Dokumentvorlage, Version vom 20.01.2011

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Vandetanib (Caprelsa®)

AstraZeneca GmbH

Modul 4 A

Medulläres Schilddrüsenkarzinom

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Stand: 07.03.2013

Inhaltsverzeichnis

Seite

Tabellenverzeichnis **4**
Abbildungsverzeichnis **15**
Abkürzungsverzeichnis **18**
4.1 Zusammenfassung der Inhalte von Modul 4 23
4.2 Methodik 41
4.2.1 Fragestellung 41
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung 46
4.2.3 Informationsbeschaffung 51
4.2.3.1 Studien des pharmazeutischen Unternehmers 51
4.2.3.2 Bibliografische Literaturrecherche 51
4.2.3.3 Suche in Studienregistern 52
4.2.3.4 Selektion relevanter Studien 53
4.2.4 Bewertung der Aussagekraft der Nachweise 53
4.2.5 Informationssynthese und -analyse 56
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien 56
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien 57
4.2.5.3 Meta-Analysen 66
4.2.5.4 Sensitivitätsanalysen 67
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren 68
4.2.5.6 Indirekte Vergleiche 71
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen 72
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel 73
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel 73
4.3.1.1.1 Studien des pharmazeutischen Unternehmers 73
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche 74
4.3.1.1.3 Studien aus der Suche in Studienregistern 77
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 78
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 79
4.3.1.2.1 Studiendesign und Studienpopulationen 79
4.3.1.2.2 Verzerrungspotenzial auf Studienebene 89
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien 91
4.3.1.3.1 Endpunkte – RCT 92
4.3.1.3.1.1 Progressionsfreies Überleben (PFS) 94
4.3.1.3.1.2 Zeit bis zur Schmerzprogression (TWP) – RCT 104
4.3.1.3.1.3 Objektive Ansprechrate (ORR) – RCT 112
4.3.1.3.1.4 Krankheitskontrollrate (DCR) 119
4.3.1.3.1.5 Dauer des objektiven Ansprechens (DOR) 123

4.3.1.3.1.6 Explorativer Morbiditätsendpunkt Gewichtszunahme 129
4.3.1.3.1.7 Explorativer Morbiditätsendpunkt Einfluss der Behandlung auf Diarrhöen 132
4.3.1.3.1.8 Explorativer Morbiditätsendpunkt WHO-Performancestatus (WHO-PS) 141
4.3.1.3.1.9 Explorativer Morbiditätsendpunkt Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline) 146
4.3.1.3.1.10 Explorativer Morbiditätsendpunkt Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline) 151
4.3.1.3.1.11 Surrogatendpunkt: Biochemisches Ansprechen (CTN; CEA) – RCT 156
4.3.1.3.1.12 Gesamtüberleben (OS) 163
4.3.1.3.1.13 Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G – RCT 168
4.3.1.3.1.14 Unerwünschte Ereignisse (UE) – RCT 177
4.3.1.3.2 Subgruppenanalysen – RCT 186
4.3.1.3.2.1 Progressionsfreies Überleben (PFS) 192
4.3.1.3.2.2 Zeit bis zur Schmerzprogression (TWP) 203
4.3.1.3.2.3 RET-Mutations-Subgruppen 216
4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien 223
4.3.2 Weitere Unterlagen 228
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 228
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 229
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 229
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 229
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 229
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 232
4.3.2.2 Nicht randomisierte vergleichende Studien 232
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 232
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 232
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 233
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien 233
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 234
4.3.2.3 Weitere Untersuchungen 234
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 234
4.3.2.3.1.1 Studien des pharmazeutischen Unternehmers 235
4.3.2.3.1.2 Studien aus der bibliografischen Literaturrecherche 236
4.3.2.3.1.3 Studien aus der Suche in Studienregistern 238
4.3.2.3.1.4 Resultierender Studienpool: weitere Untersuchungen 238
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 238
4.3.2.3.2.1 Studiendesign und Studienpopulationen 238
4.3.2.3.2.2 Verzerrungspotenzial auf Studienebene 241

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen ... 242
4.3.2.3.3.1 Objektive Ansprechrate (ORR) – weitere Untersuchungen ... 242
4.3.2.3.3.2 Progressionsfreies Überleben (PFS) – weitere Untersuchungen ... 243
4.3.2.3.3.3 WHO *Performance Status* (WHO PS) – weitere Untersuchungen 244
4.3.2.3.3.4 Gesundheitsbezogene Lebensqualität (HRQoL) – Weitere Untersuchungen ... 245
4.3.2.3.3.5 Unerwünschte Ereignisse (UE) – weitere Untersuchungen ... 246
4.3.2.3.3.6 Subgruppenanalysen – weitere Untersuchungen ... 247
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen ... 247
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens ... 248
4.4.1 Beurteilung der Aussagekraft der Nachweise ... 248
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß ... 251
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht ... 266
4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens ... 267
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte ... 268
4.5.1 Begründung für die Vorlage indirekter Vergleiche ... 268
4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen ... 268
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen ... 269
4.5.4 Verwendung von Surrogatendpunkten ... 269
4.6 Liste der eingeschlossenen Studien ... 271
4.7 Referenzliste ... 272
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche ... 281
Anhang 4-B : Suchstrategien – Suche in Studienregistern ... 288
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund ... 294
Anhang 4-D Liste der abgebrochenen Studien ... 295
Anhang 4-E : Liste der laufenden Studien ... 296
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT ... 297
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten ... 341

Tabellenverzeichnis

Seite

Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Vandetanib 48

Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 73

Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel..... 74

Tabelle 4-4: Ein-/Ausschlusskriterien für die bibliografische Literaturrecherche..... 76

Tabelle 4-5: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel 78

Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel..... 79

Tabelle 4-7: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 80

Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel..... 82

Tabelle 4-9: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 1)..... 82

Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 2: Altersgruppen) (Alle Angaben in %) 82

Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 3a: Bestes Ansprechen bei vorangegangener Chemotherapie gegen MTC) (Alle Angaben in % bezogen auf die Anzahl der Patienten im Studienarm)..... 83

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 3b: Systemische Vortherapie gegen MTC) (Alle Angaben in %)..... 83

Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 4: Primärtumor) (Alle Angaben in %)..... 83

Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 5: Regionäre Lymphknoten) (Alle Angaben in %) 84

Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 6: Fernmetastasen/ Stadienklassifikation) (Alle Angaben in %) 84

Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 7: Patienten mit hereditärer Erkrankung) (Alle Angaben in %)..... 84

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 8a: RET-Mutationsstatus positiv, sporadisch) (Alle Angaben in %) (Gesamtpopulation)..... 85

Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 8b: RET-Mutationsstatus positiv, sporadisch) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)..... 85

Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9a: RET-Mutationsstatus hereditär) (Alle Angaben in %) (Gesamtpopulation) ... 85

Tabelle 4-20: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9b: RET-Mutationsstatus hereditär) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 86

Tabelle 4-21: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9a: RET-Mutationsstatus unbekannt) (Alle Angaben in %) (Gesamtpopulation) ... 86

Tabelle 4-22: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9b: RET-Mutationsstatus unbekannt) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 86

Tabelle 4-23: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10a: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %) (Gesamtpopulation) ... 87

Tabelle 4-24: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10b: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 87

Tabelle 4-25: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 11: WHO-PS bei Studieneintritt) (Alle Angaben in %) ... 87

Tabelle 4-26: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel ... 89

Tabelle 4-27: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel ... 92

Tabelle 4-28: Matrix der explorativen Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel ... 92

Tabelle 4-29: Interaktionstests für die Populationen „progressiv und symptomatisch“ und „nicht progressiv und symptomatisch“ ... 93

Tabelle 4-30: Operationalisierung von Progressionsfreies Überleben (PFS) ... 95

Tabelle 4-31: Bewertung des Verzerrungspotenzials für Progressionsfreies Überleben (PFS) in RCT mit dem zu bewertenden Arzneimittel ... 95

Tabelle 4-32: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 97

Tabelle 4-33: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 97

Tabelle 4-34: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 98

Tabelle 4-35: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 98

Tabelle 4-36: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode) (Gesamtpopulation) 99

Tabelle 4-37: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 99

Tabelle 4-38: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel: Progressionsstatus zum Zeitpunkt des Daten *Cut-off* (Gesamtpopulation) 99

Tabelle 4-39: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel: Progressionsstatus zum Zeitpunkt des Daten *Cut-off* (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 100

Tabelle 4-40: Operationalisierung von TWP 105

Tabelle 4-41: Bewertung des Verzerrungspotenzials für Verzögerung der TWP in RCT mit dem zu bewertenden Arzneimittel 105

Tabelle 4-42: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 107

Tabelle 4-43: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 108

Tabelle 4-44: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 108

Tabelle 4-45: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 109

Tabelle 4-46: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 109

Tabelle 4-47: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 109

Tabelle 4-48: Operationalisierung der objektiven Ansprechrate (ORR) 113

Tabelle 4-49: Bewertung des Verzerrungspotenzials für die objektive Ansprechrate (ORR) in RCT mit dem zu bewertenden Arzneimittel 113

Tabelle 4-50: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 115

Tabelle 4-51: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 115

Tabelle 4-52: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 116

Tabelle 4-53: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem

zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 116

Tabelle 4-54: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 117

Tabelle 4-55: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 117

Tabelle 4-56: Operationalisierung von Krankheitskontrollrate (DCR) ... 119

Tabelle 4-57: Bewertung des Verzerrungspotenzials für Krankheitskontrollrate (DCR) in RCT mit dem zu bewertenden Arzneimittel ... 120

Tabelle 4-58: Ergebnisse für die Krankheitskontrollrate (DCR) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 121

Tabelle 4-59: Ergebnisse für die Krankheitskontrollrate (DCR) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 122

Tabelle 4-60: Ergebnisse (explorativ) für die Krankheitskontrollrate (DCR) **ohne** offene Behandlungsphase aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 122

Tabelle 4-61: Ergebnisse (explorativ) für die Krankheitskontrollrate (DCR) **ohne** offene Behandlungsphase aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation) ... 122

Tabelle 4-62: Operationalisierung von Dauer des objektiven Ansprechens (DOR) ... 124

Tabelle 4-63: Bewertung des Verzerrungspotenzials für Dauer des objektiven Ansprechens (DOR) in RCT mit dem zu bewertenden Arzneimittel ... 124

Tabelle 4-64: Ergebnisse für die geschätzte mediane Dauer des objektiven Ansprechens (DOR) aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der Gesamtpopulation) ... 126

Tabelle 4-65: Ergebnisse für die geschätzte mediane Dauer des objektiven Ansprechens (DOR) aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 127

Tabelle 4-66: Ergebnisse für die mediane Zeit von der Randomisierung bis zum Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der Gesamtpopulation) ... 127

Tabelle 4-67: Ergebnisse für die mediane Zeit von der Randomisierung bis zum Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der genäherten Zielpopulation) ... 128

Tabelle 4-68: Operationalisierung von Gewichtszunahme ... 130

Tabelle 4-69: Bewertung des Verzerrungspotenzials für Gewichtszunahme in RCT mit dem zu bewertenden Arzneimittel ... 130

Tabelle 4-70: Ergebnisse für die Rate der Gewichtszunahme aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 131

Tabelle 4-71: Ergebnisse für die Gewichtszunahme aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 131

Tabelle 4-72: Operationalisierung von Einfluss der Behandlung auf Diarrhöen ... 133

Tabelle 4-73: Bewertung des Verzerrungspotenzials für Einfluss der Behandlung auf Diarrhöen in RCT mit dem zu bewertenden Arzneimittel ... 133

Tabelle 4-74: Ergebnisse für die Stuhlfrequenz aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 135

Tabelle 4-75: Ergebnisse für die Stuhlfrequenz aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation) ... 136

Tabelle 4-76: Veränderung für die Stuhlfrequenz zum Ausgangswert - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 137

Tabelle 4-77: Veränderung der Stuhlfrequenz gegenüber dem Ausgangswert - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation) ... 139

Tabelle 4-78: Operationalisierung des WHO-Performancestatus (WHO-PS) ... 142

Tabelle 4-79: Bewertung des Verzerrungspotenzials für den WHO-Performancestatus (WHO-PS) in RCT mit dem zu bewertenden Arzneimittel ... 142

Tabelle 4-80: Ergebnisse für die Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 143

Tabelle 4-81: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 143

Tabelle 4-82: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 144

Tabelle 4-83: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 145

Tabelle 4-84: Operationalisierung des TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline) ... 147

Tabelle 4-85: Bewertung des Verzerrungspotenzials für die Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline) in RCT mit dem zu bewertenden Arzneimittel ... 148

Tabelle 4-86: Ergebnisse für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 149

Tabelle 4-87: Ergebnisse für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 149

Tabelle 4-88: Operationalisierung der Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline) ... 153

Tabelle 4-89: Bewertung des Verzerrungspotenzials für Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (*Baseline*) in RCT mit dem zu bewertenden Arzneimittel 154

Tabelle 4-90: Ergebnisse für die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 155

Tabelle 4-91: Ergebnisse für die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation) 155

Tabelle 4-92: Ergebnisse für die mediane Dauer der Reduktion des Opioid-Gebrauchs - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 155

Tabelle 4-93: Ergebnisse für die mediane Dauer der Reduktion des Opioid-Gebrauchs - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation) 156

Tabelle 4-94: Operationalisierung des biochemischen Ansprechens (CTN; CEA) 157

Tabelle 4-95: Bewertung des Verzerrungspotenzials für biochemisches Ansprechen (CTN; CEA) in RCT mit dem zu bewertenden Arzneimittel 158

Tabelle 4-96: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 159

Tabelle 4-97: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 159

Tabelle 4-98: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 160

Tabelle 4-99: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 160

Tabelle 4-100: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 161

Tabelle 4-101: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 161

Tabelle 4-102: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 161

Tabelle 4-103: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) 162

Tabelle 4-104: Operationalisierung von Gesamtüberleben (OS) 164

Tabelle 4-105: Bewertung des Verzerrungspotenzials für Gesamtüberleben (OS) in RCT mit dem zu bewertenden Arzneimittel 164

Tabelle 4-106: Ergebnisse für das Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 166

Tabelle 4-107: Ergebnisse für das Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 167

Tabelle 4-108: Operationalisierung des Endpunktes Gesundheitsbezogene Lebensqualität . 169

Tabelle 4-109: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel ... 170

Tabelle 4-110: Ergebnisse für die gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 172

Tabelle 4-111: Ergebnisse für die gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 173

Tabelle 4-112: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) gegenüber dem Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 174

Tabelle 4-113: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) gegenüber dem Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 175

Tabelle 4-114: Operationalisierung von Unerwünschte Ereignisse (UE) ... 178

Tabelle 4-115: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel ... 178

Tabelle 4-116: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation) ... 180

Tabelle 4-117: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 180

Tabelle 4-118: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation) ... 181

Tabelle 4-119: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) ... 181

Tabelle 4-120: Ergebnisse für den Anteil an Patienten mit unerwünschten Ereignissen (nach MedDRA) mit einer Häufigkeit von >10% aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation und der genäherten Zielpopulation) ... 182

Tabelle 4-121: Ergebnisse für das Auftreten von QTc-bedingten Ereignissen und UE der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation) ... 183

Tabelle 4-122: Ergebnisse für das Auftreten von QTc-bedingten Ereignissen und UEs der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA) aus RCT

mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) 184

Tabelle 4-123: Individuelle Interaktionstests (*post-hoc*) – PFS - Gesamtpopulation 189

Tabelle 4-124: Individuelle Interaktionstests (*post-hoc*) – PFS – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 190

Tabelle 4-125: Individuelle Interaktionstests (*post-hoc*) – TWP- Gesamtpopulation 191

Tabelle 4-126: Individuelle Interaktionstests (*post-hoc*) – TWP – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 191

Tabelle 4-127: PFS Subgruppenanalysen nach *Baseline*-Charakteristika - Gesamtpopulation 192

Tabelle 4-128: PFS Subgruppenanalysen nach *Baseline*-Charakteristika - genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 193

Tabelle 4-129: PFS Subgruppenanalysen nach Biomarkern - Gesamtpopulation 194

Tabelle 4-130: PFS Subgruppenanalysen nach Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 195

Tabelle 4-131: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung) - Gesamtpopulation 196

Tabelle 4-132: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung) – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 197

Tabelle 4-133: PFS Subgruppenanalysen nach Plasma Biomarkern - Gesamtpopulation..... 198

Tabelle 4-134: PFS Subgruppenanalysen nach Plasma Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 199

Tabelle 4-135: PFS Subgruppenanalyse nach Altersgruppe – Gesamtpopulation..... 200

Tabelle 4-136: PFS Subgruppenanalyse nach Altersgruppe – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 200

Tabelle 4-137: TWP Subgruppenanalysen nach *Baseline*-Charakteristika - Gesamtpopulation 204

Tabelle 4-138: TWP Subgruppenanalysen nach *Baseline*-Charakteristika - genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 205

Tabelle 4-139: TWP Subgruppenanalysen nach Biomarkern - Gesamtpopulation 206

Tabelle 4-140: TWP Subgruppenanalyse nach Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 207

Tabelle 4-141: TWP Subgruppenanalyse nach Krankheitsstatus - Gesamtpopulation..... 208

Tabelle 4-142: TWP Subgruppenanalyse nach Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 209

Tabelle 4-143: TWP Subgruppenanalysen nach Plasma Biomarkern - Gesamtpopulation... 210

Tabelle 4-144: TWP Subgruppenanalysen nach Plasma Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 211

Tabelle 4-145: TWP Subgruppenanalyse nach Altersgruppe - Gesamtpopulation..... 212

Tabelle 4-146: TWP Subgruppenanalyse nach Altersgruppe – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 212
Tabelle 4-147: PFS Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – Gesamtpopulation (ITT Population) 217
Tabelle 4-148: PFS Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung (ITT Population) 218
Tabelle 4-149: ORR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – Gesamtpopulation (ITT Population) 219
Tabelle 4-150: ORR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung (ITT Population) 220
Tabelle 4-151: DOR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - Gesamtpopulation 221
Tabelle 4-152: DOR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung 221
Tabelle 4-153: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche 229
Tabelle 4-154: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 230
Tabelle 4-155: Operationalisierung von <Endpunkt xxx> 230
Tabelle 4-156: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche 231
Tabelle 4-157: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche 231
Tabelle 4-158: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien 233
Tabelle 4-159: Operationalisierung von <Endpunkt xxx> 233
Tabelle 4-160: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien 234
Tabelle 4-161: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen 235
Tabelle 4-162: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – weitere Untersuchungen 236
Tabelle 4-163: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen mit dem zu bewertenden Arzneimittel 238
Tabelle 4-164: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen 239
Tabelle 4-165: Charakterisierung der Intervention – weitere Untersuchungen 240
Tabelle 4-166: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 1) 240
Tabelle 4-167: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 2) 240

