass ein UE berichtet wird, im Nintedanib+Docetaxel-Arm durch die Verlängerung der Therapiedauer, aufgrund der Verlängerung der Zeit bis zur Progression, erhöht. Ein quantitativer Vergleich der beiden Behandlungsgruppen auf der Basis des relativen Risikos ist deshalb verzerrt und somit für den Nachweis und die Beurteilung des Ausmaßes eines Zusatznutzens weniger geeignet. Daher wurde – der Empfehlung des G-BA im Beratungsgespräch folgend (Gemeinsamer Bundesausschuss, 2014d) – für den Vergleich der UE zwischen beiden Behandlungsarmen die Zeit bis zum ersten Auftreten eines UE betrachtet und für die Bewertung des Zusatznutzens herangezogen.

Darüber hinaus waren die Endpunkterheber über die gesamte Studiendauer verblindet und es wurden keine weiteren, das Verzerrungspotenzial beeinflussenden Aspekte identifiziert. Aus diesem Grund ist das Verzerrungspotenzial für die Analysen der Zeit bis zum ersten Auftreten von UE als niedrig einzustufen.

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

Die Ergebnisse zu den UE aus der Studie LUME-Lung 1 sind in Tabelle 4-42 dargestellt.

Tabelle 4-42: Ergebnisse für die Zeit bis zum ersten Auftreten eines unerwünschten Ereignisses aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Tage)	N	n (%)	Median (Tage)	HR [95%-KI]	p-Wert
LUME-Lung 1	320	308 (96,3)	7,0	333	314 (94,3)	8,0	1,10 [0,94; 1,29]	0,2067
2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS								
HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Ereignis; TS: Treated Set.								
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert <0,05 fett geschrieben								

Bezüglich der Zeit bis zum ersten Auftreten eines UE zeigte sich kein statistisch signifikanter Unterschied zwischen einer Behandlung mit Nintedanib+Docetaxel und einer Behandlung mit Placebo+Docetaxel (Tabelle 4-42). Bei 96,3% der Patienten mit der Nintedanib+Docetaxel-Kombinationstherapie trat ein UE auf. Unter der Placebo+Docetaxel-Vergleichstherapie wurde für 94,3% der Patienten ein UE beobachtet.

Bei dieser Operationalisierung zeigte sich kein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen (HR [95%-KI]: 1,10 [0,94; 1,29]; p=0,2067).

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.8 Schwerwiegende unerwünschte Ereignisse – RCT

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

Tabelle 4-43: Operationalisierung von schwerwiegenden unerwünschten Ereignissen

Studie	Operationalisierung
LUME-Lung 1	Zeit bis zum ersten Auftreten eines SUE: Definiert als die Zeit von der Randomisierung bis zum ersten Auftreten eines SUE.
SUE: Schwerwiegendes 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 4-44: Bewertung des Verzerrungspotenzials für schwerwiegende unerwünschte Ereignisse 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die Analysen der SUE gelten dieselben Überlegungen wie für die Gesamtrate der UE in Abschnitt 4.3.1.3.1.7. Deshalb wurde auch für diesen Endpunkt von einem niedrigen Verzerrungspotenzial ausgegangen.

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

Die Ergebnisse zu den SUE aus der Studie LUME-Lung 1 sind in Tabelle 4-45 dargestellt.

Tabelle 4-45: Ergebnisse für die Zeit bis zum ersten Auftreten eines schwerwiegenden unerwünschten Ereignisses aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Tage)	N	n (%)	Median (Tage)	HR [95%-KI]	p-Wert
LUME-Lung 1	320	111 (34,7)	n.b.	333	107 (32,1)	n.b.	1,01 [0,78; 1,32]	0,9316

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS

HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Ereignis; n.b.: nicht berechenbar; TS: Treated Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben

Bei Betrachtung der SUE (Tabelle 4-45) zeigte sich bezüglich der Zeit bis zum ersten Auftreten eines SUE kein statistisch signifikanter Unterschied zwischen Nintedanib+Docetaxel und Placebo+Docetaxel (HR [95%-KI]: 1,01 [0,78; 1,32]; p=0,9316). Trotz der unterschiedlichen Behandlungsdauern (und damit auch Beobachtungsdauern) war die Häufigkeit der Ereignisse in beiden Behandlungsgruppen ähnlich. Für 34,7% der Patienten mit der Kombinationstherapie aus Nintedanib+Docetaxel und für 32,1% der Vergleichstherapie-Patienten wurde ein SUE dokumentiert.

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.9 Unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 – RCT

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

Tabelle 4-46: Operationalisierung von unerwünschten Ereignissen mit CTCAE-Grad ≥ 3

Studie	Operationalisierung
LUME-Lung 1	Zeit bis zum ersten Auftreten eines UE mit CTCAE-Grad ≥ 3 : Definiert als die Zeit von der Randomisierung bis zum ersten Auftreten eines UE mit CTCAE-Grad ≥ 3 .
CTCAE: Common Terminology Criteria for Adverse Events; 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 4-47: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die Analysen der Zeit bis zum ersten Auftreten eines UE mit CTCAE-Grad ≥ 3 gelten dieselben Überlegungen wie für die Gesamtrate der UE in Abschnitt 4.3.1.3.1.7. Deshalb wurde auch für diesen Endpunkt von einem niedrigen Verzerrungspotenzial ausgegangen.

Stellen Sie die Ergebnisse für den Endpunkt unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Die Ergebnisse zu den UE mit CTCAE-Grad ≥ 3 aus der Studie LUME-Lung 1 sind in Tabelle 4-48 dargestellt.

Tabelle 4-48: Ergebnisse für die Zeit bis zum ersten Auftreten eines unerwünschten Ereignisses mit CTCAE-Grad ≥ 3 aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Tage)	N	n (%)	Median (Tage)	HR [95%-KI]	p-Wert
LUME- Lung 1	320	243 (75,9)	27,0	333	228 (68,5)	29,0	1,10 [0,92; 1,32]	0,2655
2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS								
HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Ereignis; TS: Treated Set.								
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben								

Bezüglich der UE vom CTCAE-Grad ≥ 3 ergab sich anhand der Zeit bis zum ersten Auftreten eines solchen UE kein statistisch signifikanter Unterschied zwischen einer Behandlung mit Nintedanib+Docetaxel und einer Behandlung mit Placebo+Docetaxel (Tabelle 4-48). Anhand der Zeit bis zum ersten Auftreten eines UE vom CTCAE-Grad ≥ 3 zeigte sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen Nintedanib+Docetaxel und Placebo+Docetaxel (HR [95%-KI]: 1,10 [0,92; 1,32]; p=0,2655).

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.10 Unerwünschte Ereignisse, die zum Therapieabbruch führten – RCT

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

Tabelle 4-49: Operationalisierung von UE, die zum Therapieabbruch führten

Studie	Operationalisierung
LUME-Lung 1	Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte: Definiert als die Zeit von der Randomisierung bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte.
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 4-50: Bewertung des Verzerrungspotenzials für 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die Analysen der UE, die zum Therapieabbruch führten, gelten dieselben Überlegungen wie für die Gesamtrate der UE in Abschnitt 4.3.1.3.1.7. Deshalb wurde auch für diesen Endpunkt von einem niedrigen Verzerrungspotenzial ausgegangen.

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.

Die Ergebnisse zu den UE, die zum Therapieabbruch führten, aus der Studie LUME-Lung 1 sind in Tabelle 4-51 dargestellt.

Tabelle 4-51: Ergebnisse für die Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, aus RCT mit dem zu bewertenden Arzneimittel

Studie	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Tage)	N	n (%)	Median (Tage)	HR [95%-KI]	p-Wert
LUME- Lung 1	320	67 (20,9)	n.b.	333	59 (17,7)	n.b.	1,08 [0,76; 1,54]	0,6557

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS
 HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Ereignis; n.b.: nicht berechenbar; TS: Treated Set.
 Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben

Für die Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, ergab sich anhand des HR kein statistisch signifikanter Unterschied zwischen Nintedanib+Docetaxel und Placebo+Docetaxel (Tabelle 4-51). Trotz der längeren Behandlungsdauern (und damit auch Beobachtungsdauern) im Nintedanib+Docetaxel-Arm, wurden UE, die zum Therapieabbruch führten für Nintedanib+Docetaxel-Patienten nur 3,2% häufiger dokumentiert (Nintedanib+Docetaxel: 20,9% vs. Placebo+Docetaxel: 17,7%). Bei Betrachtung der Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, zeigten sich keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsgruppen (HR [95%-KI]: 1,08 [0,76; 1,54]; p=0,6557).

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.1.11 Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ – RCT

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

Tabelle 4-52: Operationalisierung von Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$

Studie	Operationalisierung
LUME-Lung 1	Zeit bis zum ersten Auftreten von Nebenwirkungen mit einem CTCAE-Grad ≥ 3 , die in $\geq 5\%$ der Patienten eines Behandlungsarms auftraten, waren definiert als die Zeit von der Randomisierung bis zum ersten Auftreten einer der folgenden Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$: Diarrhö (PT), Dyspnoe (PT), febrile Neutropenie (PT), Neutropenie (SSC), Neutropenie (PT), Verminderung der neutrophilen Granulozyten (PT).
CTCAE: Common Terminology Criteria for Adverse Events; PT: Preferred Term; SSC: Special Search Category.	

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 4-53: Bewertung des Verzerrungspotenzials für Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ 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
LUME-Lung 1	Niedrig	Ja	Ja	Ja	Ja	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die Analysen der Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ gelten dieselben Überlegungen wie für die Gesamtrate der UE in Abschnitt 4.3.1.3.1.7. Deshalb wurde auch für diesen Endpunkt von einem niedrigen Verzerrungspotenzial ausgegangen.

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

Die Ergebnisse zu den Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der Studie LUME-Lung 1 sind in Tabelle 4-54 dargestellt.

Aufgrund der Zugehörigkeit der Ereignisse febrile Neutropenie, Neutropenie und Verminderung der neutrophilen Granulozyten zum medizinischen Konzept Neutropenie wird

zusätzlich die Special Search Category (SSC) „Neutropenie“ dargestellt, die alle Ereignisse umfasst.

Tabelle 4-54: Ergebnisse für die Zeit bis zum ersten Auftreten einer Nebenwirkung mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ (Studie LUME-Lung 1) aus RCT mit dem zu bewertenden Arzneimittel

Ereignis	Nintedanib+Docetaxel			Placebo+Docetaxel			Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel	
	N	n (%)	Median (Tage)	N	n (%)	Median (Tage)	HR [95%-KI]	p-Wert
Diarrhö (PT)	320	20 (6,3)	n.b.	333	12 (3,6)	n.b.	1,73 [0,85; 3,55]	0,1260
Dyspnoe (PT)	320	15 (4,7)	n.b.	333	20 (6,0)	n.b.	0,71 [0,36; 1,39]	0,3117
Neutropenie (SSC)	320	161 (50,3)	101,0	333	155 (46,5)	159,0	1,06 [0,85; 1,32]	0,5856
Febrile Neutropenie (PT)	320	23 (7,2)	n.b.	333	15 (4,5)	n.b.	1,56 [0,82; 3,00]	0,1732
Neutropenie (PT)	320	38 (11,9)	n.b.	333	45 (13,5)	n.b.	0,84 [0,54; 1,29]	0,4178
Verminderung der neutrophilen Granulozyten (PT)	320	116 (36,3)	n.b.	333	116 (34,8)	n.b.	1,03 [0,79; 1,33]	0,8245

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS
 HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl Patienten in der Analyse; n: Anzahl Patienten mit Ereignis; n.b.: nicht berechenbar; PT: Preferred Term; SSC: Special Search Category; TS: Treated Set.
 Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; p-Wert: aus Cox-Regression; p-Wert <0,05 fett geschrieben

Bei den Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ wurden alle UE berücksichtigt, die einen CTCAE-Grad ≥ 3 aufwiesen und in mindestens einer Behandlungsgruppe bei mindestens 5% der Patienten auftraten. Nach dieser Definition wurden die folgenden Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ identifiziert: Diarrhö (PT), Dyspnoe (PT), febrile Neutropenie (PT), Neutropenie (PT) sowie die Verminderung der neutrophilen Granulozyten (PT). Für keine dieser Nebenwirkungen zeigte sich ein statistisch signifikanter Effekt für die Zeit bis zum ersten Auftreten der Nebenwirkung (Tabelle 4-54).

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.

Eine Meta-Analyse wurde nicht durchgeführt.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.1.

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 Patientenzahlen innerhalb der Subgruppenkategorien, ein Ungleichgewicht zwischen den Subgruppenkategorien und die steigende Anzahl statistischer Tests erhöhen die Wahrscheinlichkeit für falsch positive signifikante Ergebnisse (siehe hierzu auch das IQWiG Methodenpapier 4.1, S. 123ff. [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013b]). Diese Aspekte müssen bei der Bewertung von Ergebnissen aus Subgruppenanalysen berücksichtigt werden.

Die betrachteten Subgruppenmerkmale umfassen Alter (<65 vs. ≥65 Jahre), Geschlecht (männlich vs. weiblich), geografische Region (Asien vs. Südafrika vs. Europa), ECOG-PS (0 vs. 1), Hirnmetastasen (Ja vs. Nein), Bevacizumab-Vorbehandlung (Ja vs. Nein), Zeit seit Beginn der Erstlinien-Therapie (<9 Monate vs. ≥9 Monate) sowie Raucherstatus (Nieraucher vs. aktiver Raucher/Ex-Raucher). Im Folgenden (Abschnitt 4.3.1.3.2.1 bis Abschnitt 4.3.1.3.2.6) werden die Ergebnisse der Subgruppenanalysen zusammengefasst. Im Sinne der Übersichtlichkeit werden zunächst die p-Werte der Interaktionstests dargestellt. Einzelergebnisse der untersuchten Subgruppen werden für diejenigen Endpunkte und

Subgruppen tabellarisch dargestellt, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation durch das jeweilige Merkmal ergab (p-Wert der Interaktionstests $<0,2$). Im Text werden speziell die Subgruppen diskutiert, bei denen sich Belege für Effektmodifikationen ($p < 0,05$) durch den subgruppenbildenden Faktor ergeben haben. Bei Hinweisen auf eine Effektmodifikation werden für solche Endpunkte Subgruppenresultate im Text beschrieben, bei denen ein signifikanter Unterschied zwischen den Behandlungsarmen in der gesamten Zulassungspopulation vorlag.

Im Hinblick auf die Interpretation der Subgruppenresultate seien folgende Bemerkungen zum Problem der Multiplizität vorangestellt:

Multiplizität

Die bereits oben erwähnte Einschränkung der Aussagekraft von Subgruppenanalysen, die sich aus dem multiplen Testen ergibt, soll anhand einiger Zahlen aus den im Folgenden beschriebenen Subgruppenanalysen konkretisiert werden. Berücksichtigt man die Anzahl der durchgeführten Subgruppenanalysen für alle Endpunkte, so kommt man auf eine Gesamtzahl von 512 durchgeführten Interaktionstests. Hiervon waren 97 zum Niveau 0,2 signifikant (davon 22 zum Niveau 0,05). Diese Zahlen liegen sogar unter den rein zufallsbedingt zu erwartenden Zahlen falsch positiver Ergebnisse. Bei 512 durchgeführten Tests werden 102 zum Niveau von 0,2 (davon 26 zum Niveau von 0,05) falsch positive Tests erwartet, wenn keinerlei Subgruppeneffekte vorliegen (Tabelle 4-55).

Tabelle 4-55: Anzahl der beobachteten und der zu erwartenden zufällig signifikanten Ergebnisse von Interaktionstests

Durchgeführte Interaktionstests N=448	Beobachtete Interaktionen in Kategorie (N)	Erwartete Interaktionen in Kategorie (N)	Differenz
p-Wert $<0,2$	97	102	- 5
p-Wert $<0,05$	22	26	- 4

N: Anzahl der Interaktionstests.

4.3.1.3.2.1 Subgruppenanalysen zu OS und PFS – RCT

In Tabelle 4-56 sind die p-Werte der Interaktionstests für die Endpunkte OS und PFS aus der LUME-Lung 1 Studie dargestellt. Die Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests $<0,2$), sind in Tabelle 4-57 zu finden.

Tabelle 4-56: p-Werte der Interaktionstests für die Endpunkte OS und PFS aus der Studie LUME-Lung 1

Endpunkt	Alter ¹	Ge- schlecht ²	Geogra- fische Region ³	ECOG- PS ⁴	Hirn- meta- stasen ⁵	Bevac- zumab- Vorbe- handlung ⁶	Zeit seit Beginn der Erstlinien- Therapie ⁷	Raucher- status ⁸
OS ⁹	0,7556	0,7942	0,4980	0,7854	0,1247	0,2412	0,4186	0,5112
PFS ⁹	0,4850	0,7990	0,4471	0,8733	0,3423	0,9412	0,0032	0,9170

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten ECOG-PS zur Baseline, Bevacizumab-Vorbehandlung und Hirnmetastasen zur Baseline
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Endpunkt fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der gesamten Zulassungspopulation ergab sich sowohl für das OS als auch für das PFS ein statistisch signifikanter Vorteil für die Nintedanib+Docetaxel-Therapie gegenüber der Docetaxel-Vergleichstherapie (siehe Abschnitt 4.3.1.3.1.1 und Abschnitt 4.3.1.3.1.2).

Die Subgruppenanalysen (Tabelle 4-56) zeigten für das OS ausschließlich einen Hinweis auf eine Interaktion bezüglich des Vorhandenseins von Hirnmetastasen zur Baseline (Interaktionstest: p=0,1247). Die Subgruppenergebnisse zum OS zeigten in der kleinen Subgruppe (8,1% [Nintedanib+Docetaxel] bzw. 6,8% [Placebo+Docetaxel]) der Patienten mit Hirnmetastasen zur Baseline keinen statistisch signifikanten Unterschied zwischen beiden Behandlungsgruppen (Tabelle 4-57). Für die Patienten ohne Hirnmetastasen ergab sich ein statistisch signifikanter Überlebensvorteil der Nintedanib+Docetaxel-Therapie gegenüber der Vergleichstherapie Placebo+Docetaxel (HR [95%-KI]: 0,80 [0,67; 0,96]). Dieser Vorteil für die Nintedanib+Docetaxel-Therapie ist mit dem Überlebensvorteil in der gesamten Zulassungspopulation vergleichbar.

Für das PFS ergab sich ein Beleg auf eine Effektmodifikation anhand der Subgruppe Zeit seit Beginn der Erstlinien-Therapie (Tabelle 4-56; Interaktionstest: p=0,0032). Für die Patienten, deren Beginn der Erstlinien-Therapie weniger als 9 Monate zurücklag, zeigte sich ein statistisch signifikanter Effekt zugunsten der Kombinationstherapie Nintedanib+Docetaxel (HR [95%-KI]: 0,68 [0,54; 0,84]). Für Patienten, deren Beginn der Erstlinien-Therapie 9 Monate oder länger zurücklag, war kein statistisch signifikanter Unterschied im progressionsfreien Überleben zwischen beiden Behandlungsgruppen erkennbar.

Alle Einzelergebnisse der Subgruppenanalysen für die Endpunkte OS und PFS, für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-57 dargestellt.

Tabelle 4-57: Einzelergebnisse der untersuchten Subgruppen für die Endpunkte OS und PFS aus der Studie LUME-Lung 1

Endpunkt	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+ Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
OS	<i>Hirnmetastasen (Interaktionstest: p=0,1247)</i>					
	Ja	26	24 (92,3)	23	18 (78,3)	1,27 [0,67; 2,38]
	Nein	296	235 (79,4)	313	258 (82,4)	0,80 [0,67; 0,96]
PFS	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,0032)</i>					
	<9 Monate	206	164 (79,6)	199	164 (82,4)	0,68 [0,54; 0,84]
	≥9 Monate	112	88 (78,6)	134	101 (75,4)	1,11 [0,83; 1,48]

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

HR: Hazard Ratio aus (stratifizierter) Cox-Regression (Stratifizierungsvariablen: ECOG-PS zur Baseline, Bevacizumab-Vorbehandlung, Hirnmetastasen zur Baseline); KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); RS: Randomized Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)

4.3.1.3.2.2 Subgruppenanalysen zu ORR und DCR – RCT

In Tabelle 4-58 sind die p-Werte der Interaktionstests für die Endpunkte ORR und DCR aus der LUME-Lung 1 Studie dargestellt. Die Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests <0,2), sind in Tabelle 4-59 zu finden.

Tabelle 4-58: p-Werte der Interaktionstests für die Endpunkte ORR und DCR aus der Studie LUME-Lung 1

Endpunkt	Alter ¹	Geschlecht ²	Geografische Region ³	ECOG-PS ⁴	Hirnmetastasen ⁵	Bevacizumab-Vorbehandlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucherstatus ⁸
ORR ⁹	0,1107	0,6949	0,6773	0,4132	0,6421	1,0000	0,0626	0,5965
DCR ⁹	0,2125	0,1612	0,1735	0,5870	0,4767	0,2982	0,0007	0,9906

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

DCR: Disease Control Rate (Krankheitskontrollrate); ECOG: Eastern Cooperative Oncology Group; ORR: Objective Response Rate (objektive Ansprechrate); PS: Performance Status (Allgemeinzustand); RS: Randomized Set.

¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ p-Werte des Q-Tests unter Benutzung des Mantel-Haenszel-Schätzers des RR

p-Wert <0,05 fett geschrieben
p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben

Endpunkt fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Die Ergebnisse für die gesamten Zulassungspopulation zeigten für die ORR keinen statistisch signifikanten Unterschied zwischen den beiden Behandlungsgruppen und für die DCR einen statistisch signifikanten Vorteil für die Nintedanib+Docetaxel-Therapie (siehe Abschnitt 4.3.1.3.1.3 und Abschnitt 4.3.1.3.1.4).

Die Subgruppenanalysen für den Endpunkt ORR ergaben anhand der p-Werte der Interaktionstests (Tabelle 4-58) Hinweise auf eine Interaktion bezüglich des Alters der Patienten (Interaktionstest: $p=0,1107$) sowie bezüglich der Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: $p=0,0626$). In der gesamten Zulassungspopulation lag bezüglich der ORR kein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen vor.

Für die DCR zeigten sich Hinweise auf eine Interaktion bezüglich des Geschlechts (Interaktionstest: $p=0,1612$) und der geografischen Region (Interaktionstest: $p=0,1735$). Zudem lag ein Beleg auf eine Effektmodifikation nach der Zeit seit Beginn der Erstlinien-Therapie vor (Interaktionstest: $p=0,0007$). Bezüglich der Subgruppe Geschlecht ergab sich ein statistisch signifikanter Vorteil der Nintedanib+Docetaxel-Therapie für die männlichen Patienten (RR [95%-KI]: 1,49 [1,23; 1,81]). Der Vorteil der Kombinationstherapie Nintedanib+Docetaxel zeigte sich numerisch auch in der weiblichen Subgruppe. Der Unterschied erwies sich aber nicht als statistisch signifikant. Die Subgruppenergebnisse der DCR bezüglich der geografischen Region deuteten ebenfalls in allen Regionen auf einen Vorteil der Nintedanib+Docetaxel-Behandlung hin. Der Effekt war aber nur in der größten Region (Europa) statistisch signifikant zugunsten von Nintedanib+Docetaxel (RR [95%-KI]: 1,50 [1,25; 1,81]). Anhand der Subgruppen nach der Zeit seit Beginn der Erstlinien-Therapie zeigte sich ein statistisch signifikanter Vorteil für die Nintedanib+Docetaxel-Therapie für Patienten, deren Beginn der Erstlinien-Therapie weniger als 9 Monate zurücklag (RR [95%-KI]: 1,79 [1,42; 2,24]). In der Subgruppe der Patienten, deren Beginn der Erstlinien-Therapie 9 Monate oder länger zurücklag, ergab sich numerisch ebenfalls ein Vorteil für die Nintedanib+Docetaxel-Patienten. Dieser war jedoch nicht statistisch signifikant. Insgesamt bestätigen die Subgruppenanalysen zur DCR den statistisch signifikanten Effekt zugunsten von Nintedanib+Docetaxel in der gesamten Zulassungspopulation.

Alle Einzelergebnisse der Subgruppenanalysen für die Endpunkte ORR und DCR, für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-59 dargestellt.

Tabelle 4-59: Einzelergebnisse der untersuchten Subgruppen für die Endpunkte ORR und DCR aus der Studie LUME-Lung 1

Endpunkt	Subgruppe Subgruppenkategorie	Nintedanib + Docetaxel		Placebo + Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel RR [95%-KI]
		N	n (%)	N	n (%)	
ORR	<i>Alter (Interaktionstest: p=0,1107)</i>					
	<65 Jahre	232	13 (5,6)	240	7 (2,9)	1,92 [0,78; 4,73]
	≥65 Jahre	90	2 (2,2)	96	5 (5,2)	0,43 [0,08; 2,14]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,0626)</i>					
	<9 Monate	206	10 (4,9)	199	3 (1,5)	3,22 [0,90; 11,53]
	≥9 Monate	112	5 (4,5)	134	9 (6,7)	0,66 [0,23; 1,93]
DCR	<i>Geschlecht (Interaktionstest: p=0,1612)</i>					
	Weiblich	119	69 (58,0)	128	62 (48,4)	1,20 [0,95; 1,51]
	männlich	203	125 (61,6)	208	86 (41,3)	1,49 [1,23; 1,81]
	<i>Geografische Region (Interaktionstest: p=0,1735)</i>					
	Asien	86	47 (54,7)	96	48 (50,0)	1,09 [0,83; 1,44]
	Europa	229	141 (61,6)	234	96 (41,0)	1,50 [1,25; 1,81]
	Südafrika	7	6 (85,7)	6	4 (66,7)	1,29 [0,68; 2,44]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,0007)</i>					
<9 Monate	206	122 (59,2)	199	66 (33,2)	1,79 [1,42; 2,24]	
	≥9 Monate	112	71 (63,4)	134	80 (59,7)	1,06 [0,87; 1,29]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
DCR: Disease Control Rate (Krankheitskontrollrate); KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; ORR: Objective Response Rate (objektive Ansprechrage); RR: Risk Ratio; RS: Randomized Set.						
Ein RR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

4.3.1.3.2.3 Subgruppenanalysen zur Symptomatik – RCT

In Tabelle 4-60 (EORTC QLQ-C30) und Tabelle 4-62 (EORTC QLQ-LC13) sind die p-Werte der Interaktionstests für die Verbesserung der Symptomatik aus der Studie LUME-Lung 1 dargestellt. Die dazugehörigen Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests <0,2), sind in Tabelle 4-61 und Tabelle 4-63 zu finden. Für die Operationalisierung Zeit bis zur Verschlechterung der Symptomatik werden die Subgruppenergebnisse analog in Tabelle 4-64 bis Tabelle 4-67 präsentiert.