Tabelle 4-168: Verzerrungspotenzial auf Studienebene – Weitere Untersuchungen mit dem zu bewertenden Arzneimittel ... 241
Tabelle 4-169: Operationalisierung der objektiven Ansprechrate ... 242
Tabelle 4-170: Ergebnisse für die objektive Ansprechrate (ORR) – Weitere Untersuchungen ... 243
Tabelle 4-171: Operationalisierung des progressionsfreien Überleben ... 243
Tabelle 4-172: Ergebnisse für das Progressionsfreies Überleben (PFS) – Weitere Untersuchungen ... 244
Tabelle 4-173: Operationalisierung des WHO *Performance Status* ... 244
Tabelle 4-174: Ergebnisse des WHO *Performance Status* (WHO PS) – Weitere Untersuchungen ... 244
Tabelle 4-175: Operationalisierung der gesundheitsbezogenen Lebensqualität ... 245
Tabelle 4-176: Ergebnisse zur Gesundheitsbezogenen Lebensqualität – Weitere Untersuchungen ... 245
Tabelle 4-177: Operationalisierung der unerwünschten Ereignisse ... 246
Tabelle 4-178: Unerwünschte Ereignisse – Weitere Untersuchungen ... 246
Tabelle 4-179: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens ... 266
Tabelle 4-180: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens ... 268
Tabelle 4-181: Für die Nutzenbewertung berücksichtigte Studien und Untersuchungen ... 272
Tabelle 4-182 (Anhang): Im *International Clinical Trials Registry Platform Search Portal* identifizierte Studien ... 289
Tabelle 4-183 (Anhang): In clinicaltrials.gov identifizierte Studien ... 291
Tabelle 4-184 (Anhang): im Volltext gesichtete und ausgeschlossene Publikationen mit Ausschlussgrund ... 294
Tabelle 4-185 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel ... 295
Tabelle 4-186 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel ... 296
Tabelle 4-187 (Anhang): Studiendesign und -methodik für Studie D4200C00058 ... 298
Tabelle 4-188 (Anhang): Studiendesign und -methodik für Studie D4200C00008 ... 327
Tabelle 4-189 (Anhang): Studiendesign und -methodik für Studie D4200C00068 ... 334
Tabelle 4-190 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00058 ... 341
Tabelle 4-191 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00008 ... 390
Tabelle 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00068 ... 411

Abbildungsverzeichnis

Seite

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel 75

Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit Vandetanib 77

Abbildung 3: Forest Plot für Progressionsfreies Überleben (PFS) für Gesamtpopulation, genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung und Restpopulation..... 98

Abbildung 4: Kaplan-Meier-Kurve für Progressionsfreies Überleben (Gesamtpopulation → ITT-Population)..... 101

Abbildung 5: Kaplan-Meier-Kurve für Progressionsfreies Überleben (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)..... 101

Abbildung 6: Grafische Darstellung des Weibull-Modells, angepasst an PFS mittels Kaplan-Meier-Kurve (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)..... 102

Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 104

Abbildung 8: Forest Plot für TWP nach Symptomen und Progressionsstatus bei *Baseline* (Gesamtpopulation)..... 110

Abbildung 9: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population innerhalb der Gesamtpopulation)..... 110

Abbildung 10: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)..... 111

Abbildung 11: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 112

Abbildung 12: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 118

Abbildung 13: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 123

Abbildung 14: Kaplan-Meier-Kurve für die Dauer des objektiven Ansprechens (DOR) (Gesamtpopulation → ITT-Population), Zeit vom ersten Ansprechen bis zur Progression .. 126

Abbildung 15: Kaplan-Meier-Kurve für Dauer des objektiven Ansprechens (DOR) (genäherte Zielpopulation → ITT-Population) 127

Abbildung 16: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 128

Abbildung 17: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> 132

Abbildung 18: Stuhlfrequenz über die Zeit – Veränderung der Stuhlfrequenz zum Ausgangswert (Gesamtpopulation) ... 138

Abbildung 19: Stuhlfrequenz über die Zeit – Veränderung der Stuhlfrequenz gegenüber dem Ausgangswert (genäherte Zielpopulation) ... 140

Abbildung 20: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 141

Abbildung 21: Kaplan-Meier-Kurve für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) (Gesamtpopulation → ITT-Population) ... 144

Abbildung 22: Kaplan-Meier-Kurve für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population) ... 145

Abbildung 23: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 146

Abbildung 24: *Forest-Plot* für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn – Gesamtpopulation ... 150

Abbildung 25: *Forest*-Plot für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ... 150

Abbildung 26: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 151

Abbildung 27: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 156

Abbildung 28: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 163

Abbildung 29: Kaplan-Meier-Kurve für das Gesamtüberleben (Gesamtpopulation → ITT-Population) ... 166

Abbildung 30: Kaplan-Meier-Kurve für das Gesamtüberleben (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population) ... 167

Abbildung 31: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 168

Abbildung 32: FACT-G über die Zeit – Veränderung der gesundheitsbezogenen Lebensqualität gegenüber dem Ausgangswert (Gesamtpopulation) ... 176

Abbildung 33: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 177

Abbildung 34: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> ... 186

Abbildung 35: *Forest-Plot* der Subgruppenanalyse für PFS – *Baseline*- Charakteristiken und Krankheitsstatus – Gesamtpopulation ... 200

Abbildung 36: *Forest-Plot* der Subgruppenanalyse für PFS – *Baseline*- Charakteristiken und Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ... 201

Abbildung 37: *Forest-Plot* der Subgruppenanalyse für PFS – Biomarker - Gesamtpopulation 201
Abbildung 38: *Forest-Plot* der Subgruppenanalyse für PFS – Biomarker – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 202
Abbildung 39: *Forest-Plot* der Subgruppenanalyse für PFS – Altersgruppen - Gesamtpopulation 202
Abbildung 40: *Forest-Plot* der Subgruppenanalyse für PFS – Altersgruppen – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 203
Abbildung 41: *Forest-Plot* der Subgruppenanalyse für TWP – *Baseline*- Charakteristiken und Krankheitsstatus – Gesamtpopulation 213
Abbildung 42: *Forest-Plot* der Subgruppenanalyse für TWP – *Baseline*- Charakteristiken und Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 213
Abbildung 43: *Forest-Plot* der Subgruppenanalyse für TWP – Biomarker - Gesamtpopulation 214
Abbildung 44: *Forest-Plot* der Subgruppenanalyse für TWP – Biomarker – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 214
Abbildung 45: *Forest-Plot* der Subgruppenanalyse für TWP – Altersgruppen – Gesamtpopulation 215
Abbildung 46: *Forest-Plot* der Subgruppenanalyse für TWP – Altersgruppen –genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung 215
Abbildung 47: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen 237
Abbildung 48: Flussdiagramm Patientenverteilung Studie D4200C00058 324
Abbildung 49: Flussdiagramm Studiendesign Studie D4200C00058 325
Abbildung 50: Flussdiagramm Teilpopulation mit symptomatischer und progressiver Erkrankung in Studie D4200C00058 326
Abbildung 51: Flussdiagramm Studiendesign Studie D4200C00008 333
Abbildung 52: Flussdiagramm Studiendesign Studie D4200C00068 340

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALP	Alkaline Phosphatase
ALT	Alaninaminotransferase
AM-NutzenV	Arzneimittelnutzenverordnung
ARMS	Amplification Refractory Mutation System
AST	Aspartataminotransferase
ATA	American Thyroid Association
bFGF	Basis- Fibroblastenwachstumsfaktor
BOR	Bestes objektives Ansprechen - nach RECIST (Best Objective Response)
BPI-SF	Patientenberichtete Schmerzerfassung (Brief Pain Inventory – Short Form)
BSC	Best Supportive Care
Ca	Calcium
CEA	Carcinoembryonales Antigen
CHMP	Committee for Medicinal Products for Human Use
CONSORT	Consolidated Standards of Reporting Trials
Cox-PH-Modell	Cox Proportional Hazards Model
CR	Komplettes Ansprechen (Complete Response)
CSP	Klinisches Studienprotokoll (Clinical Study Protocol)
CT	Computertomographie
CTCAE	Common Terminology Criteria for Adverse Events
CTN	Calcitonin
CYP3A4	Cytochrom P450, Isoform 3A4
DCR	Krankheitskontrollrate (Disease Control Rate)
DGHO	Deutsche Gesellschaft für Hämatologie und Onkologie
DIMDI	Deutsches Institut für Medizinische Dokumentation
DNA	Desoxyribonukleinsäure (Desoxyribonucleic acid)
DOR	Dauer des Ansprechens (Duration of Response)
eCRF	electronic Case Report Form
EG	Europäische Gemeinschaft
EGF	Epidermaler Wachstumsfaktor (Epidermal Growth Factor)

Abkürzung	Bedeutung
EGFR	Rezeptor für den epidermalen Wachstumsfaktor (Epidermal Growth Factor Receptor)
ECHO	Echokardiogramm
EKG	Elektrokardiogramm
EMA	European Medicines Agency
ESMO	European Society for Medical Oncology
EWB	Emotionales Wohlbefinden
FACIT-D	Functional Assessment of Chronic Illness Therapy – Diarrhea
FACT-G	Functional Assessment of Cancer Therapy General Scale
FDA	U.S. Food and Drug Administration
FMTC	Familiäres Medulläres Schilddrüsenkarzinom
FWB	Funktionelles Wohlbefinden
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
HIV	Humanes Immundefizienz-Virus
HR	Hazard Ratio
HRQoL	Gesundheitsbezogene Lebensqualität (Health Related Quality of Life)
ICTRP	International Clinical Trials Registry Platform
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitwesen
ITT	Intention to Treat
IVRS	Interactive Voice Response Services
KI	Konfidenzintervall
LVEF	linksventrikuläre Ejektionsfraktion
M918T	Aminosäureaustausch Methionin→Threonin an Position 918
MedDRA	Medical Dictionary for Regulatory Activities
MEN	Multiple endokrine Neoplasie
MRT	Magnetresonanztomographie
MTC*	Medulläres Schilddrüsenkarzinom (Medullary Thyroid Cancer)
mtc*	Mixed Treatment Comparison
MUGA	Multi-Gated Acquisition Scan
NC	nicht berechenbar (not calculable)

Abkürzung	Bedeutung
NCCN	U.S. National Comprehensive Cancer Network
NYHA	New York Heart Association
OR	Odds Ratio (Quotenverhältnis)
ORR	Objektive Ansprechrate (Objective Response Rate)
OS	Gesamtüberleben (Overall Survival)
PCR	Polymerase-Kettenreaktion (Polymerase Chain Reaction)
PD	Progressive Disease
PFS	Progressionsfreies Überleben (Progression-Free Survival)
PK	Pharmakokinetik
PP	Per Protokoll
PR	Partielles Ansprechen (Partial Response)
PRES	Posteriores reversibles Enzephalopathie-Syndrom
PRO	Patientenberichtete Zielgröße (Patient Reported Outcome)
PVC	Vorzeitige ventrikuläre Kontraktion (Premature Ventricular Contraction)
PWB	Physisches Wohlbefinden
QTc	Zeitintervall zwischen dem Start der Q-Welle und dem Ende der T-Welle (korrigiert gegen die Herzrate)
RCT	Randomisierte kontrollierte Studie (Randomized Controlled Trial)
RECIST	Response Evaluation Criteria for Solid Tumors
RET	Rearranged During Transfection
RMP	Risk Management-Plan
RR	Relatives Risiko
SAE	Schwerwiegendes unerwünschtes Ereignis (Serious Adverse Event)
SD	Standardabweichung (Standard Deviation)
SGB	Sozialgesetzbuch
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
SWB	Soziales Wohlbefinden
TDPS	Zeit bis zur Verschlechterung des WHO Performance Status (Time to Decline in Patient WHO PS)

Abkürzung	**Bedeutung**
TKI	Tyrosinkinase-Inhibitor
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TWP	Zeit bis zur Schmerzprogression (Time to Worsening of Pain)
UE	Unerwünschte Ereignisse
ULRR	Obergrenze des Referenzbereichs (Upper Limit of Reference Range)
VEGF	Vaskulärer endothelialer Wachstumsfaktor (Vascular Endothelial Growth Factor)
VEGFR	Rezeptor für den vaskulären endothelialen Wachstumsfaktor (Vascular Endothelial Growth Factor Receptor)
WHO	Weltgesundheitsorganisation (World Health Organization)
WHO PS	WHO Leistungsstatus (WHO Performance Status)

* Die Dokumentenvorlage des G-BA für Modul 4 sieht die Bezeichnung MTC als Abkürzung für Mixed Treatment Comparison vor. Im vorliegenden Dossier ist MTC jedoch generell über alle Module hinweg für das Anwendungsgebiet Medulläres Schilddrüsenkarzinom belegt. Aus Gründen der Einheitlichkeit wird für Mixed Treatment Comparison daher hier ersatzweise die Abkürzung mtc verwendet.

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.

Für Arzneimittel, die zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt. Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, müssen jedoch vorgelegt werden. Zu diesem Zweck enthält die Dokumentvorlage den Abschnitt 4.4.4, der ausschließlich in Dossiers für die oben genannten Arzneimittel zur Behandlung eines seltenen Leidens bearbeitet werden soll. In diesem Abschnitt sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich der begründenden Nachweise vorzulegen. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

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

Vandetanib erhielt am 17.02.2012 durch die EMA die europaweite Zulassung zur Behandlung von aggressivem und symptomatischem medullärem Schilddrüsenkarzinom (MTC) bei Patienten mit nicht resektabler, lokal fortgeschrittener oder metastasierter Erkrankung. Das MTC ist eine seltene Ausprägung innerhalb der Gesamtheit der Schilddrüsenkarzinome, die ihrerseits eine gegenüber anderen soliden Tumorarten vergleichsweise seltene Erkrankung darstellen.

Ziel der vorliegenden Bewertung ist, das Ausmaß des Zusatznutzens von Vandetanib bei der Behandlung der Zielpopulation von Patienten mit aggressivem und symptomatischem MTC in Relation zur zweckmäßigen Vergleichstherapie zu quantifizieren.

Im Rahmen der Beratung durch den Gemeinsamen Bundesausschuss (G-BA) vom 29. September 2011 wurde *Best Supportive Care* (BSC) als zweckmäßige Vergleichstherapie festgelegt ([1], vgl. Modul 3A, Abschnitt 3.1). Im strengen Sinne der Verfahrensordnung des Gemeinsamen Bundesausschusses und nach den Standards der evidenzbasierten Medizin existiert keine zweckmäßige Vergleichstherapie im Indikationsgebiet. Unter dem Oberbegriff BSC wird die bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität als die einzige für die Patienten im durch die Indikation beschriebenen Krankheitsstadium zum Zeitpunkt der Durchführung der Zulassungsstudie verfügbare Behandlungsmöglichkeit verstanden. Es ist dabei darauf hinzuweisen, dass die Behandlung mit dem zu bewertenden Arzneimittel Vandetanib den Einsatz von BSC nicht ersetzt, sondern zusätzlich erfolgt.

Die Zulassung von Vandetanib basiert auf der pivotalen Phase III Studie D4200C00058 („Studie 58“), die erwachsene Patienten mit nicht resektablem, lokal fortgeschrittenem oder metastasiertem MTC einschloss, und wurde unterstützt durch Daten aus zwei Phase II Studien. Im Laufe des Zulassungsverfahrens wurden *post-hoc* diverse Subpopulationen ausgewertet und betrachtet (Patienten mit symptomatischer, mit progressiver, mit progressiver und / oder symptomatischer Erkrankung etc.). Analysen spezifisch für die Zulassungspopulation, wie in der Indikation final formuliert (Patienten mit „aggressiver und symptomatischer“ Erkrankung) wurden nicht generiert. Die Festlegung der Indikation auf Patienten mit aggressiver Erkrankung erfolgte erst spät im Verfahren. Die EMA legte sich hiermit auf eine klinische Bewertung der Patientensituation durch den behandelnden Arzt fest, eine Population von Patienten mit aggressiver Erkrankung läßt sich aus den Studiendaten nicht sauber abgrenzen. Entsprechend wurde für die Bewertung der Wirksamkeit die gesamte Studienpopulation der Studie 58 herangezogen, die Einschränkung der Zielpopulation erfolgte aufgrund einer Nutzen-Risiko-Abwägung der EMA bei einer onkologischen Erkrankung mit verhältnismäßig langem Verlauf [2].

Im Rahmen des Verfahrens der frühen Nutzenbewertung wurde am 14.03.2012 basierend auf den Daten für die Gesamtpopulation der Studie bereits ein Nutzendossier für Caprelsa®

eingereicht. Dieses wurde vom G-BA mit dem Auftrag zur Bewertung an das IQWiG übergeben. Das IQWiG beurteilte das Dossier als inhaltlich unvollständig. Die inhaltliche Unvollständigkeit ist lt. IQWiG „maßgeblich dadurch bedingt, dass keine Daten zu patientenrelevanten Endpunkten innerhalb der Zulassungspopulation vorliegen, welche eine Bewertung des Zusatznutzens von Vandetanib ermöglicht hätte“ [3]. Das IQWiG akzeptierte, dass eine Subgruppenanalyse für Patienten mit aggressivem und symptomatischem MTC mangels klarer Definition der Kriterien für die Zulassungspopulation innerhalb der Population der Zulassungsstudie zu Vandetanib nicht vorgenommen werden kann. Gleichzeitig schlug das IQWiG in seiner Nutzenbewertung vor, dass AstraZeneca die Teilpopulation der Patienten mit symptomatischem und **progressivem** MTC als adäquate Näherung an die Zulassungspopulation (Patienten mit **aggressivem** und symptomatischem MTC) hätte auswerten können [3]. Für AstraZeneca war es im Vorfeld nicht voraussehbar, dass das IQWiG eine Vergleichbarkeit von ‚aggressiv und symptomatisch‘ mit ‚progressiv und symptomatisch‘ sehen würde. Entsprechend waren dieser Standpunkt und die Nicht-Bewertung durch das IQWiG für AstraZeneca unerwartet. Die vorgeschlagene Vorgehensweise wurde von AstraZeneca jedoch gern aufgenommen. Nach Veröffentlichung der IQWiG-Bewertung wurden die Endpunktanalysen für die Näherungspopulation im Rahmen des Stellungnahmeverfahrens (Fristende der Einreichung war der 06.07.2012) umgehend von AstraZeneca zur Verfügung gestellt. Nach mündlicher Anhörung am 24.07.2012 erfolgte die Beschlussfassung des G-BA am 06.09.2012 (Veröffentlichung im Bundesanzeiger am 21.09.2012) [4]. Die im Rahmen der Stellungnahme von AstraZeneca nachgereichten Daten wurden vom G-BA aus Verfahrensgründen nicht bewertet. Der G-BA kam zu dem Schluss, dass der Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie nicht belegt sei, da die erforderlichen Nachweise nicht vollständig vorgelegt worden seien und folgte damit in seinem Beschluss der Bewertung des IQWiG.

AstraZeneca beantragte daraufhin auf Grundlage von § 35a Abs. 5b SGB V (neu) [5] am 14.11.2012 die Neueinreichung eines Dossiers für Vandetanib. Diesem Antrag wurde vom G-BA am 06.12.2012 stattgegeben.

In der vorliegenden Neufassung des Dossiers sind neben den Ergebnissen für die gesamte Studienpopulation alle Ergebnisse auch für Teilpopulation mit symptomatischer und progressiver Erkrankung als Näherung für die Zielpopulation „aggressiv und symptomatisch“ dargestellt (im Folgenden daher als „genäherte Zielpopulation“ bezeichnet). Die im Studienprotokoll der für die Nutzenbewertung herangezogenen RCT geplanten Analysen waren allerdings nicht für diese genäherte Zielpopulation geplant; durch die Beschränkung der Analysen auf die genäherte Zielpopulation sinkt die statistische *Power*. Durch Interaktionstests konnte aber gezeigt werden, dass die in der gesamten Studienpopulation beobachteten Ergebnisse auf die nach Auffassung des IQWiG bewertungsrelevante genäherte Zielpopulation „progressiv und symptomatisch“ übertragbar sind.

Die Quantifizierung des Zusatznutzens wird daher in der vorliegenden Neufassung des Dossiers 1) gemäß statistischem Analyseplan der RTC für die Gesamtpopulation der relevanten Phase III-Studie ebenso wie, 2) für die entsprechend dem Vorschlag des IQWiG

als Subgruppe aus der relevanten Phase III-Studie ausgewertete genäherte Zielpopulation der Patienten mit symptomatischer und progressiver Erkrankung anhand der folgenden patientenrelevanten Endpunkte vorgenommen:

- **Morbidität**
  - Progressionsfreies Überleben (*Progression-Free Survival*, PFS)
  - Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)
  - Objektive Ansprechrate (*Objective Response Rate*, ORR)
  - Krankheitskontrollrate (*Disease Control Rate*, DCR)
  - Dauer des objektiven Ansprechens (*Duration of Response*, DOR)
  - Explorative Morbiditätsendpunkte:
    - Gewichtszunahme
    - Einfluss der Behandlung auf Diarrhöen
    - WHO-Performancestatus (WHO-PS)
    - Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (*Baseline*)
    - Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (*Baseline*)
- **Mortalität**
  - Gesamtüberleben (*Overall Survival*, OS)
- **Gesundheitsbezogene Lebensqualität**
  - Gesundheitsbezogene Lebensqualität (*Health Related Quality of Life*, HRQoL, anhand des FACT-G Fragebogens)
- **Nebenwirkungen**
  - Unerwünschte Ereignisse (UE)

Eine detaillierte Beschreibung der Endpunkte sowie deren Validität und Patientenrelevanz wird in Abschnitt 4.2.5.2 gegeben. Zusätzlich wurden als Surrogate innerhalb der Kategorie

Morbidität als Unterstützung des klinischen Endpunktes Progression die Parameter CTN und CEA für das biochemische Ansprechen untersucht.

Datenquellen

MTC ist mit einer europaweiten Prävalenz von 7/100.000 [6] eine Erkrankung, die die Kriterien der Verordnung (EG) Nr.141/2000 des EU-Parlaments für seltene Leiden [7] in vollem Umfang erfüllt, da es sich bei MTC um eine schwere und seltene Erkrankung ohne vorbestehende zugelassene Therapie handelt, so dass eine erstmalige Therapieoption für diese Patienten von erheblichem Nutzen sein wird. Der *Orphan Drug* Status wurde für Vandetanib am 24.1.2006 von der EMA anerkannt. Zusätzlich wurde 2008 für Vandetanib ein pädiatrisches Prüfkonzept verabschiedet, das die Durchführung einer klinischen Studie vorsah [8]. Um die Verlängerung des ergänzenden Schutzzertifikates in Anspruch nehmen zu können, entschied sich AstraZeneca nach interner Prüfung für das pädiatrische Prüfkonzept. Der *Orphan Drug* Status wurde deshalb auf Antrag von AstraZeneca von der EMA zurückgezogen [9]. Aus diesem Grund besitzt Vandetanib gemäß § 12 VerfO des G-BA [10] nicht den Status eines Arzneimittels für seltene Leiden (*Orphan Drug*), wenngleich Vandetanib praktisch alle dafür notwendigen Vorraussetzungen erfüllt.

Die Bewertung soll auch bei einer äußerst seltenen Indikation wie dem MTC, soweit möglich, auf der Basis randomisierter kontrollierter Studien (RCT) erfolgen, da diese, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet sind. Grundlage für die Quantifizierung des Zusatznutzens von Vandetanib gegenüber der zweckmäßigen Vergleichstherapie BSC im vorliegenden Dossier sind daher im Wesentlichen die Ergebnisse zur Gesamtpopulation und der genäherten Zielpopulation der Phase-III-RCT D4200C00058 („Studie 58“), die von AstraZeneca im Rahmen der Vorbereitung der amerikanischen und europäischen Zulassung an 331 (Gesamtpopulation) bzw. 186 (genäherte Zielpopulation) Patienten durchgeführt worden ist. Studie 58 ist weltweit die erste große Phase-III-Studie, die im Indikationsgebiet MTC durchgeführt wurde. Daneben werden aufgrund der Seltenheit der Erkrankung nichtkontrollierte Phase-II-Studien als unterstützende Evidenz präsentiert, um die Grundlage der Zulassungsentscheidung abzubilden und die Vollständigkeit der zu Vandetanib vorhandenen Evidenz in dieser Indikation zu dokumentieren.

Um sicherzustellen, dass alle verfügbaren Studiendaten zu Vandetanib beim MTC für die Bewertung berücksichtigt werden, sind darüber hinaus eine systematische Literaturrecherche sowie eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien mit dem zu bewertenden Arzneimittel Vandetanib in den gängigen medizinischen Studienregistern durchgeführt worden (Abschnitt 4.2.3.2 und 4.2.3.3).

Ein-/Ausschlusskriterien für Studien

Die Ein- und Ausschlusskriterien für Studien dienen dem Zweck, diejenigen Studien zu identifizieren, die grundsätzlich für die Beantwortung der Fragestellung gemäß 4.2.1 geeignet sind. Die in Abschnitt 4.2.2 in Tabelle 4-1 spezifizierten Selektionskriterien grenzen die

einzuschließenden Studien in Bezug auf folgende Kriterien ein: Population, Intervention, Komparator, Endpunkte (*Outcome*) und Studiendesign.

Wie in ausführlicher Form am Anfang des Abschnitts begründet, wird in der vorliegenden Neufassung des Dossiers zusätzlich zur Gesamtpopulation auch die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung betrachtet und anhand dieser auch gesondert eine Quantifizierung des Zusatznutzens von Vandetanib vorgenommen. Die erforderlichen Daten zur genäherten Zielpopulation wurden hierzu aus den Studienaufzeichnungen extrahiert [11] und *post-hoc* statistisch aufbereitet.

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

Die zur Quantifizierung des Zusatznutzens eingeschlossene Studie wurde hinsichtlich ihrer Aussagekraft bewertet. Diese Bewertung erfolgte im Rahmen des vorliegenden Dossiers anhand der zur Verfügung stehenden Informationen aus dem Studienbericht der Studie. Zusätzlich hat AstraZeneca im Rahmen der Erstellung des vorliegenden Dossiers weitere *post-hoc*-Analysen für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung durchgeführt, die über die Analysen hinausgehen, die der EMA und der FDA im Rahmen des Zulassungsprozesses vorlagen. Ergänzend hierzu wurden Interaktionstests für die Studienendpunkte durchgeführt, um die Vergleichbarkeit der Ergebnisse der Gesamtpopulation und der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung und daraus folgend die Übertragbarkeit auf die Zulassungspopulation zu belegen.

<u>Bewertung des Verzerrungspotenzials</u>

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte wurden untersucht. Bei Verzerrungen auf Endpunktebene wurde jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Entsprechend der Verfahrensordnung des G-BA zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V vom 20. Januar 2011 wurden dabei folgende Aspekte bewertet:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnisgesteuerte Berichterstattung
- Sonstige das Verzerrungspotential beeinflussende endpunktübergreifende Punkte

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des *Intention-To-Treat* (ITT)-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte: Behandlung fehlender Werte und *Drop-outs*

Die Bewertungen des Verzerrungspotentials der Einzelstudien wurden tabellarisch zusammengefasst. Gemäß Verfahrensordnung wurde das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, 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.

Abschließend wurde unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem „niedrigen“ Verzerrungspotential wurde ausgegangen, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotentials für jede Studie und jeden Endpunkt richtete sich nach den Hinweisen der Bewertungsbögen in Anhang 4-G.

Eine zusammenfassende Bewertung der Verzerrungsaspekte erfolgte ausschließlich für die RCT.

Im Rahmen der Erstellung der vorliegenden Neufassung des Moduls 4 A erfolgte unter Berücksichtigung der IQWiG-Kommentare *post-hoc* eine Reihe von Auswertungen aus der RCT zu den erhobenen Endpunkten innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung als näherungsweises Abbild der Indikation „aggressiv und symptomatisch“. Dabei wurden auch die Verzerrungsaspekte, sowohl auf Studienebene, als auch auf Endpunktebene vor dem Hintergrund dieser Neuauswertung noch einmal beurteilt. Sowohl die *post-hoc*-Analysen, als auch die Analysemethodik an sich beruhen auf den prä-spezifizierten ITT-Daten bzw. der im Studienbericht dargestellten präspezifizierten Analysemethodik. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichterstattung, noch sonstige das Verzerrungspotenzial beeinflussende Punkte identifiziert werden, die auf Studien- oder Endpunktebene ein Abweichen von der im Rahmen der Evaluierung der Gesamtpopulation erfolgten Einschätzung notwendig gemacht hätten.

Die für die Nutzenbewertung von Vandetanib herangezogene randomisierte kontrollierte Studie wird in Abschnitt 4.3.1.2 sowie in Anhang 4-F anhand von Design-Charakteristika

beschrieben. Die Beschreibung erfolgte gemäß den Punkten 2b bis 14 des CONSORT-Statements. Der Patientenfluss wurde gemäß CONSORT-*Flow-Chart* dargestellt.

Über die Informationen aus der RCT hinausgehend wurden unterstützende Informationen aus zwei nicht-randomisierten Studien herangezogen. Diese Studien sind gemäß den in Abschnitt 4.2.2 angegeben Kriterien nicht geeignet, die Fragestellung gemäß 4.2.1 zu beantworten. Sie wurden dennoch in Abschnitt 4.3.2.3 sowie ebenfalls in Anhang 4-F dargestellt, um die Evidenz zu Vandetanib im MTC zu vervollständigen. Auch die in diesen Studien erhobenen Endpunkte wurden der Vollständigkeit halber in Anhang 4-G hinsichtlich ihres Verzerrungspotentials bewertet. Ihre Resultate wurden allerdings **nicht** für die Bewertung des medizinischen Nutzens und des Zusatznutzens von Vandetanib berücksichtigt. Aus diesem Grund werden in den Bewertungsbögen zu Studie D4200C00008 (Tabelle 4-191) und Studie D4200C00068 (Tabelle 4-192) patientenrelevante Endpunkte im Sinne von Tabelle 4-1 nur aufgeführt, insoweit sie in diesen Studien auch erhoben wurden.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die Patienten der durch die Zulassung für Vandetanib festgelegten Zulassungspopulation mit aggressivem und symptomatischem MTC und nicht resektabler, lokal fortgeschrittener oder metastasierter Erkrankung, leiden an einer unheilbaren, sich fortschreitend verschlechternden, bösartigen Erkrankung und versterben fast immer am MTC. Diesem stand bislang keine effektive Therapie gegenüber, die bestehenden Behandlungsmöglichkeiten dienten lediglich einer Linderung der Symptome der Erkrankung.

Der langfristige Erhalt einer akzeptablen Lebensqualität steht für die betroffenen Patienten an erster Stelle. Eine neue Therapie bedeutet folglich dann für die Patienten einen erheblichen Zusatznutzen, wenn sie in der Lage ist, das Fortschreiten der Erkrankung und damit ihrer Begleitsymptome relevant hinauszuzögern, ein Tumoransprechen zu erreichen, und sie sollte darüber hinaus auch bei Einsatz über einen langen Zeitraum hinweg gut verträglich sein.

Vandetanib ist ein Multi-Tyrosinkinaseinhibitor und wirkt selektiv als hochpotenter Antagonist des vaskulären endothelialen Wachstumsfaktor-Rezeptors 2 (VEGFR-2) und, in geringerer Wirkstärke, des VEGFR-1 und -3, des Rezeptors für den epidermalen Wachstumsfaktor (EGFR) und insbesondere auch der RET (*Rearranged During Transfection*)-Tyrosinkinase. Vandetanib hemmt sowohl den Wildtyp als auch die meisten aktiven Mutationen von RET [12, 13], des Protoonkogens, das bei der Entstehung aller erblichen Varianten von MTC und bis zu 80% aller Fälle von sporadischem MTC ursächlich beteiligt ist. *In vitro* unterbindet Vandetanib signifikant die Proliferation von MTC-Zelllinien. Vandetanib hemmt die VEGF-stimulierte endotheliale Zellmigration, Zellproliferation, das Überleben von Zellen und die Bildung neuer Blutgefäße in *in-vitro*-Modellen der Angiogenese. Darüber hinaus inhibiert Vandetanib die EGF-stimulierte EGF Rezeptor-Tyrosinkinase in Tumor- und Endothelzellen. Vandetanib unterbindet *in vitro* die EGFR-abhängige Zellproliferation und verhindert so das Überleben der Zelle. Eine ausführliche

Darstellung des Wirkmechanismus von Vandetanib findet sich in Modul 2 des vorliegenden Dossiers.

Im Rahmen der randomisierten, kontrollierten Studie 58 konnte für die patientenrelevanten Endpunkte „Progressionsfreies Überleben“ (*Progression-Free Survival*, PFS), „Zeit bis zur Schmerzprogression“ (*Time to Worsening of Pain*, TWP), „Objektive Ansprechrate“ (*Objective Response Rate*, ORR) und Krankheitskontrollrate (*Disease Controle Rate*, DCR) sowohl innerhalb der Gesamtpopulation der Studie als auch innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ein statistisch signifikanter Vorteil von Vandetanib gegenüber der mit Placebo behandelten Kontrollgruppe gezeigt werden, wobei beide Behandlungsgruppen zusätzlich „*Best Supportive Care*“ (BSC) erhielten. Zusätzlich wurden als Surrogate für den klinischen Endpunkt Progression die Parameter CTN und CEA für das biochemische Ansprechen untersucht. Vorteile für Vandetanib gegenüber Placebo zeichneten sich auch in einigen weiteren, explorativ erhobenen patientenrelevanten Morbiditätsendpunkten ab, die in direktem Bezug zu den Symptomen der Erkrankung stehen (Gewicht, weitere Schmerzendpunkte, Performancestatus). Da zum Zeitpunkt der Studie keine zugelassene Therapie für lokal fortgeschrittenes oder metastasiertes MTC existierte, wurde allen Patienten in beiden Behandlungsarmen nach Progression eine *Open-Label*-Behandlung mit Vandetanib ermöglicht (*Cross-Over* Design). Aufgrund der hohen Patientenrelevanz der Morbiditätsendpunkte in der Palliativsituation und der eingeschränkten Aussagekraft der Mortalitätsdaten wegen der erlaubten *Open-Label*-Behandlung werden die Morbiditätsendpunkte im vorliegenden Nutzendossier vorrangig diskutiert und folgerichtig maßgeblich für die Bestimmung des Ausmaßes des Zusatznutzens herangezogen (siehe auch die ausführliche Diskussion zum Thema „Patientenrelevanz“ in Abschnitt 4.2.5.2).

Morbidität

Verbesserung des Progressionsfreien Überlebens (PFS)

In der palliativen onkologischen Therapie stellt ein Hinauszögern des Fortschreitens der Erkrankung und der damit verbundenen Symptomatik eines der wichtigsten Therapieziele dar [14]; dies gilt insbesondere auch bei einem zu erwartenden verhältnismäßig langen Gesamtüberleben [15-17]. Deshalb zählt das progressionsfreie Überleben (*Progression-Free Survival*, PFS) in der palliativen Situation als Funktion der Wachstumsrate des Tumors und der Wirksamkeit des untersuchten Präparates zu den wichtigsten patientenrelevanten Endpunkten [18], um die Effektivität einer eingesetzten Behandlung während des beobachteten Therapiezeitraums zu bewerten. Die Progression der Erkrankung geht mit einer zunehmenden Morbidität und häufig stark zunehmenden Schmerzen des Patienten einher [19].

Für PFS, den primären Endpunkt der Studie 58, wurde in der Vandetanib-Gruppe bezogen auf die Gesamtpopulation ein statistisch signifikanter Vorteil gegenüber Placebo beobachtet (HR: 0,46; 95%-KI: [0,31;0,69]). Die entsprechende *post-hoc* Auswertung der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ergab ebenfalls einen

statistisch signifikanten Vorteil von Vandetanib gegenüber Placebo (HR: 0,47; 95%-KI: [0,29;0,77]).

Das bedeutet für die Patienten eine Verzögerung ihrer Progression und damit fortschreitender Krankheitslast um 11 Monate in der Gesamt- bzw. um 12 Monate in der genäherten Zielpopulation. Der durchgeführte Interaktionstest bestätigt die Übertragbarkeit der Ergebnisse der Gesamtpopulation auf die genäherte Zielpopulation klar (p=0,9617) (Tabelle 4-29).

Die Überlegenheit von Vandetanib zeigt sich auch im zeitlichen Verlauf. Die Wahrscheinlichkeit, keine Progression zu erleiden (Kaplan-Meier Schätzer) war in der Vandetanib-Gruppe in der Gesamtpopulation zu allen Zeitpunkten um ca. 20% höher: im Monat 6 mit 91,4% versus 74,5%, Monat 12 mit 83,3% versus 62,7% und im Monat 24 mit 58,8% versus 41,7%. In der genäherten Zielpopulation zeigte sich dies genauso: im Monat 6 mit 89,4% versus 72,9%, Monat 12 mit 80,8% versus 57,1% und im Monat 24 waren noch 51,7% der Patienten progressionsfrei versus 36,0% in der Placebo-Gruppe.

Aus der erheblichen Verlängerung des progressionsfreien Überlebens und der besonderen Relevanz dieses Endpunktes im Hinblick auf die Morbidität und Lebensqualität des Patienten lässt sich für diesen patientenrelevanten Endpunkt gemäß der Klassifizierung des Zusatznutzens nach §5 Punkt 7 Am-NutzenV [20] der Beleg eines erheblichen Zusatznutzens ableiten (siehe Diskussion zu Ausmaß und Nachweissicherheit des Zusatznutzens in Abschnitt 4.4.2).

Zeit bis zur Schmerzprogression (TWP)

Schmerz stellt eine bedeutsame Komponente der Symptomatik des aggressiven, nicht resektablen, lokal fortgeschrittenen oder metastasierten MTC dar. Dies bestätigt aus klinischer Sicht die DGHO (Deutsche Gesellschaft für Hämatologie und Onkologie), nach deren Aussage ein Großteil der Patienten extreme Schmerzen hat [21]. Für die Patienten ist eine möglichst langfristige Verzögerung der Schmerzprogression von großer Bedeutung, da Schmerzen nicht nur fortschreitende Morbidität signalisieren, sondern auch die Lebensqualität stark beeinträchtigen und der Einsatz von Schmerzmitteln wiederum von entsprechenden Nebenwirkungen begleitet ist. Die Vermeidung bzw. das Hinauszögern von Tumorschmerzen muss daher als herausragendes Therapieziel bei der Behandlung dieser Patienten in der Palliativsituation gelten.