Operationalisierung: Verbesserung der Symptomatik (EORTC QLQ-C30, Responderanalysen)

Tabelle 4-60: p-Werte der Interaktionstests für die Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1

Symptom	Alter ¹	Ge- schlecht ²	Geogra- fische Region ³	ECOG- PS ⁴	Hirn- meta- stasen ⁵	Bevac- zumab- Vorbe- handlung ⁶	Zeit seit Beginn der Erstlinien- Therapie ⁷	Raucher- status ⁸
Appetitverlust ⁹	0,9157	0,4838	0,6455	0,9369	0,0468	0,7635	0,3059	0,2103
Diarrhö ⁹	0,4508	0,3683	0,7799	0,6021	0,5157	0,3379	0,7497	0,3248
Dyspnoe ⁹	0,1778	0,9846	0,4528	0,6976	0,0378	0,0844	0,0609	0,0257
Fatigue ⁹	0,6182	0,6042	0,1140	0,2987	0,9359	0,1447	0,0210	0,7521
Schlaflosigkeit ⁹	0,5982	0,6814	0,3883	0,2579	0,0952	0,8410	0,4092	0,5047
Schmerzen⁹	0,2634	0,5252	0,1112	0,3860	0,8076	0,1018	0,2771	0,7994
Übelkeit und Erbrechen ⁹	0,4815	0,5223	0,4827	0,8740	0,4197	0,7049	0,4653	0,8774
Verstopfung ⁹	0,0768	0,4711	0,9738	0,1867	0,2463	0,9649	0,6024	0,5960

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ p-Werte des Q-Tests unter Benutzung des Mantel-Haenszel-Schätzers des RR
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Symptom fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der gesamten Zulassungspopulation ergab die Responderanalyse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30, nur für das Symptom Schmerzen ein statistisch signifikantes Ergebnis. Bezüglich des Symptoms Schmerzen zeigte sich ein statistisch signifikanter Vorteil zugunsten der Kombinationstherapie Nintedanib+Docetaxel (siehe Abschnitt 4.3.1.3.1.5).

Die Subgruppenanalysen zur Verbesserung der Symptomatik (EORTC QLQ-C30) ergaben anhand der p-Werte der Interaktionstests für das Symptom Diarrhö und für das Symptom Übelkeit und Erbrechen keine Hinweise auf eine Interaktion. Für alle anderen Symptome des Fragebogens EORTC QLQ-C30 lagen Hinweise oder Belege auf eine Interaktion vor (Tabelle 4-60).

Appetitverlust

Für das Symptom Appetitverlust zeigte sich ein Beleg auf eine Interaktion anhand der Subgruppe Hirnmetastasen zur Baseline (Interaktionstest: p=0,0468). Für Patienten mit Hirnmetastasen zur Baseline lag ein statistisch signifikanter Effekt zum Nachteil von

Nintedanib+Docetaxel gegenüber Placebo+Docetaxel vor (RR [95%-KI]: 0,28 [0,09; 0,90]). Für die Gruppe der Patienten ohne Hirnmetastasen, wie auch für die gesamte Zulassungspopulation, war kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen zu beobachten (Tabelle 4-61).

Dyspnoe

Bezüglich des Symptoms Dyspnoe ergaben sich Belege für eine Effektmodifikation anhand der Subgruppen Hirnmetastasen zur Baseline (Interaktionstest: $p=0,0378$) und Raucherstatus (Interaktionstest: $p=0,0257$). Die Einzelergebnisse zur Subgruppe Hirnmetastasen zeigten einen statistisch signifikanten Effekt in der Subgruppe ohne Hirnmetastasen zum Nachteil von Nintedanib+Docetaxel. In der Subgruppe mit Hirnmetastasen als auch in der gesamten Zulassungspopulation war der Effekt nicht statistisch signifikant (Tabelle 4-61). In der Gruppe der Nieraucher ergab sich ebenfalls ein statistisch signifikanter Nachteil für Nintedanib+Docetaxel. Zudem lagen für Dyspnoe Hinweise auf eine Effektmodifikation bezüglich der Subgruppen Alter (Interaktionstest: $p=0,1778$), Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0844$) und Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: $p=0,0609$) vor.

Fatigue

Für das Symptom Fatigue ergaben die Subgruppenanalysen einen Hinweis auf Interaktion anhand der Subgruppen geografische Region (Interaktionstest: $p=0,1140$) und Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1447$). Ein Beleg für eine Effektmodifikation lag bezüglich der Subgruppe der Zeit seit Beginn der Erstlinien-Therapie vor (Interaktionstest: $p=0,0210$). Die Einzeleffekte in den Subgruppen waren jedoch, wie auch schon der Effekt für die gesamte Zulassungspopulation, nicht statistisch signifikant (Tabelle 4-61).

Schmerzen

Für das Symptom Schmerzen zeigte sich bei Betrachtung der gesamten Zulassungspopulation ein statistisch signifikanter Vorteil für die Patienten mit Nintedanib+Docetaxel-Therapie. Hinweise auf eine Modifikation dieses Effektes liegen für die Subgruppenanalysen nach geografischer Region (Interaktionstest: $p=0,1112$) und Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1018$) vor. In der Region Europa ergab sich, wie in der gesamten Zulassungspopulation, ein statistisch signifikanter Vorteil für Nintedanib+Docetaxel-Patienten bezüglich des Symptoms Schmerzen (RR [95%-KI]: 1,96 [1,29; 2,98]). Für die Region Asien zeigte sich ebenfalls ein numerischer Vorteil für Nintedanib+Docetaxel, dieser war jedoch nicht statistisch signifikant. In der sehr kleinen Subpopulation Südafrika zeigte sich kein statistisch signifikanter Effekt. Bezüglich der Bevacizumab-Vorbehandlung lag in der Gruppe der nicht vorbehandelten Patienten ebenfalls ein statistisch signifikanter Vorteil für die Kombinationstherapie Nintedanib+Docetaxel vor (RR [95%-KI]: 1,67 [1,20; 2,32]), während der Effekt in der kleinen Gruppe der mit Bevacizumab vorbehandelten Patienten keine statistische Signifikanz aufwies (Tabelle 4-61).

Zudem zeigte sich für das Symptom Schlaflosigkeit ein Hinweis auf eine Interaktion bezüglich vorhandener Hirnmetastasen (Interaktionstest: $p=0,0952$). Für das Symptom

Verstopfung lag ein Hinweis auf eine Interaktion für die Subgruppen nach Alter (Interaktionstest: $p=0,0768$) und ECOG-PS (Interaktionstest: $p=0,1867$) vor. Für beide Symptome hatten die Analysen für die gesamte Zulassungspopulation keine statistisch signifikanten Ergebnisse gezeigt.

Die Einzelergebnisse der Subgruppenanalysen zur Verbesserung der Symptome (EORTC QLQ-C30), für die sich ein Hinweis oder Beleg ergab, werden alle in Tabelle 4-61 dargestellt.

Tabelle 4-61: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel RR [95%-KI]
		N	n (%)	N	n (%)	
Appetitverlust	<i>Hirnmetastasen (Interaktionstest: $p=0,0468$)</i>					
	Ja	25	3 (12,0)	21	9 (42,9)	0,28 [0,09; 0,90]
	Nein	263	49 (18,6)	270	52 (19,3)	0,97 [0,68; 1,37]
Dyspnoe	<i>Alter (Interaktionstest: $p=0,1778$)</i>					
	<65 Jahre	209	45 (21,5)	209	63 (30,1)	0,71 [0,51; 0,99]
	≥65 Jahre	80	22 (27,5)	83	21 (25,3)	1,09 [0,65; 1,82]
	<i>Hirnmetastasen (Interaktionstest: $p=0,0378$)</i>					
	Ja	25	9 (36,0)	22	3 (13,6)	2,64 [0,82; 8,54]
	Nein	264	58 (22,0)	270	81 (30,0)	0,73 [0,55; 0,98]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0844$)</i>					
	Ja	23	2 (8,7)	16	6 (37,5)	0,23 [0,05; 1,01]
	Nein	266	65 (24,4)	276	78 (28,3)	0,86 [0,65; 1,15]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: $p=0,0609$)</i>					
	<9 Monate	180	37 (20,6)	168	53 (31,5)	0,65 [0,45; 0,94]
	≥9 Monate	105	30 (28,6)	121	31 (25,6)	1,12 [0,73; 1,71]
<i>Raucherstatus (Interaktionstest: $p=0,0257$)</i>						
Nieraucher	98	16 (16,3)	100	33 (33)	0,49 [0,29; 0,84]	
aktiver Raucher/Ex-Raucher	191	51 (26,7)	192	51 (26,6)	1,01 [0,72; 1,40]	
Fatigue	<i>Geografische Region (Interaktionstest: $p=0,1140$)</i>					
	Asien	74	15 (20,3)	82	24 (29,3)	0,69 [0,39; 1,22]
	Europa	208	41 (19,7)	204	29 (14,2)	1,39 [0,90; 2,14]
	Südafrika	7	1 (14,3)	6	2 (33,3)	0,43 [0,05; 3,64]

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel RR [95%-KI]
		N	n (%)	N	n (%)	
Fatigue	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1447)</i>					
	Ja	23	3 (13)	16	5 (31,3)	0,42 [0,12; 1,50]
	Nein	266	54 (20,3)	276	50 (18,1)	1,12 [0,79; 1,58]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,0210)</i>					
	<9 Monate	180	36 (20)	168	21 (12,5)	1,60 [0,97; 2,63]
	≥9 Monate	105	21 (20)	121	34 (28,1)	0,71 [0,44; 1,15]
Schlaf- losigkeit	<i>Hirnmetastasen (Interaktionstest: p=0,0952)</i>					
	Ja	25	6 (24,0)	22	1 (4,5)	5,28 [0,69; 40,53]
	Nein	264	69 (26,1)	270	77 (28,5)	0,92 [0,69; 1,21]
Schmerzen	<i>Geografische Region (Interaktionstest: p=0,1112)</i>					
	Asien	74	23 (31,1)	82	21 (25,6)	1,21 [0,74; 2,00]
	Europa	208	54 (26,0)	204	27 (13,2)	1,96 [1,29; 2,98]
	Südafrika	7	0 (0,0)	6	2 (33,3)	0,17 [0,01; 2,99]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1018)</i>					
	Ja	23	3 (13)	16	4 (25)	0,52 [0,13; 2,02]
	Nein	266	74 (27,8)	276	46 (16,7)	1,67 [1,20; 2,32]
Verstopfung	<i>Alter (Interaktionstest: p=0,0768)</i>					
	<65 Jahre	206	34 (16,5)	200	31 (15,5)	1,06 [0,68; 1,66]
	≥65 Jahre	79	26 (32,9)	82	13 (15,9)	2,08 [1,15; 3,74]
	<i>ECOG-PS (Interaktionstest: p=0,1867)</i>					
	0	89	18 (20,2)	84	8 (9,5)	2,12 [0,98; 4,62]
	1	196	42 (21,4)	198	36 (18,2)	1,18 [0,79; 1,76]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
ECOG: Eastern Cooperative Oncology Group; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; PS: Performance Status (Allgemeinzustand); RR: Relatives Risiko; RS: Randomized Set.						
Ein RR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

Operationalisierung: Verbesserung der Symptomatik (EORTC QLQ-LC13, Responderanalysen)

Tabelle 4-62: p-Werte der Interaktionstests für die Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalysen) aus der LUME-Lung 1 Studie

Symptom	Alter ¹	Ge- schlecht ²	Geogra- fische Region ³	ECOG- PS ⁴	Hirn- meta- stasen ⁵	Bevaci- zumab- Vorbe- handlung ⁶	Zeit seit Beginn der Erstlinien- Therapie ⁷	Raucher- status ⁸
Bluthusten ⁹	0,5619	0,3436	0,0225	0,9415	0,5821	0,3417	0,8130	0,9729
Dyspnoe ⁹	0,2871	0,9772	0,2952	0,0221	0,2121	0,1983	0,1305	0,6953
Haarausfall ⁹	0,8318	0,8215	0,8873	0,7235	0,1464	0,9986	0,6401	0,7659
Husten ⁹	0,5835	0,1217	0,0699	0,7847	0,8608	0,9249	0,2084	0,8455
Mund- schmerzen ⁹	0,5300	0,1515	0,4371	0,9142	0,8769	0,1650	0,6918	0,3948
Periphere Neuropathie ⁹	0,2457	0,4865	0,2327	0,9696	0,6811	0,6710	0,2376	0,2932
Schluck- beschwerden ⁹	0,5083	0,8146	0,6041	0,7402	0,6650	0,6880	0,7178	0,2393
Schmerzen (Arm/Schulter) ⁹	0,4206	0,4714	0,4307	0,3959	0,3742	0,4371	0,5143	0,0115
Schmerzen (Brust) ⁹	0,7350	0,3396	0,8012	0,5372	0,1125	0,7112	0,1030	0,0862
Schmerzen (andere) ⁹	0,9323	0,2570	0,1438	0,0945	0,9923	0,5694	0,5081	0,3491

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ p-Werte des Q-Tests unter Benutzung des Mantel-Haenszel-Schätzers des RR
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Symptom fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der gesamten Zulassungspopulation zeigte sich kein statistisch signifikanter Effekt bezüglich der Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (siehe Abschnitt 4.3.1.3.1.5).

Bei Betrachtung der Subgruppen ergaben sich für die Symptome periphere Neuropathie, Schluckbeschwerden und Schmerzen (Arm/Schulter) keine Hinweise auf Effektmodifikationen. Für alle anderen Symptome des Fragebogens EORTC QLQ-LC13 lag mindestens ein Hinweis auf Interaktion vor (Tabelle 4-62).

Bluthusten

Für das Symptom Bluthusten zeigte sich ein Beleg auf eine Effektmodifikation bezüglich der geografischen Region (Interaktionstest: $p=0,0225$). Für die einzelnen Subgruppen ergaben sich, wie auch in der gesamten Zulassungspopulation, keine statistisch signifikanten Effekte (Tabelle 4-63).

Dyspnoe

Für das Symptom Dyspnoe lagen Hinweise auf eine Effektmodifikation bezüglich der Subgruppen Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1983$) und Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: $p=0,1305$) vor. Der Interaktionstest bezüglich der Subgruppe ECOG-PS wies darüber hinaus auf einen Beleg für eine Effektmodifikation hin (Interaktionstest: $p=0,0221$). Für die Patienten mit einem ECOG-PS von 0 ergab sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (RR [95%-KI]: 0,46 [0,22; 0,95]). Der Effekt in der Gruppe der Patienten mit einem ECOG-PS von 1 weist numerisch eine stärkere Verbesserung für Nintedanib+Docetaxel-Patienten aus. Dieser Effekt zeigte aber, ebenso wie in der gesamten Zulassungspopulation, keine statistische Signifikanz (Tabelle 4-63).

Schmerzen (Arm/Schulter)

Für das Symptom Schmerzen (Arm/Schulter) ergab sich ein Beleg für eine Effektmodifikation bezüglich der Subgruppen Raucherstatus (Interaktionstest: $p=0,0115$). In der Gruppe der Nieraucher war ein statistisch signifikant besseres Ergebnis für Nintedanib+Docetaxel zu beobachten. Es lag jedoch kein statistisch signifikanter Effekt in der gesamten Zulassungspopulation vor.

Zudem lag für das Symptom Haarausfall ein Hinweis auf Effektmodifikation anhand der Subgruppe Hirnmetastasen zur Baseline vor (Interaktionstest: $p=0,1464$) und für das Symptom Husten wies der p-Wert des Interaktionstests bezüglich der geografischen Region auf eine Interaktion hin (Interaktionstest: $p=0,0699$). Bezüglich der Mundschmerzen wurde ein Hinweis auf eine Effektmodifikation für das Geschlecht (Interaktionstest: $p=0,1515$) und die Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1650$) identifiziert. Für die Schmerzen (Brust) wies der Interaktionstest auf eine Effektmodifikation nach der Subgruppe Hirnmetastasen (Interaktionstest: $p=0,1125$), bezüglich der Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: $p=0,1030$) und bezüglich des Raucherstatus (Interaktionstest: $p=0,0862$) hin. Für das Symptom Schmerzen (andere) lag ein Hinweis bezüglich der geografischen Region (Interaktionstest: $p=0,1438$) und des ECOG-PS (Interaktionstest: $p=0,0945$) vor. Bei Betrachtung der gesamten Zulassungspopulation hatte sich jedoch für keines dieser Symptome ein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen gezeigt.

Alle Einzelergebnisse der Subgruppenanalysen zur Verbesserung der Symptome (EORTC QLQ-LC13), für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-63 dargestellt.

Tabelle 4-63: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 (Responderanalysen) aus der Studie LUME-Lung 1

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel RR [95%-KI]
		N	n (%)	N	n (%)	
Bluthusten	<i>Geografische Region (Interaktionstest: p=0,0225)</i>					
	Asien	74	7 (9,5)	82	2 (2,4)	3,88 [0,83; 18,09]
	Europa	205	3 (1,5)	201	8 (4,0)	0,37 [0,10; 1,37]
	Südafrika	7	0 (0,0)	6	0 (0,0)	n.b.
Dyspnoe	<i>ECOG-PS (Interaktionstest: p=0,0221)</i>					
	0	88	9 (10,2)	85	19 (22,4)	0,46 [0,22; 0,95]
	1	196	40 (20,4)	204	34 (16,7)	1,22 [0,81; 1,85]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1983)</i>					
	Ja	23	2 (8,7)	16	4 (25,0)	0,35 [0,07; 1,68]
	Nein	261	47 (18,0)	273	49 (17,9)	1,00 [0,70; 1,44]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,1305)</i>					
	<9 Monate	175	25 (14,3)	165	31 (18,8)	0,76 [0,47; 1,23]
	≥9 Monate	105	24 (22,9)	121	21 (17,4)	1,32 [0,78; 2,22]
Haarausfall	<i>Hirnmetastasen (Interaktionstest: p=0,1464)</i>					
	Ja	23	4 (17,4)	22	6 (27,3)	0,64 [0,21; 1,96]
	Nein	262	40 (15,3)	267	26 (9,7)	1,57 [0,99; 2,49]
Husten	<i>Geschlecht (Interaktionstest: p=0,1217)</i>					
	Weiblich	105	42 (40,0)	110	30 (27,3)	1,47 (1,00; 2,15)
	Männlich	181	49 (27,1)	181	50 (27,6)	0,98 (0,70; 1,37)
	<i>Geografische Region (Interaktionstest: p=0,0699)</i>					
	Asien	74	22 (29,7)	82	28 (34,1)	0,87 [0,55; 1,38]
	Europa	205	69 (33,7)	203	49 (24,1)	1,39 [1,02; 1,90]
	Südafrika	7	0 (0,0)	6	3 (50,0)	0,12 [0,01; 1,98]
Mund- schmerzen	<i>Geschlecht (Interaktionstest: p=0,1515)</i>					
	Weiblich	104	10 (9,6)	110	13 (11,8)	0,81 [0,37; 1,77]
	männlich	181	15 (8,3)	181	8 (4,4)	1,88 [0,82; 4,31]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1650)</i>					
	Ja	23	5 (21,7)	16	0 (0,0)	7,72 [0,46; 130,33]
	Nein	262	20 (7,6)	275	21 (7,6)	1,00 [0,55; 1,80]

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel RR [95%-KI]
		N	n (%)	N	n (%)	
Schmerzen (Arm/ Schulter)	<i>Raucherstatus (Interaktionstest: p=0,0115)</i>					
	Nieraucher	97	25 (25,8)	100	11 (11,0)	2,34 [1,22; 4,50]
	aktiver Raucher/Ex- Raucher	188	47 (25,0)	190	52 (27,4)	0,91 [0,65; 1,28]
Schmerzen (Brust)	<i>Hirnmetastasen (Interaktionstest: p=0,1125)</i>					
	Ja	24	7 (29,2)	22	2 (9,1)	3,21 [0,74; 13,83]
	Nein	261	51 (19,5)	268	55 (20,5)	0,95 [0,68; 1,34]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,1030)</i>					
	<9 Monate	177	39 (22,0)	167	28 (16,8)	1,31 [0,85; 2,03]
	≥9 Monate	104	18 (17,3)	120	28 (23,3)	0,74 [0,44; 1,26]
	<i>Raucherstatus (Interaktionstest: p=0,0862)</i>					
	Nieraucher	96	28 (29,2)	100	20 (20,0)	1,46 [0,88; 2,41]
	aktiver Raucher/Ex- Raucher	189	30 (15,9)	190	37 (19,5)	0,82 [0,53; 1,26]
Schmerzen (andere)	<i>Geografische Region (Interaktionstest: p=0,1438)</i>					
	Asien	69	17 (24,6)	76	17 (22,4)	1,10 [0,61; 1,98]
	Europa	194	49 (25,3)	196	35 (17,9)	1,41 [0,96; 2,08]
	Südafrika	7	1 (14,3)	6	4 (66,7)	0,21 [0,03; 1,43]
	<i>ECOG-PS (Interaktionstest: p=0,0945)</i>					
	0	85	20 (23,5)	81	9 (11,1)	2,12 [1,03; 4,37]
	1	185	47 (25,4)	197	47 (23,9)	1,06 [0,75; 1,51]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
ECOG: Eastern Cooperative Oncology Group; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; n.b.: nicht berechenbar; PS: Performance Status (Allgemeinzustand); RR: Relatives Risiko; RS: Randomized Set.						
Ein RR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

Operationalisierung: Zeit bis zur Verschlechterung der Symptomatik (EORTC QLQ-C30)

Tabelle 4-64: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 aus der Studie LUME-Lung 1

Symptom	Alter ¹	Ge-schlecht ²	Geogra-fische Region ³	ECOG-PS ⁴	Hirn-metastasen ⁵	Bevacizumab-Vor-behandlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucher-status ⁸
Appetitverlust ⁹	0,4485	0,8280	0,0719	0,9898	0,0429	0,5251	0,4430	0,5902
Diarrhö⁹	0,2856	0,5332	0,4928	0,2272	0,2577	0,4162	0,5658	0,3533
Dyspnoe ⁹	0,6574	0,2426	0,1482	0,3047	0,8679	0,4293	0,3720	0,1038
Fatigue ⁹	0,1665	0,8194	0,7409	0,1767	0,4840	0,0669	0,8074	0,1611
Schlaflosigkeit ⁹	0,5054	0,1351	0,2920	0,0996	0,4869	0,2435	0,2087	0,6551
Schmerzen ⁹	0,2646	0,8011	0,2490	0,0555	0,7877	0,6321	0,2593	0,4320
Übelkeit und Erbrechen⁹	0,6189	0,0607	0,6122	0,9243	0,0311	0,5986	0,6050	0,2921
Verstopfung ⁹	0,1019	0,1487	0,2250	0,9985	0,0723	0,0132	0,9091	0,5873

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten Behandlung, Subgruppe und Behandlung nach Subgruppe
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Symptom fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der gesamten Zulassungspopulation bezüglich der Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30, zeigte sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel für die Symptome Diarrhö sowie Übelkeit und Erbrechen (siehe Abschnitt 4.3.1.3.1.5).

Bei Betrachtung der Subgruppen ergab sich für das Symptom Diarrhö kein Hinweis oder Beleg auf Effektmodifikationen. Für alle anderen Symptome des Fragebogens EORTC QLQ-C30 lag mindestens ein Hinweis auf Interaktion vor (Tabelle 4-64).

Appetitverlust

Ein Beleg auf Interaktion zeigte sich für das Symptom Appetitverlust bezüglich der Subgruppe Hirnmetastasen zur Baseline (Interaktionstest: p=0,0429). Anhand der Einzelergebnisse für die Subgruppen zeigte sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel in der Subgruppe der Patienten mit Hirnmetastasen (HR [95%-KI]: 2,35 [1,12; 4,94]). In der Gruppe der Patienten ohne Hirnmetastasen lag kein

statistisch signifikanter Unterschied zwischen beiden Behandlungsgruppen vor (Tabelle 4-65). Auch in der gesamten Zulassungspopulation hatte sich kein statistisch signifikanter Effekt gezeigt. Für das Symptom Appetitverlust lag zudem ein Hinweis auf eine Effektmodifikation nach der geografischen Region vor (Interaktionstest: $p=0,0719$).

Übelkeit und Erbrechen

Ein Beleg auf eine Effektmodifikation anhand der Subgruppe Hirnmetastasen zur Baseline ergab sich für das Symptom Übelkeit und Erbrechen (Interaktionstest: $p=0,0311$). In der Subgruppe der Patienten mit Hirnmetastasen zur Baseline war die Zeit bis zur Verschlechterung des Symptoms Übelkeit und Erbrechen im Nintedanib+Docetaxel-Arm kürzer als im Placebo+Docetaxel-Arm. Der Effekt mit einem HR [95%-KI] von 2,72 [1,26; 5,86] war statistisch signifikant. Für Patienten ohne Hirnmetastasen zeigte sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (Tabelle 4-65). Für das Symptom Übelkeit und Erbrechen lag zudem ein Hinweis auf eine Effektmodifikation bezüglich des Geschlechts vor (Interaktionstest: $p=0,0607$). In der weiblichen Subgruppe war die Zeit bis zur Verschlechterung dieses Symptoms in der Nintedanib+Docetaxel-Gruppe gegenüber der Placebo+Docetaxel-Gruppe statistisch signifikant kürzer (HR [95%-KI]: 1,56 [1,12; 2,17]). Für männliche Patienten war der Effekt nicht statistisch signifikant.

Verstopfung

Für das Symptom Verstopfung wies der p-Wert des Interaktionstests auf einen Beleg für eine Effektmodifikation nach der Subgruppe Bevacizumab-Vorbehandlung hin (Interaktionstest: $p=0,0132$). Die Auswertung der Einzelergebnisse für diese Subgruppen ergab einen statistisch signifikanten Effekt zugunsten von Nintedanib+Docetaxel für vorbehandelte Patienten (HR [95%-KI]: 0,33 [0,13; 0,82]). Der Effekt war in der Gruppe der nicht vorbehandelten Patienten gleichgerichtet, aber nicht statistisch signifikant (Tabelle 4-65). Zudem zeigten die Interaktionstests Hinweise für eine Effektmodifikation bezüglich der Subgruppen Alter (Interaktionstest: $p=0,1019$), Geschlecht (Interaktionstest: $p=0,1487$) und Hirnmetastasen zur Baseline (Interaktionstest: $p=0,0723$). In der gesamten Zulassungspopulation lag jedoch kein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen bezüglich des Symptoms Verstopfung vor.