Die Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP), abgeleitet aus dem höchsten Schmerzwert der Schmerzfragebögen in Verbindung mit dem von den Patienten angegebenen Gebrauch von opioidhaltigen Schmerzmitteln, war unter Vandetanib im Median mehr als doppelt so lang wie in der Vergleichsgruppe, sowohl für die Gesamtpopulation (7,85 vs. 3,25 Monate), als auch in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung (11,1 vs. 3,4 Monate). Der Unterschied zwischen den Behandlungsgruppen war sowohl innerhalb der Gesamtpopulation (HR: 0,61; 95%-KI:

[0,43;0,87]) als auch innerhalb der genäherten Zielpopulation (HR: 0,62 [0,39;0,99]) statistisch signifikant.

Der Endpunkt TWP zeigt damit eine nachhaltige und mit den bisherigen Behandlungsmethoden nicht erreichte große Verbesserung des therapiebedingten Nutzens. Die für die Palliativsituation als langfristig zu bewertende Freiheit von progredienten Schmerzen stellt somit einen erheblichen Zusatznutzen für die betroffenen Patienten gemäß §5 Punkt 7 AM-NutzenV [20] dar (siehe Diskussion zu Ausmaß und Nachweissicherheit des Zusatznutzens in Abschnitt 4.4.2).

Objektive Ansprechrate (ORR)

Bei der Behandlung onkologischer Erkrankungen in der metastasierten Situation ist die Objektive Ansprechrate (*Objective Response Rate*, ORR) sowohl ein objektivierbares Maß zur Kontrolle des Ansprechens der Therapie (gemessen nach RECIST, modifiziert für MTC [22, 23]) als auch ein für die Patienten bedeutsames Therapieziel. Die ORR misst dabei die Verringerung der Tumorlast und umfasst sowohl komplettes als auch partielles Ansprechen, allerdings nicht den Stillstand der Erkrankung [24-26]. Sie ist damit ein wichtiger Parameter für reduzierte Morbidität.

Anhand der ORR konnte eine Verringerung der Tumorlast durch Vandetanib gezeigt werden. In der Gesamtpopulation erreichten 45% der Patienten in der Vandetanib-Gruppe eine partielle Remission (gemäß RECIST Kriterien) im Vergleich zu 13% der Patienten unter BSC (wobei für 12 von 13 Patienten in der Placebo-Gruppe ein Ansprechen erst nach dem Wechsel auf Vandetanib nach Ende der verblindeten Behandlung verzeichnet wurde). In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung erreichten 44,4% der Patienten in der Vandetanib-Gruppe eine partielle Remission im Vergleich zu 16,7% der Patienten unter BSC (wobei für 9 von 10 Patienten in der Placebo-Gruppe ein Ansprechen erst nach dem Wechsel auf Vandetanib nach Ende der verblindeten Behandlung erreicht wurde). In beiden Behandlungsgruppen und Populationen erreichte keiner der Patienten eine komplette Remission.

Das Wissen um ein Tumoransprechen und die Metastasenreduktion bringen dem Patienten eine psychische Entlastung in einer für ihn extrem kritischen Lebenssituation. Für den Patienten stellt sich dies als therapierelevanter Nutzen dar. Unter Berücksichtigung dessen und angesichts des statistisch signifikanten Ergebnisses und der Effektgröße für das objektive Ansprechen (OR: 5,48; 95%-KI: [2,99;10,79] bzw. 4,00 [1,93;9,01] für die genäherte Zielpopulation) sowie der Aussagekraft des erbrachten Nachweises ergibt sich für diesen Endpunkt der Beleg eines erheblichen Zusatznutzens (siehe Diskussion zu Ausmaß und Nachweissicherheit des Zusatznutzens in Abschnitt 4.4.2).

Das Resultat der Sensitivitätsanalyse ohne Berücksichtigung der Daten aus der *Open-Label* Phase unterstreicht das obige signifikante Ergebnis und zeigt einen noch deutlicheren Effekt für Vandetanib. Das objektive Ansprechen erlebten 101/231 (55/126 in der genäherten

Zielpopulation) Patienten unter Vandetanib und nur 1/100 (1/60) Patienten unter Placebo (OR [95%-KI]: 76,91 [16,68;1366] bzw. 45,70 [9,58;819,7]).

Krankheitskontrollrate (DCR)

Die Krankheitskontrollrate (*Disease Control Rate*, DCR) ist ein Maß für die Krankheitsstabilisierung und berücksichtigt über die Patienten mit partieller oder kompletter Remission hinaus auch diejenigen Patienten, die durch die Therapie über einen längeren Zeitraum eine Stabilisierung ihrer Erkrankung erleben dürfen. Damit stellt die Krankheitskontrollrate ein weiteres Maß für den klinischen Nutzen einer Therapie dar, indem sie all jene Patienten einbezieht, bei denen über einen relevanten Zeitraum kein Progress vorliegt und die damit von einer Krankheitskontrolle durch die Therapie profitieren.

Die Krankheitskontrollrate beinhaltet den Anteil der Patienten, deren bestes objektives Ansprechen (B*est Objective Response*, BOR) einem kompletten Ansprechen (*Complete Response* CR), einem partiellen Ansprechen (*Partial Response*, PR) oder einer Krankheitsstabilität für mindestens 24 Wochen gemäß RECIST Kriterien entspricht. Die Anzahl der Patienten mit Krankheitskontrolle war in beiden evaluierten Populationen (Gesamt- bzw. genäherte Zielpopulation) in der Vandetanib-Gruppe im Vergleich zu Placebo höher (86,6% vs. 71,0% für die Gesamtpopulation bzw. 84,1% vs. 68,3% für die genäherte Zielpopulation). Bezüglich der Krankheitskontrollrate wurde in der Vandetanib-Gruppe sowohl für die Gesamtpopulation als auch für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ein statistisch signifikantes Ergebnis gegenüber Placebo beobachtet (OR: 2,64; 95%-KI: [1,48;4,69] bzw. OR: 2,46; 95%-KI: [1,19;5,09]). Unter Ausschluss der offenen Behandlungsphase (12 von 13 Patienten der Gesamtpopulation bzw. 9 von 10 Patienten der genäherten Zielpopulation in der Placebo-Gruppe hatten erst in der offenen Behandlungsphase ein beginnendes Ansprechen) betrug der Anteil der Patienten mit Krankheitskontrolle in der Vandetanib-Gruppe 84,8% bzw. 82,5% gegenüber 61% bzw. 56,7% in der Placebo-Gruppe. Es konnte eine statistisch signifikante Verbesserung der Krankheitskontrollrate für die Vandetanib-Gruppe gezeigt werden (OR: 3,58; 95%-KI: [2,09;6,17] bzw. OR: 3,61; 95%-KI: [1,83;7,26]).

Für diesen Endpunkt ergibt sich somit als Ausmaß des Zusatznutzens der Beleg eines erheblichen Zusatznutzens (siehe Diskussion zu Ausmaß und Nachweissicherheit des Zusatznutzens in Abschnitt 4.4.2).

Dauer des objektiven Ansprechens (DOR)

Die Dauer des objektiven Ansprechens (*Duration of Response*, DOR) misst den Zeitraum, für den eine erreichte Tumorantwort (komplette oder partielle Remission) bestehen bleibt. Für die Betroffenen spielt es selbstverständlich eine erhebliche Rolle, ob ein einmal erreichtes Ansprechen ihrer Erkrankung auch möglichst lange aufrechterhalten werden kann. Die DOR ist damit ein weiteres Maß für das Aufhalten des Progresses und damit der Verzögerung zunehmenden Leidens. Das Ziel jeder onkologischen Therapie in der Palliation muss sein, dass der Patient möglichst lange von der Therapie profitiert.

Die Dauer des objektiven Ansprechens (DOR) wurde basierend auf dem zentralen, unabhängigen *Review* der Bildgebung nach den RECIST Kriterien bewertet. Die Abbildungen der Kaplan-Meier-Kurven beider Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigten für die DOR für die Zeit vom Ansprechen bis zur ersten Progression in beiden Behandlungsarmen ein lange anhaltendes Ansprechen, wobei ein Ansprechen für 12 bzw. 9 von 13 bzw. 10 Patienten der Placebo-Gruppe erst in der offenen Behandlungsphase unter Vandetanib einsetzte. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens lagen unter Vandetanib in der Gesamtpopulation bei 22,2 Monaten und unter Placebo bei 16,3 Monaten. Für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ließ sich der gleiche Trend beobachten. Hier lagen die geschätzten medianen Werte für die Dauer des objektiven Ansprechens in der Vandetanib-Gruppe bei 21,8 Monaten und bei 13,0 Monaten in der Placebo-Gruppe. Die Dauer des objektiven Ansprechens ist für Patienten, die auf Vandetanib ansprachen, im Trend damit länger als bei Patienten unter Placebo/BSC. Die Ergebnisse für die Dauer des objektiven Ansprechens (DOR) untermauern die Belege für einen erheblichen Zusatznutzen der bereits dargestellten Endpunkte.

Weitere explorative patientenrelevante Morbiditätsendpunkte:

Die Ergebnisse der nachfolgend aufgeführten explorativen Morbiditätsendpunkte unterstützen die Ergebnisse der bereits dargestellten Morbiditätsendpunkte, welche übereinstimmend eine Überlegenheit von Vandetanib gegenüber *Best Supportive Care* zeigen und den sich daraus resultierenden erheblichen Zusatznutzen belegen:

Gewichtszunahme

Für die klinische Gewichtszunahme zeigte sich in beiden Populationen (Gesamt- bzw. genäherte Zielpopulation) ein deutliches Ergebnis zugunsten von Vandetanib (OR: 3,24; 95%-KI: [1,69;6,76] bzw. OR: 4,03; 95%-KI: [1,71;11,17]). Der Anteil der Patienten mit klinischer Gewichtszunahme betrug in der Vandetanib-Gruppe 28,6% bzw. 31,0% verglichen zur Placebo-Gruppe mit 11,0% bzw. 10,0%.

Einfluss der Behandlung auf Diarrhö

Der Einfluss der Behandlung auf Diarrhö wurde anhand der Stuhlfrequenz bestimmt. Die Angaben der Patienten bezüglich ihrer Stuhlfrequenzen im Hinblick auf den Einfluss der Behandlung auf Diarrhö zeigten in der Gesamtpopulation und auch der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung zwischen den untersuchten Behandlungsgruppen Vandetanib und Placebo keine auffälligen Unterschiede. Ein Anhaltspunkt für einen Einfluss der Behandlung mit Vandetanib auf Diarrhö ergab sich aus der exploratorischen Analyse in beiden betrachteten Populationen nicht.

WHO-Perfomancestatus (WHO-PS)

Die Daten für die Zeit bis zum ersten Nachweis einer Verschlechterung des WHO-PS wurden relativ zum Zeitpunkt der Randomisierung ausgewertet. Eine Verschlechterung des WHO-PS wurde definiert als Erhöhung um mindestens einen Punkt im Vergleich zum Ausgangswert. Für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigte sich statistisch kein deutlicher Unterschied in den Ergebnissen (HR: 0,74; 95%-KI: [0,48;1,14] bzw. HR: 0,71; 95%-KI: [0,42; 1,21]). Für die Gesamtpopulation konnte die mediane Zeit bis zur Verschlechterung des WHO-PS wegen der geringen Anzahl an Ereignissen in der Vandetanib-Gruppe zum Daten *Cut-off* nicht errechnet werden. Für die Placebo-Gruppe betrug die mediane Zeit bis zur Verschlechterung des WHO-PS 24,9 Monate. In der genäherten Zielpopulation lag die mediane Zeit bis zur Verschlechterung des WHO-PS für die Vandetanib-Gruppe bei 22,6 Monaten, verglichen mit 15,4 Monaten in der Placebo-Gruppe, was einer Verlängerung um ca. 7 Monate entspricht.

Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline)

Der Endpunkt „Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn“ stellt in seiner Form eine Subgruppenanalyse zum TWP dar. Sowohl für die Gesamt-, als auch für die genäherte Zielpopulation spiegelt die visuelle Prüfung des *Forest-Plots* die Konsistenz des TWP über die zwei Behandlungsgruppen mit bekanntem Opioidanalgetika-Gebrauch wider. Für die Gesamtpopulation liegt in der Gruppe „≥10 mg/Tag Opioidanalgetika-Gebrauch“ das *Hazard Ratio* bei 0,36 (95%-KI: [0,16;0,84], p=0,0176). In der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika-Gebrauch“, liegt das *Hazard Ratio* bei 0,68 mit einem 95%-KI [0,44;1,03] und einem p-Wert von 0,0698. In der genäherten Zielpopulation ließen sich sowohl für die Gruppe „≥10 mg/Tag Opioidanalgetika- Gebrauch“ mit einem *Hazard Ratio* von 0,50 (95%-KI: [0,20;1,30], p=0,1573), als auch in der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika-Gebrauch“ mit einem *Hazard Ratio* von 0,69 (95%-KI [0,39;1,22]; p=0,1978) keine statistisch signifikanten Unterschiede zeigen. Die Ergebnisse sind aufgrund der teils geringen Gruppengrößen, insbesondere in der genäherten Zielpopulation, mit Vorsicht zu interpretieren.

Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

Die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn stellt in seiner Form eine Subgruppenanalyse zum TWP dar. Für die Reduktion des Opioid-Gebrauchs wurden nur Patienten mit einem Schmerzmittelgebrauch von ≥10 mg/Tag Morphinsulfat-Äquivalenten berücksichtigt. Sowohl in der Gesamt-, als auch in der genäherten Zielpopulation zeigte sich ein Trend in der Reduktion des Schmerzmittelgebrauchs zugunsten von Vandetanib (Gesamtpopulation OR: 2,78; 95%-KI: [0,77;10,80] bzw. genäherte Zielpopulation OR: 2,25; 95%-KI: [0,50;11,27]). Die Reduktion der Häufigkeit des Schmerzmittelgebrauchs betrug in der Vandetanib-Gruppe 62,5% bzw. 52,9% verglichen mit 37,5% bzw. 33,3% in der Placebo-Gruppe. Die mediane Dauer der

Reduktion des Schmerzmittelgebrauchs betrug für die Vandetanib-Gruppe 11,01 bzw. 11,56 Monate und für die Placebo-Gruppe in beiden Populationen 6,51 Monate. Alle Ergebnisse sind aufgrund der geringen Anzahl an Ereignissen mit Vorsicht zu interpretieren.

Surrogatendpunkte für Morbidität

Biochemisches Ansprechen (CTN; CEA)

Das biochemische Ansprechen, gemessen an den Serumspiegeln von CTN und CEA, stellt gerade im Falle des MTC gut geeignete Marker für den klinischen Endpunkt Progression zur Verfügung und dient auf Grund seiner Signalwirkung für ein Ansprechen auf die Therapie wie auch für eine Progredienz der Erkrankung als Surrogat für patientenrelevante Endpunkte des Bereichs Morbidität (krankheitsbedingte Symptome). Durchfall und Flush-Syndrom treten häufig infolge humoraler Faktoren auf (krankheitsbedingt hohe CTN-Werte). Das biochemische Ansprechen der CTN- und CEA-Werte zeigt sich in beiden betrachteten Populationen. Für CTN ergibt sich eine Spiegelreduzierung für 69,3%/67,5% (Gesamtpopulation/genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) der mit Vandetanib behandelten Patienten im Vergleich zu 3,0%/3,3% bei Behandlung gemäß BSC (OR für CTN-Ansprechen: 72,86; 95%-KI: [26,22;303,2]/OR für CTN-Nichtansprechen: 0,01; 95%-KI: [0,00;0,04] in der Gesamtpopulation bzw. OR für CTN-Ansprechen 60,12; 95%-KI: [17,56;378,0]/OR für CTN-Nichtansprechen: 0,02; 95%-KI: [0,00;0,06] in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung).

Für CEA ergibt sich eine Spiegelreduzierung für 51,5%/46,0% der Patienten in der Vandetanib-Gruppe im Vergleich zu 2,0%/0,0% in der Placebo-Gruppe (OR für CEA-Ansprechen: 52,03; 95%-KI: [15,95;320,3]/OR für CEA-Nichtansprechen: 0,02; 95%-KI: [0,00;0,08] in der Gesamtpopulation; für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung wurde aufgrund des fehlenden Ansprechens in der Placebo-Gruppe kein OR berechnet).

Mortalität

Gesamtüberleben (OS)

Aufgrund der geringen Ereignisraten zum Zeitpunkt des *Cut-off* für die primäre Analyse und darüber hinaus des möglichen Einflusses des Wechsels auf *Open-Label* Vandetanib nach Beendigung der verblindeten Behandlung, wodurch eine Verzerrung der Ergebnisse wahrscheinlich ist, wurde die statistische Auswertung des Gesamtüberlebens nicht für die Nutzenbewertung herangezogen. Das Gesamtüberleben wurde dennoch in Abschnitt 4.3.1.3.1.12 dargestellt, um die Evidenz zu Vandetanib zu vervollständigen. Eine erste Analyse der Gesamtüberlebensrate zeigte bei 48 eingetretenen Ereignissen ein *Hazard Ratio*

von 0,89 zugunsten von Vandetanib. Der Unterschied war nicht statistisch signifikant (KI: [0,28;2,85], p=0,713). Für die genäherte Zielpopulation zeigte eine erste Analyse der Gesamtüberlebensrate bei 31 eingetretenen Ereignissen ein *Hazard Ratio* von 1,06 (KI: [0,26; 4,38], p=0,8791).

Gesundheitsbezogene Lebensqualität

Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G

Daten zur gesundheitsbezogenen Lebensqualität wurden im Rahmen der Zulassungsstudie rein exploratorisch ausgewertet. Die Angaben der Patienten bezüglich ihrer Lebensqualität im Hinblick auf ihre Gesundheit in den FACT-G Fragebögen zeigten jedoch in der Gesamtpopulation und auch der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung zwischen den untersuchten Behandlungsgruppen Vandetanib und BSC keine auffälligen Unterschiede. Es ist anzunehmen, dass das objektive Ansprechen auf die Behandlung in Verbindung mit einer Schmerzverminderung einigen der Patienten in der palliativen Situation ein gewisses Maß an Zuversicht und Lebensmut gibt. Ein Anhaltspunkt für eine Verminderung der gesundheitsbezogenen Lebensqualität durch die Therapie mit Vandetanib ergab sich aus der exploratorischen Auswertung in beiden betrachteten Populationen nicht. Ein Zusatznutzen bezüglich des Endpunkts „gesundheitsbezogene Lebensqualität“ ist für Vandetanib damit für diesen Endpunkt nicht zu belegen.

Sicherheit

Unerwünschte Ereignisse

Bei nahezu allen Patienten beider Gruppen und beider Populationen (Gesamtpopulation bzw. genäherte Zielpopulation) trat mindestens ein unerwünschtes Ereignis auf. Der Anteil der Patienten mit unerwünschten Ereignissen in der Vandetanib- bzw. Placebo-Gruppe betrug in der Gesamtpopulation 99,6% bzw. 90,9%, in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung 100% bzw. 94,9%. Für von schwerwiegenden unerwünschten Ereignissen, Abbrüchen aufgrund unerwünschter Ereignisse und von unerwünschten Ereignissen vom Schweregrad CTCAE Grad 3 oder höher betroffene Patienten ergaben sich in der Vandetanib-Gruppe signifikant höhere Prozentsätze als für die Placebo-Gruppe (CTCAE Grad 3 oder höher 55,4% vs. 24,2% für die Gesamtpopulation bzw. 61,1% vs. 23,7% für die genäherte Zielpopulation; schwere unerwünschte Ereignisse (SUE) 30,7% vs. 13,1% für die Gesamtpopulation bzw. 31,7% vs. 16,9% für die genäherte Zielpopulation). Der Anteil der Patienten mit unerwünschten Ereignissen, die zum Absetzen der Studienbehandlung führten, war unter Vandetanib im Vergleich zu Placebo größer (Gesamtpopulation 12,1% vs. 3,0% – OR: 4,41; 95%-KI: [1,31;14,8], genäherte Zielpopulation 11,9% vs. 1,7% – OR: 7,84; 95%-KI: [1,01;60,82]). Die Anzahl der Ereignisse muss allerdings immer in Relation zur Expositionsdauer gesehen werden: Die mediane

Behandlungsdauer auf verblindeter Therapie war in der Vandetanib-Gruppe erheblich länger als in der Placebo-Gruppe (90,1 vs. 39,9 Wochen in der Gesamtpopulation bzw. 88,6 vs. 37,1 Wochen in der genäherten Zielpopulation), wodurch sich für die Vandetanib-Gruppe eine sehr viel längere Expositionsdauer ergibt.

Die mit Vandetanib am häufigsten beobachteten Nebenwirkungen Diarrhö, Hautausschlag, Übelkeit, Bluthochdruck und Kopfschmerzen ließen sich durch medizinische Interventionen gut beherrschen. Das potenzielle Risiko einer QTc-Verlängerung sowie von Torsade de Pointes ist sowohl durch Warnhinweise in der Fachinformation, als auch durch Maßnahmen im Risk-Management-Plan für Vandetanib angemessen berücksichtigt worden (siehe Modul 3 A, Abschnitt 3.4.3). Die *post-hoc* Auswertungen für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung bestätigen alle Erkenntnisse der Gesamtpopulation.

In der Zusammenfassung muss für die genannten Sicherheitsendpunkte somit ein größerer Schaden (siehe Diskussion zu Ausmaß und Nachweissicherheit des Zusatznutzens/-schadens in Abschnitt 4.4.2) als belegt gelten, dessen Ausmaß jedoch im Kontext der Gesamtraten unerwünschter Ereignisse in beiden Behandlungsgruppen, der guten Beherrschbarkeit der meisten unerwünschten Ereignisse, der langen Therapiedauer unter Vandetanib sowie unter Berücksichtigung des Schweregrades der Erkrankung betrachtet werden muss.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Vandetanib stellt eine medizinische Innovation dar. Es ist die erste wirksame medikamentöse Behandlungsmöglichkeit für Patienten mit aggressivem und symptomatischem, nicht resektablem, fortgeschrittenem oder metastasiertem MTC.

In der palliativen Situation weist die Therapie mit Vandetanib einen bedeutsamen Zusatznutzen für die betroffenen Patienten auf. Ihre Morbidität wird durch das um mehr als 11 Monate, in der genäherten Zielpopulation um etwa ein Jahr, verlängerte progressionsfreie Überleben, eine Verringerung der Tumorlast und eine Verbesserung der Schmerzsituation im Vergleich mit BSC relevant reduziert. Diese Resultate waren in der Gesamtpopulation der Studie wie auch in der als Näherung für die Zielpopulation der Studie betrachteten Population der Studienteilnehmer mit progressiver und symptomatischer Erkrankung konsistent. Dem gegenüber steht ein Beleg für einen im Vergleich zu Placebo größeren, nicht quantifizierbaren Schaden in Bezug auf SUE und Abbrüche aufgrund unerwünschter Ereignisse. Die im Rahmen der RCT unter Vandetanib beobachteten unerwünschten Ereignisse erwiesen sich in der Regel jedoch als transient und durch eine geeignete medikamentöse Behandlung gut kontrollierbar. Das Risiko von QTc-Verlängerungen und Torsade de Pointes wurde durch entsprechende Warnhinweise in der Fachinformation und Maßnahmen im Risk-Management-Plan adressiert.

Angesichts der Schwere der Erkrankung und des erheblichen *medical need* in der Zielpopulation, für die derzeit keine Behandlungsalternative in Deutschland existiert, ist der Zusatznutzen von Vandetanib auch nach Abwägung der Vorteile (erheblicher Zusatznutzen in fast jedem der betrachteten Endpunkte) gegenüber dem Schaden in der Gesamtschau zum gegenwärtigen Zeitpunkt als **erheblich** einzustufen.

In der Vandetanib-Gruppe trat zwar eine größere als unter BSC gemessene Häufigkeit unerwünschter Ereignisse auf, dennoch erscheint es nicht angemessen, in der Gesamtschau den aufgezeigten positiven Nutzen deshalb abzuwerten. Dafür sprechen im Wesentlichen drei Argumente:

1. Aus klinisch-therapeutischer Sicht dürfte unter allen direkt gemessenen Endpunkten die mindestens 11-monatige, in der genährten Zielpopulation ca. 1-jährige, Verlängerung des progressionsfreien Überlebens für diese unheilbar kranken Patienten von vorrangigem Wert sein.

2. Die Zeit bis zur Schmerzprogression (TWP) konnte in beiden Populationen maßgeblich verlängert werden, in der genäherten Zielpopulation in der Vandetanib-Gruppe sogar um fast 8 Monate im Vergleich zur Placebo-Gruppe. Die Vermeidung/Hinauszögerung von Tumorschmerzen muss als herausragendes Therapieziel bei der Behandlung von Krebspatienten in der Palliativsituation gelten.

3. Die unerwünschten Ereignisse waren unter Vandetanib in der Regel hinreichend gut beherrschbar und nur 12% der Patienten sowohl der Studien- als auch der genäherten Zielpopulation brachen die Behandlung wegen unerwünschter Ereignisse ab. Zudem waren auch die im Kontrollarm eingesetzten Maßnahmen der BSC bei über 90 % der Patienten von Nebenwirkungen begleitet.

EMA und FDA haben die Effektivität von Vandetanib im Rahmen der europäischen und der US-amerikanischen Zulassung auf der Basis der gesamten Studienpopulation bewertet. Die für die EMA *post-hoc* für die Teilpopulationen „progressiv“, „symptomatisch“, „progressiv oder symptomatisch“ und „progressiv und symptomatisch“ für den primären Endpunkt PFS, sowie den sekundären Endpunkt ORR durchgeführten Subgruppenanalysen bestätigten die Konsistenz der Daten über die betrachteten Teilpopulationen. Die Beschränkung der Indikation auf die Patienten mit aggressiver und symptomatischer Erkrankung durch die EMA erfolgte aufgrund einer Nutzen-Risiko-Abwägung, nicht aufgrund unterschiedlicher Wirksamkeit [2]. Der Begriff „aggressiv“ wurde dabei bewusst als klinische Festlegung gewählt und ist innerhalb der Studienpopulation nicht klar abgrenzbar. Als genäherte Zielpopulation wurde daher wie vom IQWiG vorgeschlagen die Gruppe der Patienten mit progressiver und symptomatischer Erkrankung ausgewertet. Durch Interaktionstests konnte gezeigt werden, dass sich die Ergebnisse für die Gesamtpopulation der Zulassungsstudie auf die genäherte Zielpopulation der Pateinten mit symptomatischer und progressiver Erkrankung übertragen lassen (siehe Tabelle 4-29).

Unter Berücksichtigung obiger Erwägungen und in Anbetracht der zusätzlichen Tatsache, dass es sich bei Vandetanib um die erste und bisher einzige zugelassene zielgerichtete Behandlung für die betreffende MTC-Teilpopulation mit aggressiver und symptomatischer Erkrankung handelt, beansprucht AstraZeneca für dieses Arzneimittel für die gesamte Zielpopulation einen **erheblichen** Zusatznutzen im Vergleich zu BSC.

AstraZeneca stellt an dieser Stelle zur Diskussion, ob die Vorgaben der AM-NutzenV bzw. diejenigen der G-BA VerfO zur Quantifizierung des Zusatznutzen-Ausmaßes eine Einstufung des Zusatznutzens als ‚erheblich' für neue Arzneimittel, welche – wie Vandetanib – als erste zur gezielten Behandlung einer nicht mehr heilbaren, fortschreitenden Erkrankung zugelassen werden, grundsätzlich überhaupt erlauben oder ob hier nicht ggf. die Gefahr einer systematischen Benachteiligung solcher Arzneimittel vorliegen kann:

- Erstens ist bei Patienten im Palliativstadium einer Krebserkrankung keine Heilung mehr zu erwarten; stattdessen besteht das Ziel darin, die Lebensqualität und die Selbstbestimmung dieser schwerkranken Menschen zu erhalten, zu fördern und zu verbessern. Im besten Fall wird sich, wie im Falle Vandetanib, der fortschreitende Verlauf der Erkrankung für eine begrenzte zusätzliche Zeitspanne aufhalten lassen.

- Zweitens dürfte für ein solches Arzneimittel auch kaum jemals eine nachhaltige große Verbesserung nachweisbar sein, wenn der G-BA – wie aus der zusammenfassenden Dokumentation zum Ticagrelor-Beschluss vom 15.12.2011 [27] hervorzugehen scheint – mit dem Kriterium der Nachhaltigkeit grundsätzlich den Anspruch auf langfristige symptomfreie Zeiträume von mehreren Jahren verbindet. Die Patienten aus der Zielpopulation von Vandetanib haben nach derzeitigem Kenntnisstand nur noch eine mediane verbleibende Lebenserwartung von durchschnittlich 2 bis 3 Jahren; vor diesem Hintergrund müssen eine PFS-Verlängerung um ca. ein Jahr und eine Verlängerung der Zeitdauer bis zur Schmerzprogression um ca. 5 Monate (in der genäherten Zielpopulation um fast 8 Monate) aus Sicht der Patienten und der sie behandelnden Ärzte als bereits sehr nachhaltig angesehen werden.

4.2 Methodik

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

4.2.1 Fragestellung

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

– Patientenpopulation
– Intervention
– Vergleichstherapie
– Endpunkte
– Studientypen

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.

Als Schilddrüsenkrebs (Schilddrüsenkarzinom, lat. *Struma maligna*) wird eine bösartige Neubildung der Schilddrüse bezeichnet [28]. Diese stellt den am häufigsten auftretenden malignen Tumor des endokrinen Systems dar. Etwa 5% bis 10% aller Schilddrüsenkrebserkrankungen entfallen auf das medulläre Schilddrüsenkarzinom (*Medullary Thyroid Cancer*, MTC), das auch als C-Zellkarzinom bezeichnet wird und hereditären (ca. 25%) oder sporadischen Ursprungs (ca. 75%) sein kann [29].

Die Prognose bei MTC ist im Allgemeinen günstig, wenn die Erkrankung bereits in einem frühen Stadium behandelt wird; die fünf-Jahres-Überlebensrate nach Behandlung liegt dann durchschnittlich bei 60-70% [30]. Bei Patienten mit Fernmetastasen liegt sie jedoch nur bei etwa 40% mit einer medianen Gesamtüberlebensrate von 2-3 Jahren [31, 32]. Bei geschätzten 35-50% der Erkrankten erstreckt sich der Tumor bereits bei der Erstdiagnose über den Bereich der Schilddrüse hinaus und involviert auch lokale Lymphknoten; 10-15% haben bereits Fernmetastasen [30-32]. Bei etwa 50% aller MTC-Patienten flammt die Erkrankung nach Behandlung (kurative Operation/Thyreoidektomie) erneut auf [30]. Das metastasierte MTC breitet sich in der Mehrzahl aller Fälle auf regionale Lymphknoten aus, anschließend auf Leber, Lunge und Knochen; bei den Patienten tritt eine Vielzahl an Symptomen auf, die entweder mit humoralen Faktoren zusammenhängen, welche der Tumor selbst produziert, oder mit der Lage des Tumors. Lokale Symptome umfassen Schluckbeschwerden, Halsschmerzen, Heiserkeit und Atemnot. Knochenmetastasen können oft Schmerzen verursachen, die mit Opiaten oder Bisphosphonaten behandelt werden. Durchfall und *Flush-*

Syndrom treten häufig infolge humoraler Faktoren auf (krankheitsbedingt hohe CTN-Werte). Metastasen können außerdem Funktionsstörungen der Leber oder Kompressionen der Wirbelsäule verursachen. Annähernd 90% aller Patienten mit metastasierter Erkrankung sterben letztendlich an diesem fortgeschrittenen Krebs [31, 33].

In der Mehrzahl der Fälle ist das MTC durch unspezifische (typischerweise aufgrund einer Mutation erfolgte) Aktivierung des RET (*Rearranged During Transfection*) Protoonkogens, eines transmembranösen Rezeptors mit Tyrosinkinaseaktivität, charakterisiert [34].

Zum gegenwärtigen Zeitpunkt ist in Deutschland neben Vandetanib keine effektive oder anerkannte Behandlungsoption für fortgeschrittenes MTC verfügbar. Die überwiegende Mehrheit der betroffenen Patienten verstirbt letztendlich an MTC. Der langfristige Erhalt einer akzeptablen Lebensqualität steht für diese Patienten an erster Stelle. Da sich mit hoher Wahrscheinlichkeit früher oder später Tumor- und/oder CTN-bedingte Symptome (siehe Modul3A Abschnitt 3.2.1) manifestieren und spätestens bei Auftreten von Knochenmetastasen Schmerzen entstehen, deren Intensität selbst durch den Einsatz von Opiaten oft nur noch unzureichend begegnet werden kann, gibt es einen hohen unbefriedigten medizinischen Bedarf für eine Behandlungsoption. Eine solche Therapie sollte in der Lage sein, das Fortschreiten der Erkrankung und ihrer Begleitsymptome hinauszuzögern, ein Tumoransprechen zu erreichen und sie sollte darüber hinaus auch bei Einsatz über einen langen Zeitraum hinweg gut verträglich sein.

Fragestellung:

Fragestellung und Zielsetzung dieser frühen Nutzenbewertung sind der Nachweis und die Quantifizierung des Zusatznutzens von Vandetanib gegenüber der zweckmäßigen Vergleichstherapie bei der Behandlung von erwachsenen Patienten mit aggressivem und symptomatischem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem MTC. Die Bewertung des medizinischen Zusatznutzens erfolgt gegenüber der zweckmäßigen Vergleichstherapie im Indikationsgebiet. Im Rahmen der Beratung durch den Gemeinsamen Bundesausschuss (G-BA) vom 29.09.2011 wurde *Best Supportive Care* (BSC) als zweckmäßige Vergleichstherapie festgelegt ([1], siehe Modul 3A, Abschnitt 3.1.2). Die Quantifizierung des medizinischen Nutzens und des medizinischen Zusatznutzens erfolgt anhand patientenrelevanter Endpunkte (siehe unten) auf der Basis der Ergebnisse einer randomisierten kontrollierten Studie (RCT).

<u>Patientenpopulation</u>

Die Zielpopulation der zu bewertenden Intervention sind erwachsene Patienten mit aggressivem und symptomatischem MTC in nicht resektablem, lokal fortgeschrittenem oder metastasiertem Stadium. Näherungsweise wird für diese Patienten die Teilpopulation der Patienten mit symptomatischem und progressivem Verlauf innerhalb der Zulassungsstudie von Vandetanib als genäherte Zielpopulation dargestellt, da sich die Gruppe der aggressiven Erkrankungen klinisch definiert und nicht genau abgrenzen lässt [3]. Im Dossier werden

jeweils die Ergebnisse für die Gesamtpopulation der Studie sowie für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung dargestellt.

Intervention

Das zu bewertende Arzneimittel ist Vandetanib, 300 mg Tablette (Handelsname Caprelsa®), in der zugelassenen Dosierung von 300 mg einmal täglich oral. Für eventuell erforderliche Dosisreduktionen steht auch eine 100 mg Tablette zur Verfügung.

Vergleichstherapie

Die Standardtherapie beim MTC ist die Operation in kurativer Intention. Für das fortgeschrittene bzw. metastasierte MTC existierte vor Zulassung von Vandetanib keine zugelassene bzw. allgemein empfohlene Therapie.

Die aktuellen Behandlungsempfehlungen und Guidelines der *European Society for Medical Oncology* (ESMO; [35]), der European Thyroid Association [36] und der *American Thyroid Association* (ATA; [37]) raten vom Einsatz zytotoxischer Chemotherapie zur Behandlung von Patienten mit fortgeschrittenem MTC aufgrund der unbefriedigenden Resultate ab. Als Erstlinientherapie empfehlen die Fachgesellschaften stattdessen die Behandlung der Patienten mit TKI im Rahmen klinischer Studien und bezeichnen zum Teil ebenso wie die aktuelle Leitlinie (Version 1.2013) des *National Comprehensive Cancer Network* (NCCN) zum Schilddrüsenkarzinom [38] explizit Vandetanib als in Betracht zu ziehende zugelassene Behandlungs-Option.

Da bis zum Zeitpunkt der Zulassung von Vandetanib keine andere zugelassene Behandlungsoption zur Verfügung stand, erhielten MTC-Patienten der Zielpopulation bis dato anstatt einer zielgerichteten Therapie unterstützende Behandlungen im Rahmen von *Best Supportive Care* (BSC). Dies umfasst die nach dem Stand der medizinischen Wissenschaften bestmöglichen unterstützenden Behandlungsmaßnahmen, also die individuelle Behandlung physischer sowie psychischer Beschwerden mit den Zielen [39]:

- Linderung krankheitsbedingter Symptome
- Erhalt oder Steigerung der Lebensqualität des Patienten
- Verminderung des Auftretens unerwünschter Nebenwirkungen von potenziell toxischen Therapien
- Prognostische Verbesserung der Behandlungsergebnisse

In Abhängigkeit vom Verlauf der Erkrankung und der primär betroffenen Organe erhalten Patienten eine symptomatische oder palliative Behandlung. Diese ist in ihren Maßnahmen nicht einheitlich definiert, sondern umfasst die individuell erforderliche Kontrolle von z. B. gastrointestinalen Störungen und Schmerzen entsprechend der klinischen Situation.

Gemäß der Beratung durch den G-BA (Vorgangsnummer 2011-B-25) wird das Ausmaß des medizinischen Zusatznutzens von Vandetanib gegenüber der zweckmäßigen Vergleichstherapie BSC bemessen („Die zweckmäßige Vergleichstherapie für Vandetanib zur Behandlung von Patienten mit nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom ist *Best Supportive Care*.“) [1].

Die im Entwicklungsprogramm von Vandetanib in der Indikation durchgeführte RCT ist placebokontrolliert. Die Wahl der Kontrollgruppe sichert eine robuste Bewertung des primären Endpunktes und war ethisch vertretbar, da für die Studienpopulation keine Standardtherapie verfügbar war. Die Studienpopulation war zwar hinsichtlich systemischer Krebstherapien, nicht aber hinsichtlich unterstützender Maßnahmen und Arzneimittel eingeschränkt. Medikamente, die für die Sicherheit und das Wohlbefinden des Patienten erforderlich waren, konnten nach dem Ermessen des behandelnden Arztes eingesetzt werden [22]. Sowohl die Patienten in der Vandetanib-Gruppe als auch in der Placebo-Gruppe erhielten eine als BSC eingestufte Begleitbehandlung. Somit bildet die Placebo-Gruppe BSC im Versorgungsalltag der Zielpopulation von Vandetanib weitestgehend ab. Für weitere Ausführungen zu BSC wird auf Modul 3, Abschnitt 3.1 verwiesen.