Weitere Hinweise auf eine Effektmodifikation ergaben sich für die Dyspnoe bezüglich der geografischen Region (Interaktionstest: $p=0,1482$) und des Raucherstatus (Interaktionstest: $p=0,1038$), für die Schlaflosigkeit bezüglich des Geschlechts (Interaktionstest: $p=0,1351$) und des ECOG-PS (Interaktionstest: $p=0,0996$) sowie für die Schmerzen bezüglich des ECOG-PS (Interaktionstest: $p=0,0555$). Für das Symptom Fatigue ergab sich ein Hinweis auf Effektmodifikation nach dem Alter (Interaktionstest: $p=0,1665$), dem ECOG-PS (Interaktionstest: $p=0,1767$), der Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0669$) und dem Raucherstatus (Interaktionstest: $p=0,1611$). Bei Betrachtung der gesamten Zulassungspopulation hatte sich für keines dieser Symptome ein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen gezeigt. Alle Einzelergebnisse der Subgruppenanalysen zur Zeit bis zur Verschlechterung der Symptome (EORTC QLQ-C30), für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-65 dargestellt.

Tabelle 4-65: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30 aus der Studie LUME-Lung 1

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel
		N	n (%)	N	n (%)	HR [95%-KI]
Appetitverlust	<i>Geografische Region (Interaktionstest: p=0,0719)</i>					
	Asien	86	60 (69,8)	96	50 (52,1)	1,67 [1,14; 2,44]
	Europa	229	129 (56,3)	234	128 (54,7)	0,99 [0,78; 1,27]
	Südafrika	7	5 (71,4)	6	2 (33,3)	1,96 [0,38; 10,17]
	<i>Hirnmastasen (Interaktionstest: p=0,0429)</i>					
	Ja	26	21 (80,8)	23	11 (47,8)	2,35 [1,12; 4,94]
Nein	296	173 (58,4)	313	169 (54,0)	1,08 [0,87; 1,33]	
Dyspnoe	<i>Geografische Region (Interaktionstest: p=0,1482)</i>					
	Asien	86	49 (57)	96	45 (46,9)	1,47 [0,98; 2,21]
	Europa	229	122 (53,3)	234	117 (50,0)	0,92 [0,72; 1,19]
	Südafrika	7	6 (85,7)	6	4 (66,7)	1,87 [0,46; 7,55]
	<i>Raucherstatus (Interaktionstest: p=0,1038)</i>					
	Nieraucher	115	68 (59,1)	115	56 (48,7)	1,31 [0,92; 1,87]
aktiver Raucher/Ex-Raucher	207	109 (52,7)	221	110 (49,8)	0,91 [0,70; 1,19]	
Fatigue	<i>Alter (Interaktionstest: p=0,1665)</i>					
	<65 Jahre	232	158 (68,1)	240	151 (62,9)	1,05 [0,84; 1,32]
	≥65 Jahre	90	72 (80,0)	96	60 (62,5)	1,40 [1,00; 1,98]
	<i>ECOG-PS (Interaktionstest: p=0,1767)</i>					
	0	96	73 (76,0)	99	61 (61,6)	1,41 [1,00; 1,99]
	1	226	157 (69,5)	237	150 (63,3)	1,06 [0,84; 1,32]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0669)</i>					
	Ja	24	22 (91,7)	21	11 (52,4)	2,07 [0,96; 4,43]
	Nein	298	208 (69,8)	315	200 (63,5)	1,09 [0,90; 1,32]
	<i>Raucherstatus (Interaktionstest: p=0,1611)</i>					
Nieraucher	115	71 (61,7)	115	79 (68,7)	0,94 [0,68; 1,3]	
aktiver Raucher/Ex-Raucher	207	159 (76,8)	221	132 (59,7)	1,26 [1,00; 1,58]	
Schlaflosigkeit	<i>Geschlecht (Interaktionstest: p=0,1351)</i>					
	Weiblich	119	64 (53,8)	128	64 (50,0)	1,22 [0,86; 1,73]
	Männlich	203	112 (55,2)	208	107 (51,4)	0,87 [0,66; 1,13]

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel
		N	n (%)	N	n (%)	HR [95%-KI]
Schlaf- losigkeit	<i>ECOG-PS (Interaktionstest: p=0,0996)</i>					
	0	96	58 (60,4)	99	48 (48,5)	1,28 [0,87; 1,88]
	1	226	118 (52,2)	237	123 (51,9)	0,87 [0,68; 1,12]
Schmerzen	<i>ECOG-PS (Interaktionstest: p=0,0555)</i>					
	0	96	68 (70,8)	99	58 (58,6)	1,24 [0,87; 1,77]
	1	226	138 (61,1)	237	138 (58,2)	0,83 [0,65; 1,05]
Übelkeit und Erbrechen	<i>Geschlecht (Interaktionstest: p=0,0607)</i>					
	Weiblich	119	78 (65,5)	128	67 (52,3)	1,56 [1,12; 2,17]
	Männlich	203	119 (58,6)	208	106 (51,0)	1,06 [0,81; 1,38]
	<i>Hirnetastasen (Interaktionstest: p=0,0311)</i>					
	Ja	26	20 (76,9)	23	11 (47,8)	2,72 [1,26; 5,86]
	Nein	296	177 (59,8)	313	162 (51,8)	1,15 [0,93; 1,42]
Verstopfung	<i>Alter (Interaktionstest: p=0,1019)</i>					
	<65 Jahre	232	100 (43,1)	240	104 (43,3)	0,80 [0,61; 1,06]
	≥65 Jahre	90	45 (50,0)	96	36 (37,5)	1,21 [0,78; 1,89]
	<i>Geschlecht (Interaktionstest: p=0,1487)</i>					
	Weiblich	119	55 (46,2)	128	49 (38,3)	1,12 [0,76; 1,65]
	Männlich	203	90 (44,3)	208	91 (43,8)	0,78 [0,58; 1,04]
	<i>Hirnetastasen (Interaktionstest: p=0,0723)</i>					
	Ja	26	17 (65,4)	23	10 (43,5)	1,70 [0,77; 3,76]
	Nein	296	128 (43,2)	313	130 (41,5)	0,85 [0,67; 1,09]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0132)</i>					
Ja	24	8 (33,3)	21	11 (52,4)	0,33 [0,13; 0,82]	
Nein	298	137 (46,0)	315	129 (41,0)	0,97 [0,76; 1,24]	
Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
ECOG: Eastern Cooperative Oncology Group; HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; PS: Performance Status (Allgemeinzustand), RS: Randomized Set.						
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

Operationalisierung: Zeit bis zur Verschlechterung der Symptomatik (EORTC QLQ-LC13)

Tabelle 4-66: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 aus der Studie LUME-Lung 1

Symptom	Alter ¹	Geschlecht ²	Geografische Region ³	ECOG-PS ⁴	Hirnmetastasen ⁵	Bevacizumab-Vorbehandlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucherstatus ⁸
Bluthusten ⁹	0,1168	0,7224	0,1710	0,4255	0,2354	0,4526	0,6778	0,4519
Dyspnoe ⁹	0,8896	0,6714	0,4751	0,0410	0,2733	0,9261	0,8998	0,9705
Haarausfall ⁹	0,6492	0,4361	0,9804	0,4329	0,3408	0,6618	0,2335	0,5019
Husten ⁹	0,6367	0,3297	0,6318	0,8076	0,8332	0,7488	0,8407	0,0993
Mundschmerzen ⁹	0,2517	0,4616	0,5604	0,6007	0,5770	0,0231	0,7196	0,7163
Periphere Neuropathie ⁹	0,3954	0,9313	0,2123	0,6129	0,2418	0,0341	0,6113	0,9567
Schluckbeschwerden ⁹	0,5065	0,5549	0,7480	0,2193	0,4294	0,1249	0,8118	0,4325
Schmerzen (Arm/Schulter)⁹	0,1927	0,9276	0,3853	0,3689	0,2702	0,7661	0,4296	0,2718
Schmerzen (Brust) ⁹	0,8605	0,4008	0,0791	0,4582	0,5760	0,0167	0,6001	0,8344
Schmerzen (andere) ⁹	0,2606	0,9711	0,0837	0,7551	0,8124	0,5106	0,7551	0,4481

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS

ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.

¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten Behandlung, Subgruppe und Behandlung nach Subgruppe

p-Wert <0,05 fett geschrieben

p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben

Symptom fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der gesamten Zulassungspopulation bezüglich der Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13, zeigte sich ein statistisch signifikanter Effekt zugunsten von Nintedanib+Docetaxel für das Symptom Schmerzen (Arm/Schulter). Die Zeit bis zur Verschlechterung dieses Symptoms war im Nintedanib+Docetaxel-Arm länger als in der Placebo+Docetaxel-Behandlungsgruppe (siehe Abschnitt 4.3.1.3.1.5).

Die Subgruppenanalysen ergaben für die Symptome Haarausfall und Husten keine Hinweise oder Belege auf eine Effektmodifikation. Für alle anderen Symptome des Fragebogens EORTC QLQ-LC13 lag mindestens ein Hinweis auf Interaktion vor (Tabelle 4-66).

Dyspnoe

Ein Beleg für eine Effektmodifikation des Symptoms Dyspnoe zeigte sich anhand des p-Werts des Interaktionstests der Subgruppe ECOG-PS (Interaktionstest: $p=0,0410$). Der Effekt war in beiden Subgruppen als auch für die gesamte Zulassungspopulation statistisch nicht signifikant (Tabelle 4-67).

Mundschmerzen

Für das Symptom Mundschmerzen ergab sich ein Beleg auf eine Effektmodifikation bezüglich der Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0231$). Die Einzelergebnisse der vorbehandelten sowie der nicht vorbehandelten Patienten zeigten, wie auch schon die Ergebnisse der gesamten Zulassungspopulation, keine statistische Signifikanz (Tabelle 4-67).

Periphere Neuropathie

Ein weiterer Beleg auf eine Effektmodifikation bezüglich der Bevacizumab-Vorbehandlung lag für das Symptom periphere Neuropathie vor (Interaktionstest: $p=0,0341$). Auch die Subgruppen-Einzelergebnisse nach Bevacizumab-Vorbehandlung für die periphere Neuropathie wiesen keine statistische Signifikanz auf (Tabelle 4-67). In der gesamten Zulassungspopulation lag ebenfalls kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen bezüglich der peripheren Neuropathie vor.

Schmerzen (Arm/Schulter)

Ein Hinweis auf eine Effektmodifikation für das Symptom Schmerzen (Arm/Schulter) lag bezüglich des Alters vor (Interaktionstest: $p=0,1927$). Für jüngere Patienten (<65 Jahre) zeigte sich ein statistisch signifikanter Vorteil der Nintedanib+Docetaxel-Therapie (HR [95%-KI]: 0,72 [0,55; 0,95]). In der Gruppe der älteren Patienten (≥ 65 Jahre) war der Effekt nicht entgegengerichtet, war aber nicht statistisch signifikant (Tabelle 4-67).

Schmerzen (Brust)

Für das Symptom Schmerzen (Brust) zeigte sich ein weiterer Beleg auf eine Interaktion anhand der Subgruppe Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0167$). Für vorbehandelte Patienten wiesen die Einzelergebnisse nach Subgruppen einen statistisch signifikanten Effekt zugunsten der Kombinationstherapie Nintedanib+Docetaxel aus (HR [95%-KI]: 0,38 [0,15; 0,98]). Für nicht vorbehandelte Patienten war der Effekt, ebenso wie in der gesamten Zulassungspopulation, nicht statistisch signifikant (Tabelle 4-67). Zudem zeigte sich für die Schmerzen (Brust) ein Hinweis auf eine Effektmodifikation anhand der geografischen Region (Interaktionstest: $p=0,0791$).

Weitere Hinweise auf Effektmodifikationen zeigten sich anhand der Interaktionstests für das Symptom Bluthusten bezüglich des Alters (Interaktionstest: $p=0,1168$) und der geografischen Region (Interaktionstest: $p=0,1710$), für das Symptom Husten bezüglich des Raucherstatus (Interaktionstest: $p=0,0993$), für das Symptom Schluckbeschwerden bezüglich der Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1249$) sowie für das Symptom Schmerzen (andere) bezüglich der geografischen Region (Interaktionstest: $p=0,0837$). In der

gesamten Zulassungspopulation hatte sich für keines dieser Symptome ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen gezeigt.

Alle Einzelergebnisse der Subgruppenanalysen zur Zeit bis zur Verschlechterung der Symptome (EORTC QLQ-LC13), für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-67 dargestellt.

Tabelle 4-67: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13 aus der Studie LUME-Lung 1

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
Bluthusten	<i>Alter (Interaktionstest: p=0,1168)</i>					
	<65 Jahre	232	65 (28,0)	240	74 (30,8)	0,77 [0,55; 1,08]
	≥65 Jahre	90	33 (36,7)	96	26 (27,1)	1,24 [0,74; 2,08]
	<i>Geografische Region (Interaktionstest: p=0,1710)</i>					
	Asien	86	32 (37,2)	96	25 (26,0)	1,37 [0,81; 2,32]
	Europa	229	63 (27,5)	234	72 (30,8)	0,77 [0,55; 1,08]
	Südafrika	7	3 (42,9)	6	3 (50,0)	0,82 [0,13; 5,06]
Dyspnoe	<i>ECOG-PS (Interaktionstest: p=0,0410)</i>					
	0	96	74 (77,1)	99	63 (63,6)	1,40 [1,00; 1,96]
	1	226	148 (65,5)	237	157 (66,2)	0,91 [0,73; 1,15]
Husten	<i>Raucherstatus (Interaktionstest: p=0,0993)</i>					
	Nieraucher	115	58 (50,4)	115	52 (45,2)	1,22 [0,84; 1,77]
	aktiver Raucher/Ex-Raucher	207	110 (53,1)	221	114 (51,6)	0,84 [0,65; 1,10]
Mundschmerzen	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0231)</i>					
	Ja	24	10 (41,7)	21	12 (57,1)	0,45 [0,19; 1,05]
	Nein	298	146 (49,0)	315	133 (42,2)	1,09 [0,86; 1,38]
Periphere Neuropathie	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0341)</i>					
	Ja	24	11 (45,8)	21	11 (52,4)	0,47 [0,20; 1,11]
	Nein	298	165 (55,4)	315	150 (47,6)	1,09 [0,87; 1,36]
Schluckbeschwerden	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1249)</i>					
	Ja	24	8 (33,3)	21	10 (47,6)	0,45 [0,18; 1,15]
	Nein	298	124 (41,6)	315	126 (40,0)	0,96 [0,75; 1,23]
Schmerzen (Arm/Schulter)	<i>Alter (Interaktionstest: p=0,1927)</i>					
	<65 Jahre	232	95 (40,9)	240	114 (47,5)	0,72 [0,55; 0,95]
	≥65 Jahre	90	48 (53,3)	96	46 (47,9)	0,98 [0,65; 1,48]
Schmerzen (Brust)	<i>Geografische Region (Interaktionstest: p=0,0791)</i>					
	Asien	86	51 (59,3)	96	41 (42,7)	1,48 [0,98; 2,23]
	Europa	229	113 (49,3)	234	114 (48,7)	0,89 [0,69; 1,16]
	Südafrika	7	6 (85,7)	6	2 (33,3)	1,81 [0,35; 9,42]

Symptom	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
Schmerzen (Brust)	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0167)</i>					
	Ja	24	8 (33,3)	21	10 (47,6)	0,38 [0,15; 0,98]
	Nein	298	162 (54,4)	315	147 (46,7)	1,11 [0,88; 1,38]
Schmerzen (andere)	<i>Geografische Region (Interaktionstest: p=0,0837)</i>					
	Asien	86	43 (50,0)	96	38 (39,6)	1,06 [0,69; 1,64]
	Europa	229	101 (44,1)	234	113 (48,3)	0,75 [0,58; 0,99]
	Südafrika	7	6 (85,7)	6	2 (33,3)	6,05 [0,72; 50,56]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
ECOG: Eastern Cooperative Oncology Group; HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.						
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

4.3.1.3.2.4 Subgruppenanalysen zur gesundheitsbezogenen Lebensqualität – RCT

In Tabelle 4-68 bzw. Tabelle 4-70 sind die p-Werte der Interaktionstests für die Verbesserung der gesundheitsbezogenen Lebensqualität bzw. für die Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität aus der LUME-Lung 1 Studie gemessen anhand EORTC QLQ-C30 dargestellt. Die entsprechenden Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests <0,2), sind in Tabelle 4-69 bzw. Tabelle 4-71 zu finden. Zudem werden die Subgruppenanalysen für den EQ-5D und die EQ-VAS in Tabelle 4-72 und Tabelle 4-73 dargestellt.

Operationalisierung: Verbesserung der gesundheitsbezogenen Lebensqualität (EORTC QLQ-C30, Responderanalysen)

Tabelle 4-68: p-Werte der Interaktionstests für die Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1

Domäne	Alter ¹	Ge- schlecht ²	Geogra- fische Region ³	ECOG -PS ⁴	Hirn- meta- stasen ⁵	Bevaci- zumab- Vorbe- handlung ⁶	Zeit seit Beginn der Erstlinien- Therapie ⁷	Raucher- status ⁸
Globaler Gesundheits- status⁹	0,3562	0,5439	0,9424	0,4196	0,8434	0,0419	0,6365	0,6954
Emotionale Funktion ⁹	0,8434	0,1359	0,2949	0,1710	0,4269	0,6119	0,7432	0,7231
Kognitive Funktion⁹	0,7096	0,2589	0,2608	0,6379	0,3363	0,8309	0,9855	0,2074
Körperliche Funktion ⁹	0,9493	0,3164	0,7439	0,5855	0,2162	0,1347	0,4967	0,4367
Rollenfunk- tion ⁹	0,2678	0,1034	0,6014	0,2520	0,6202	0,6065	0,6673	0,4676
Soziale Funk- tion ⁹	0,5361	0,8827	0,7488	0,8379	0,8161	0,7511	0,2549	0,7817

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ p-Werte des Q-Tests unter Benutzung des Mantel-Haenszel-Schätzers des RR
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Domäne fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Bei Betrachtung der Verbesserung der gesundheitsbezogenen Lebensqualität in der gesamten Zulassungspopulation zeigte sich für den globalen Gesundheitsstatus und die kognitive Funktion ein statistisch signifikanter Vorteil für die Kombinationstherapie Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel. Für alle anderen Aspekte der gesundheitsbezogenen Lebensqualität unterschieden sich die Ergebnisse beider Behandlungsgruppen nicht statistisch signifikant voneinander (siehe Abschnitt 4.3.1.3.1.6).

Die Subgruppenanalysen ergaben nur wenige Hinweise und einen Beleg für eine Interaktion anhand der p-Werte der Interaktionstests (Tabelle 4-68).

Globaler Gesundheitsstatus

Für den globalen Gesundheitsstatus zeigte der Interaktionstest bezüglich der Bevacizumab-Vorbehandlung einen Beleg auf Interaktion an (Interaktionstest: $p=0,0419$). Die Subgruppenergebnisse zum globalen Gesundheitsstatus (Tabelle 4-69) nach Bevacizumab-Vorbehandlung wiesen für Patienten ohne Bevacizumab-Vorbehandlung einen statistisch signifikanten Vorteil für Nintedanib+Docetaxel aus. Das RR [95%-KI] betrug in dieser Subgruppe 1,76 [1,16; 2,67] und war damit statistisch signifikant zugunsten der Kombinationstherapie Nintedanib+Docetaxel. In der Gruppe der mit Bevacizumab vorbehandelten Patienten ergab sich kein statistisch signifikanter Unterschied zwischen der Nintedanib+Docetaxel- und der Placebo+Docetaxel-Therapie.

Für die emotionale Funktion zeigte sich ein Hinweis auf Interaktion für die Subgruppen Geschlecht (Interaktionstest: $p=0,1359$) und ECOG-PS (Interaktionstest: $p=0,1710$). Zudem ergab sich ein Hinweis auf Interaktion für die Bevacizumab-Vorbehandlung beim Endpunkt körperliche Funktion (Interaktionstest: $p=0,1347$) und ein weiterer Hinweis bezüglich des Geschlechts beim Endpunkt Rollenfunktion (Interaktionstest: $p=0,1034$). Für keines dieser Symptome hatten sich statistisch signifikante Effekte in der gesamten Zulassungspopulation gezeigt.

Die Einzelergebnisse der Subgruppen, für die sich ein Hinweis oder Beleg ergab, werden in Tabelle 4-69 dargestellt.

Tabelle 4-69: Einzelergebnisse der untersuchten Subgruppen zur Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 (Responderanalysen) aus der Studie LUME-Lung 1

Domäne	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel
		N	n (%)	N	n (%)	RR [95%-KI]
Globaler Gesundheitsstatus	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,0419)</i>					
	Ja	23	2 (8,7)	15	4 (26,7)	0,33 [0,07; 1,56]
	Nein	265	51 (19,2)	274	30 (10,9)	1,76 [1,16; 2,67]
Emotionale Funktion	<i>Geschlecht (Interaktionstest: p=0,1359)</i>					
	Weiblich	107	19 (17,8)	109	24 (22,0)	0,81 [0,47; 1,38]
	Männlich	182	41 (22,5)	181	30 (16,6)	1,36 [0,89; 2,08]
	<i>ECOG-PS (Interaktionstest: p=0,1710)</i>					
	0	89	18 (20,2)	84	21 (25,0)	0,81 [0,46; 1,41]
	1	200	42 (21,0)	206	33 (16,0)	1,31 [0,87; 1,98]
Körperliche Funktion	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1347)</i>					
	Ja	23	0 (0,0)	16	2 (12,5)	0,14 [0,01; 2,74]
	Nein	266	39 (14,7)	275	29 (10,5)	1,39 [0,89; 2,18]
Rollenfunktion	<i>Geschlecht (Interaktionstest: p=0,1034)</i>					
	Weiblich	107	18 (16,8)	110	26 (23,6)	0,71 [0,42; 1,22]
	Männlich	182	37 (20,3)	181	29 (16,0)	1,27 [0,82; 1,97]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
ECOG: Eastern Cooperative Oncology Group; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; RR: Relatives Risiko; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.						
Ein RR >1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

Operationalisierung: Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität (EORTC QLQ-C30, Responderanalysen)

Tabelle 4-70: p-Werte der Interaktionstests für die Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 aus der Studie LUME-Lung 1

Domäne	Alter ¹	Geschlecht ²	Geografische Region ³	ECOG-PS ⁴	Hirnmetastasen ⁵	Bevacizumab-Vorbehandlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucherstatus ⁸
Globaler Gesundheitsstatus ⁹	0,5264	0,9748	0,7966	0,8469	0,3666	0,2520	0,4267	0,6226
Emotionale Funktion ⁹	0,2745	0,1098	0,0794	0,4808	0,2151	0,2293	0,5540	0,1451
Kognitive Funktion ⁹	0,9926	0,8288	0,4464	0,2897	0,1821	0,3427	0,2199	0,9383
Körperliche Funktion ⁹	0,8093	0,2718	0,8136	0,5961	0,9530	0,4690	0,9292	0,9437
Rollenfunktion ⁹	0,2901	0,3837	0,6869	0,8221	0,4035	0,5766	0,6436	0,6644
Soziale Funktion ⁹	0,3582	0,7102	0,9866	0,5239	0,4681	0,1588	0,9295	0,3300

2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS
 ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); RS: Randomized Set.
¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten Behandlung, Subgruppe und Behandlung nach Subgruppe
 p-Wert <0,05 fett geschrieben
 p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben
 Domäne fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Für die Operationalisierung Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität zeigte sich in der gesamten Zulassungspopulation kein statistisch signifikanter Unterschied zwischen beiden Behandlungsgruppen (siehe Abschnitt 4.3.1.3.1.6).

Die Subgruppenanalysen ergaben anhand der p-Werte der Interaktionstests nur vereinzelte Hinweise auf eine Interaktion bezüglich der untersuchten Subgruppen (Tabelle 4-70).

Für die emotionale Funktion lag ein Hinweis auf Interaktion für die Subgruppen Geschlecht (Interaktionstest: p=0,1098), geografische Region (Interaktionstest: p=0,0794) und Raucherstatus (Interaktionstest: p=0,1451) vor. Weitere Hinweise auf Interaktion zeigten sich für die kognitive Funktion bezüglich der Hirnmetastasen zur Baseline (Interaktionstest: p=0,1821) und für die soziale Funktion bezüglich der Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1588). Bei Betrachtung der gesamten Zulassungspopulation hatten sich

anhand dieser Symptome keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen gezeigt.

Tabelle 4-71: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität, gemessen anhand der Funktionsskalen des Fragebogens EORTC QLQ-C30 aus der Studie LUME-Lung 1

Domäne	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
Emotionale Funktion	<i>Geschlecht (Interaktionstest: p=0,1098)</i>					
	Weiblich	119	66 (55,5)	128	60 (46,9)	1,31 [0,92; 1,86]
	Männlich	203	109 (53,7)	208	104 (50,0)	0,93 [0,71; 1,21]
	<i>Geografische Region (Interaktionstest: p=0,0794)</i>					
	Asien	86	52 (60,5)	96	41 (42,7)	1,52 [1,01; 2,28]
	Europa	229	117 (51,1)	234	120 (51,3)	0,92 [0,71; 1,19]
	Südafrika	7	6 (85,7)	6	3 (50,0)	2,10 [0,50; 8,73]
Kognitive Funktion	<i>Raucherstatus (Interaktionstest: p=0,1451)</i>					
	Nieraucher	115	59 (51,3)	115	51 (44,3)	1,29 [0,89; 1,88]
	aktiver Raucher/Ex- Raucher	207	116 (56,0)	221	113 (51,1)	0,94 [0,72; 1,22]
Soziale Funktion	<i>Hirnmetastasen (Interaktionstest: p=0,1821)</i>					
	Ja	26	20 (76,9)	23	14 (60,9)	1,68 [0,85; 3,34]
Soziale Funktion	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1588)</i>					
	Ja	24	13 (54,2)	21	12 (57,1)	0,62 [0,28; 1,38]
	Nein	298	182 (61,1)	315	177 (56,2)	1,01 [0,82; 1,24]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: RS						
HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; RS: Randomized Set.						
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

Operationalisierung: Mittelwertdifferenzen der gesundheitsbezogenen Lebensqualität (EQ-5D UK Utility Scores und EQ-VAS)

Tabelle 4-72: p-Werte der Interaktionstests für die gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS aus der Studie LUME-Lung 1

Domäne	Alter ¹	Geschlecht ²	Geografische Region ³	ECOG-PS ⁴	Hirnmetastasen ⁵	Bevacizumab-Vorbehandlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucherstatus ⁸
EQ-5D UK Utility ⁹	0,469 6	0,5816	0,0622	0,3687	0,2397	0,6213	0,7153	0,2440
EQ-VAS ⁹	0,378 4	0,9051	0,4578	0,8123	0,6669	0,3341	0,0855	0,6672

ECOG: Eastern Cooperative Oncology Group; EQ-5D: EuroQol-5 Dimensions (Patientenfragebogen zur Lebensqualität); EQ-VAS: EuroQol-Visual Analogue Scale; PS: Performance Status (Allgemeinzustand).

¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ p-Werte des Q-Tests unter Benutzung des Mantel-Haenszel-Schätzers des RR

p-Werte des Q-Tests mit der Differenz und Standardfehler der Differenz als Effekt- und Varianzschätzer

p-Wert <0,05 fett geschrieben
p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben

Domäne fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Für die gesundheitsbezogene Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS, zeigte sich in der gesamten Zulassungspopulation kein statistisch signifikanter Unterschied zwischen beiden Behandlungsgruppen (siehe Abschnitt 4.3.1.3.1.6).

Die Subgruppenanalysen ergaben anhand der p-Werte der Interaktionstests für den EQ-5D UK Utility Score einen Hinweis auf eine Effektmodifikation bezüglich der Subgruppen nach geografischer Region (Interaktionstest: p=0,0622). Für die EQ-VAS lag ein Hinweis auf eine Effektmodifikation nach der Zeit seit Beginn der Erstlinien-Therapie vor (Tabelle 4-72, Interaktionstest: p=0,0855).

Die Einzelergebnisse zum EQ-5D UK Utility Score in den drei geografischen Regionen – Asien, Europa, Südafrika – zeigten ebenso wie die Subgruppen nach der Zeit seit Beginn der Erstlinien-Therapie für die EQ-VAS keine statistisch signifikanten Effekte (Tabelle 4-73).

Tabelle 4-73: Einzelergebnisse der untersuchten Subgruppen zur gesundheitsbezogenen Lebensqualität, gemessen anhand des EQ-5D UK Utility Scores sowie der EQ-VAS aus der Studie LUME-Lung 1

Domäne	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel
		N	Mittelwert	N	Mittelwert	Mittelwertdifferenz [95%-KI]
	<i>Geografische Region (Interaktionstest: p=0,0622)</i>					
EQ-5D	Asien	76	0,658	88	0,687	-0,029 [-0,090; 0,033]
UK Utility	Europa	216	0,695	226	0,673	0,022 [-0,009; 0,052]
	Südafrika	7	0,707	6	0,811	-0,104 [-0,224; 0,015]
	<i>Zeit seit Beginn der Erstlinien-Therapie (Interaktionstest: p=0,0855)</i>					
EQ-VAS	<9 Monate	188	64,7	185	65,7	-1,0 [-3,1; 1,1]
	≥9 Monate	112	69,9	134	68,1	1,7 [-0,5; 4,0]
EQ-5D: EuroQol-5 Dimensions (Patientenfragebogen zur Lebensqualität); EQ-VAS: EuroQol-Visual Analogue Scale; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis.						
Eine Mittelwertdifferenz >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 0 nicht)						

4.3.1.3.2.5 Subgruppenanalysen zu unerwünschten Ereignissen – RCT

In Tabelle 4-74 sind die p-Werte der Interaktionstests für die Endpunkte unerwünschte Ereignisse aus der LUME-Lung 1 Studie dargestellt. Die Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests <0,2), sind in Tabelle 4-75 zu finden.

Tabelle 4-74: p-Werte der Interaktionstests für die Zeit bis zum ersten Auftreten unerwünschter Ereignisse aus der Studie LUME-Lung 1

UE	Alter ¹	Ge- schlecht ²	Geogra- fische Region ³	ECOG- PS ⁴	Hirn- meta- stasen ⁵	Bevac- zumab- Vorbe- handlung ⁶	Zeit seit Beginn der Erstlinien- Therapie ⁷	Raucher- status ⁸
UE ⁹	0,4929	0,5451	0,9120	0,5259	0,7710	0,2569	0,2652	0,0977
SUE ⁹	0,8168	0,2102	0,1851	0,8084	0,1892	0,6024	0,5481	0,1716
UE CTCAE-Grad ≥3 ⁹	0,0874	0,0238	0,6803	0,6754	0,3213	0,1098	0,2179	0,0566
UE, die zum Therapieabbruch führten ⁹	0,3452	0,3046	0,6728	0,6896	0,6498	0,0921	0,5460	0,0968

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS

CTCAE: Common Terminology Criteria for Adverse Events; ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); SUE: Schwerwiegende unerwünschte Ereignisse; TS: Treated Set; UE: Unerwünschte Ereignisse.

¹ <65 Jahre vs. ≥65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten Behandlung, Subgruppe und Behandlung nach Subgruppe

p-Wert <0,05 fett geschrieben
p-Wert ≥0,05 und <0,2 fett und kursiv geschrieben

UE fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Die Auswertungen zu den unerwünschten Ereignissen für die gesamte Zulassungspopulation ergaben keine statistisch signifikanten Unterschiede zwischen Nintedanib+Docetaxel und der Vergleichstherapie Placebo+Docetaxel (siehe Abschnitt 4.3.1.3.1.7 bis Abschnitt 4.3.1.3.1.11).

Die Subgruppenanalysen zu den UE, den SUE, den UE mit CTCAE-Grad ≥3 und den UE, die zum Therapieabbruch führten zeigten anhand der p-Werte der Interaktionstests nur vereinzelte Hinweise und ein Beleg auf eine Interaktion (Tabelle 4-74).

UE, SUE, UE mit CTCAE-Grad ≥ 3 und UE, die zum Therapieabbruch führten

Bei den UE zeigte sich ein Hinweis auf Effektmodifikation für den Raucherstatus (Interaktionstest: $p=0,0977$). Die Analysen zu den UE vom CTCAE-Grad ≥ 3 ergaben einen Beleg auf eine Interaktion für den subgruppenbildenden Faktor Geschlecht (Interaktionstest: $p=0,0238$) und jeweils einen Hinweis auf eine Interaktion für die Faktoren Alter (Interaktionstest: $p=0,0874$), Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,1098$) und Raucherstatus (Interaktionstest: $p=0,0566$). Die Subgruppenergebnisse für die UE vom CTCAE-Grad ≥ 3 bezüglich des Geschlechts ergaben in der weiblichen Subgruppe einen statistisch signifikanten Nachteil für Nintedanib+Docetaxel (HR [95%-KI]: 1,44 [1,08; 1,91]). Für die Subgruppe der Männer lagen, wie auch schon in der gesamten Zulassungspopulation, keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen vor (Tabelle 4-75).

Für die SUE betrug der p-Wert des Interaktionstests bezüglich der geografischen Region $p=0,1851$, bezüglich der Hirnmetastasen zur Baseline $p=0,1892$ und bezüglich des Raucherstatus $p=0,1716$. Für die SUE lag somit ein Hinweis auf eine Interaktion bezüglich der geografischen Region, der Hirnmetastasen zur Baseline und des Raucherstatus vor. Für die UE, die zum Therapieabbruch führten, zeigten die p-Werte der Interaktionstests einen Hinweis auf eine Interaktion anhand der Subgruppen nach Bevacizumab-Vorbehandlung (Interaktionstest: $p=0,0921$) und anhand des Raucherstatus (Interaktionstest: $p=0,0968$). In der gesamten Zulassungspopulation hatten sich bezüglich dieser UE keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsarmen gezeigt.

Tabelle 4-75: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zum ersten Auftreten unerwünschter Ereignisse aus der Studie LUME-Lung 1

UE	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
UE	<i>Raucherstatus (Interaktionstest: p=0,0977)</i>					
	Nieraucher	115	110 (95,7)	114	102 (89,5)	1,31 [1,00; 1,72]
	aktiver Raucher/Ex-Raucher	205	198 (96,6)	219	212 (96,8)	1,00 [0,82; 1,21]
SUE	<i>Geografische Region (Interaktionstest: p=0,1851)</i>					
	Asien	86	29 (33,7)	95	21 (22,1)	1,51 [0,86; 2,65]
	Europa	227	76 (33,5)	232	83 (35,8)	0,86 [0,63; 1,17]
	Südafrika	7	6 (85,7)	6	3 (50,0)	0,88 [0,20; 3,85]
	<i>Hirnmetastasen (Interaktionstest: p=0,1892)</i>					
	Ja	26	13 (50,0)	23	7 (30,4)	2,26 [0,85; 6,01]
	Nein	294	98 (33,3)	310	100 (32,3)	0,96 [0,72; 1,27]
	<i>Raucherstatus (Interaktionstest: p=0,1716)</i>					
	Nieraucher	115	33 (28,7)	114	24 (21,1)	1,38 [0,81; 2,33]
	aktiver Raucher/Ex-Raucher	205	78 (38,0)	219	83 (37,9)	0,89 [0,66; 1,22]
UE CTCAE-Grad ≥3	<i>Alter (Interaktionstest: p=0,0874)</i>					
	<65 Jahre	230	167 (72,6)	239	165 (69,0)	1,00 [0,80; 1,24]
	≥65 Jahre	90	76 (84,4)	94	63 (67,0)	1,43 [1,02; 1,99]
	<i>Geschlecht (Interaktionstest: p=0,0238)</i>					
	Weiblich	119	103 (86,6)	126	89 (70,6)	1,44 [1,08; 1,91]
	Männlich	201	140 (69,7)	207	139 (67,1)	0,95 [0,75; 1,20]
	<i>Bevacizumab-Vorbehandlung (Interaktionstest: p=0,1098)</i>					
	Ja	24	18 (75,0)	21	17 (81,0)	0,67 [0,34; 1,32]
	Nein	296	225 (76,0)	312	211 (67,6)	1,14 [0,94; 1,37]
	<i>Raucherstatus (Interaktionstest: p=0,0566)</i>					
Nieraucher	115	94 (81,7)	114	75 (65,8)	1,37 [1,01; 1,85]	
aktiver Raucher/Ex-Raucher	205	149 (72,7)	219	153 (69,9)	0,98 [0,78; 1,23]	
UE, die zum Therapieabbruch führten	<i>Bevacizumab Vorbehandlung (Interaktionstest: p=0,0921)</i>					
	Ja	24	3 (12,5)	21	6 (28,6)	0,33 [0,08; 1,34]
	Nein	296	64 (21,6)	312	53 (17,0)	1,18 [0,82; 1,69]

UE	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
UE, die zum Therapieabbruch führten	<i>Raucherstatus (Interaktionstest: p=0,0968)</i>					
	Nieraucher	115	26 (22,6)	114	15 (13,2)	1,67 [0,88; 3,15]
	aktiver Raucher/Ex- Raucher	205	41 (20,0)	219	44 (20,1)	0,87 [0,57; 1,33]
2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS						
CTCAE: Common Terminology Criteria for Adverse Events; HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; SUE: Schwerwiegende unerwünschte Ereignisse; TS: Treated Set; UE: Unerwünschte Ereignisse.						
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)						

4.3.1.3.2.6 Subgruppenanalysen zu Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ – RCT

In Tabelle 4-76 sind die p-Werte der Interaktionstests für die Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der LUME-Lung 1 Studie dargestellt. Die Einzelergebnisse der untersuchten Subgruppen für die Endpunkte, bei denen sich mindestens ein Hinweis auf eine Effektmodifikation ergeben hat (p-Wert des Interaktionstests $< 0,2$), sind in Tabelle 4-77 zu finden.

Tabelle 4-76: p-Werte der Interaktionstests für die Zeit bis zum ersten Auftreten der Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der Studie LUME-Lung 1

Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$	Alter ¹	Ge-schlecht ²	Geogra-fische Region ³	ECOG -PS ⁴	Hirn-meta-stasen ⁵	Bevaci-zumab-Vorbe-handlung ⁶	Zeit seit Beginn der Erstlinien-Therapie ⁷	Raucher-status ⁸
Diarrhö (PT) ⁹	0,6285	0,2416	0,6582	0,2775	0,6817	0,2066	0,8991	0,4695
Dyspnoe (PT) ⁹	0,6545	0,7746	0,8644	0,0977	0,9880	0,9760	0,6354	0,6128
Neutropenie (SSC) ⁹	0,7435	0,0282	0,8312	0,1679	0,5935	0,9240	0,6927	0,0388
febrile Neutropenie (PT) ⁹	0,8918	0,0507	0,4815	0,1575	0,9883	0,9892	0,8890	0,1386
Neutropenie (PT) ⁹	0,5045	0,0690	0,5050	0,7063	0,3475	0,3366	0,9944	0,7949
Verminderung der neutrophilen Granulozyten (PT) ⁹	0,7208	0,1746	0,2752	0,0467	0,8748	0,4583	0,6545	0,0226

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS

CTCAE: Common Terminology Criteria for Adverse Events; ECOG: Eastern Cooperative Oncology Group; PS: Performance Status (Allgemeinzustand); PT: Preferred Term; SSC: Special Search Category; TS: Treated Set.

¹ <65 Jahre vs. ≥ 65 Jahre; ² männlich vs. weiblich; ³ Asien vs. Südafrika vs. Europa; ⁴ Performance Status 0 vs. 1; ⁵ Ja vs. Nein; ⁶ Ja vs. Nein; ⁷ <9 Monate vs. ≥ 9 Monate; ⁸ Nieraucher vs. aktiver Raucher/Ex-Raucher; ⁹ Test auf Interaktion aus Cox-Regression mit den Kovariaten Behandlung, Subgruppe und Behandlung nach Subgruppe

p-Wert <0,05 fett geschrieben
p-Wert $\geq 0,05$ und <0,2 fett und kursiv geschrieben

Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ fett bedeutet statistisch signifikante Behandlungsunterschiede in der gesamten Zulassungspopulation.

Auch für die Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ zeigten sich in der gesamten Zulassungspopulation keine statistisch signifikanten Unterschiede zwischen der Behandlung mit Nintedanib+Docetaxel und der Behandlung mit Placebo+Docetaxel (siehe Abschnitt 4.3.1.3.1.11).

Anhand der p-Werte der Interaktionstests zeigten sich für die Diarrhö keine Hinweise auf Interaktionen zwischen den Subgruppen. Für die anderen Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ ergaben sich wenige Hinweise bzw. Belege auf eine Interaktion der untersuchten Subgruppen (Tabelle 4-76).

Neutropenie (SSC)

Die Subgruppenanalysen zur Neutropenie (SSC) zeigten für die Subgruppe Geschlecht einen Beleg für eine Interaktion (Interaktionstest: $p=0,0282$). Die Ergebnisse in der Subgruppe der weiblichen Patienten zeigten einen statistisch signifikanten Nachteil für Nintedanib+Docetaxel gegenüber der Placebo+Docetaxel-Vergleichstherapie (HR [95%-KI]; 1,43 [1,02; 2,02]). In der Subgruppe der Männer sowie bereits in der gesamten Zulassungspopulation war der Unterschied zwischen den beiden Behandlungsgruppen

statistisch nicht signifikant (Tabelle 4-77). Beim Raucherstatus zeigte sich ebenfalls ein Beleg (Interaktionstest: $p=0,0388$), jedoch keine statistisch signifikanten Effekte in den einzelnen Subgruppen. Für ECOG-PS ergab sich zudem ein Hinweis auf Interaktion (Interaktionstest: $p=0,1679$).

Verminderung der neutrophilen Granulozyten (PT)

Ein Beleg für Interaktion ergab sich auch für die Verminderung der neutrophilen Granulozyten bezüglich des ECOG-PS (Interaktionstest: $p=0,0467$). Hier war der Unterschied zwischen beiden Behandlungsgruppen weder in der gesamten Zulassungspopulation noch in einer der beiden Subgruppen (ECOG-PS=0 vs. ECOG-PS=1) statistisch signifikant (Tabelle 4-77). Das gleiche Bild zeigte sich bezüglich des Raucherstatus (Interaktionstest: $p=0,0226$). Auch hier waren die Unterschiede innerhalb der Subgruppen (Nieraucher vs. aktive Raucher/Ex-Raucher) nicht statistisch signifikant. Für die Subgruppen nach Geschlecht lag zudem ein Hinweis auf Interaktion vor (Interaktionstest: $p=0,1746$).

Darüber hinaus zeigte sich für die Dyspnoe ein Hinweis auf Interaktion bezüglich der ECOG-PS Subgruppen (Interaktionstest: $p=0,0977$). Die Subgruppenanalysen zum Endpunkt febrile Neutropenie ergaben einen Hinweis für die Subgruppen Geschlecht (Interaktionstest: $p=0,0507$), ECOG-PS (Interaktionstest: $p=0,1575$) und Raucherstatus (Interaktionstest: $p=0,1386$). Bezüglich des Geschlechts zeigte sich auch der einzige Hinweis auf eine Interaktion (Interaktionstest: $p=0,0690$) für die Nebenwirkung Neutropenie (PT) mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$. Bei Betrachtung dieser Effekte in der gesamten Zulassungspopulation hatten sich keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsarmen gezeigt.

Tabelle 4-77: Einzelergebnisse der untersuchten Subgruppen zur Zeit bis zum ersten Auftreten der Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ aus der Studie LUME-Lung 1

Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel
		N	n (%)	N	n (%)	HR [95%-KI]
Dyspnoe (PT)	<i>ECOG-PS (Interaktionstest: p=0,0977)</i>					
	0	96	6 (6,3)	96	3 (3,1)	1,88 [0,47; 7,54]
	1	224	9 (4,0)	237	17 (7,2)	0,50 [0,22; 1,13]
Neutropenie (SSC)	<i>Geschlecht (Interaktionstest: p=0,0282)</i>					
	Weiblich	119	75 (63,0)	126	60 (47,6)	1,43 [1,02; 2,02]
	Männlich	201	86 (42,8)	207	95 (45,9)	0,87 [0,65; 1,17]
	<i>ECOG-PS (Interaktionstest: p=0,1679)</i>					
	0	96	43 (44,8)	96	49 (51,0)	0,83 [0,55; 1,26]
	1	224	118 (52,7)	237	106 (44,7)	1,17 [0,90; 1,53]
	<i>Raucherstatus (Interaktionstest: p=0,0388)</i>					
Nieraucher	115	71 (61,7)	114	53 (46,5)	1,43 [1,00; 2,04]	
aktiver Raucher/Ex-Raucher	205	90 (43,9)	219	102 (46,6)	0,89 [0,67; 1,18]	
Febrile Neutropenie (PT)	<i>Geschlecht (Interaktionstest: p=0,0507)</i>					
	Weiblich	119	10 (8,4)	126	2 (1,6)	5,42 [1,19; 24,73]
	Männlich	201	13 (6,5)	207	13 (6,3)	0,97 [0,45; 2,11]
	<i>ECOG-PS (Interaktionstest: p=0,1575)</i>					
	0	96	8 (8,3)	96	2 (2,1)	4,05 [0,86; 19,08]
	1	224	15 (6,7)	237	13 (5,5)	1,19 [0,56; 2,49]
	<i>Raucherstatus (Interaktionstest: p=0,1386)</i>					
Nieraucher	115	12 (10,4)	114	4 (3,5)	3,01 [0,97; 9,34]	
aktiver Raucher/Ex-Raucher	205	11 (5,4)	219	11 (5,0)	1,02 [0,44; 2,37]	
Neutropenie (PT)	<i>Geschlecht (Interaktionstest: p=0,0690)</i>					
	Weiblich	119	20 (16,8)	126	16 (12,7)	1,33 [0,69; 2,57]
	Männlich	201	18 (9,0)	207	29 (14,0)	0,58 [0,32; 1,05]
Verminderung der neutrophilen Granulozyten (PT)	<i>Geschlecht (Interaktionstest: p=0,1746)</i>					
	Weiblich	119	51 (42,9)	126	44 (34,9)	1,28 [0,85; 1,91]
	Männlich	201	65 (32,3)	207	72 (34,8)	0,89 [0,64; 1,25]

Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$	Subgruppe Subgruppenkategorie	Nintedanib+ Docetaxel		Placebo+ Docetaxel		Behandlungseffekt Nintedanib+Docetaxel vs. Placebo+Docetaxel HR [95%-KI]
		N	n (%)	N	n (%)	
		Verminderung der neutrophilen Granulozyten (PT)	<i>ECOG-PS (Interaktionstest: $p=0,0467$)</i>			
	0	96	32 (33,3)	96	43 (44,8)	0,70 [0,44; 1,11]
	1	224	84 (37,5)	237	73 (30,8)	1,23 [0,90; 1,68]
	<i>Raucherstatus (Interaktionstest: $p=0,0226$)</i>					
	Nieraucher	115	54 (47,0)	114	39 (34,2)	1,50 [0,99; 2,26]
	aktiver Raucher/Ex-Raucher	205	62 (30,2)	219	77 (35,2)	0,81 [0,58; 1,13]

2. Datenschnitt vom 15. Februar 2013; Analyse Set: TS

CTCAE: Common Terminology Criteria for Adverse Events; ECOG: Eastern Cooperative Oncology Group; HR: Hazard Ratio aus Cox-Regression; KI: Konfidenzintervall; N: Anzahl der Patienten in der Auswertung; n: Anzahl der Patienten mit Ereignis; PS: Performance Status (Allgemeinzustand); PT: Preferred Term; SSC: Special Search Category; TS: Treated Set.

Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an; Effektschätzer fett bei signifikantem Ergebnis (KI enthält die 1 nicht)

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Der Nutzen von Nintedanib+Docetaxel wurde bereits durch die Zulassungsbehörden bestätigt. Im Dossier insgesamt wie auch in der folgenden Zusammenfassung der Ergebnisse aus RCT wird daher explizit der Zusatznutzen gegenüber der ZVT betrachtet. Im Folgenden werden die Ergebnisse der Studie LUME-Lung 1 zusammengefasst. Die Zusammenfassung der Ergebnisse beschreibt nur Effektmodifikationen, für die der Endpunkt in der gesamten Zulassungspopulation statistisch signifikant war und für die sich ein Beleg auf Effektmodifikation (Interaktions p-Wert <0,05) ergab.

Mortalität

Gesamtüberleben

Für Patienten mit Adenokarzinom war das Gesamtüberleben (OS) unter der Behandlung mit Nintedanib+Docetaxel statistisch signifikant länger als unter der Behandlung mit Placebo+Docetaxel ($p=0,0359$). Das Sterberisiko für Patienten, die Nintedanib in Kombination mit Docetaxel erhielten, verringerte sich um 17 Prozent, verglichen mit

Patienten, die mit Placebo+Docetaxel behandelt wurden (HR [95%-KI]: 0,83 [0,70; 0,99]). Dies entsprach einer Verlängerung des medianen OS um 2,3 Monate von 10,3 auf 12,6 Monate.

Morbidität

Progressionsfreies Überleben

Das PFS war unter der Kombinationstherapie mit Nintedanib+Docetaxel gegenüber einer Placebo+Docetaxel-Vergleichstherapie (4,2 Monate versus 2,8 Monate) statistisch signifikant verlängert (HR [95%-KI]: 0,84 [0,71; 1,00]; $p=0,0485$). Patienten mit der Kombinationstherapie Nintedanib+Docetaxel waren somit 1,4 Monate länger progressionsfrei als Patienten mit der Placebo+Docetaxel-Vergleichstherapie. Der PFS-Vorteil zugunsten der Nintedanib+Docetaxel-Kombinationstherapie war auch schon beim 1. Datenschnitt zum Zeitpunkt der primären PFS-Analyse gegeben (HR [95%-KI]: 0,77 [0,62; 0,96]; $p=0,0193$).

Die Subgruppenanalysen zum PFS zeigten nach der Zeit seit Beginn der Erstlinien-Therapie einen Beleg auf eine Effektmodifikation an (Interaktionstest: $p=0,0032$). Für die Patienten, deren Beginn der Erstlinien-Therapie weniger als 9 Monate zurücklag, ergab sich ein statistisch signifikanter Effekt zugunsten der Therapie Nintedanib+Docetaxel (HR [95%-KI]: 0,68 [0,54; 0,84]). Für Patienten, deren Beginn der Erstlinien-Therapie 9 Monate oder länger zurücklag, war kein statistisch signifikanter Unterschied im progressionsfreien Überleben zwischen beiden Behandlungsgruppen erkennbar.

Objektive Ansprechrates

Die Ergebnisse zur ORR zeigten keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie Nintedanib+Docetaxel und der Placebo+Docetaxel-Therapie ($p=0,4823$). In der Nintedanib+Docetaxel-Gruppe wiesen 4,7% und in der Placebo+Docetaxel-Gruppe 3,6% der Patienten ein bestätigtes objektives Ansprechen auf (RR [95%-KI]: 1,30 [0,62; 2,74] / OR [95%-KI]: 1,32 [0,61; 2,86] / RD [95%-KI]: 1,09 [-1,95; 4,13]).

Krankheitskontrollrate

Hinsichtlich DCR konnte eine statistisch signifikante Verbesserung unter der Kombinationstherapie Nintedanib+Docetaxel gezeigt werden im Vergleich zur Placebo+Docetaxel-Therapie (RR [95%-KI]: 1,37 [1,18; 1,59] / OR [95%-KI]: 1,93 [1,41; 2,63] / RD [95%-KI]: 16,20 [8,67; 23,73]; $p<0,0001$). Im Nintedanib+Docetaxel-Behandlungsarm erreichten 60,2% der Patienten eine Krankheitskontrolle, im Placebo+Docetaxel Arm nur 44,0%.

Für die Subgruppen nach der Zeit seit Beginn der Erstlinien-Therapie lag für die DCR ein Beleg auf eine Effektmodifikation vor (Interaktionstest: $p=0,0007$). Es zeigte sich ein statistisch signifikanter Vorteil für die Nintedanib+Docetaxel-Therapie für Patienten, deren Beginn der Erstlinien-Therapie weniger als 9 Monate zurücklag (RR [95%-KI]: 1,79 [1,42; 2,24]). In der Subgruppe der Patienten, deren Beginn der Erstlinien-Therapie

9 Monate oder länger zurücklag, ergab sich numerisch ebenfalls ein Vorteil für die Nintedanib+Docetaxel-Patienten. Dieser war jedoch nicht statistisch signifikant.