Das Studienprotokoll erlaubte eine palliative Strahlentherapie zur Behandlung von Symptomen während der Studie. Ausgeschlossen war jedoch die Bestrahlung von Zielläsionen, die vorab zur Kontrolle des Tumorwachstums nach RECIST bestimmt worden waren. Diese Einschränkung der Bestrahlung von Zielläsionen ist im Rahmen der Durchführung von klinischen Studien aus Sicht von AstraZeneca erforderlich, um zu gewährleisten, dass die Tumorkontrolluntersuchungen nach RECIST den Effekt der zu untersuchenden Behandlung abbilden und nicht den Effekt einer anderen zusätzlichen Behandlung. Metastasen, die nicht als Zieltumoren für die Tumorkontrolluntersuchungen bestimmt worden waren, konnten während der Studie palliativ bestrahlt werden. Hier gab das Protokoll lediglich vor, dass dies nur so lange innerhalb der verblindeten Studienbehandlung erlaubt war, wie diese Nicht-Zielläsionen nicht erheblich größer geworden waren und damit gemäß den Tumorkontrollkriterien nach RECIST einen Progress konstituierten. Patienten, bei denen ein Tumorprogress vorlag, fielen gemäß dem Protokoll aus dem aktiven Teil der Studie heraus und konnten im Anschluss an die verblindete Behandlung nach der Entscheidung des behandelnden Arztes bestrahlt werden. Sechs Patienten in der Zulassungsstudie erhielten als erste Folgetherapie nach der Studie eine Strahlentherapie.

Da zum Zeitpunkt der Studie keine zugelassene Therapie vorhanden war und bereits klare Hinweise auf eine Wirksamkeit von Vandetanib in der geprüften Indikation vorlagen, wurde allen Patienten in beiden Behandlungsarmen nach Progression eine *Open-Label*-Behandlung mit Vandetanib ermöglicht. In der Gesamtpopulation erhielten 102 Patienten (30,8%) in der offenen Behandlungsphase Vandetanib, aufgeteilt auf 44 Patienten (19%) des Vandetanib-Arms und 58 Patienten (58%) des Placebo-Arms. In der genäherten Zielpopulation traf dies auf 64 Patienten (34,4%) zu, wobei 26 Patienten des Vandetanib-Arms (20,6%) und 38 Patienten des Placebo-Arms (63,3%) Vandetanib in der offenen Behandlungsphase erhielten.

Patientenrelevante Endpunkte

Die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens gegenüber BSC erfolgt auf der Grundlage der für die Indikation etablierten patientenrelevanten Endpunkte. Patientenrelevanz und Validität bezüglich der nachfolgend genannten Endpunkte werden in Abschnitt 4.2.5.2 ausführlich dargestellt. Zur Quantifizierung des therapeutisch bedeutsamen Zusatznutzens werden betrachtet:

- Progressionsfreies Überleben (*Progression-Free Survival*, PFS)
- Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)
- Objektive Ansprechrate (*Objective Response Rate*, ORR)
- Krankheitskontrollrate (*Disease Controle Rate*, DCR)
- Dauer der objektiven Ansprechens (*Duration of Response*, DOR)
- Explorative Morbiditätsendpunkte:
  - Gewichtszunahme
  - Einfluss der Behandlung auf Diarrhöen
  - WHO-Performancestatus (WHO-PS)
  - Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (*Baseline*)
  - Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (*Baseline*)
- Gesamtüberleben (*Overall Survival*, OS)
- Gesundheitsbezogene Lebensqualität (*Health Related Quality of Life*, HRQoL, anhand des FACT-G Fragebogens)
- Unerwünschte Ereignisse (UE)

Neben den patientenrelevanten Endpunkten werden zusätzlich die Parameter CTN und CEA für das biochemische Ansprechen dargestellt. CTN und CEA sind geeignete Marker für den klinischen Endpunkt Progression und dienen auf Grund ihrer Signalwirkung für eine progrediente Erkrankung als Surrogat für patientenrelevante Endpunkte des Bereichs Morbidität (krankheitsbedingte Symptome) [40]. Hier werden sie insbesondere zur Unterstützung der Ergebnisse zu PFS und ORR dargestellt.

Studientypen

Alle oben aufgeführten patientenrelevanten Endpunkte lassen sich in randomisierten kontrollierten Studien (RCT) untersuchen. RCT sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung entsprechend durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des medizinischen Nutzens einer medizinischen Intervention. Daher erfolgt die Nutzenbewertung von Vandetanib auf der Basis einer randomisierten kontrollierten Studie.

Über die Informationen aus der RCT hinausgehend werden unterstützende Informationen aus zwei nicht randomisierten Studien herangezogen, um die Evidenz für Vandetanib im MTC zu vervollständigen. Diese Studien sind gemäß den in Abschnitt 4.2.2 angegeben Kriterien nicht geeignet, die Fragestellung gemäß 4.2.1 zu beantworten. Sie werden dennoch in Abschnitt 4.3.2.3 sowie ebenfalls in Anhang 4-F dargestellt. Auch die in diesen Studien erhobenen Endpunkte werden der Vollständigkeit halber in Anhang 4-G hinsichtlich ihres Verzerrungspotentials bewertet. Ihre Resultate sind allerdings mit hoher Ergebnisunsicherheit behaftet und wurden daher **nicht** für die Bewertung des medizinischen Nutzens und des Zusatznutzens von Vandetanib berücksichtigt. Aus diesem Grund werden in den Bewertungsbögen zu Studie D4200C00008 (Tabelle 4-191) und Studie D4200C00068 (Tabelle 4-192) patientenrelevante Endpunkte im Sinne von Tabelle 4-1 (siehe unten) nur aufgeführt, insoweit sie in diesen Studien auch erhoben wurden.

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.

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.

Ziel der Ein- und Ausschlusskriterien ist es, diejenigen Studien zu charakterisieren, die grundsätzlich für die Beantwortung der Fragestellung gemäß 4.2.1 geeignet sind. Die Selektionskriterien grenzen die einzuschließenden Studien in Bezug auf folgende Kriterien ein: Population, Intervention, *Outcome* und Studiendesign [41].

Die zu bewertende Intervention ist Vandetanib (Handelsname: Caprelsa®) in der zulassungsrelevanten Dosierung (300 mg einmal täglich oral). Die Vergleichsintervention ist *Best Supportive Care*. In die Nutzenbewertung wurde eine Studie mit erwachsenen Patienten beiderlei Geschlechts mit medullärem Schilddrüsenkarzinom (MTC) (*Medullary Thyroid Carcinoma)* in nicht resektablem lokal fortgeschrittenem oder metastasiertem Stadium,

eingeschlossen, die auch Patienten der genäherten Zielpopulation mit progressivem und symptomatischem MTC umfasste, und anhand patientenrelevanter Endpunkte ausgewertet. Aufgrund der Seltenheit der Erkrankung ist die in die Nutzenbewertung eingeschlossene Studie weltweit die erste Phase-III-Studie, sowie die erste Studie dieser Größe (Einschluss von 331 Patienten in der Gesamtpopulation, davon 186 in der genäherten Zielpopulation), die zur Behandlung des MTC durchgeführt wurde.

In die Nutzenbewertung sollen ausschließlich randomisierte kontrollierte Studien (RCT) einfließen, die Resultate zu patientenrelevanten Endpunkten berichten. Eine Übersicht über die Ein- bzw. Ausschlusskriterien findet sich in Tabelle 4-1.

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

Kategorie	**Prädefinierte Einschlusskriterien**	**Prädefinierte Ausschlusskriterien**
Patienten (P)	Studienpopulation umfasst erwachsene Patienten beiderlei Geschlechts mit aggressivem und symptomatischem medullärem Schilddrüsenkarzinom (MTC) in nicht resektablem lokal fortgeschrittenem oder metastasiertem Stadium	Patienten unter 18 Jahren; Indikationen außerhalb des MTC; Stadium der Erkrankung außerhalb der Indikation
Intervention (I)	Behandlung mit Vandetanib 300 mg einmal täglich oral	Behandlung mit anderen Wirkstoffen oder mit abweichenden Dosierungen von Vandetanib
Kontrollgruppe (C)	Behandlung mit Placebo/*Best Supportive Care*	Vom Einschlusskriterium abweichende/keine Kontrollgruppe
Endpunkte (O)	Mindestens **einer** der patientenrelevanten Endpunkte wird berichtet: • *Mortalität* o Gesamtüberleben (OS) • *Morbidität* o Progressionsfreies Überleben (PFS) o Zeit bis zur Schmerzprogression (TWP) o Objektive Ansprechrate (ORR) o Krankheitskontrollrate (DCR) o Dauer des objektiven Ansprechens (DOR) o Gewichtszunahme o Einfluss der Behandlung auf Diarrhöen o WHO-Performancestatus (WHO-PS) o Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (*Baseline*) o Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline) • *Lebensqualität* o Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G • *Sicherheit* o Unerwünschte Ereignisse (UE)	Keine Resultate zur Wirksamkeit/Sicherheit bzw. abweichende Endpunkte
Studientyp	Randomisierte kontrollierte Studie (RCT)	Vom Einschlusskriterium abweichendes Studiendesign

Kategorie	Prädefinierte Einschlusskriterien	Prädefinierte Ausschlusskriterien
Studiendauer	Keine Einschränkungen; ein medianes *Follow-Up* für primäre Endpunktanalyse von möglichst etwa 24 Monaten (Hinweis: *Follow-Up* für Gesamtüberleben bei Eintreten von 50% Ereignissen) erscheint jedoch sinnvoll hinsichtlich valider Aussagen zu patientenrelevanten Endpunkten.	Keine
Publikationstyp	Vollpublikation oder Bericht verfügbar, der den Kriterien des CONSORT Statements genügt und so eine Einschätzung der Studienergebnisse ermöglicht. Doppelpublikationen ohne relevante Zusatzinformation werden ausgeschlossen.	Studien in der Rekrutierungsphase; ohne verfügbare Vollpublikation oder Bericht

Anhand dieser prädefinierten Einschlusskriterien wurden relevante Studien selektiert, anhand derer das Ausmaß des Zusatznutzens von Vandetanib gegenüber der zweckmäßigen Vergleichstherapie, *Best Supportive Care*, mit hoher Beweissicherheit bewertet wurde. Bei Nichterfüllung eines oder mehrerer Einschlusskriterien oder bei Erfüllung eines Ausschlusskriteriums wurden entsprechende Studien ausgeschlossen.

Die Wahl der Einschlusskriterien wird im Folgenden begründet. Die Ausschlusskriterien ergeben sich zum überwiegenden Teil durch Nichterfüllung der Einschlusskriterien.

Die Kriterien hinsichtlich Population und Behandlung richten sich nach der durch die Zulassung festgelegten Zielpopulation des zu bewertenden Arzneimittels Vandetanib (siehe Fachinformation [42]). Studien mit signifikant davon abweichender Patientenpopulation wurden ausgeschlossen.

Das Einschlusskriterium bezüglich der Population (vgl. Tabelle 4-1) beinhaltet die Zulassungspopulation von Vandetanib. Gemäß Zulassung ist eine „aggressive und symptomatische“ Erkrankung gefordert. Dieses Kriterium wurde als Rahmenbedingung im europäischen Zulassungsprozess definiert, insbesondere, um dem behandelnden Arzt die Möglichkeit zu geben, den Zeitpunkt für den Therapiebeginn auf die individuellen Bedürfnisse seiner Patienten abzustimmen [2]. Eine klare Definition des Begriffs „aggressiv“ erfolgte an dieser Stelle nicht; es wurden nur vorschlagsweise einige Parameter spezifiziert, anhand derer die Orientierung bei der Einstufung erfolgen könnte. Als „symptomatisch“ sind nach den an gleicher Stelle benannten Kriterien alle Patienten mit krankheitstypischen Symptomen (Durchfälle, Schmerzen, Erschöpfung, etc.) anzusehen.

Im Rahmen des europäischen Zulassungsverfahrens wurden *post-hoc* zusätzliche Auswertungen für verschiedene Teilpopulationen generiert („symptomatisch“, „progressiv“, „symptomatisch oder progressiv“, „symptomatisch und progressiv“, „nur symptomatisch“, „nur progressiv“). Nach abschließender Diskussion entschied sich die EMA jedoch für eine anders lautende, klinisch definierte Zielpopulation, nämlich Patienten mit aggressiver und symptomatischer Erkrankung [2].

Da sich der Anteil an Patienten mit aggressiver Erkrankung mangels präziser Definition nicht genau abgrenzen lässt, folgt AstraZeneca der Empfehlung des IQWiG [3], die Population der Patienten mit „symptomatischer und progressiver“ Erkrankung als Näherung an die tatsächliche Zielpopulation heranzuziehen. Entsprechend werden im Rahmen der vorliegenden Neufassung des Nutzendossiers alle relevanten Auswertungen sowohl für die Gesamtpopulation der RCT (Auswertungen gemäß Studienprotokoll und statistischem Analyseplan) als auch für die genäherte Zielpopulation der Patienten mit symptomatischer und progressiver Erkrankung (*post-hoc* Auswertungen) vorgelegt. Anhand von Interaktionstests wird ein hohes Maß an Übereinstimmung der Ergebnisse für beide Populationen gezeigt und damit belegt, dass von einer Übertragbarkeit der Ergebnisse von der Gesamtpopulation auf die genäherte Zielpopulation der Patienten mit progressiver und symptomatischer Erkrankung ausgegangen werden kann.

Als Einschlusskriterium für die Teilpopulation „symptomatisch und progressiv“ wurden folgende Kriterien angelegt: Als „progressiv“ galten Patienten mit Progression innerhalb der letzten 12 Monate vor *Baseline* (gemäß Angabe durch den Prüfarzt). Für das Kriterium „symptomatisch“ wurden typische MTC-Symptome herangezogen: Diarrhö, *Flushing*, Erschöpfung, Schmerzen, Atemwegssymptome, Übelkeit, Gewichtsverlust, Dysphagie oder Dysphonie zur *Baseline*. Diese Kriterien wurden auch für die Auswertungen für die genäherte Zielpopulation verwendet [11].

Die Wahl des Kriteriums für die Kontrollgruppe erfolgte gemäß Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA und spiegelt zugleich die aktuelle medizinische Praxis wider, da neben der symptomatischen und palliativen Symptomkontrolle (*Best Supportive Care*) keine geeigneten Therapien im Indikationsgebiet zur Verfügung stehen. Eine Placebo-Gruppe ist als Kontrollgruppe geeignet, solange in dieser Gruppe begleitende Maßnahmen zur Symptomkontrolle erlaubt sind. Die Patienten in einer solchen Placebokontrollgruppe bilden damit die klinische Praxis der *Best Supportive Care* ab. Studien mit abweichenden Kontrollarmen wurden ausgeschlossen.

Die im Rahmen dieser Nutzenbewertung betrachteten Endpunkte stellen in der Indikation etablierte patientenrelevante Endpunkte dar. Eine ausführliche Begründung für die Wahl der patientenrelevanten Endpunkte findet sich im Abschnitt 4.2.5.2.

Da diese Endpunkte alle im Rahmen von RCT beurteilt werden können, schränkt das Einschlusskriterium hinsichtlich Studientyp die Nutzenbewertung auf RCT ein, weil RCT, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt werden, mit der geringsten Ergebnisunsicherheit behaftet sind. Entsprechend wurden die beiden identifizierten nicht-kontrollierten Studien nur illustrativ in die Darstellung mit aufgenommen, ohne ihre Ergebnisse für die Quantifizierung des Zusatznutzens von Vandetanib heranzuziehen.

Hinsichtlich der Beobachtungsdauer gibt es keine Einschränkungen, allerdings werden Studien mit einer Beobachtungsdauer von mindestens 24 Monaten bevorzugt. Diese Beobachtungsdauer gilt als ausreichend, um patientenrelevante Endpunkte bzw. Ereignisse zu

beobachten, insbesondere das progressionsfreie Überleben sowie die krankheits- und therapiebezogene Morbidität. Das Kriterium hinsichtlich des Publikationstyps soll eine zuverlässige Einschätzung der Studienergebnisse ermöglichen.

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 4.2.1 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 in folgenden Fällen erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt:

– Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.

– Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

Sofern eine bibliografische Literaturrecherche erforderlich ist, soll diese mindestens in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken 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. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Geben Sie dabei an, ob Sie eine gemeinsame Suche für mehrere Fragestellungen durchgeführt haben (z. B. gemeinsame Suche nach Studien mit dem zu bewertenden Arzneimittel und Studien zur zweckmäßigen Vergleichstherapie für indirekte Vergleiche). Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Eine bibliografische Literaturrecherche für das zu bewertende Arzneimittel war im Rahmen der Erstellung des vorliegenden Dossiers erforderlich, da die Zulassung von Vandetanib in den USA (06.04.2011,[43]) zum Zeitpunkt der Neueinreichung länger als 12 Monate zurück liegt. Zwei weitere Zulassungen von Vandetanib wurden in Kanada am 12.01.2012 und in der Schweiz am 01.05.2012 erteilt (siehe Modul 2, Abschnitt 2.2.3). Literaturrecherchen für etwaige indirekte Vergleiche entfielen, da die Nutzenbewertung anhand eines direkten Vergleichs mit der zweckmäßigen Vergleichstherapie durchgeführt werden konnte und somit keine indirekten Vergleiche zur Bewertung des medizinischen Zusatznutzens von Vandetanib erforderlich waren.

Die bibliografische Literaturrecherche zur Identifizierung von Studien zur Beantwortung der unter Abschnitt 4.2.1 dargelegten Fragestellung für das vorliegende Nutzendossier wurde am 01.10.2012 gemäß den Vorgaben der VerfO getrennt in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt und am 11.01.2013 aktualisiert. Diese Recherche wurde nur zur Beantwortung der unter 4.2.1 beschriebenen Fragestellung durchgeführt. Die Treffer wurden auf deutsch- und englischsprachige Publikationen sowie auf die Publikationsfelder „Mensch“ und „Krebserkrankungen“ beschränkt. Eine zeitliche Beschränkung wurde nicht angewandt. Suchstrategien und primäre Trefferzahlen sind in Anhang 4-A ausführlich dargestellt, die Ergebnisse der Recherche in Abschnitt 4.3.1.1.2.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien von Dritten vollständig identifiziert werden.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov, clinicalstudyresults.org sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden.

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. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob 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 Vandetanib für die Zusatznutzenbewertung berücksichtigt werden, ist eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien in den gängigen medizinischen Studienregistern *clinicaltrials.gov*, *clinicalstudyresults.org* sowie über das *International Clinical Trials Registry Platform Search Portal* (ICTRP *Search Portal*, Suchportal der WHO) durchgeführt worden. Ein Update der ursprünglich am 05.01.2012 durchgeführten Suchen erfolgte am 07.01.2013 für die Erstellung des vorliegenden Dossiers. 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 Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die durch die gemäß Abschnitt 4.2.3.2 durchgeführte bibliografische Literaturrecherche sowie durch die gemäß Abschnitt 4.2.3.3 durchgeführte Studienregister-Abfrage identifizierten Studien wurden unter Zuhilfenahme der in Abschnitt 4.2.1 prädefinierten Kriterien zum Ein- bzw. Ausschluss relevanter Studien (Tabelle 4-1) von zwei unabhängigen Personen selektiert. Auf Basis aller verfügbaren Informationen aus den Abstracts bzw. Volltexten bzw. dem Studienregistereintrag und eventuell darin hinterlegten Dokumenten wurden Studien bei Erfüllung aller Kriterien als relevant eingestuft oder bei Nichterfüllung eines oder mehrerer Kriterien begründet ausgeschlossen. Bei Diskrepanzen im Selektionsprozess wurde die Einschätzung einer unabhängigen dritten Person in die Selektion einbezogen. In einem Selektionsschritt wurden die durch die Recherche identifizierten Publikationen und Registereinträge auf Grundlage der vorhandenen Informationen als relevant oder nicht relevant bewertet.

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-G):

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 des Behandlers
– 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 des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die zur Quantifizierung des Zusatznutzens eingeschlossene Studie wurde hinsichtlich ihrer Aussagekraft bewertet. Die Bewertung der eingeschlossenen Studie im Rahmen des vorliegenden Dossiers erfolgte anhand der zur Verfügung stehenden Informationen aus dem Studienbericht der Studie.

Bewertung des Verzerrungspotenzials

Sowohl allgemeine (endpunktübergreifende) als auch endpunktspezifische Gesichtspunkte wurden untersucht. Bei Verzerrungen auf Endpunktebene wurde jeder Endpunkt und gegebenenfalls jede Operationalisierung getrennt betrachtet. Entsprechend der Verfahrensordnung des G-BA zur Bewertung des Nutzens von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V vom 20. Januar 2011 [10] wurden dabei folgende Aspekte bewertet:

Verzerrungsaspekte der Ergebnisse auf Studienebene:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des behandelnden Arztes
- Ergebnisgesteuerte Berichterstattung
- Sonstige das Verzerrungspotential beeinflussende endpunktübergreifende Punkte

Verzerrungsaspekte der Ergebnisse auf Endpunktebene:

- Verblindung der Endpunkterheber
- Umsetzung des *Intention-To-Treat* (ITT)-Prinzips
- Ergebnisgesteuerte Berichterstattung
- Sonstige Aspekte: Behandlung fehlender Werte und *Drop-outs*

Die Bewertung des Verzerrungspotentials der Einzelstudie wurde tabellarisch zusammengefasst. Gemäß Verfahrensordnung wurde das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, 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.

Abschließend wurde unter Berücksichtigung der oben genannten Aspekte das Verzerrungspotenzial auf Ebene der Endpunkte als „hoch“ oder „niedrig“ eingestuft. Von einem „niedrigen“ Verzerrungspotential wurde ausgegangen, wenn mit großer

Wahrscheinlichkeit ausgeschlossen werden konnte, dass die Ergebnisse relevant verzerrt sind. Eine relevante Verzerrung liegt vor, falls sich die Aussagen und Schlussfolgerungen bezüglich der Ergebnisse bei Behebung des verzerrenden Aspektes verändern würden. Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

Die Vorgehensweise bei der Bewertung des Verzerrungspotentials für die Studie und jeden Endpunkt (Operationalisierung) richtete sich nach den Hinweisen der Bewertungsbögen in Anhang 4-G.

Für die unter „Weitere Untersuchungen“ dargestellten nichtkontollierten Studien wurde ebenfalls eine Bewertung des Verzerrungspotentials vorgenommen, auch wenn deren Ergebnisse nicht in die Nutzenbewertung einfließen. Generell wird das studienspezifische Verzerrungspotential für nichtkontrollierte Studien als „hoch“ eingestuft.

Eine zusammenfassende Bewertung der Verzerrungsaspekte erfolgte ausschließlich für die RCT.

Im Rahmen der Erstellung der vorliegenden Neufassung des Dossiers erfolgten *post-hoc* eine Reihe von Auswertungen zu den erhobenen Endpunkten innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung als näherungsweises Abbild der Indikation „aggressiv und symptomatisch“. Dabei wurden auch die Verzerrungsaspekte, sowohl auf Studienebene, als auch auf Endpunktebene gemäß obiger dargestellter Methodik vor dem Hintergrund dieser Neuauswertung noch einmal beurteilt.

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

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

epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[2] bzw. STROBE-Statements[3] 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 für die Nutzenbewertung von Vandetanib herangezogene randomisierte kontrollierte Studie wurde im Abschnitt 4.3.1.2 sowie im Anhang 4-F anhand von Design-Charakteristika beschrieben. Die Beschreibung erfolgte gemäß den Punkten 2b bis 14 des CONSORT-Statements. Der Patientenfluss wurde gemäß CONSORT-*Flow-Chart* dargestellt.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

Benennen Sie die Patientencharakteristika und Endpunkte, zu denen Ergebnisse dargestellt werden. 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). Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Die gesamte Studien- sowie die genäherte Zielpopulation (vgl. Abschnitt 4.2.2) der in die Nutzenbewertung eingehenden RCT wurden anhand folgender Charakteristika beschrieben:

- Alter
- Geschlecht
- Krankheitsspezifische *Baseline*-Charakteristika der MTC-Erkrankung (Details siehe 4.3.1.2)

[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.

Die Ergebnisse der Studie wurden nach Zielgrößen geordnet gegenübergestellt und vergleichend beschrieben. Dabei wurden die folgenden aussagekräftigen und patientenrelevanten Endpunkte quantifiziert:

- Progressionsfreies Überleben *(Progression-Free Survival*, PFS)
- Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)
- Objektive Ansprechrate (*Objective Response Rate*, ORR)
- Krankheitskontrollrate (*Disease Control Rate*, DCR)
- Dauer des objektiven Ansprechens (*Duration of Response*, DOR)
- *Explorativ*:
    - Gewichtszunahme
    - Einfluss der Behandlung auf Diarrhöen
    - WHO-Performancestatus (WHO-PS)
    - Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (*Baseline*)
    - Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (*Baseline*)
- Gesamtüberleben (*Overall Survival*, OS)
- Gesundheitsbezogene Lebensqualität (*Health Related Quality of Life*, HRQoL) gemessen anhand des FACT-G Fragebogens
- Unerwünschte Ereignisse (UE)

Alle aufgeführten Zielgrößen stellen nach Überzeugung von AstraZeneca patientenrelevante Endpunkte dar. Die Patientenrelevanz begründet sich wie folgt:

Progressionsfreies Überleben

In der palliativen onkologischen Therapie stellt ein Herauszögern des Fortschreitens der Erkrankung und der damit verbundenen Symptomatik eines der wichtigsten Therapieziele dar [14]; dies gilt insbesondere auch bei einem zu erwartenden langen Gesamtüberleben [15-17]. Deshalb zählt das progressionsfreie Überleben (*Progression-Free Survival*, PFS) in der palliativen Situation als Funktion der Wachstumsrate des Tumors und der Wirksamkeit des

untersuchten Präparates zu den patientenrelevanten Endpunkten [18], um die Effektivität einer eingesetzten Behandlung während des beobachteten Therapiezeitraums zu bewerten. Die Progression der Erkrankung geht mit zunehmender Morbidität des Patienten einher. PFS erfasst die Dauer, in der kein relevantes Tumorwachstum vorliegt, beziehungsweise der Tumor schrumpft und damit die Wirkung der Therapie gegeben ist und die gegenwärtige Morbiditätslast des Patienten bezogen auf den Tumor stabil ist. Dem durch die fortgeschrittene Erkrankung auch psychisch meist schwer belasteten Patienten wird zudem der Erfolg der Behandlung durch den Stillstand der Erkrankung vor Augen geführt, was umso mehr von Bedeutung sein kann, als die Erfolgsaussichten einer medikamentösen Intervention mit Antidepressiva für psychische Anpassungsstörungen bei fortgeschrittener Tumorerkrankung bzw. damit einhergehenden Komplikationen gering sind [44]. Durch Vandetanib kann über einen längeren Zeitraum eine stabile Situation hergestellt werden, in der der Patient eine kontinuierliche, wirksame Therapie erhält. Dies ist von besonderer Bedeutung, da es in der beschriebenen Indikation bislang keinen Therapiestandard gibt und den Patienten außerhalb von klinischen Studien kaum eine Therapie über BSC hinaus angeboten werden kann. Einem Beitrag zur G-BA Diskussionsrunde „Frühe Nutzenbewertung onkologischer Arzneimittel“ im November 2010 zufolge lässt sich der Nutzen eines verlängerten progressionsfreien Überlebens für den Patienten wie folgt zusammenfassen [45]:

1. Die Reduzierung tumorbedingter Symptome aufgrund einer Verringerung der Tumorgröße oder Unterdrückung des fortschreitenden Wachstums
2. Das Hinauszögern einer möglichen Folgebehandlung mit der Gefahr von zusätzlichen Nebenwirkungen
3. Die für den Patienten subjektiv als positiv erlebte Stabilisierung der Erkrankung („die Behandlung wirkt“).

Auch die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. weist in ihrem Kommentar zum IQWiG-Bericht „Aussagekraft von Surrogatendpunkten in der Onkologie“ darauf hin, dass PFS durchaus klinisch relevant sein kann, da eine relevant verlängerte krankheitsfreie Zeit für den Patienten wertvoll ist [17]. Hier steht nicht nur das Hinauszögern einer drohenden Pflegebedürftigkeit und der gefürchteten terminalen Krankheitsphase im Vordergrund, sondern auch der möglichst lange Erhalt eines Lebens in relativer Selbstbestimmung, und die zusätzliche Zeit für eine Auseinandersetzung mit und Akzeptanz der neuen Lebenssituation sind für den Patienten von großer Bedeutung. Besonders bedeutsam für den Patienten wird die Verzögerung einer Progression auch vor dem Hintergrund, dass das Fortschreiten der Krebserkrankung immer auch mit einer Therapieumstellung und damit potentiellen weiteren Erschwernis der individuellen Lebensumstände einhergeht [46]. Entsprechend konnte in einer retrospektiven Untersuchung an Patientinnen mit Brustkrebs ein direkter Zusammenhang zwischen einer Progression und klinisch relevanten Symptomatiken und somit auch den Ergebnissen patientenberichteter Erhebungen zur gesundheitsbezogenen Lebensqualität gezeigt werden [47]. Auch das IQWiG betont in einer aktuellen Pressemitteilung zum Thema „Brustkrebs mit Metastasen“ den Zusammenhang zwischen einem Hinauszögern der Progression und der Lebensqualität der Erkrankten [14]. In der deutschen Rechtsprechung wird ebenfalls die Symptomverzögerung

bzw. eine günstige Beeinflussung des Krankheitsverlaufs als Therapieziel bei fortgeschrittenen bzw. lebensbedrohlichen Erkrankungen als gleichberechtigt neben einer Verlängerung des Überlebens behandelt [48, 49].

Das PFS ist definiert als der Zeitraum von der Randomisierung bis zur Progression der Erkrankung oder bis zum Tod aus jeglicher Ursache. Die Bewertung der Progression erfolgt gemäß der vorhergehenden Festlegung im Studienprotokoll anhand modifizierter RECIST Kritierien [22, 23]. PFS eignet sich besonders für fortgeschrittene Krankheitsstadien [25] und bietet den Vorteil kleinerer Patientenzahlen in den klinischen Studien und einer kürzeren Beobachtungszeit im Vergleich zum Gesamtüberleben [16], welches insbesondere im Falle einer langen Überlebenszeit nach der Progression oft keine statistisch signifikanten Aussagen mehr über die untersuchte Therapie erlaubt [50]. Darüber hinaus ist PFS ein valides Messinstrument, da es auch unter Verblindung auf objektiven, quantitativen Messungen beruht und so die Beurteilung des Therapie-spezifischen Nutzens ohne Beeinflussung durch Effekte eventueller Folgetherapien ermöglicht [45]. Da in onkologischen Studiendesigns ein Wechsel der Patienten in den Behandlungsarm mit aktiver Medikation (*Cross-Over*) meist erst nach Progression vorgesehen ist, hat der Wechsel von Placebo auf Verum keinen verzerrenden Einfluss auf die Auswertung des PFS, wie dies hingegen in der Regel beim Gesamtüberleben gegeben ist [24, 51-54]. Die Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO) weist ebenfalls darauf hin, dass bei Studien mit substantiellem *Cross-Over* aus dem Kontroll- in den Behandlungsarm (im Falle von Progression) der potenzielle Endpunkt „Überleben“ nicht mehr anwendbar ist. Als verzerrungsfrei auswertbare Endpunkte kommen in derartigen Situationen daher nur PFS oder Ansprechrate infrage [16], wobei die Übertragbarkeit von Ergebnissen zum PFS auf das OS in einer Reihe onkologischer Erkrankungen als hinreichend gesichert gilt [55-57]. Bei der Beurteilung von PFS sollte allerdings immer berücksichtigt werden, dass eine Verzögerung der Progression und der damit einhergehenden zunehmenden Morbidität nur dann einen Vorteil für die Patienten darstellt, wenn gleichzeitig das Nebenwirkungsprofil der Therapie für die Patienten akzeptabel ist und somit die Lebensqualität so wenig wie möglich beeinträchtigt wird [24, 54]. FDA und EMA haben bereits mehrere neue Medikamente auf Basis einer PFS-Verbesserung zugelassen [24, 54, 58-60].

Zeit bis zur Schmerzprogression

Schmerz stellt eine bedeutsame Komponente der Symptomatik des aggressiven, nicht resektablen, lokal fortgeschrittenen oder metastasierten MTC dar. Dies bestätigt aus klinischer Sicht die DGHO, nach deren Aussage ein Großteil der Patienten extreme Schmerzen hat [21]. Im Rahmen der Nutzenbewertung wird unter anderem der patientenberichtete Parameter „Schmerz“, operationalisiert als Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP), als Morbiditätskriterium betrachtet. Die Zeit bis zur Schmerzprogression wurde aus dem höchsten Schmerzwert des validierten Schmerzfragebogens (*Brief Pain Inventory* (*Short Form*), BPI-SF) und dem vom Patienten angegebenen Gebrauch opioidhaltiger Schmerzmittel abgeleitet. Für die Patienten ist eine möglichst langfristige Verzögerung der

Schmerzprogression von großer Bedeutung, da Schmerzen nicht nur fortschreitende Morbidität signalisieren, sondern auch die Lebensqualität stark beeinträchtigen und auch der Einsatz von Schmerzmitteln wieder von entsprechenden Nebenwirkungen begleitet ist. Die Vermeidung/Hinauszögerung von Tumorschmerzen muss daher als herausragendes Therapieziel bei der Behandlung von Krebspatienten im Endstadium gelten.

Der BPI-SF wurde 1983 von Daut und Mitarbeitern entwickelt und dient zur Messung des Schmerzes und zur Einstufung der Schmerzfolgen. Sensorische Schmerzkomponenten und schmerzbedingte Beeinträchtigung bilden den Schwerpunkt des Fragebogens. Am häufigsten im angloamerikanischen Raum eingesetzt, wurde der BPI-SF von Radbruch und Mitarbeitern 1999 in die deutsche Sprache übersetzt und validiert [61]. Er wird von der *European Association for Palliative Care* als Messinstrument für klinische Studien der Tumorschmerztherapie bei Tumorpatienten empfohlen, kann aber auch für andere schmerzbehaftete Erkrankungen verwendet werden [62].

Objektive Ansprechrate

Bei der Behandlung onkologischer Erkrankungen in der metastasierten Situation stellt die Objektive Ansprechrate (*Objective Response Rate*, ORR) als Maß für die Kontrolle des Tumorwachstums und die Beobachtung des Ansprechens der Therapie sowohl ein valides Messinstrument als auch ein bedeutsames Therapieziel dar. Für die Messung des Tumoransprechens werden bildgebende Verfahren (CT/MRT) verwendet und gemäß den standardisierten und etablierten RECIST-Kriterien bewertet, um die Veränderungen der Tumorgröße im zeitlichen Verlauf darzustellen [22, 23]. Die ORR misst dabei die Verringerung der Tumorlast und umfasst sowohl komplettes als auch partielles Ansprechen, allerdings nicht den Stillstand der Erkrankung [24-26] und ist damit ein Parameter für reduzierte Morbidität. Das Wissen um ein Tumoransprechen und die Metastasenreduktion stellen eine psychische Entlastung und damit einen direkten Nutzen für den Patienten dar. Darüber hinaus dient auch dieser Endpunkt in vielen Studien als primärer Endpunkt für die Wirksamkeit als Zulassungsgrundlage bei der FDA [24, 26], und man geht davon aus, dass eine verbesserte ORR durch die Verkleinerung des Tumors die Symptome verringern kann und so gegebenenfalls auch ohne Verlängerung der Gesamtüberlebensrate einen patientenrelevanten klinischen Vorteil darstellt [53, 63]. In einigen onkologischen Indikationen gibt es auch Hinweise darauf, dass sich ein gutes Ansprechen positiv auf die Überlebenszeit auswirkt [64].

Krankheitskontrollrate

Die Krankheitskontrollrate (*Disease Control Rate*, DCR) ist ein Maß für die Krankheitsstabilisierung und berücksichtigt über die Patienten mit partieller oder kompletter Remission hinaus auch diejenigen Patienten, die durch die Therapie über einen längeren

Zeitraum eine Stabilisierung ihrer Erkrankung erleben dürfen. Damit stellt die Krankheitskontrollrate ein weiteres Maß für den klinischen Nutzen einer Therapie dar, indem sie all jene Patienten einbezieht, bei denen über mindestens ein halbes Jahr kein Progress vorliegt und die damit für einen gewissen Zeitraum von einer Krankheitskontrolle durch die Therapie profitieren. Dies stellt eine psychische Entlastung dar und bedeutet, dass Arzt und Patient nicht vor die Situation gestellt werden, den nächstfolgenden Behandlungsschritt entscheiden zu müssen. Auch die DCR ist gut und konsistent erhebbar und wird von nachfolgenden Therapiewechseln nicht beeinflusst [65], weshalb sie in breitem Maß in Evaluierungs- und Zulassungsstudien onkologischer Wirkstoffe zum Einsatz kommt [66-68].

Dauer des objektiven Ansprechens

Die Dauer des objektiven Ansprechens (*Duration of Response*, DOR) misst den Zeitraum, für den eine erreichte Tumorantwort (komplette oder partielle Remission) bestehen bleibt. Die Patientenrelevanz dieses Endpunktes ergibt sich unmittelbar aus der Patientenrelevanz des Ansprechens selbst (siehe oben, ORR): für die Betroffenen spielt es selbstverständlich eine Rolle, dass ein einmal erreichtes Ansprechen ihrer Erkrankung auch möglichst lange aufrechterhalten werden kann. Auch die DOR ist ein Endpunkt in vielen Evaluierungs- und Zulassungsstudien onkologischer Wirkstoffe [69-72] und wird als wichtiges Kriterium zur Beurteilung der Wirksamkeit herangezogen [73, 74].