Symptomatik

EORTC QLQ-C30

Bei Betrachtung der Responderanalyse zur Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30, ergab sich bezüglich des Symptoms **Schmerzen** ein statistisch signifikanter Vorteil zugunsten der Kombinationstherapie Nintedanib+Docetaxel (RR [95%-KI]: 1,56 [1,13; 2,14] / OR [95%-KI]: 1,76 [1,18; 2,62] / RD [95%-KI]: 9,52 [2,84; 16,20]; p=0,005).

Bei Betrachtung der Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-C30, zeigte sich ein statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel für die Symptome **Diarrhö** (HR [95%-KI]: 1,90 [1,54; 2,34]; p<0,0001) sowie **Übelkeit und Erbrechen** (HR [95%-KI]: 1,23 [1,00; 1,50]; p=0,0467).

Die Subgruppenanalysen ergaben für das Symptom Übelkeit und Erbrechen einen Beleg auf eine Effektmodifikation anhand der Subgruppe Hirnmetastasen zur Baseline (Interaktionstest: p=0,0311). In der Subgruppe der Patienten mit Hirnmetastasen zur Baseline war die Zeit bis zur Verschlechterung des Symptoms Übelkeit und Erbrechen im Nintedanib+Docetaxel-Arm statistisch signifikant kürzer als im Placebo+Docetaxel-Arm (HR [95%-KI]: 2,72 [1,26; 5,86]). Für Patienten ohne Hirnmetastasen zeigte sich kein statistisch signifikanter Unterschied zwischen beiden Behandlungsarmen.

EORTC QLQ-LC13

Bezüglich der Verbesserung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13, zeigte sich kein statistisch signifikanter Effekt.

Bei Betrachtung der Zeit bis zur Verschlechterung der Symptomatik, gemessen anhand der Symptomskalen aus dem Fragebogen EORTC QLQ-LC13, zeigte sich für das Symptom **Schmerzen (Arm/Schulter)** ein statistisch signifikanter Effekt zugunsten von Nintedanib+Docetaxel (HR [95%-KI]: 0,80 [0,63; 1,00]; p=0,0456). Die Zeit bis zur Verschlechterung der Schmerzen (Arm/Schulter) war im Nintedanib+Docetaxel-Arm länger als in der Placebo+Docetaxel-Behandlungsgruppe.

Gesundheitsbezogene Lebensqualität

EORTC QLQ-C30

Bei Betrachtung der Verbesserung der gesundheitsbezogenen Lebensqualität in der gesamten Zulassungspopulation zeigte sich für den **globalen Gesundheitsstatus** (RR [95%-KI]: 1,56 [1,05; 2,33] / OR [95%-KI]: 1,69 [1,06; 2,69] / RD [95%-KI]: 6,64 [0,82; 12,45]; p=0,026) und die **kognitive Funktion** (RR [95%-KI]: 1,64 [1,12; 2,41] / OR [95%-KI]: 1,81 [1,15; 2,84] / RD [95%-KI]: 8,00 [2,00; 14,00]; p=0,009) ein statistisch signifikanter

Vorteil für die Kombinationstherapie Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel.

Die Subgruppenanalysen zum globalen Gesundheitsstatus zeigten bezüglich der Bevacizumab-Vorbehandlung einen Beleg auf eine Effektmodifikation an (Interaktionstest: $p=0,0419$). Für Patienten ohne Bevacizumab-Vorbehandlung lag ein statistisch signifikanter Vorteil für Nintedanib+Docetaxel vor (RR [95%-KI]: 1,76 [1,16; 2,67]). In der Gruppe der mit Bevacizumab vorbehandelten Patienten ergab sich kein statistisch signifikanter Unterschied zwischen der Nintedanib+Docetaxel- und der Placebo+Docetaxel-Therapie.

Für die Operationalisierung Zeit bis zur Verschlechterung der gesundheitsbezogenen Lebensqualität zeigte sich kein statistisch signifikanter Unterschied zwischen beiden Behandlungsgruppen.

EQ-5D

Anhand der Auswertung der Ergebnisse des EQ-5D UK Utility Scores sowie der EQ-VAS zeigten sich keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie mit Nintedanib+Docetaxel sowie der Placebo+Docetaxel-Vergleichstherapie.

Nebenwirkungen

UE

Bei Betrachtung der Zeit bis zum ersten Auftreten irgendeines UE zeigte sich kein statistisch signifikanter Unterschied zwischen der Kombinationstherapie Nintedanib+Docetaxel und der Vergleichstherapie Placebo+Docetaxel (HR [95%-KI]: 1,10 [0,94; 1,29]; $p=0,2067$).

SUE

Für die Zeit bis zum ersten Auftreten eines SUE ergaben sich keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsgruppen Nintedanib+Docetaxel und Placebo+Docetaxel (HR [95%-KI]: 1,01 [0,78; 1,32]; $p=0,9316$).

UE mit CTCAE-Grad ≥ 3

Bezüglich der UE vom CTCAE-Grad ≥ 3 ergab sich anhand der Zeit bis zum ersten Auftreten eines solchen UE kein statistisch signifikanter Unterschied zwischen einer Behandlung mit Nintedanib+Docetaxel und einer Behandlung mit Placebo+Docetaxel (HR [95%-KI]: 1,10 [0,92; 1,32]; $p=0,2655$).

UE, die zum Therapieabbruch führten

Für die Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, ergab sich anhand des HR kein statistisch signifikanter Unterschied zwischen Nintedanib+Docetaxel und Placebo+Docetaxel (HR [95%-KI]: 1,08 [0,76; 1,54]; $p=0,6557$).

Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$

Für die Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ – Diarrhö (PT), Dyspnoe (PT), Neutropenie (SSC), febrile Neutropenie (PT), Neutropenie (PT) und

Verminderung der neutrophilen Granulozyten (PT) – zeigten sich für die Zeit bis zum ersten Auftreten der Nebenwirkung keine statistisch signifikanten Unterschiede zwischen der Behandlung mit Nintedanib+Docetaxel und der Behandlung mit Placebo+Docetaxel.

Zusammenfassende methodische Bewertung

Die Interaktionstests für Subgruppeneffekte sind aufgrund des multiplen Testens nur von eingeschränkter Aussagekraft. Es wurden weniger Belege und Hinweise auf Interaktion entdeckt als durch reinen Zufall erwartet worden wären. Darüber hinaus ergab sich auch innerhalb der Subgruppen kein einheitliches Bild, da es zumeist Endpunkte mit Verbesserung und Verschlechterung der Effekte in den jeweiligen Subgruppen gab.

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 4-78: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.1.**

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

Tabelle 4-79: 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 4-80: 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 4-81: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

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

Begründen Sie für jede Studie die abschließende Einschätzung.

nicht 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 4-82: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1.1)
<Studie 1>	

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 der üblichen 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 4.3.1.3.2.

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 4-83: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

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

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

nicht 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 4-84: 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 4-85: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

nicht 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.

4.3.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

nicht 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 4-86: 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 4.3.1.3.2.**

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

Zusammenfassung des Abschnitts 4.4

Mit der Phase-III-Studie LUME-Lung 1 liegt eine RCT mit niedrigem Verzerrungspotential auf Studienebene vor, welche der höchsten Evidenzstufe zuzuordnen ist.

Auf Basis dieser zulassungsrelevanten Phase-III-Studie lassen sich Hinweise auf einen Zusatznutzen ableiten, da die patientenrelevanten Endpunkte ein niedriges Verzerrungspotential aufweisen.

Ein Zusatznutzen mit beträchtlichem Ausmaß wird nach AM-NutzenV dann ausgewiesen, wenn eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapielevanten 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.

Fazit

Mit einer Verlängerung des medianen Gesamtüberlebens um 2,3 Monate auf 12,6 Monate ist mit Nintedanib in Kombination mit Docetaxel als Therapieoption für Adenokarzinom-Patienten ein relevanter Fortschritt in der Zweitlinien-Therapie des fortgeschrittenen NSCLC erreicht worden. Die Daten der LUME-Lung 1 Studie zum wichtigsten sekundären Endpunkt OS konnten erstmals seit fast 10 Jahren einen signifikanten Überlebensvorteil im vorliegenden Anwendungsgebiet zeigen.

Aus dieser klinisch relevanten Verlängerung des Gesamtüberlebens, der Verbesserung der aus Patientensicht wichtigen Endpunkte PFS, DCR und Lebensqualität bei gut kontrollierbaren Nebenwirkungen resultiert nach der oben dargestellten Anforderung der AM-NutzenV ein **beträchtlicher Zusatznutzen** für Nintedanib.

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.

Das OS zählt aufgrund seiner eindeutigen und klaren Bestimmbarkeit zu den harten klinischen Endpunkten, da es verlässlich und unabhängig von subjektiven Einschätzungen durch Patienten oder andere Beurteilende erhoben werden kann. Sowohl Patienten als auch behandelnde Personen waren verblindet, die Auswertung erfolgte gemäß Randomisierung und die Folgetherapien waren in beiden Behandlungsgruppen ausgeglichen. Eine Verzerrung des Ergebnisses für den Endpunkt OS ist daher nicht wahrscheinlich.

PFS, der primäre Endpunkt der LUME-Lung 1 Studie, sowie das Tumoransprechen sind durch die unabhängige und verblindete Bewertung radiografischer Scans unverzerrt messbar.

Die Endpunkte Symptomatik und HRQoL wurden mittels validierter Messinstrumente von den Patienten direkt vor jedem Arztbesuch bewertet. Dadurch hatte der aktuelle Krankheitsstatus, wie z. B. das Auftreten eines Progresses keinen Einfluss auf die Bewertung durch die Patienten, da eine solche Information erst im darauf folgenden Arztgespräch mitgeteilt wurde. Aufgrund der Verblindung der Patienten und behandelnden Personen kann von einer unverzerrten Messung der Symptomatik und HRQoL ausgegangen werden. Angaben zur Validität der Messinstrumente sowie zur Validität der Schwellenwerte für klinisch relevante Verbesserungen sind in Abschnitt 4.2.5.2 nachzulesen. Der Zusatznutzen für diese Endpunkte wurde separat anhand der Responderanalysen sowie der Zeit bis zur Verschlechterung der Symptome bzw. HRQoL abgeleitet.

Das Auftreten von UE unter Therapie wurde anhand der CTCAE-Kategorien berichtet. UE stellen patientenrelevante Endpunkte dar. Das Verzerrungspotential kann als niedrig

eingestuft werden, da für den Vergleich die Zeit bis zum ersten Auftreten eines UE betrachtet wird und keine das Verzerrungspotential beeinflussenden Aspekte vorliegen.

Mit der LUME-Lung 1 Studie liegt eine RCT mit niedrigem Verzerrungspotential auf Studienebene vor, welche der höchsten Evidenzstufe zuzuordnen ist.

Auf der Basis der LUME-Lung 1 Studie lassen sich Hinweise auf einen Zusatznutzen ableiten, da die patientenrelevanten Endpunkte ein niedriges Verzerrungspotential aufweisen (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013b).

Wie bereits in Abschnitt 4.3.1.3.2 beschrieben, ist es wahrscheinlich, dass aufgrund multiplen Testens falsch positive statistisch signifikante Effektmodifikationen bei einigen Subgruppen auftraten. Die Wahrscheinlichkeit dafür steigt mit der Zahl der durchgeführten Tests, unabhängig davon, ob eine tatsächliche Effektmodifikation vorliegt.

Die Interaktionstests für Subgruppeneffekte sind aufgrund des multiplen Testens nur von eingeschränkter Aussagekraft. Es wurden weniger Belege und Hinweise auf Interaktion entdeckt als durch reinen Zufall erwartet worden wären. Darüber hinaus ergab sich auch innerhalb der Subgruppen kein einheitliches Bild, da es zumeist Endpunkte mit Verbesserung und Verschlechterung der Effekte in den jeweiligen Subgruppen gab.

Aus diesen Gründen wurden signifikante Unterschiede zwischen den Subgruppen im Dossier beschrieben und diskutiert, jedoch wurden die Subgruppenergebnisse nicht für die Ableitung des Zusatznutzens herangezogen.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

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

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

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

- *erheblicher Zusatznutzen*

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

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

Nintedanib gehört zur Wirkstoffklasse der TKI (Tyrosinkinase-Inhibitoren) und ist zugelassen in Kombination mit Docetaxel für die Anwendung bei erwachsenen Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Im Folgenden werden die Ergebnisse der LUME-Lung 1 Studie zu den einzelnen Endpunkten zusammengefasst und die Schlussfolgerungen zum Ausmaß des Zusatznutzens und dessen Wahrscheinlichkeit für erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiviertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie abgeleitet.

Wie bereits in Abschnitt 4.4.1 erläutert, lässt sich auf Grundlage der LUME-Lung 1 Studie durch das niedrige Verzerrungspotential aller dargestellten Endpunkte jeweils ein Hinweis für einen Zusatznutzen ableiten. Eine tabellarische Zusammenfassung findet sich in Abschnitt 4.1 (Tabelle 4-2).

Eine Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist aus den in Abschnitt 4.3.1.2.1 genannten Gründen sowie durch ein Studienprotokoll, das sich an den relevanten deutschen und europäischen Leitlinien orientiert, gegeben.

Gesamtüberleben

Ergebnisse aus LUME-Lung 1	<p>Gesamtüberleben [4.3.1.3.1.1]</p> <p><u>Effektschätzer</u> HR [95%-KI]: 0,83 [0,70; 0,99]; p=0,0359</p> <p><u>Medianes OS</u> 12,6 Monate (Nintedanib+Docetaxel) vs. 10,3 Monate (Placebo+Docetaxel)</p>
Erläuterungen	<ul style="list-style-type: none"> • Nach dem SGB V – beispielsweise in der AM-NutzenV operationalisiert – wird ein patientenrelevanter Nutzen durch Verlängerung des Überlebens bestimmt. • OS ist ein vom IQWiG anerkannter patientenrelevanter Endpunkt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011). • Die FDA empfiehlt die Verwendung des Gesamtüberlebens neben dem PFS bei Studien zum NSCLC als Standard zur Beschreibung des klinischen Nutzens (Food and Drug Administration, 2011).
Bewertung des Zusatznutzens	<p>Die moderate Verlängerung der Überlebensdauer durch Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel stellt eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens dar.</p> <p>Insgesamt ergibt sich für den Endpunkt OS (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]):</p> <ul style="list-style-type: none"> • ein Hinweis auf einen beträchtlichen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel aufgrund einer moderaten Verlängerung der Überlebensdauer.
Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.	

Progressionsfreies Überleben

Ergebnisse aus LUME-Lung 1	Progressionsfreies Überleben [4.3.1.3.1.2] <u>Effektschätzer</u> HR [95% KI]: 0,84 [0,71; 1,00]; p=0,0485 <u>Medianes PFS</u> 4,2 Monate (Nintedanib+Docetaxel) vs. 2,8 Monate (Placebo+Docetaxel)
Erläuterungen	<ul style="list-style-type: none"> • Die EMA erkennt das PFS neben dem OS als einen akzeptablen, eigenständigen und patientenrelevanten primären Endpunkt an (European Medicines Agency, 2005; European Medicines Agency, 2008; European Medicines Agency, 2012). • Die FDA empfiehlt die Verwendung des PFS neben dem Gesamtüberleben bei Studien zum NSCLC als Standard zur Beschreibung des klinischen Nutzens (Food and Drug Administration, 2011). • Bridges et al. zeigen in einer Patientenpräferenzstudie mit NSCLC-Patienten mittels einer wahlbasierten Conjointanalyse, dass aus Patientensicht eine Verlängerung des PFS als direkter Vorteil wahrgenommen wird (Bridges et al., 2012). • In einem Discrete Choice Experiment mit 211 NSCLC-Patienten in Deutschland machten Mühlbacher et al. die Beobachtung, dass bei der Betrachtung von 7 Attributen aus Patientensicht die Zeit ohne Tumorprogress sowie die Behandlung von tumorassoziierten Symptomen die zwei wichtigsten Entscheidungskriterien für die Wahl einer Therapie sind* (Mühlbacher und Bethge, 2014).
Bewertung des Zusatznutzens	Die Verlängerung des PFS durch Nintedanib+Docetaxel stellt eine gegenüber der Vergleichstherapie Placebo+Docetaxel bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapielevanten Nutzens dar. Insgesamt ergibt sich für den Endpunkt PFS (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]): <ul style="list-style-type: none"> • ein Hinweis auf einen geringen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel aufgrund einer Verlängerung der Zeit ohne Tumorprogression/Tod.
<p>* Die Studien von Bridges et al. und Mühlbacher und Bethge widerlegten die Nullhypothese, dass die unerwünschten Effekte eine vergleichbare Relevanz wie tumorassoziierte Symptome oder PFS haben. Obwohl die Forschungsfrage in beiden Studien OS nicht berücksichtigte, ist von einer Umkehr der Ränge nicht auszugehen. Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.</p>	

Objektive Ansprechrate und Krankheitskontrollrate

Ergebnisse aus LUME-Lung 1	<p>Objektive Ansprechrate [4.3.1.3.1.3]</p> <p><u>Effektschätzer</u> RR [95%-KI]: 1,30 [0,62; 2,74] / OR [95%-KI]: 1,32 [0,61; 2,86] / RD [95%-KI]: 1,09 [-1,95; 4,13].</p> <p><u>Objektives Ansprechen</u> 4,7% der Patienten (Nintedanib+Docetaxel) versus 3,6% der Patienten (Placebo+ Docetaxel); p=0,4823</p> <p>Krankheitskontrollrate [4.3.1.3.1.4]</p> <p><u>Effektschätzer</u> RR [95%-KI]: 1,37 [1,18; 1,59] / OR [95%-KI]: 1,93 [1,41; 2,63] / RD [95%-KI]: 16,20 [8,67; 23,73]</p> <p><u>Krankheitskontrolle</u> 60,2% der Patienten (Nintedanib+Docetaxel) versus 44,0% der Patienten (Placebo+ Docetaxel); p<0,0001</p>
Erläuterungen	<ul style="list-style-type: none"> • Tumoransprechen sowie Tumorstabilisierung sind eng mit der Progression der Erkrankung verbunden. Eine unmittelbare Kontrolle der Progression ist bei Tumoren, die lebenswichtige Organe bedrohen, nötig und viele Patienten und ihre Betreuer sehen daher eine schnelle und effektive Tumorkontrolle sowie das Verzögern der Progression des Tumors als Vorteil an, unabhängig von der Verlängerung des Gesamtüberlebens (Zhuang et al., 2009). • Bei Patienten mit fortgeschrittenem NSCLC erwies sich die DCR als ein stärkerer Prädiktor für das Überleben als die ORR (Lara et al., 2008).
Bewertung des Zusatznutzens	<p>Für die objektive Ansprechrate ist weder ein Zusatznutzen noch ein Nachteil belegbar. Jedoch liegt durch die signifikant verbesserte Krankheitskontrollrate unter der Kombinationstherapie mit Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel eine gegenüber der ZVT bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens vor.</p> <p>Insgesamt ergibt sich für den Endpunkt DCR (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]):</p> <ul style="list-style-type: none"> • ein Hinweis auf einen beträchtlichen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel aufgrund einer bedeutsamen Stabilisierung der Erkrankung.
<p>Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.</p>	

Symptomatik

Im folgenden Abschnitt wird die Bewertung des Zusatznutzens für alle Symptomskalen aus den Fragebögen EORTC QLQ-C30 und EORTC QLQ-LC13 anhand der Ergebnisse von zwei Auswertungen abgeleitet:

1. Responderanalyse des Anteils an Patienten, die durchschnittlich eine Verbesserung des jeweiligen Scores um mindestens 10 Punkte erreicht haben (Osoba et al., 1998).
2. Zeit bis zur Verschlechterung des Scores der jeweiligen Symptome um mindestens 10 Punkte (Osoba et al., 1998)

Diese Form der Auswertungen stellt laut IQWiG-Methodenpapier 4.1 ein adäquates Vorgehen zur Interpretation von Skalen dar (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013b).

Nachfolgend wird nur auf Symptome mit signifikanten Ergebnissen eingegangen. Hierzu zählen Schmerzen, Diarrhö sowie Übelkeit und Erbrechen. Für alle anderen Symptome liegen keine signifikanten Unterschiede vor.

Schwerwiegende Symptome

<p>Ergebnisse aus LUME-Lung 1</p>	<p>Schmerzen (EORTC QLQ-C30) – Responderanalyse [4.3.1.3.1.5]</p> <p><u>Effektschätzer</u> RR [95%-KI]: 1,56 [1,13; 2,14], OR [95%-KI]: 1,76 [1,18; 2,62], RD [95%-KI]: 9,52 [2,84; 16,20]</p> <p><u>Schmerzen - Responderanalyse</u> Statistisch signifikanter Vorteil für Nintedanib+Docetaxel versus Placebo+Docetaxel (p=0,005)</p> <p>Schmerzen (Arm/Schulter), (EORTC QLQ-LC13) – Zeit bis zur Verschlechterung [4.3.1.3.1.5]</p> <p><u>Effektschätzer</u> HR [95%-KI]:0,80 [0,63; 1,00]</p> <p><u>Schmerzen (Arm/Schulter) - Zeit bis zur Verschlechterung</u> Statistisch signifikanter Vorteil für Nintedanib+Docetaxel versus Placebo+Docetaxel (p=0,0456)</p>
<p>Erläuterungen</p>	<ul style="list-style-type: none"> • Innerhalb der EORTC QLQ-C30 Scores wird eine Differenz von mindestens 10 Punkten (auf der 0-100 Skala) als klinisch relevant auf Patientenebene angesehen (Osoba et al., 1998; Cocks et al., 2011). Diese Skalierung ist auch für den QLQ-LC13 Score anerkannt (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2013a). • Schmerzen haben einen signifikanten negativen Einfluss auf die von NSCLC-Patienten berichtete gesundheitsbezogene Lebensqualität (Iyer et al., 2013) und werden von Lungenkrebs-Patienten häufig als schwerwiegend wahrgenommen (Krech et al., 1992). • NSCLC-Patienten stuften in einer Patientenpräferenzstudie Schmerz (56%) nach Übelkeit und Erbrechen (63%) als zweitwichtigste Nebenwirkung ein, die es zu vermeiden gilt (Osoba et al., 2006).

	<ul style="list-style-type: none"> • Die Conjointanalyse von Bridges et al. zeigte, dass NSCLC-Patienten eine Verlängerung des PFS zusammen mit einer Verbesserung der Schwere von Symptomen (Schmerzen, Husten und Dyspnoe) am wichtigsten ist (Bridges et al., 2012). • Aus Sicht von NSCLC-Patienten im Stadium IV (UICC 7. Auflage) stellten die Zeit ohne Fortschreiten des Tumors sowie die Behandlung von tumorassoziierten Symptomen (Dyspnoe, Husten, Schmerz) die zwei wichtigsten Entscheidungskriterien für die Wahl einer Therapie dar (Mühlbacher und Bethge, 2014).
Bewertung des Zusatznutzens	<p>Die signifikante Verbesserung des Symptoms Schmerzen (EORTC QLQ-C30) durch Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel stellt eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens dar.</p> <p>Die verlängerte Zeit bis zur Verschlechterung von Schmerzen in Arm/Schulter (EORTC QLQ-LC13) stellt eine bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapielevanten Nutzens im Sinne einer Verringerung eines schwerwiegenden Symptoms dar.</p> <p>Insgesamt ergibt sich für den Endpunkt Schmerzen (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]):</p> <ul style="list-style-type: none"> • ein Hinweis auf einen beträchtlichen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich der Verbesserung der Schmerzen aufgrund einer Abschwächung des schwerwiegenden Symptoms. • ein Hinweis auf einen geringen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich der Zeit bis zur Verschlechterung der Schmerzen in Arm/Schulter aufgrund der Verringerung eines schwerwiegenden Symptoms.
<p>Ein RR/OR >1, HR <1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.</p>	

Weitere Symptome

Ergebnisse aus LUME-Lung 1	<p>Diarrhö (EORTC QLQ-C30) – Zeit bis zur Verschlechterung [4.3.1.3.1.5]</p> <p><u>Effektschätzer</u> HR [95%-KI]: 1,90 [1,54; 2,34]</p> <p><u>Diarrhö – Zeit bis zur Verschlechterung</u> Statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (p<0,0001)</p> <p>Übelkeit und Erbrechen (EORTC QLQ-C30) – Zeit bis zur Verschlechterung [4.3.1.3.1.5]</p> <p><u>Effektschätzer</u> HR [95%-KI]: 1,23 [1,00; 1,50]</p> <p><u>Übelkeit und Erbrechen – Zeit bis zur Verschlechterung</u> Statistisch signifikanter Effekt zum Nachteil von Nintedanib+Docetaxel (p=0,0467)</p>
Erläuterungen	<ul style="list-style-type: none"> • In einem Discrete Choice Experiment mit 211 NSCLC-Patienten in Deutschland machten Mühlbacher et al. die Beobachtung, dass bei der Betrachtung von 7 Attributen aus Patientensicht Übelkeit und Erbrechen an drittwichtigster und Diarrhö an fünftwichtigster Stelle der Entscheidungskriterien für die Wahl einer Therapie genannt wurde (Mühlbacher und Bethge, 2014). Das Experiment zeigt, dass Zeit ohne Tumorprogress sowie die Behandlung von tumorassoziierten Symptomen die zwei wichtigsten Entscheidungskriterien für die Wahl einer Therapie sind. • Empfehlungen für geeignete Maßnahmen und proaktives Therapiemanagement finden sich in der Fach- und Gebrauchsinformation für Nintedanib wieder.
Bewertung des Zusatznutzens	<p>Insgesamt ergibt sich für den Endpunkt Diarrhö bzw. Übelkeit und Erbrechen (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]):</p> <ul style="list-style-type: none"> • ein Hinweis für einen Nachteil von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich der Zeit bis zur Verschlechterung des Symptoms Diarrhö aufgrund einer signifikant kürzeren Zeit bis zur Verschlechterung*. • ein Hinweis für einen Nachteil von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich der Zeit bis zur Verschlechterung des Symptoms Übelkeit und Erbrechen aufgrund einer signifikant kürzeren Zeit bis zur Verschlechterung*.
<p>* Bei Auftreten dieser Symptome ist eine Unterbrechung der Therapie oder eine Herabsetzung der Dosis von Nintedanib+Docetaxel möglich (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014b). Ein HR <1 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.</p>	

Gesundheitsbezogene Lebensqualität

Ergebnisse aus LUME-Lung 1	<p>Globaler Gesundheitsstatus (EORTC QLQ-C30) – Responderanalyse [4.3.1.3.1.6]</p> <p><u>Effektschätzer</u> RR [95%-KI]: 1,56 [1,05; 2,33], OR [95%-KI]: 1,69 [1,06; 2,69], RD [95%-KI]: 6,64 [0,82; 12,45]</p> <p><u>Globaler Gesundheitsstatus – Responderanalyse</u> Statistisch signifikanter Vorteil für Nintedanib+Docetaxel versus Placebo+Docetaxel (p=0,026)</p> <p>Kognitive Funktion (EORTC QLQ-C30) – Responderanalyse [4.3.1.3.1.6]</p> <p><u>Effektschätzer</u> RR [95%-KI]: 1,64 [1,12; 2,41], OR [95%-KI]: 1,81 [1,15; 2,84], RD [95%-KI]: 8,00 [2,00; 14,00].</p> <p><u>Kognitive Funktion – Responderanalyse</u> Statistisch signifikanter Vorteil für Nintedanib+Docetaxel versus Placebo+Docetaxel (p=0,009)</p>
Erläuterungen	<ul style="list-style-type: none"> • Dem Erhalt der Lebensqualität durch die Kontrolle von Symptomen kommt eine besondere Bedeutung zu (McCannon und Temel, 2012). • Zum Nachweis der therapeutischen Verbesserung sind nach VerFO und AM-NutzenV vorrangig klinische Studien zu berücksichtigen, die unter anderem insbesondere Lebensqualität als patientenrelevanter Endpunkt darstellen (Bundesministerium der Justiz, 2010; Gemeinsamer Bundesausschuss, 2014b). • Bei Patienten mit fortgeschrittenem Lungenkrebs stellen die krankheitsspezifischen Symptome eine große Belastung dar, was sich wiederum negativ auf die Funktionalität und die Lebensqualität der Tumorerkrankten auswirken kann (McCannon und Temel, 2012).
Bewertung des Zusatznutzens	<p>Durch den statistisch signifikanten Vorteil des globalen Gesundheitsstatus durch Nintedanib+Docetaxel gegenüber der Vergleichstherapie Placebo+Docetaxel ergibt sich eine bisher nicht erreichte deutliche Verbesserung der Lebensqualität.</p> <p>Der statistisch signifikante Vorteil für die kognitive Funktion stellt eine bisher nicht erreichte deutliche Verbesserung der Lebensqualität dar.</p> <p>Insgesamt ergibt sich für den Endpunkt globaler Gesundheitszustand bzw. kognitive Funktion (gemäß AM-NutzenV [Bundesministerium der Justiz, 2010]):</p> <ul style="list-style-type: none"> • ein Hinweis auf einen beträchtlichen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich einer deutlichen Verbesserung der Lebensqualität aufgrund der Verbesserung des globalen Gesundheitsstatus. • ein Hinweis auf einen beträchtlichen Zusatznutzen von Nintedanib+Docetaxel im Vergleich zu Placebo+Docetaxel bezüglich einer deutlichen Verbesserung der Lebensqualität aufgrund der Verbesserung der kognitiven Funktion.
<p>Ein RR/OR >1 bzw. eine RD >0 zeigt einen Behandlungseffekt zugunsten der Nintedanib+Docetaxel-Therapie an.</p>	

Fragebogen EQ-5D

Anhand der Auswertung der Ergebnisse des EQ-5D UK Utility Scores sowie der EQ-VAS zeigten sich keine statistisch signifikanten Unterschiede zwischen der Kombinationstherapie mit Nintedanib+Docetaxel sowie der Placebo+Docetaxel-Vergleichstherapie. Es liegt daher weder ein Zusatznutzen noch ein Nachteil vor.

Nebenwirkungen

Gesamtrate UE

Bei Betrachtung der Zeit bis zum ersten Auftreten irgendeines UE zeigte sich kein statistisch signifikanter Unterschied zwischen der Kombinationstherapie Nintedanib+Docetaxel und der Vergleichstherapie Placebo+Docetaxel. Es liegt daher weder ein Zusatznutzen noch ein Nachteil für Nintedanib vor.

Schwerwiegende unerwünschte Ereignisse

Für die Zeit bis zum ersten Auftreten eines SUE ergab sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen Nintedanib+Docetaxel und Placebo+Docetaxel. Daher kann für diesen Endpunkt weder ein Zusatznutzen noch ein Nachteil für Nintedanib abgeleitet werden.

Unerwünschte Ereignisse CTCAE-Grad ≥ 3

Bezüglich der UE vom CTCAE-Grad ≥ 3 ergab sich anhand der Zeit bis zum ersten Auftreten eines solchen UE kein statistisch signifikanter Unterschied zwischen einer Behandlung mit Nintedanib+Docetaxel und einer Behandlung mit Placebo+Docetaxel. Es ist daher weder ein Zusatznutzen noch ein Nachteil für Nintedanib ableitbar.

UE, die zum Therapieabbruch führten

Für die Zeit bis zum ersten Auftreten eines UE, das zum Therapieabbruch führte, ergab sich anhand des HR kein statistisch signifikanter Unterschied zwischen Nintedanib+Docetaxel und Placebo+Docetaxel. Für diesen Endpunkt liegt somit weder ein Zusatznutzen noch ein Nachteil von Nintedanib vor.

Besondere Nebenwirkungen (CTCAE-Grad ≥ 3 und Häufigkeit $> 5\%$)

Auch für die besonderen Nebenwirkungen – Diarrhö (PT), Dyspnoe (PT), Neutropenie (SSC), febrile Neutropenie (PT), Neutropenie (PT) und Verminderung der neutrophilen Granulozyten (PT) – zeigten sich für die Zeit bis zum ersten Auftreten der Nebenwirkung keine statistisch signifikanten Unterschiede zwischen der Behandlung mit Nintedanib+Docetaxel und der Behandlung mit Placebo+Docetaxel. Für die besonderen Nebenwirkungen kann daher weder ein Zusatznutzen noch ein Nachteil von Nintedanib abgeleitet werden.

Gesamtfazit

Aus den dargestellten Daten geht hervor, dass die Behandlung mit Nintedanib+Docetaxel für erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem NSCLC mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie eine wirksame und im Allgemeinen gut verträgliche Therapieoption darstellt. Mit einer Verlängerung des medianen Gesamtüberlebens um 2,3 Monate auf 12,6 Monate ist mit Nintedanib in Kombination mit Docetaxel als Therapieoption für Adenokarzinom-Patienten ein relevanter Fortschritt in der Zweitlinien-Therapie des fortgeschrittenen NSCLC erreicht worden. Die Daten der LUME-Lung 1 Studie zum wichtigsten sekundären Endpunkt OS konnten erstmals seit fast 10 Jahren einen signifikanten Überlebensvorteil im vorliegenden Anwendungsgebiet zeigen.

Aus dieser klinisch relevanten Verlängerung des Gesamtüberlebens, der Verbesserung der aus Patientensicht wichtigen Endpunkte PFS, DCR und Lebensqualität bei gut kontrollierbaren Nebenwirkungen ein **beträchtlicher Zusatznutzen** für Nintedanib.

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 4-87: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit lokal fortgeschrittenem, metastasiertem oder lokal rezidiertem nicht-kleinzelligen Lungenkarzinom (NSCLC) mit Adenokarzinom-Histologie nach Erstlinien-Chemotherapie	beträchtlich

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

nicht zutreffend

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

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

nicht 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 2005²⁶, Molenberghs 2010²⁷). 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 2006²⁸) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten

²⁶ Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

²⁷ 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.

²⁸ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

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 2006²⁹) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

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

nicht zutreffend

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).

²⁹ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006; 25(2): 183-203.

Tabelle 4-88: Liste der für die Nutzenbewertung berücksichtigten Studien und Untersuchungen

Studie	Titel	Datenquelle
LUME-Lung 1	LUME-Lung 1; 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	<p>Studienprotokoll (Boehringer Ingelheim Pharma GmbH & Co. KG, 2008)</p> <p>Statistischer Analyseplan (Boehringer Ingelheim Pharma GmbH & Co. KG, 2011)</p> <p>Statistischer Analyseplan Addendum (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013b)</p> <p>Studienbericht Primäre PFS-Analyse (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b)</p> <p>Studienbericht Finale OS-Analyse (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a)</p> <p>Biomarker Report (Boehringer Ingelheim Pharma GmbH & Co. KG, 2013c)</p> <p>Statistischer Analyseplan 1199.14 (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012a)</p> <p>Health Economic Report (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014d)</p> <p>Subgruppenanalysen EQ-5D und EQ-VAS (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014c)</p> <p>Weitere Analysen (Boehringer Ingelheim Pharma GmbH & Co. KG, 2014e)</p>

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., Ahmedzai, S., Bergman, B., et al. 1993. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*, 85(5), 365-76.
- [2] Aventis Pharma S.A. 2014. Fachinformation Taxotere® 20 mg/1 ml, 80 mg/4 ml, 160 mg/8 ml Konzentrat zur Herstellung einer Infusionslösung (Docetaxel). Stand: April 2014. Verfügbar: <http://www.fachinfo.de/> [Aufgerufen am 18.12.2014].
- [3] Bergman, B., Aaronson, N. K., Ahmedzai, S., et al. 1994. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. *Eur J Cancer*, 30A(5), 635-42.
- [4] Boehringer Ingelheim Pharma GmbH & Co. KG. 2008. Clinical Trial Protocol: 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 (LUME-Lung 1). Document Number: U08-1569-01. (Data on file) 01.07.2008.
- [5] Boehringer Ingelheim Pharma GmbH & Co. KG. 2011. Trial Statistical Analysis Plan: 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 (LUME-Lung 1). Document Number: T11-1028-03. (Data on file) 08.07.2011 revised.
- [6] Boehringer Ingelheim Pharma GmbH & Co. KG. 2012a. Trial Statistical Analysis Plan: Multicenter, randomized, 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 first line chemotherapy (LUME-Lung 2). Document Number: T12-1023-01. (Data on file) 05.07.2012.
- [7] Boehringer Ingelheim Pharma GmbH & Co. KG. 2012b. Clinical Trial Report (Primary PFS Analysis): 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 (LUME-Lung 1) Document Number: U10-1991-01. (Data on file) 19.09.2012.
- [8] Boehringer Ingelheim Pharma GmbH & Co. KG. 2013a. Clinical Trial Report (Final OS Analysis): 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 (LUME-Lung 1). Document Number: U13-1504-01. (Data on file) 26.08.2013.
- [9] Boehringer Ingelheim Pharma GmbH & Co. KG. 2013b. Trial Statistical Analysis Plan Addendum to TSAP T11-1028-03: 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 (LUME-Lung 1). (Data on file) 23.01.2013.
- [10] Boehringer Ingelheim Pharma GmbH & Co. KG. 2013c. Biomarker Report: Exploratory analysis to identify and confirm a clinical biomarker defining a subgroup of second line non small cell lung cancer patients with a potential treatment benefit from nintedanib when given in combination with pemetrexed or docetaxel (LUME-Lung 1 und LUME-Lung 2). Document Number: U13-1128-02. (Data on file) 14.02.2013.
- [11] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014a. Risk Management Plan (Version 2.0). Document Number: s00020173-07. (Data on file) 02.12.2014.
- [12] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014b. Fachinformation Vargatef® 100 mg/150 mg Weichkapseln (Nintedanib). Stand: November 2014.
- [13] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014c. Subgruppenanalysen EQ-5D und EQ-VAS (LUME-Lung 1). (Data on file) 30.05.2014 und 18.12.2014.
- [14] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014d. Health Economic Report: Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared with placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non-small cell lung cancer after failure of first-line chemotherapy (LUME-Lung 1). (Data on file) 13.03.2014.
- [15] Boehringer Ingelheim Pharma GmbH & Co. KG. 2014e. Weitere Analysen der LUME-Lung 1 Studie. (Data on file) 25.04.2014, 15.05.2014, 04.06.2014, 27.06.2014, 10.10.2014 und 16.12.2014.
- [16] Braun, D. P., Gupta, D. & Staren, E. D. 2011. Quality of life assessment as a predictor of survival in non-small cell lung cancer. *BMC Cancer*, 11(11), 353.
- [17] 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, 169-80.
- [18] Bridges, J. F., Mohamed, A. F., Finnern, H. W., et al. 2012. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer*, 77, 224-31.
- [19] Bundesministerium der Justiz. 1988. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung, Stand: 07.08.2013. Verfügbar: http://www.gesetze-im-internet.de/bundesrecht/sgb_5/gesamt.pdf [Aufgerufen am 29.10.2014].
- [20] Bundesministerium der Justiz. 2010. Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV), Stand: 28.12.2010. Verfügbar: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf> [Aufgerufen am 11.11.2014].
- [21] Campos, M. P., Hassan, B. J., Riechelmann, R., et al. 2011. Cancer-related fatigue: a practical review. *Ann Oncol*, 22(6), 1273-9.
- [22] ClinicalTrials.gov. 2014a. LUME-Columbus: Nintedanib Plus Docetaxel in Advanced Non-small Cell Lung Cancer With Translational Research (NCT02231164) [Online]. Adresse: <http://clinicaltrials.gov/ct2/show/NCT02231164>. Zuletzt geändert am: 20.10.2014. [Aufgerufen am 27.10.2014].
- [23] ClinicalTrials.gov. 2014b. LUME-Lung 1: BIBF 1120 Plus Docetaxel as Compared to Placebo Plus Docetaxel in 2nd Line Non Small Cell Lung Cancer (NCT00805194) [Online]. Adresse: <https://clinicaltrials.gov/ct2/show/NCT00805194?term=lume+lung+1&rank=1>. Zuletzt geändert am: 07.10.2014. [Aufgerufen am 27.10.2014].

- [24] Cocks, K., King, M. T., Velikova, G., et al. 2011. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol*, 29(1), 89-96.
- [25] Currow, D. C., Smith, J., Davidson, P. M., et al. 2010. Do the trajectories of dyspnea differ in prevalence and intensity by diagnosis at the end of life? A consecutive cohort study. *J Pain Symptom Manage*, 39(4), 680-90.
- [26] de Boer, A. G., van Lanschot, J. J., Stalmeier, P. F., 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*, 13, 311-20.
- [27] Devlen, J., Maguire, P., Phillips, P., et al. 1987a. Psychological problems associated with diagnosis and treatment of lymphomas. II: Prospective study. *Br Med J* 295, 955-7.
- [28] Devlen, J., Maguire, P., Phillips, P., et al. 1987b. Psychological problems associated with diagnosis and treatment of lymphomas. I: Retrospective study. *Br Med J* 295, 953-4.
- [29] Efficace, F., Therasse, P., Piccart, M. J., et al. 2004. Health-related quality of life parameters as prognostic factors in a nonmetastatic breast cancer population: an international multicenter study. *J Clin Oncol*, 22(16), 3381-8.
- [30] Efficace, F., Bottomley, A., Smit, E. F., et al. 2006. Is a patient's self-reported health-related quality of life a prognostic factor for survival in non-small-cell lung cancer patients? A multivariate analysis of prognostic factors of EORTC study 08975. *Ann Oncol*, 17, 1698-704.
- [31] Eisenhauer, E. A., Therasse, P., Bogaerts, J., et al. 2009. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*, 45, 228-47.
- [32] EU Clinical Trials Register. 2014. 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 (2007-004803-36) [Online]. Adresse: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2007-004803-36/DE>. Zuletzt geändert. [Aufgerufen am 30.10.2014].
- [33] European Medicines Agency. 1996. ICH Topic E 3 Structure and Content of Clinical Study Reports. Verfügbar: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002832.pdf [Aufgerufen am 11.11.2014].
- [34] European Medicines Agency. 2005. Guideline on the Evaluation of Anticancer Medicinal Products in Man. Verfügbar: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/12/WC500017748.pdf [Aufgerufen am 11.11.2014].
- [35] European Medicines Agency. 2008. Appendix 1 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man; Methodological Considerations for Using Progression-free Survival (PFS) as Primary Endpoint in Confirmatory Trials for Registration. Verfügbar: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2009/12/WC500017749.pdf [Aufgerufen am 11.11.2014].
- [36] European Medicines Agency. 2012. Guideline on the Evaluation of Anticancer Medicinal Products in Man. Verfügbar:

- http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf [Aufgerufen am 11.11.2014].
- [37] EuroQoL Group. 2013. EQ-5D-3L User Guide. Version 5.0. Verfügbar: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf [Aufgerufen am 11.11.2014].
- [38] Fang, F.-M., Liu, Y.-T., Tang, Y., et al. 2004. Quality of life as a survival predictor for patients with advanced head and neck carcinoma treated with radiotherapy. *Cancer*, 100(2), 425-32.
- [39] Folkman, J. 1995. Seminars in Medicine of the Beth Israel Hospital, Boston. Clinical applications of research on angiogenesis. *N Engl J Med*, 333(26), 1757-63.
- [40] Food and Drug Administration. 2007. Guidance for industry – Clinical trial endpoints for the approval of cancer drugs and biologics. Verfügbar: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071590.pdf> [Aufgerufen am 11.11.2014].
- [41] Food and Drug Administration. 2011. Guidance for industry – Clinical trial endpoints for the approval of non-small cell lung cancer drugs and biologics (Draft Guidance). Verfügbar: <http://www.fda.gov/downloads/Drugs/Guidances/UCM259421.pdf> [Aufgerufen am 11.11.2014].
- [42] French, C. L., Irwin, R. S., Curley, F. J., et al. 1998. Impact of chronic cough on quality of life. *Arch Intern Med*, 158, 1657-61.
- [43] Gemeinsamer Bundesausschuss. 2012. 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 - Ipilimumab. Verfügbar: https://www.g-ba.de/downloads/40-268-2010/2012-08-02_AM-RL-XII_Ipilimumab_TrG.pdf [Aufgerufen am 11.11.2014].
- [44] Gemeinsamer Bundesausschuss. 2013a. 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 - Axitinib. Verfügbar: https://www.g-ba.de/downloads/40-268-2240/2013-03-21_AM-RL-XII_Axitinib_TrG.pdf [Aufgerufen am 29.10.2014].
- [45] Gemeinsamer Bundesausschuss. 2013b. 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 - Crizotinib. Verfügbar: http://www.g-ba.de/downloads/40-268-2301/2013-05-02_AM-RL-XII_Crizotinib_TrG.pdf [Aufgerufen am 11.11.2014].
- [46] Gemeinsamer Bundesausschuss. 2014a. Änderung der zweckmäßigen Vergleichstherapie. Beratungsanforderung 2013-B-119. Nintedanib zur Behandlung des Lungenkarzinoms (Data on file) 14.11.2014.
- [47] Gemeinsamer Bundesausschuss. 2014b. Verfahrensordnung des Gemeinsamen Bundesausschusses, Stand: 06. September 2014. Verfügbar: https://www.g-ba.de/downloads/62-492-922/VerfO_2014-06-19.pdf [Aufgerufen am 20.11.2014].
- [48] Gemeinsamer Bundesausschuss. 2014c. 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 - Afatinib. Verfügbar: https://www.g-ba.de/downloads/40-268-2301/2014-06-19_AM-RL-XII_Afatinib_TrG.pdf

- [ba.de/downloads/40-268-2792/2014-05-08_AM-RL-XII_Afatinib_2013-11-15-D-082_TrG.pdf](http://www.ba.de/downloads/40-268-2792/2014-05-08_AM-RL-XII_Afatinib_2013-11-15-D-082_TrG.pdf) [Aufgerufen am 11.11.2014].
- [49] Gemeinsamer Bundesausschuss. 2014d. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV. Beratungsanforderung 2013-B-119. Nintedanib zur Behandlung des Lungenkarzinoms (Data on file) 05.05.2014.
- [50] Goeckenjan, G., Sitter, H., Thomas, M., et al. 2010. Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms. *Pneumologie*, 64 Suppl 2, e1-164.
- [51] Griesinger, F., Eberhard, W., Früh, M., et al. 2012. Lungenkarzinom, nicht-kleinzellig (NSCLC). Leitlinie. Stand: Oktober 2012. Verfügbar: <http://www.dgho-onkopedia.de/de/onkopedia/leitlinien/lungenkarzinom-nicht-kleinzellig-nsclc/lungenkarzinom-nicht-kleinzellig-nsclc.pdf> [Aufgerufen am 11.11.2014].
- [52] Groenvold, M., Klee, M. C., Sprangers, M. A., et al. 1997. Validation of the EORTC QLQ-C30 quality of life questionnaire through combined qualitative and quantitative assessment of patient-observer agreement. *J Clin Epidemiol*, 50(4), 441-50.
- [53] Grond, S., Zech, D., Diefenbach, C., et al. 1996. Assessment of cancer pain: a prospective evaluation in 2266 cancer patients referred to a pain service. *Pain*, 64, 107-14.
- [54] Gutstein, H. B. 2001. The biologic basis of fatigue. *Cancer*, 92(6), 1678-83.
- [55] Heigener, D. F., Deppermann, K.-M., Sebastian, M., et al. 2011. Stadieneinteilung des nichtkleinzelligen Lungenkarzinoms. Aktueller Algorithmus. *Onkologe*, 17, 679-683.
- [56] Hensch, I., Bergman, B., Gustafsson, M., et al. 2007. The impact of symptoms, coping capacity, and social support on quality of life experience over time in patients with lung cancer. *J Pain Symptom Manage*, 34(4), 370-9.
- [57] Herschbach, P., Keller, M., Knight, L., et al. 2004. Psychological problems of cancer patients: a cancer distress screening with a cancer-specific questionnaire. *Br J Cancer*, 91, 504-11.
- [58] Hjermstad, M. J., Fossa, S. D., Bjordal, K., et al. 1995. Test/retest study of the european organization for research and treatment of cancer core quality-of-life questionnaire. *J Clin Oncol*, 13, 1249-54.
- [59] Hofman, M., Ryan, J. L., Figueroa-Moseley, C. D., et al. 2007. Cancer-related fatigue: the scale of the problem. *Oncologist*, 12 Suppl 1, 4-10.
- [60] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. 2011. Aussagekraft von Surrogatendpunkten in der Onkologie Version 1.1. Stand: 21.11.2011. Verfügbar: https://www.iqwig.de/download/A10-05_Rapid_Report_Version_1-1_Surrogatendpunkte_in_der_Onkologie.pdf [Aufgerufen am 11.11.2014].
- [61] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. 2013a. Addendum zum Auftrag A12-15 (Crizotinib). Verfügbar: https://www.iqwig.de/download/A13-13_Addendum-zum-Auftrag-A12-15_Crizotinib.pdf [Aufgerufen am 11.11.2014].
- [62] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. 2013b. Allgemeine Methoden Version 4.1. Stand: 28.11.2013. Verfügbar: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf [Aufgerufen am 11.11.2014].
- [63] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. 2014. Afatinib - Nutzenbewertung gemäß § 35a SGB V. Verfügbar: https://www.g-ba.de/downloads/92-975-403/2014-02-13_A13-41_Afatinib_Nutzenbewertung-35a-SGB-V.pdf [Aufgerufen am 11.11.2014].
- [64] Irwin, R. S., French, C. T. & Fletcher, K. E. 2002. Quality of life in coughers. *Pulm Pharmacol Ther*, 15, 283-6.

- [65] Iyer, S., Taylor-Stokes, G. & Roughley, A. 2013. Symptom burden and quality of life in advanced non-small cell lung cancer patients in France and Germany. *Lung Cancer*, 81(2), 288-93.
- [66] Kaiser, R., Barrueco, J., Reck, M., et al. 2013. Identification of a Clinical Biomarker for Second-line Combination Treatment with Nintedanib in Adenocarcinoma Non-small Cell Lung Cancer (NSCLC) Patients in Two Phase III Trials (#P388). European Cancer Congress (ESCO-ESMO-ESTRO) Amsterdam, The Netherlands, 27.09.-01.10.2013.
- [67] Krech, R. L., Davis, J., Walsch, D., et al. 1992. Symptoms of lung cancer. *Palliative Medicine*, 6, 309-315.
- [68] Lara, P. N., Jr., Redman, M. W., Kelly, K., et al. 2008. Disease control rate at 8 weeks predicts clinical benefit in advanced non-small-cell lung cancer: results from Southwest Oncology Group randomized trials. *J Clin Oncol*, 26(3), 463-7.
- [69] Larsson, M., Ljung, L. & Johansson, B. B. 2012. Health-related quality of life in advanced non-small cell lung cancer: correlates and comparisons to normative data. *Eur J Cancer Care*, 21, 642-9.
- [70] Lutz, S., Norrell, R., Bertucio, C., et al. 2001. Symptom frequency and severity in patients with metastatic or locally recurrent lung cancer: a prospective study using the Lung Cancer Symptom Scale in a community hospital. *J Palliat Med*, 4(2), 157-65.
- [71] Maisey, N. R., Norman, A., Watson, M., et al. 2002. Baseline quality of life predicts survival in patients with advanced colorectal cancer. *Eur J Cancer*, 38, 1351-7.
- [72] McCannon, J. & Temel, J. 2012. Comprehensive management of respiratory symptoms in patients with advanced lung cancer. *J Support Oncol*, 10(1), 1-9.
- [73] Mercadante, S., Armata, M. & Salvaggio, L. 1994. Pain characteristics of advanced lung cancer patients referred to a palliative care service. *Pain*, 59, 141-5.
- [74] Meyerowitz, B. E., Watkins, I. K. & Sparks, F. C. 1983. Quality of life for breast cancer patients receiving adjuvant chemotherapy. *Am J Nurs*, 83, 232-5.
- [75] Miller, A. B., Hoogstraten, B., Staquet, M., et al. 1981. Reporting results of cancer treatment. *Cancer*, 47(1), 207-14.
- [76] Moher, D., Hopewell, S., Schulz, K. F., et al. 2010. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Br Med J* 340, c869.
- [77] Mühlbacher, A. C. & Bethge, S. 2014. Patients' preferences: a discrete-choice experiment for treatment of non-small-cell lung cancer. *Eur J Health Econ*, DOI 10.1007/s10198-014-0622-4.
- [78] National Cancer Institute. 2006. Common terminology criteria for adverse events, Version 3.0 (CTCAE). Verfügbar: http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcae3.pdf [Aufgerufen am 11.11.2014].
- [79] 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.
- [80] Osoba, D., Hsu, M.-A., Copley-Merriman, C., et al. 2006. Stated preferences of patients with cancer for health-related quality-of-life (HRQOL) domains during treatment. *Qual Life Res*, 15, 273-83.
- [81] Padilla, G. V. & Grant, M. M. 1985. Quality of life as a cancer nursing outcome variable. *ANS Adv Nurs Sci*, 8(1), 45-60.
- [82] Pazdur, R. 2008. Endpoints for assessing drug activity in clinical trials. *Oncologist*, 13 Suppl 2, 19-21.