Weitere patientenrelevante Endpunkte, die explorativ untersucht wurden:

Gewichtszunahme

Kachexie ist sowohl ein bedeutendes Symptom bei vielen Tumorerkrankungen als auch eine häufige Nebenwirkung onkologischer Therapien und somit ein direktes Anzeichen zunehmender Morbidität [75]. Insbesondere bei fortgeschrittenen Krebserkrankungen haben Patienten einen deutlichen und oft immer weiter zunehmenden Gewichtsverlust, der durch Ernährung nicht mehr kompensierbar ist. Die Kachexie entsteht zum Teil durch eine verminderte Nahrungsaufnahme, wobei die Appetitlosigkeit nicht die Ursache der Kachexie ist, sondern Folge der metabolischen Störungen, welche mit Entzündungsvorgängen zusammenhängen, die mit dem Tumorgeschehen einhergehen. Als sicht- und spürbarer Ausdruck ihrer Erkrankung ist diese sog. „Krebskachexie“ für viele Patienten und ihre Angehörigen eine erhebliche Belastung. Durch einen erhöhten Umsatz von Glucose (Einfachzucker) und einen vermehrten Abbau von Eiweißen führt die Kachexie zu einem Verlust an Muskelmasse und geht einher mit verstärkter allgemeiner körperlicher Schwäche, zunehmendem Kräfteverlust, Mutlosigkeit und damit verbunden einem zunehmenden Verlust an Lebensqualität. Bei vielen Krebspatienten stellt die Kachexie letzten Endes die eigentliche Todesursache dar [76]. Zur Gewichtsabnahme tragen auch psychologische Faktoren wie Angst und Depression bei. Die normalen Funktionen im Magen-Darm-Bereich, die für die

Resorption der Nahrung erforderlich sind, nehmen ab. Ein verminderter Appetit bei Geschmacksstörungen, verursacht z. B. durch die Chemotherapie, ist ebenfalls häufig.

Eine Gewichtsstabilisierung ist ein wichtiges Morbiditätskriterium für eine neue Therapie, da nur bei einem stabilen Gewicht der zunehmende Verlust an Lebensqualität aufgehalten werden kann [75].

Einfluss der Behandlung auf Diarrhöen

Diarrhöen bzw. eine erhöhte Stuhlfrequenz sind typische Nebenwirkungen des MTC. Die Diarrhö wird sehr häufig durch die tumortypischen erhöhten Calcitoninspiegel ausgelöst und kann in fortgeschrittenen Stadien der Erkrankung lebensbedrohliche Ausmaße erreichen [77] (siehe auch Modul 3, Abschnitt 3.2.1). Zudem gehören gastrointestinale Nebenwirkungen zu den häufigsten Nebenwirkungen einer medikamentösen Tumortherapie. Form und Schweregrad von Nebenwirkungen sind dabei sehr individuell. Sie hängen vom verwendeten Zytostatikum, sowie dessen Dosierung und Applikationsart ab. Außerdem spielt eine patientenspezifische Prädisposition eine Rolle. Die Befindlichkeit von Tumorpatienten kann durch die Auswirkungen der Diarrhö erheblich beeinträchtigt sein. Die Bereitschaft eines Patienten zur aktiven Mitwirkung an therapeutischen Maßnahmen kann vermindert und die Fortsetzung einer zeitgerechten Therapie verzögert werden. Potenziell lebensbedrohlich sind Komplikationen wie Dehydration, Elektrolytverlust, Bikarbonatverlust und eine krankhafte Übersäuerung des Körpers. Durch Diarrhö bedingte Verdauungsinsuffizienz und Malabsorptionen können zu einer Beeinträchtigung der Nährstoffausnutzungen und letztlich zu einer Mangelernährung führen. Diese bedeutet Defizite an Makro- und Mikronährstoffen und geht einher mit einem Energieverlust. Damit verbunden ist eine Abnahme an Wohlbefinden und Lebensqualität. Deshalb ist der Endpunkt „Einfluss der Behandlung auf Diarrhö“ ein wichtiger patientenrelevanter Endpunkt für eine neue Therapie, weil bereits durch die Krankheit an sich hervorgerufene Diarrhöen nicht noch zusätzlich durch die Nebenwirkung einer medikamentösen Therapie verstärkt werden sollten.

WHO-Performancestatus (WHO-PS)

Die in der Onkologie bekanntesten und in der Praxis am häufigsten angewendeten Paramater für die Beschreibung des funktionellen Status des Patienten sind zum einen der anhand des 1949 von David Karnofsky publizierte [78] und 1984 durch Mor validierte Index [79] und zum anderen der 1982 durch Oken et al. veröffentlichte WHO- oder ECOG Performance Status [80], wobei eine Studie von Buccheri et al. (1996) die Gleichwertigkeit der Instrumente gezeigt hat [81]. Der WHO Performance Status beschreibt den Allgemeinzustand und die Leistungsfähigkeit eines Patienten im täglichen Leben. Er ist als patientenrelevanter Endpunkt zu sehen, da er als Indikator für die Lebensqualität dient.

Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline)

In diesem explorativen Endpunkt wird basierend auf patientenberichteten Daten (anhand von Schmerztagebüchern) untersucht, ob der Opioid-Gebrauch zu Studienbeginn als Merkmal für die Effektstärke von bestehendem Schmerz einen Einfluss auf die Zeit bis zur Schmerzprogression (TWP) hat. Da Schmerzen die Lebensqualität stark beeinträchtigen, ist eine möglichst langfristige Verzögerung der Schmerzprogression für die Patienten von großer Bedeutung. Die Bedeutung der TWP für die Patienten im Allgemeinen wurde im vorliegenden Abschnitt bereits weiter vorne diskutiert (siehe dort).

Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

Im Verlauf einer Tumorerkrankung treten in vielen Fällen z. T. unerträgliche Schmerzen auf. Zur Verbesserung der Lebensqualität vieler tumorkranker Patienten ist eine erfolgreiche Schmerztherapie eine notwendige Voraussetzung, wobei Opioide die wichtigste Medikamentengruppe zur Therapie starker Schmerzen in der Tumorschmerztherapie darstellen. Trotz des Einsatzes von adäquaten Schmerzmitteln kann im Verlauf der Erkrankung eine Steigerung der Opioiddosis bzw. ein Wechsel auf ein stärkeres Opioid oder eine Rotation des Opioidregimens erforderlich werden [82, 83], was Ausdruck einer Toleranzentwicklung sein kann, jedoch meist auf eine Progredienz der Krankheit hindeutet. Insbesondere ist beim Gebrauch von Opioiden darauf zu achten, dass sich besonders bei langfristiger und/oder höchstdosierter Opioidtherapie eine opioidinduzierte Hyperalgesie ausbilden kann. Deshalb ist der Endpunkt „Reduktion des Opioidgebrauchs gegenüber dem Ausgangswert“ ein wichtiger patientenrelevanter Endpunkt für eine neue Therapie, mit dem Ziel, den Opioid-Verbrauch möglichst langfristig niedrig zu halten.

Biochemisches Ansprechen (Surrogat)

Als Surrogat innerhalb der patientenrelevanten klinischen Endpunktkategorie „krankheitsbedingte Symptome“ (Morbidität) wird zusätzlich das biochemische Ansprechen, gemessen an den Serumspiegeln von CTN und CEA, herangezogen (siehe Abschnitt 4.5.4). CTN und CEA haben gerade bei MTC eine Signalwirkung hinsichtlich der Progredienz der Erkrankung [40]. Hier werden sie insbesondere zur Unterstützung der Ergebnisse zu PFS und ORR dargestellt.

Gesamtüberleben

Bei der Behandlung fortgeschrittener onkologischer Erkrankungen stellt die Verlängerung des Gesamtüberlebens (*Overall Survival*, OS) des Patienten ein herausragendes Therapieziel dar.

Aufgrund der eindeutigen Definition, der Belastbarkeit des gemessenen Ergebnisses sowie der Tatsache, dass bereits implizit eine Risiko/Nutzen-Abwägung enthalten ist, wurde und wird das Gesamtüberleben in vielen Bereichen der Onkologie als zuverlässigster Endpunkt eingeordnet [26]. Eine Abbildung des Gesamtüberlebens ist in Studien zu onkologischen Interventionen jedoch oft kein geeigneter Parameter, da der Einfluss nachfolgender Therapien nicht bestimmt und so der individuelle Beitrag der Studienmedikation als einzelne Komponente des Therapieschemas der Patienten auf die Überlebenszeit nicht quantifiziert werden kann [51]. Zudem müsste für die Erhebung des Gesamtüberlebens ein weitaus längerer Beobachtungszeitraum als für das progressionsfreie Überleben (*Progression-Free Survival*, PFS) in der klinischen Studie angesetzt werden; dies gilt insbesondere für onkologische Populationen, die trotz progredienten Verlaufs der Erkrankung eine vergleichsweise lange Überlebensdauer haben [25, 53, 63, 84].

Für klinische Zulassungsstudien bei onkologischen Therapien schreiben Ethikkommissionen zudem häufig die Möglichkeit eines Wechsels zu *Open-Label*-Verum-Therapie im Studiendesign vor, insbesondere, wenn Therapiealternativen im Anschluss an die Studienmedikation fehlen. Dies bedeutet, dass Patienten verblindet in die Therapiearme randomisiert werden und bei Vorliegen von definierten Kriterien den Therapiearm wechseln können. In Zulassungsstudien zur onkologischen Therapie kann ein Therapiewechsel entweder stattfinden, wenn ein Effekt der zu untersuchenden Intervention nachgewiesen worden ist oder wenn z. B. eine Progression einsetzt. In diesem Fall können Patienten des Vergleichsarms der klinischen Studie aus ethischen Gründen ebenfalls die nachgewiesen effektivere Medikation erhalten. In die statistische Auswertung gehen dann bei einer – im Allgemeinen geforderten – Auswertung nach *Intention To Treat* (ITT) im Vergleichsarm auch die Effekte der medizinischen Intervention nach Wechsel der Therapie ein. Die Vergleichsgruppe zeigt somit höhere Effekte – eben die Effekte, die durch den Therapiewechsel zusätzlich erzielt wurden – als ohne ein solches *Cross-Over* Design sichtbar wären, während der inkrementelle Effekt sinkt [51]. Eine Studiendurchführung mit einem durch derartige Faktoren nicht beeinflussten Endpunkt wie z. B. PFS als primärem und OS nur als sekundärem Endpunkt gehört daher in vielen onkologischen Anwendungsgebieten zum anerkannten Standard [18, 46]. Auch aus klinischer Sicht wurde diese Vorgehensweise für Vandetanib bestätigt. Aus Sicht der DGHO ist die Möglichkeit eines frühzeitigen Wechels zu *Open-Label*-Vandetanib-Therapie aufgrund fehlender Behandlungsalternativen notwendig. Es sei unethisch, „auf *Overall Survival* zu gehen und nur wegen der Statistik eine Gruppe unbehandelt zu lassen“ [21].

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität ist eine wichtige Zielgröße bei der medizinischen Versorgung krebskranker Patienten. Die gesundheitsbezogene Lebensqualität wurde über den krankheitsunspezifischen, validierten Fragebogen FACT-G erhoben [85] und zeigt eine gute Korrelation zum klinischen Verlauf der Erkrankung. Dieses Instrument erfasst die vier Subskalen Physisches Wohlbefinden (PWB), Soziales Wohlbefinden (SWB), Emotionales

Wohlbefinden (EWB) und Funktionelles Wohlbefinden (FWB), welche in der zur Quantifizierung des Zusatznutzens von Vandetanib herangezogenen RCT durch die Patienten alle 12 Wochen bis zur Beendigung der randomisierten Behandlung ausgefüllt worden sind. Ein Erhalt oder sogar eine Verbesserung der Lebensqualität stellt – gerade in der palliativen Situation, die auf Dauer fast immer mit einer Verschlechterung der Lebensqualität einhergeht – neben den zuvor genannten Parametern ebenfalls einen patientenrelevanten Endpunkt dar, da die Therapie die Lebensqualität von Patienten möglichst verbessern, zumindest jedoch nicht verschlechtern sollte [86].

Unerwünschte Arzneimittelwirkungen

Ergänzend zum Nachweis des Nutzens muss die Erfassung unerwünschter Arzneimittelwirkungen zur Vervollständigung der Nutzen-Risiko-Abwägung stattfinden. Neben den erwünschten Wirkungen eines Arzneimittels treten in vielen Fällen auch unerwünschte auf, die den Patienten neben der eigentlichen Erkrankung physisch und psychisch belasten können. Daher muss der Nutzen einer palliativen Therapie hinsichtlich patientenrelevanter Zielkriterien wie Überleben oder PFS stets auch in Relation zur Toxizität bewertet werden.

4.2.5.3 Meta-Analysen

Sofern die vorliegenden Studien dazu geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden. Dabei sollen verschiedene Studientypen nicht in einer Meta-Analyse zusammengefasst werden. 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 Effekten[4] erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen

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

Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet[5] 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 Heterogenität[6] 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.

Die Meta-Analyse ist ein statistisches Verfahren um die Ergebnisse mehrerer Studien, die die gleiche Frage bearbeiten, quantitativ zu einem Ergebnis zusammenzufassen und dadurch die Aussagekraft (Genauigkeit der Effektschätzer) gegenüber Einzelstudien zu erhöhen.

Da nur eine RCT zur Quantifizierung des Zusatznutzens betrachtet wird, konnte keine Meta-Analyse durchgeführt werden. Daher entfällt die Beschreibung zur Durchführung von Meta-Analysen.

4.2.5.4 Sensitivitätsanalysen

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

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

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

[5] *Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*
[6] *Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.*

Sensitivitätsanalysen dienen zur Überprüfung der Robustheit der Ergebnisse und Schlussfolgerungen gegenüber der Variation methodischer Faktoren. Da keine Meta-Analyse durchgeführt worden ist, beziehen sich die durchgeführten Sensitivitätsanalysen nicht auf die Variation meta-analytischer Methodik (bei der Zusammenfassung von Resulten aus mehreren Studien wie z. B. der Verwendung eines Modells mit festen Effekten anstelle eines Modells mit zufälligen Effekten), sondern auf die Untersuchung der Robustheit der Resultate der Einzelstudie im Hinblick auf methodische Variation. Dabei ist man auf die Verfügbarkeit solcher Analysen aus der Einzelstudie angewiesen. Folgende Sensitivitätsanalysen sind im Ergebnisteil adressiert:

- Sensitivitätsanalyse hinsichtlich der Überrekrutierung (Ergebnis nur auf der Basis der gemäß Fallzahlplanung zu rekrutierenden Patienten) (4.3.1.2.2)

- Adjustierte Analysen für PFS, OS, TWP, objektive Ansprechrate und DCR mittels eines Cox-PH-Modells (*Cox Proportional Hazards Model*) bzw. einer logistischen Regression mit *Baseline*-Charakteristika als Kovariaten, um den möglichen Effekt von Unbalanciertheiten in *Baseline*-Charakteristika zwischen den Behandlungsgruppen zu untersuchen (siehe entsprechende Unterkapitel in Abschnitt 4.3.1.3). Diese Analysen wurden, außer im Fall des Gesamtüberlebens (OS), sowohl für die Gesamtpopulation der Zulassungsstudie als auch für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung durchgeführt.

- In einer Sensitivitätsanalyse für den Endpunkt „Zeit bis zur Schmerzprogression“ (TWP) wurden Modifikationen der Definition der TWP ausgewertet, um eine valide Beurteilung der Schmerzprogression unter Berücksichtigung beider Komponenten, des Schmerzes (gemessen anhand validierter Skalen) und des Schmerzmittelgebrauchs, abzugeben. Hierbei wurden unterschiedliche Strategien für die Ersetzung fehlender Werte verwendet. Diese Analysen wurden für die TWP sowohl innerhalb der Gesamtpopulation als auch der genäherten Zielpopulation durchgeführt.

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 sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

– Geschlecht
– Alter
– Krankheitsschwere bzw. -stadium

Sollten sich aus den verfügbaren Informationen Anhaltspunkte 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. 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.

Es sind Subgruppenanalysen durchgeführt worden, um zu überprüfen, ob ein Therapieeffekt über Subgruppen hinweg konstant ist, oder ob es Hinweise für eine Effektmodifikation gibt. Für den Endpunkt „progressionsfreies Überleben“ (PFS; primärer Endpunkt der Studie) wurden gemäß Studienprotokoll folgende Subgruppen betrachtet:

- Geschlecht (männlich, weiblich)
- Rasse (kaukasisch, andere)
- WHO *Performance Status* (WHO-PS), (0, ≥1)
- RET Mutationssatus (positiv, negativ, unbekannt)
- CTN Verdopplungszeit (≤24 Monate, >24 Monate, unbekannt)
- CEA Verdopplungszeit (≤24 Monate, >24 Monate, unbekannt)
- MTC-Status (hereditäre/sporadische Mutation, metastatisch, lokal fortgeschritten)
- Vorherige Therapie (vorbehandelt/nicht vorbehandelt, Ansprechen)

- Plasma-Biomarker: p-VEGF, p-VEGFR2, p-bFGF (hoch, niedrig, unbekannt)
- *a posteriori* spezifiziert : Alter (<65, ≥65)

Für die jeweiligen Subgruppen wurde jeweils ein *Hazard Ratio* mit zugehörigem 95%-Konfidenzintervall (KI) und 2-seitigem p-Wert mittels *Log-Rank* Test berechnet. Anschließend wurden die Ergebnisse mittels *Forest-Plot* dargestellt. Hinweise auf Effektmodifikationen wurden mithilfe von Interaktionstests abgeleitet. Um die Stärke einer Interaktion einzuordnen gibt es hier hinsichtlich des p-Wertes unterschiedliche Herangehensweisen. In der klinischen Statistik ist für Testentscheidungen die Verwendung eines Signifikanzniveaus von 0,05 üblich [87]. Die Cochrane Collaboration empfiehlt jedoch eine Erhöhung auf 0,1 als Ausgleich für die geringe Power, die Subgruppenanalysen meistens mit sich bringen [88]. Das Unterschreiten eines Schwellenwertes von 0,1 wurde auch in der Zulassungsstudie von Vandetanib als Hinweis auf Interaktion gehandhabt. In Analogie zu den Empfehlungen des IQWiG [89] wird jedoch im vorliegenden Dossier ein p-Wert kleiner als 0,2 als Hinweis auf eine Effektmodifikation gewertet. Alle Subgruppen wurden für den Endpunkt PFS für die Gesamt- und *a posteriori* auch für die genäherte Zielpopulation der Patienten mit symptomatischer und progressiver Erkrankung berechnet.

Für den Endpunkt „Zeit bis zur Schmerzprogression“ (TWP) wurde eine explorative Subgruppenanalyse, getrennt nach *Baseline*-Opioidanalgetika-Gebrauch, prospektiv geplant und durchgeführt und *post-hoc* auch für die genäherte Zielpopulation ausgewertet (siehe Abschnitt 4.3.1.3.1.2).

Zusätzlich wurden aufgrund der hohen Patientenrelevanz von TWP in der Vorbereitung des vorgelegten Dossiers alle – für den Endpunkt PFS im Protokoll vordefinierten – Subgruppenanalysen für die Gesamtpopulation und die genäherte Zielpopulation *post-hoc* re-analysiert (siehe Abschnitt 4.3.1.3.2.2). Die besondere Patientenrelevanz gerade dieses Endpunkts lässt sich damit begründen, dass für die Patienten eine möglichst langfristige Verzögerung der Schmerzprogression von großer Bedeutung ist, da Schmerzen die Lebensqualität stark beeinträchtigen und auch der Einsatz von Schmerzmitteln von entsprechenden Nebenwirkungen begleitet ist.

Im Rahmen des EU-Zulassungsverfahrens wurden *post-hoc* Subgruppenanalysen zum RET-Mutationsstatus durchgeführt. Diese wurden in Vorbereitung des vorliegenden Nutzendossiers auch für die genäherte Zielpopulation analysiert (siehe Abschnitt 4.3.1.3.2).