- [83] Peppercorn, J. M., Smith, T. J., Helft, P. R., et al. 2011. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. *J Clin Oncol*, 29(6), 755-60.
- [84] PharmNet.bund. 2014. 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 2007-004803-36 [Online]. Verfügbar: https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#__DEFANCHOR__ [Aufgerufen am 30.10.2014].
- [85] Plunkett, T. A., Chrystal, K. F. & Harper, P. G. 2003. Quality of life and the treatment of advanced lung cancer. *Clin Lung Cancer*, 5(1), 28-32.
- [86] Portenoy, R. K. 2000. Cancer-related fatigue: An immense problem. *Oncologist*, 5, 350-2.
- [87] Potter, J. & Higginson, I. J. 2004. Pain experienced by lung cancer patients: a review of prevalence, causes and pathophysiology. *Lung Cancer*, 43, 247-57.
- [88] Rabin, R. & de Charro, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*, 33, 337-43.
- [89] Reck, M., Kaiser, R., Mellemegaard, A., et al. 2014a. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol*, 15(2), 143-55.
- [90] Reck, M., Popat, S., Reinmuth, N., et al. 2014b. Metastatic non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 25 Suppl 3, iii27-iii39.
- [91] Rhodes, V. A., Watson, P. M. & Hanson, B. M. 1988. Patients' descriptions of the influence of tiredness and weakness on self-care abilities. *Cancer Nurs*, 11(3), 186-94.
- [92] Schuette, W., Tesch, H., Buttner, H., et al. 2012. Second-line treatment of stage III/IV non-small-cell lung cancer (NSCLC) with pemetrexed in routine clinical practice: evaluation of performance status and health-related quality of life. *BMC Cancer*, 12, 14.
- [93] Schulz, K. F., Altman, D. G., Moher, D., et al. 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Br Med J*, 340, c332.
- [94] Smith, E. L., Hann, D. M., Ahles, T. A., et al. 2001. Dyspnea, anxiety, body consciousness, and quality of life in patients with lung cancer. *J Pain Symptom Manage*, 21(4), 323-9.
- [95] Stone, P., Richardson, A., Ream, E., et al. 2000. Cancer-related fatigue: inevitable, unimportant and untreatable? Results of a multi-centre patient survey. *Ann Oncol*, 11, 971-5.
- [96] Tanaka, K., Akechi, T., Okuyama, T., et al. 2002. Prevalence and screening of dyspnea interfering with daily life activities in ambulatory patients with advanced lung cancer. *J Pain Symptom Manage*, 23(6), 484-9.
- [97] Teckle, P., Peacock, S., McTaggart-Cowan, H., et al. 2011. The ability of cancer-specific and generic preference-based instruments to discriminate across clinical and self-reported measures of cancer severities. *Health Qual Life Outcomes*, 9, 106.
- [98] Tessen, H. W. 2014. Tumorregister Lungenkarzinom (TLK). Siebte Zwischenauswertung. Datenstand 31.01.2014. Data on file

- [99] Therasse, P., Arbuck, S. G., Eisenhauer, E. A., et al. 2000. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst*, 92(3), 205-16.
- [100] Velikova, G., Coens, C., Efficace, F., et al. 2012. Health-Related Quality of Life in EORTC clinical trials – 30 years of progress from methodological developments to making a real impact on oncology practice. *EJC Supplements*, 10(1), 141-149.
- [101] Wong, S. S.-L., 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.
- [102] World Health Organization - International Clinical Trial Registry Platform. 2014a. LUME-Lung 1: BIBF 1120 plus docetaxel as compared to placebo plus docetaxel in 2nd line non small cell lung cancer (EUCTR2007-004803-36-DE) [Online]. Adresse: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2007-004803-36-DE>. Zuletzt geändert am: 19.05.2014. [Aufgerufen am 20.10.2014].
- [103] World Health Organization - International Clinical Trial Registry Platform. 2014b. Nintedanib Plus Docetaxel in Advanced Non-small Cell Lung Cancer (NCT02231164) [Online]. Adresse: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02231164>. Zuletzt geändert am: 13.10.2014. [Aufgerufen am 20.10.2014].
- [104] Zhuang, S. H., Xiu, L. & Elsayed, Y. A. 2009. Overall survival: a gold standard in search of a surrogate: the value of progression-free survival and time to progression as end points of drug efficacy. *Cancer J*, 15(5), 395-400.

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 ³⁰] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

³⁰ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

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

Datenbankname	MEDLINE (mesd)
Suchoberfläche	OVID
Datum der Suche	16.10.2014
Zeitsegment	1946 to Present with Daily Update
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 - <i>Strategy minimizing difference between sensitivity and specificity</i> (Wong et al., 2006)

#	Suchbegriffe	Ergebnis
1	(nintedanib or intedanib or "BIBF 1120" or BIBF1120).mp.	86
2	(vargatev or vargatef).mp.	2
3	(928326-83-4 or 656247-17-5 or 656247-18-6).rn.	0
4	1 or 2 or 3	86
5	randomized controlled trial.pt. or randomized.mp.	575420
6	placebo.mp.	158956
7	5 or 6	622109
8	Carcinoma, Non-Small-Cell Lung/	33710
9	(nonsmall cell\$ or non small cell\$ or NSCLC).mp.	40779
10	8 or 9	40779
11	4 and 7 and 10	6

Datenbankname MEDLINE(R) In-Process & Other Non-Indexed Citations (prem)
Suchoberfläche OVID
Datum der Suche 16.10.2014
Zeitsegment October 15, 2014
Suchfilter Keine Filter verwendet*

#	Suchbegriffe	Ergebnis
1	(nintedanib or intedanib or "BIBF 1120" or BIBF1120).mp.	24
2	(vargatev or vargatef).mp.	2
3	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	0
4	1 or 2 or 3	24
5	(random\$ or double blind or single blind or crossover or cross over or controlled trial).mp.	76353
6	placebo.mp.	9347
7	5 or 6	78697
8	(nonsmall cell\$ or non small cell\$ or NSCLC).mp.	3875
9	4 and 7 and 8	1
* In der Datenbank wurde kein kontrolliertes Vokabular verwendet /keine Indexierung durchgeführt, daher ist die Verwendung eines Filters mit Deskriptoren nicht möglich.		

Datenbankname	EMBASE (emedd)
Suchoberfläche	OVID
Datum der Suche	16.10.2014
Zeitsegment	1988 - 2014-October 15
Suchfilter	Filter für randomisierte, kontrollierte Studien nach Wong 2006 - <i>Strategy minimizing difference between sensitivity and specificity</i> (Wong et al., 2006)

#	Suchbegriffe	Ergebnis
1	nintedanib/	378
2	(nintedanib or intedanib or "BIBF 1120" or BIBF1120).mp.	591
3	(vargatev or vargatef).mp.	42
4	(928326-83-4 or 656247-17-5 or 656247-18-6).rn.	419
5	1 or 2 or 3 or 4	592
6	(random\$. or double-blind\$.)tw.	893841
7	placebo\$.mp.	286744
8	6 or 7	1031787
9	lung non small cell cancer/	61202
10	(nonsmall cell\$ or non small cell\$ or NSCLC).mp.	71667
11	9 or 10	71667
12	5 and 8 and 11	96

Datenbankname Embase Daily Alerts (emda)
Suchoberfläche OVID
Datum der Suche 16.10.2014
Zeitsegment 2014/08/18-2014/10/15
Suchfilter keine Filter verwendet*

#	Suchbegriffe	Ergebnis
1	(nintedanib or intedanib or "BIBF 1120" or BIBF1120).mp.	13
2	(vargatev or vargatef).mp.	1
3	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	0
4	1 or 2 or 3	13
5	(random\$.tw. or double-blind\$.tw.	13989
6	placebo\$.mp.	2380
7	5 or 6	14572
8	(nonsmall cell\$ or non small cell\$ or NSCLC).mp.	1087
9	4 and 7 and 8	0
* In der Datenbank wurde kein kontrolliertes Vokabular verwendet /keine Indexierung durchgeführt, daher ist die Verwendung eines Filters mit Deskriptoren nicht möglich.		

Datenbankname Cochrane Central Register of Controlled Trials
Suchoberfläche OVID
Datum der Suche 16.10.2014
Zeitsegment 1991 – September 2014
Suchfilter keine Filter verwendet

#	Suchbegriffe	Ergebnis
1	(nintedanib or intedanib or "BIBF 1120" or BIBF1120).mp.	18
2	(vargatev or vargatef).mp.	0
3	(928326-83-4 or 656247-17-5 or 656247-18-6).mp.	0
4	1 or 2 or 3	18
5	Carcinoma, Non-Small-Cell Lung/	1864
6	(nonsmall cell\$ or non small cell\$ or NSCLC).mp.	4115
7	5 or 6	4115
8	4 and 7	4

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

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	ClinicalTrials.gov
Internetadresse	http://clinicaltrials.gov/
Datum der Suche	16.10.2014
Suchstrategie	nintedanib OR intedanib OR "BIBF 1120" OR BIBF1120
Filter	Study Type: Interventional Studies, Phase 2, 3, 4 Condition: Carcinoma, Non-Small-Cell Lung OR nonsmall cells\$ OR non small cell\$ OR NSCLC
Treffer	5

Die Datenbankabfrage über ClinicalTrials.gov ergab anhand der Suchanfrage 5 Treffer.

Studienregister	WHO-ICTRP
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	16.10.2014
Suchstrategie	nintedanib OR intedanib OR BIBF 1120 OR BIBF1120
Filter	Conditions: NSCLC* OR nonsmall cell\$ or non small cell
Treffer	8
<p>*Synonyms searched: NSCLC OR "Carcinoma, Non Small Cell Lung" OR "Carcinoma, Non-Small Cell Lung" OR "Carcinoma, Non-Small-Cell Lung" OR "Carcinomas, Non-Small-Cell Lung" OR "Lung cancer, non-small cell" OR "Lung Carcinoma, Non-Small-Cell" OR "Lung Carcinomas, Non-Small-Cell" OR "Non Small Cell Lung Carcinoma" OR "non-small cell lung cancer" OR "Non-Small Cell Lung Carcinoma" OR "Non-Small-Cell Lung Carcinoma" OR "Non-Small-Cell Lung Carcinomas" OR "nonsmall cell lung cancer" OR "NONSMALL CELL LUNG CARCINOMA" OR "NSCLC"</p>	

Die Datenbankabfrage über das WHO-ICTRP ergab anhand der Suchanfrage 8 Treffer.

Studienregister	EU-CTR
Internetadresse	www.clinicaltrialsregister.eu
Datum der Suche	16.10.2014
Suchstrategie	(nintedanib OR intedanib OR "BIBF 1120" OR BIBF1120) AND (NSCLC OR "nonsmall cell\$" OR "non small cell\$").
Filter	Study Type: Phase 2, 3, 4
Treffer	5

Anhand der Datenbankabfrage des EU-CTR Registers konnten 5 zutreffende Studien ermittelt werden.

Studienregister	PharmNet.bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	16.10.2014
Suchstrategie	(((((ASUINN="nintedanib" OR ASUSUPPCODE="nintedanib" OR ASUDESCNAME="nintedanib")) OR (TI=nintedanib OR CTILAY=nintedanib OR CTABBRTI=nintedanib)) OR (ASUINN="BIBF 1120" OR ASUSUPPCODE="BIBF 1120" OR ASUDESCNAME="BIBF 1120")) OR (TI=BIBF 1120 OR CTILAY=BIBF 1120 OR CTABBRTI=BIBF 1120)) OR (ASUINN="BIBF1120" OR ASUSUPPCODE="BIBF1120" OR ASUDESCNAME="BIBF1120")) OR (TI=BIBF1120 OR CTILAY=BIBF1120 OR CTABBRTI=BIBF1120)) OR (ASUINN="intedanib" OR ASUSUPPCODE="intedanib" OR ASUDESCNAME="intedanib")) OR (TI=intedanib OR CTILAY=intedanib OR CTABBRTI=intedanib)) AND (GIPHASTWO="1" OR GIPHASTHR="1" OR GIPHASFOU="1")
Filter	Study Type: Phase 2, 3, 4
Treffer	14

Die Registersuche bei PharmNet.bund lieferte 14 Treffer.

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

#	Publikation	Ausschlussgrund
[1]	Crino, L. and G. Metro (2014). "Therapeutic options targeting angiogenesis in nonsmall cell lung cancer." <i>European Respiratory Review</i> 23(131): 79-91.	E6: Review
[2]	Filipits, M. (2014). "New developments in the treatment of squamous cell lung cancer." <i>Current Opinion in Oncology</i> 26(2): 152-158.	E6: Review; E1: Population
[3]	Hanna, N.H. et al.(2014). "LUME-Lung 2: a multicenter, randomized, double-blind, phase III study of nintedanib plus pemetrexed versus placebo plus pemetrexed in patients with advanced nonsquamous non-small cell lung cancer (NSCLC) after failure of first-line chemotherapy." <i>J Clin Oncol</i> 2013;31:abstr 8034.	E2: Intervention; E6: Kongresspublikation; A1: Außerhalb der Zulassung
[4]	Onn, A., et al. (2014). "Angiogenesis inhibition and lung-cancer therapy." <i>The Lancet Oncology</i> 15(2): 124-125.	E6: Kommentar zur Lume-Lung 1 Studie
[5]	Pirker, R. (2014). "Novel drugs against non-small-cell lung cancer." <i>Current Opinion in Oncology</i> 26(2): 145-151.	E6: Review
[6]	Reck, M., et al. (2014). "Nintedanib (BIBF 1120) + docetaxel as 2ndline therapy in patients with stage iiib/iv or recurrent NSCLC: Results of the phase iii, randomised, double-blind lume-lung 1 trial. (for the lume-lung 1 study group)." <i>Journal of Thoracic Oncology</i> 1): S39-S40.	E6: Kongresspublikation
[7]	Reck, M., et al. (2014). "Nintedanib (BIBF 1120) + docetaxel as second-line therapy in patients with stage IIIB/IV or recurrent NSCLC: Results of the phase III, randomised, double-blind LUME-Lung 1 trial." <i>Lung Cancer</i> 83: S12.	E6: Kongresspublikation
[8]	Reck, M., et al. (2013). "Impact of tumor burden on the overall survival analysis of the lume-lung 1 study: A randomized, double-blind phase 3 trial of nintedanib (BIBF 1120) + docetaxel in NSCLC patients progressing after first-line chemotherapy." <i>Journal of Thoracic Oncology</i> 8: S196.	E6: Kongresspublikation

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

#	Studienbezeichnungen	Text	Ausschlussgrund / Verletzt Einschlusskriterium
Clinicaltrials.gov			
[1]	NCT01948141	Nintedanib in Treating Patients With Advanced Non-Small Cell Lung Cancer Who Have Failed Up to Two Previous Chemotherapy Regimens. ClinicalTrials.gov. 2014 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://ClinicalTrials.gov/show/NCT01948141	E1, E2, A1
[2]	NCT02182050	Efficacy and Safety of Oral Treatment With BIBF 1120 ES in Advanced Non-small-cell Lung Cancer. ClinicalTrials.gov. 2014 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://ClinicalTrials.gov/show/NCT02182050	E2, A1
[3]	NCT00806819	Lume Lung 2: BIBF 1120 Plus Pemetrexed Compared to Placebo Plus Pemetrexed in 2nd Line Nonsquamous NSCLC. ClinicalTrials.gov. 2014 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://ClinicalTrials.gov/show/NCT00806819	E2, A1
WHO-ICTRP			
[4]	EUCTR2008-002072-10-HU	BIBF 1120 plus pemetrexed compared to placebo plus pemetrexed in 2nd line nonsquamous NSCLC. ICTRP. 2012. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-002072-10-HU	E2, A1
[5]	EUCTR2010-019707-32-GB	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. - LUME-Lung 3. ICTRP. 2012. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-019707-32-GB	E1, E2, A1

#	Studien- bezeichnungen	Text	Ausschlussgrund / Verletzt Einschluss- kriterium
[6]	EUCTR2013-004324-11-NL	Study with BIBF1120 in lung cancer patients with an abnormal fibroblast growth factor-1 receptor (NVALT-15 study). ICTRP. 2014. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-004324-11-NL	E2, A1
[7]	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. 2013. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01683682	E1, E2, A1
[8]	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. 2013. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01684111	E1, E2, A1
[9]	NCT01948141	Nintedanib in Treating Patients With Advanced Non-Small Cell Lung Cancer Who Have Failed Up to Two Previous Chemotherapy Regimens. ICTRP. 2014. [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01948141	E1, E2, A1
EU-CTR			
[10]	2005-000317-35-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. EU-CTR. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2005-000317-35	E2, A1
[11]	2008-002072-10-HU	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 first line chemotherapy. EU-CTR. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-002072-10	E2, A1

#	Studien- bezeichnungen	Text	Ausschlussgrund / Verletzt Einschluss- kriterium
[12]	2010-019707-32- GB	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: 20.10.2014]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-019707-32	E1, E2, A1
[13]	2013-004324-11- NL	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 study). EU-CTR. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-004324-11	E2, A1
PharmNet.bund			
[14]	2006-002875-42- 4033015	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[15]	2006-004528-35- 4032436	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[16]	2008-002072-10- 4036707	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1

#	Studienbezeichnungen	Text	Ausschlussgrund / Verletzt Einschlusskriterium
[17]	2008-005364-14-4035491	A phase I-II study of BIBF 1120 and Folfox compared to Bevacizumab and Folfox in first line metastatic colorectal cancer patients.. PharmNet.Bund. 0000 [Zugriffsdatum: 27.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[18]	2009-011925-14-4036085	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[19]	2009-013788-21-4036253	A phase II open label, roll over study of the long term tolerability,safety and efficacy of oral BIBF 1120 in patients with IdiopathicPulmonary Fibrosis. PharmNet.Bund. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[20]	2010-019707-32-4037348	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[21]	2010-023050-37-4037142	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[22]	2010-024251-87-4037089	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). PharmNet.Bund. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1

#	Studienbezeichnungen	Text	Ausschlussgrund / Verletzt Einschlusskriterium
[23]	2010-024252-29-4037090	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). PharmNet.Bund. 0000 [Zugriffsdatum: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[24]	2011-000921-61-4037812	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[25]	2011-002766-21-4038213	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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1
[26]	2012-005201-48-4039055	Double blind, randomised, multicentre, phase II 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: 20.10.2014]. Verfügbar unter: http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm	E1, E2, A1

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 4-89 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-89 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-89 (Anhang): Studiendesign und -methodik für Studie LUME-Lung 1

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genauere Ziele, Fragestellung und Hypothesen	Bestimmung der Wirksamkeit und Sicherheit von Nintedanib+Docetaxel gegenüber Placebo+Docetaxel bei NSCLC-Patienten im Stadium IIIB/IV oder Patienten mit rezidiviertem NSCLC nach Erstlinien-Chemotherapie-Versagen.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multinationale, multizentrische, zweiarmige, randomisierte, doppelblinde, placebokontrollierte, parallele Phase-III-Studie mit Zuteilungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Durch Amendments zum Prüfplan wurden folgende für die Methodik relevante Änderungen vorgenommen:</p> <p><u>Amendment 1 (15. Mai 2009)</u></p> <ul style="list-style-type: none"> Die Ausnahme in Ausschlusskriterium 6 (siehe 4a), die eine Bestrahlung des Gehirns innerhalb der letzten 4 Wochen vor Behandlungsbeginn mit der Studienmedikation erlaubte, wurde aufgehoben. Die Laboruntersuchung von Bilirubin wurde dahingehend spezifiziert, dass im Falle von erhöhtem Bilirubin CTCAE-Grad ≥ 1 direkte und indirekte Bilirubin-Level zu bestimmen sind. Die Screening-Phase für außergewöhnliche Fälle (z.B. bei nicht akzeptabler Qualität der Bildgebung zur Baseline und notwendiger Wiederholung) wurde von 14 Tagen auf 3 Wochen ausgeweitet. Die Bedeutung eines Monitoring von Hypersensitivitäten gegenüber Docetaxel wurde stärker hervorgehoben. Dem Data Monitoring Committee (DMC) wurde auch die Verantwortung für eine kontinuierliche Evaluierung der Sicherheit von Patienten mit Plattenepithelkarzinom übertragen. Die Prüfarzte wurden angewiesen, die Zeitpunkte der letzten Nahrungs-/ Getränkeaufnahme vor jeder Einnahme von Nintedanib bzw. Placebo für die beiden Tage vor sowie am Tag der Entnahme der Blutproben zur Pharmakokinetik im Case Report Form (CRF) zu dokumentieren. Für die Messung der Körpertemperatur wurde die Art der erlaubten Messung ausgeweitet. Neben der oralen Messung sowie der Messung im Ohr waren laut Amendment 1 auch weitere Messarten erlaubt. Die Messung sollte jedoch über alle Zeitpunkte identisch erfolgen. Bei zwei Ausschlusskriterien wurden die Messeinheiten korrigiert zu „absolute Neutrophilanzahl $< 1.500/\mu\text{L}$“ sowie zu „Blutplättchenanzahl $< 100.000/\mu\text{L}$“. Diese Änderungen beeinflussten auch die Kriterien, die vor einer erneuten Behandlung der Patienten mit Docetaxel zu

Item ^a	Charakteristikum	Studieninformation
		<p>erfüllen sind. Die Wirkstärke von Docetaxel wurde, entsprechend der zur Zeit der Protokollerstellung gültigen Produktinformation, auf "20 mg Konzentrat" angepasst.</p> <ul style="list-style-type: none"> • Ein Zeitfenster von 2 Tagen zur Erfassung von Sicherheitslaborparametern (z. B. Urinproben, Schilddrüsenhormone) beim ersten Behandlungszyklus wurde eingeführt.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p><u>Einschlusskriterien:</u></p> <ol style="list-style-type: none"> 1. Männliche oder weibliche Patienten im Alter von 18 Jahren oder älter 2. Histologisch oder zytologisch bestätigtes, lokal fortgeschrittenes und/oder metastasiertes NSCLC im Stadium IIIB/IV (UICC/AJCC 6. oder 7. Auflage) oder rezidiertes NSCLC (aller Histologien) 3. Rezidiv oder Versagen einer vorherigen Chemotherapie in der Erstlinie (im Fall einer rezidierten Erkrankung war ein zusätzliches Regime als adjuvante, neoadjuvante oder neoadjuvante plus adjuvante Therapie erlaubt.) 4. Mindestens eine Tumorkläsion, die innerhalb der letzten 3 Monate nicht bestrahlt wurde und deren größter Durchmesser zumindest in einer Dimension ≥ 20 mm (MRI [Magnetic Resonance Imaging {Magnetresonanztomographie}], konventionelles CT [Computertomographie] oder ≥ 10 mm (Spiral-CT) betrug 5. Lebenserwartung von mindestens drei Monaten 6. ECOG-PS 0 oder 1 7. Schriftliche Einverständniserklärung entsprechend der ICH-GCP Standards und lokaler gesetzlicher Vorgaben <p><u>Ausschlusskriterien:</u></p> <ol style="list-style-type: none"> 1. Mehr als ein vorheriges Chemotherapieregime für das fortgeschrittene und/oder metastasierte oder rezidierte NSCLC 2. Mehr als ein Chemotherapieregime (neoadjuvant oder adjuvant oder neoadjuvant plus adjuvant) vor der Erstlinien-Chemotherapie für das fortgeschrittene und/oder metastasierte oder rezidierte NSCLC 3. Vorherige Therapie mit anderen VEGFR-Inhibitoren (außer Bevacizumab) oder Docetaxel zur Behandlung des NSCLC 4. Andauernde, klinisch relevante und mit der Therapie in Zusammenhang stehende Toxizitäten der vorherigen Chemo- und/oder Strahlentherapie 5. Behandlung mit anderen Studienmedikationen oder Behandlung in einer anderen klinischen Studie innerhalb der letzten 4 Wochen vor Behandlungsbeginn oder begleitend zu dieser Studie 6. Chemo-, Hormon-, Strahlen- (mit Ausnahme für die Bestrahlung von Gehirn und Extremitäten) oder Immuntherapie oder Therapie mit monoklonalen Antikörpern oder niedermolaren TKI innerhalb der letzten

Item ^a	Charakteristikum	Studieninformation
		<p>4 Wochen vor Behandlungsbeginn mit der Studienmedikation; d. h. es mussten mindestens 4 Wochen zwischen der letzten Krebstherapie und der ersten Einnahme von Nintedanib liegen. (Mit Protokoll-Amendment 1 vom 15. Mai 2009 wurde die Ausnahme für die Bestrahlung des Gehirns aufgehoben.)</p> <ol style="list-style-type: none"> 7. Strahlentherapie (mit Ausnahme für die Bestrahlung von Hirn und Extremitäten) innerhalb der letzten 3 Monate vor der Bildgebung zur Baseline 8. Aktive Hirnmetastasen (z. B. stabil <4 Wochen, keine vorherige adäquate Strahlentherapie, symptomatisch, notwendige Behandlung mit Anti-Konvulsiva; Therapie mit Dexamethason war erlaubt, sofern für mindestens einen Monat vor Randomisierung in stabiler Dosierung verabreicht) oder leptomeningeale Erkrankung 9. Radiografischer Nachweis kavitärer oder nekrotischer Tumore 10. Zentral lokalisierte Tumore mit radiografischem Nachweis (CT oder MRI) einer lokalen Invasion großer Blutgefäße 11. Klinisch bedeutsamer Bluthusten innerhalb der letzten 3 Monate (mehr als ein Teelöffel frisches Blut pro Tag) 12. Therapeutische Antikoagulation (mit Ausnahme von niedrigdosiertem Heparin und/oder Heparinspülung bei Dauerkathetern) oder antithrombozytäre Therapie (mit Ausnahme von chronischer Therapie mit niedrig dosierter Acetylsalicylsäure ≤ 325 mg/Tag) 13. Bedeutsames thrombotisches oder klinisch relevantes Blutungsereignis in den letzten 6 Monaten 14. Bekannte erbliche Prädisposition für Blutungen oder Thrombosen 15. Bedeutsame kardiovaskuläre Erkrankungen (z. B. medikamentös nicht kontrollierbare Hypertonie, instabile Angina, Myokardinfarkt innerhalb der letzten 6 Monate, kongestives Herzversagen >New York Heart Association (NYHA) Stadium II, schwerwiegende Herzrhythmusstörung, Perikarderguss) 16. Serumkreatinin >1,5-fach erhöhte Obergrenze des Normalwertes 17. Proteinurie vom CTCAE-Grad ≥ 2 18. Gesamtbilirubin >Obergrenze des Normalwertes 19. ALT und/oder AST >1,5-fach erhöhte Obergrenze der Normalwerte 20. Prothrombinzeit und/oder partielle Thromboplastinzeit >50% Abweichung von den Normalwerten 21. Absolute Neutrophilenzahl (Absolute Neutrophil Count [ANC]) <1.500/μl 22. Thrombozytenzahl <100.000/μl 23. Hämoglobin <9,0 g/dL 24. Signifikanter Gewichtsverlust (>10%) innerhalb der letzten 6 Wochen vor Behandlung in der Studie 25. Periphere Neuropathie vom CTCAE-Grad 2 außer

Item ^a	Charakteristikum	Studieninformation
		<p>aufgrund von Traumata</p> <ol style="list-style-type: none"> 26. Bereits bestehende Bauchwassersucht (Aszites) und/oder klinisch signifikanter Pleuraerguss 27. Größere Verletzungen und/oder Operationen innerhalb der letzten 10 Tage vor der Randomisierung mit noch nicht abgeschlossener Wundheilung 28. Schwerwiegende Infektionen, die eine systemische Antibiotikatherapie (z. B. antiviral, antimikrobiell, antimykotisch) erfordern 29. Dekompensierter Diabetes Mellitus oder andere Kontraindikationen für eine hochdosierte Kortikosteroidtherapie 30. Gastrointestinale Störungen oder Abnormalitäten, die mit der Absorption der Studienmedikation interferieren könnten 31. Aktive oder chronische Hepatitis C- und/oder Hepatitis B-Infektion 32. Schwere Krankheit oder begleitende nicht-onkologische Erkrankung wie z. B. neurologische, psychiatrische oder infektiöse Erkrankungen oder aktive Ulzera (Gastrointestinaltrakt, Haut) oder Auffälligkeiten in den Laborwerten, die möglicherweise das mit einer Studienteilnahme oder der Verabreichung der Studienmedikation verbundene Risiko erhöhen und die den Patienten, nach Einschätzung der Prüfarzte, ungeeignet für die Studie machen. 33. Patienten, die sexuell aktiv sind und die keine medizinisch akzeptierte Form der Empfängnisverhütung (z. B. Implantate, Spritzen, kombinierte orale Kontrazeptiva, einige Spiralen oder Vasektomie des Partners bei teilnehmenden Frauen, Kondome bei teilnehmenden Männern) während der Studie und mindestens 12 Monate nach dem Ende der aktiven Behandlung anwenden möchten. 34. Schwangerschaft oder Stillen 35. Psychologische, familiäre, soziologische oder geografische Faktoren, die möglicherweise die Einhaltung des Studienprotokolls oder des Zeitplans der Follow-up-Visiten erschweren 36. Patienten, die das Protokoll nicht einhalten können 37. Aktiver Alkohol- oder Drogenmissbrauch 38. Andere maligne Erkrankung innerhalb der letzten 3 Jahre mit Ausnahme von Basalzellkarzinom oder eines Carcinoma <i>in situ</i> des Gebärmutterhalses 39. Kontraindikationen für eine Therapie mit Docetaxel 40. Schwere Überempfindlichkeitsreaktionen auf Docetaxel oder andere Medikamente, die mit Polysorbat 80 (Tween 80) formuliert wurden 41. Überempfindlichkeit gegenüber Nintedanib und/oder anderen Hilfsstoffen der Studienmedikation 42. Überempfindlichkeit gegenüber Kontrastmittel

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische, multinationale Studie mit 209 Zentren in 27 Ländern in Südafrika, Asien, und Europa.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<u>Nintedanib+Docetaxel:</u> 200 mg Nintedanib (b.i.d.) zur oralen Einnahme an den Tagen 2-21 jedes 21-tägigen Behandlungszyklus und Docetaxel als intravenöse Infusion (75 mg/m ²) jeweils am ersten Tag jedes 21-tägigen Behandlungszyklus <u>Placebo+Docetaxel:</u> Placebo als Weichgelatine kapsel zur oralen Einnahme an den Tagen 2-21 jedes 21-tägigen Behandlungszyklus und Docetaxel als intravenöse Infusion (75 mg/m ²) jeweils am ersten Tag jedes 21-tägigen Behandlungszyklus
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	Primärer Endpunkt: <ul style="list-style-type: none"> Progressionsfreies Überleben (PFS) gemäß unabhängiger, radiologischer Beurteilung (IRR) nach RECIST-Kriterien 1.0 Sekundäre Endpunkte: <ul style="list-style-type: none"> Gesamtüberleben (OS) wichtigster sekundärer Endpunkt PFS gemäß Beurteilung durch Prüfarzt Tumoransprechen gemäß RECIST-Kriterien 1.0 <ul style="list-style-type: none"> Bestätigtes objektives Ansprechen (ORR) Dauer des bestätigten, objektiven Ansprechens Zeit bis zum objektiven Ansprechen Tumorkontrolle Dauer der Tumorkontrolle Veränderung der Tumorgröße Klinische Verbesserung (Gewicht und/oder ECOG-PS) Gesundheitsbezogene Lebensqualität (HRQoL) gemessen mit den standardisierten Fragebögen EQ-5D, EORTC QLQ-C30 und dem Lungenkrebsmodul EORTC QLQ-LC13. Endpunkte Husten (QLQ-LC13 Frage 1), Dyspnoe (QLQ-LC13 Fragen 3-5, QLQ-C30 Frage 8) und Schmerzen (QLQ-C30 Fragen 9 und 19, QLQ-LC13 Fragen 10-12) UE (Inzidenz und Intensität gemäß CTCAE-Version 3.0) und Veränderungen von Sicherheitslaborparametern Pharmakokinetik von Nintedanib
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Bei der Bestimmung der Fallzahlen wurde für die Alternativhypothese angenommen, dass die Kombination von Nintedanib+Docetaxel im Vergleich zu Docetaxel das mediane PFS um 28-32% erhöht. Diese Abschätzung erfolgte unter Annahme

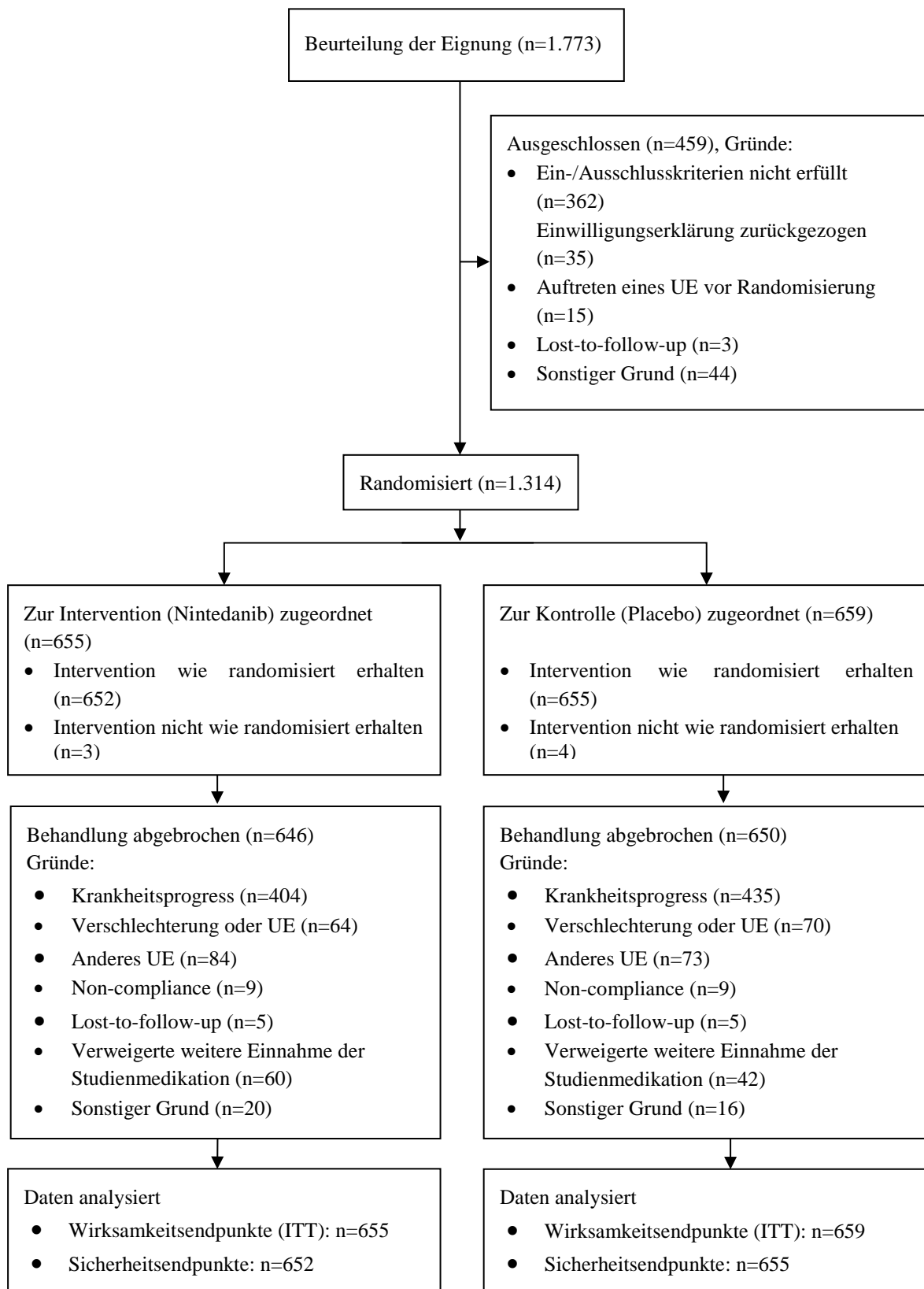
Item ^a	Charakteristikum	Studieninformation
		<p>eines medianen PFS von 4 Monaten für Patienten mit einem ECOG-PS von 0 oder 1, die ausschließlich mit Docetaxel behandelt wurden. Daraus wurde errechnet, dass 713 PFS-Ereignisse benötigt werden, um eine statistische Power von 90% zu erreichen.</p> <p>Für den wichtigsten sekundären Endpunkt OS wurde ein medianes OS von 9 Monaten unter Docetaxel und eine Verlängerung um 18% unter Nintedanib+Docetaxel angenommen. Um diesen Effekt mit einer Power von 80% nachzuweisen, mussten 1.151 Todesfälle beobachtet und 1300 Patienten rekrutiert werden.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Für den wichtigsten sekundären Endpunkt OS wurde zu dem Zeitpunkt der primären PFS-Analyse eine Zwischenauswertung durchgeführt. In der finalen OS-Analyse wurde das Alpha-Niveau für diese Zwischenauswertung adjustiert.</p> <p>Studienabbruch durch Sponsor, falls:</p> <ul style="list-style-type: none"> • Einschlussziele nicht erreicht werden • Informationen über die Sicherheit und/oder Wirksamkeit von Nintedanib bekannt werden, die die Fortführung der Studie signifikant beeinflussen können • "Futility-Analysen" oder die Empfehlung des DMC einen Abbruch nahelegen • Das Entwicklungsprogramm von Nintedanib eingestellt wird • Eine Verletzung des Protokolls, der GCP (Good Clinical Practice)-Richtlinien oder der Verträge mit einem Zentrum oder Prüfarzt auftritt, sodass die ordnungsgemäße Durchführung der Studie beeinträchtigt wird. <p>Studienabbruch durch Patienten, falls:</p> <ul style="list-style-type: none"> • Rücknahme der Einverständniserklärung • Protokollverletzung • Unfähigkeit zur Fortsetzung der Therapie bzw. keine weitere Dosisreduzierung möglich • Patienten weniger als 4 Zyklen der Therapie (Nintedanib+Docetaxel/Placebo+Docetaxel) erhalten hatten • Dokumentierte Krankheitsprogression • Einnahme von nicht erlaubter Begleitmedikation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierte Zuteilung zu einer Behandlungsgruppe.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	<p>Randomisierung in 4er-Blöcken.</p> <p>Stratifizierung innerhalb jedes Landes nach ECOG-PS (0 vs. 1), vorheriger Bevacizumab-Behandlung (Ja vs. Nein), Tumorhistologie (Plattenepithelkarzinom vs. Nicht-Plattenepithelkarzinom), Hirnmetastasen zu Untersuchungsbeginn (Ja vs. Nein).</p>

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	Der Zugang zum Randomisierungscode wurde auf einen bestimmten Personenkreis beschränkt. Er wurde kontrolliert und dokumentiert. Zuteilung von Medikationsnummern erfolgte mithilfe eines interaktiven Sprach- oder Webdialogsystems (IVRS und IWRS).
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?	s. o. Punkt 8a. Die Boehringer Ingelheim Pharma GmbH & Co. KG, Clinical Trial Support Group erstellte die Randomisierungsliste.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es waren a) die Patienten, b) die Prüfarzte und c) die durchführenden Studienmitarbeiter/Beurteiler verblindet. Alle Personen, die in die Studiendurchführung, in das Datenmanagement und in die Statistik involviert waren, blieben während der Studie verblindet. Die Zwischenanalyse erfolgte durch ein eigenes entblindetes Team. Patienten, Prüfarzte und die Studienmitarbeiter/Beurteiler, welche die Studie durchführten, blieben bis zur finalen Analyse des wichtigsten sekundären Endpunktes OS verblindet. Der leitende, koordinierende Prüfarzt wurde über die Zwischenergebnisse informiert. Für den restlichen Verlauf der Studie übernahm daraufhin ein anderer Prüfarzt die Leitung, welcher keinen Einblick in die Ergebnisse erhielt. Die Ergebnisse der Zwischenauswertung wurden erst nach der Entblindung des OS Endpunktes öffentlich publiziert. Die vorzeitige Entblindung war notfallmäßig möglich, wenn die weitere Behandlung des Patienten von der erhaltenen Intervention abhing. Der Sponsor und der Monitor wurden in diesem Fall umgehend informiert.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Placebo-Gelatinekapsel entsprach der Nintedanib-Gelatinekapsel mit Ausnahme des Wirkstoffs.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<u>Analysepopulation:</u> <ul style="list-style-type: none"> • Randomisierungsset (alle Patienten, die randomisiert wurden) • Treated Set (Alle Patienten des Randomisierungssets, die mindestens eine Dosis der Studienmedikation erhalten haben. Zuweisung der Patienten in die Behandlungsgruppen entsprechend der tatsächlich erhaltenen Behandlung.) <u>Primärer Endpunkt:</u>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • PFS: zentrale und unabhängige Beurteilung der Tumorprogression; Stratifizierter Logrank-Test mit zweiseitigem Signifikanzniveau von 0,05 <p><u>Sekundäre Endpunkte:</u></p> <ul style="list-style-type: none"> • Analyse des OS: Stratifizierter Logrank-Test mit zweiseitigem Signifikanzniveau von insgesamt 0,05. In der finalen OS-Analyse wurde das Alpha-Niveau adjustiert, um der Zwischenauswertung für OS zum Zeitpunkt der primären PFS-Analyse Rechnung zu tragen. Die finale OS-Analyse folgte einer hierarchischen Teststrategie: (1) Testung in der Population der Adenokarzinom-Patienten mit Zeit seit Beginn der Erstlinien-Therapie <9 Monate, (2) Testung in der Population aller Adenokarzinom-Patienten, (3) Testung in der Population aller Patienten der LUME-Lung 1 Studie. • Tumoransprechen gemäß RECIST-Kriterien 1.0 <ul style="list-style-type: none"> ○ Bestätigtes ORR: Logistisches Regressionsmodell ○ Dauer des bestätigten, objektiven Ansprechens: Kaplan-Meier Schätzer und 95%-KI ○ Zeit bis zum objektiven Ansprechen: Kaplan-Meier Schätzer und 95%-KI ○ Tumorkontrolle: Logistisches Regressionsmodell ○ Dauer der Tumorkontrolle: Kaplan-Meier Schätzer und 95%-KI ○ Veränderung der Tumorgroße: ANOVA (Analysis of Variance [Varianzanalyse]) • Klinische Verbesserung (Gewicht und/oder ECOG-PS): Stratifiziertes Cox Modell (Breslow-Schätzer) • HRQoL gemessen mit den standardisierten Fragebögen EQ-5D, EORTC QLQ-C30 und dem Lungenkrebsmodul EORTC QLQ-LC13. Endpunkte Husten (QLQ-LC13 Frage 1), Dyspnoe (QLQ-LC13 Fragen 3-5, QLQ-C30 Frage 8) und Schmerzen (QLQ-C30 Fragen 9 und 19, QLQ-LC13 Fragen 10-12): Stratifiziertes Cox Modell (Breslow-Schätzer) für Hazard Ratio und 95%-KI • Unerwünschte Ereignisse: Anzahl der Patienten mit UE
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Im Studienbericht dargestellte Subgruppenanalysen für PFS und OS:</p> <ul style="list-style-type: none"> • ECOG-PS zur Baseline (ECOG 0, 1) • Vorhandensein von Hirnmetastasen zur Baseline (Ja, Nein) • Vorbehandlung mit Bevacizumab (Ja, Nein) • Geschlecht (männlich, weiblich) • Alter (<65 Jahre, ≥65 Jahre) • Zeit seit Beginn der Erstlinien-Therapie (<9 Monate, ≥9 Monate) • Geografische Region (Asien, Europa, Südafrika) • Rauchverhalten (Nieraucher, aktiver Raucher, Ex-Raucher) • Ethnie (asiatisch, nicht-asiatisch) • Tumorlast gemessen anhand der Summe der größten Durchmesser der Zielläsionen zur Baseline (<7,5 cm, ≥7,5 cm) • Bestes Ansprechen auf die Erstlinien-Behandlung (CR/PR/SD, PD (Progressive Disease [fortschreitende Erkrankung])),

Item ^a	Charakteristikum	Studieninformation
		unbekannt/fehlend/nicht verfügbar (not available [NA]) <ul style="list-style-type: none"> • Vorhandensein von Lebermetastasen (Ja, Nein) • Vorhandensein von Nebennierenmetastasen (Ja, Nein) • Erkrankungsstadium bei Erstdiagnose (<IIIB/IV, IIIB, IV) • Begleittherapie mit Biphosphonaten zur Baseline (Ja, Nein) • Anzahl metastasierter Organe zur Baseline (≤ 2 metastasierte Organe, > 2 metastasierte Organe, ohne zentrale Beurteilung) • Lactatdehydrogenase (LDH)-Gehalt zur Baseline (Baseline LDH ≤ 1, Baseline LDH > 1) • Tumorhistologie (Adenokarzinom, Nicht-Adenokarzinom)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<u>Nintedanib+Docetaxel:</u> a) 655 Patienten b) 652 Patienten c) 655 Patienten (Tumorhistologie Adenokarzinom 322) <u>Placebo+Docetaxel:</u> a) 659 Patienten b) 655 Patienten c) 659 Patienten (Tumorhistologie Adenokarzinom 336)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<u>Nintedanib+Docetaxel:</u> Aufgrund UE: 2 Patienten Andere Gründe: 1 Patient <u>Placebo+Docetaxel:</u> Aufgrund UE: 2 Patienten Non-compliance mit dem Protokoll : 1 Patient Andere Gründe: 1 Patient
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Beginn der Studie: 23. Dezember 2008 (erster Patient eingeschrieben); letzter Patient randomisiert: 9. Februar 2011; Datenschnitt für primäre Analysen: 2. November 2010 (Wirksamkeitsdaten) und 20. April 2011 (Sicherheitsdaten); Datum des Berichts mit den primären Analysen: 19. September 2012; Datum des Berichts mit den finalen OS Analysen: 26. August 2013
14b	Informationen, warum die Studie endete oder beendet wurde	Laufende Studie, Rekrutierung abgeschlossen. Finale OS Analyse abgeschlossen.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.



Quelle: (Boehringer Ingelheim Pharma GmbH & Co. KG, 2012b; Boehringer Ingelheim Pharma GmbH & Co. KG, 2013a)

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

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

Der Bogen gliedert sich in zwei Teile:

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

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

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

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

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

Tabelle 4-90 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie LUME-Lung 1

Studie: LUME-Lung 1

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Clinical Trial Protocol 1199.13</u> 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 (LUME Lung 1). 2008</p>	LUME-Lung 1: Protocol
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Trial Statistical Analysis Plan 1199.13</u> 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 (LUME Lung 1). 2011</p>	LUME-Lung 1: TSAP
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Trial Statistical Analysis Plan Addendum 1199.13</u> 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 (LUME-Lung 1). 2013</p>	LUME-Lung 1: TSAP Addendum

Genauere Benennung der Quelle	Kürzel
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Clinical Trial Report 1199.13 (Primary PFS Analysis):</u> 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 (LUME Lung 1). 2012</p> <p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Clinical Trial Report 1199.13 (Final OS Analysis):</u> 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 (LUME Lung 1). 2013</p>	<p>LUME-Lung 1: Clinical Trial Report (Primary PFS Analysis)</p> <p>LUME-Lung 1: Clinical Trial Report (Final OS Analysis)</p>
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Biomarker Report 1199.13-14:</u> Exploratory analysis to identify and confirm a clinical biomarker defining a subgroup of second line non small cell lung cancer patients with a potential treatment benefit from nintedanib when given in combination with pemetrexed or docetaxel (LUME Lung 1 und LUME-Lung 2). 2013</p> <p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Trial Statistical Analysis Plan 1199.14:</u> Multicenter, randomized, 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 first line chemotherapy (LUME-Lung 2). 2012</p>	<p>LUME-Lung 1 und LUME-Lung 2: Biomarker Report</p> <p>LUME-Lung 2: Biomarker TSAP</p>
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Health Economic Report 1199.13:</u> Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared with placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non-small cell lung cancer after failure of first-line chemotherapy (LUME Lung 1). 2014</p>	<p>LUME-Lung 1: Health Economic Report</p>
<p>Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Weitere Analysen (LUME-Lung 1).</u> 2014</p>	<p>LUME-Lung 1: Weitere Analysen</p>

Genaue Benennung der Quelle	Kürzel
Boehringer Ingelheim Pharma GmbH & Co. KG. <u>Subgruppenanalysen EQ-5D und EQ-VAS (LUME-Lung 1)</u> . 2014	LUME-Lung 1: Subgruppenanalysen EQ-5D und EQ-VAS

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:

Sowohl die Patienten als auch die behandelnden Personen agierten verblindet, da die Patienten in der Kontrollgruppe mit Nintedanib entsprechenden Placebo-Kapseln behandelt wurden.

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die behandelnden sowie weiterbehandelnden Personen agierten verblindet, da die Patienten in der Kontrollgruppe mit Nintedanib entsprechenden Placebo-Kapseln behandelt wurden.

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:

Die Studie erfolgte randomisiert mit verdeckter Gruppenzuteilung und sowohl die Patienten als auch die behandelnden Personen agierten verblindet. Es erfolgte eine ergebnisunabhängige Berichterstattung und weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**Endpunkt: Gesamtüberleben****1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor. Zudem erfolgte kein Cross-over.

Endpunkt: Progressionsfreies Überleben**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Objektive Ansprechrates**1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Krankheitskontrollrate**1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 (Responderanalysen)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber (die Patienten) agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Symptomatik, gemessen anhand der Symptomskalen aus den Fragebögen EORTC QLQ-C30 sowie EORTC QLQ-LC13 (Zeit bis zur Verschlechterung)

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber (die Patienten) agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Gesundheitsbezogene Lebensqualität (Responderanalysen)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber (die Patienten) agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

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

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber (die Patienten) agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Gesundheitsbezogene Lebensqualität (EQ-5D UK Utility und EQ-VAS)**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber (die Patienten) agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Unerwünschte Ereignisse**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Die mittlere Behandlungszeit im Nintedanib+Docetaxel- und im Placebo+Docetaxel-Arm unterschied sich nur geringfügig (Nintedanib+Docetaxel: 5,4 Monate; Placebo+Docetaxel: 4,7 Monate). Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Schwerwiegende unerwünschte Ereignisse**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Die mittlere Behandlungszeit im Nintedanib+Docetaxel- und im Placebo+Docetaxel-Arm unterschied sich nur geringfügig (Nintedanib+Docetaxel: 5,4 Monate; Placebo+Docetaxel: 4,7 Monate). Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Unerwünschte Ereignisse mit CTCAE-Grad ≥ 3 **1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Die mittlere Behandlungszeit im Nintedanib+Docetaxel- und im Placebo+Docetaxel-Arm unterschied sich nur geringfügig (Nintedanib+Docetaxel: 5,4 Monate; Placebo+Docetaxel: 4,7 Monate). Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Unerwünschte Ereignisse, die zum Therapieabbruch führten**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Die mittlere Behandlungszeit im Nintedanib+Docetaxel- und im Placebo+Docetaxel-Arm unterschied sich nur geringfügig (Nintedanib+Docetaxel: 5,4 Monate; Placebo+Docetaxel: 4,7 Monate). Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

Endpunkt: Nebenwirkungen mit CTCAE-Grad ≥ 3 und einer Häufigkeit $\geq 5\%$ **1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt.

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 Endpunkterheber agierten verblindet, das ITT-Prinzip wurde adäquat umgesetzt und es erfolgte eine ergebnisunabhängige Berichterstattung. Die mittlere Behandlungszeit im Nintedanib+Docetaxel- und im Placebo+Docetaxel-Arm unterschied sich nur geringfügig (Nintedanib+Docetaxel: 5,4 Monate; Placebo+Docetaxel: 4,7 Monate). Weitere Aspekte, die zu einer Verzerrung führen könnten, lagen nicht vor.

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: endpunktübergreifend; 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:
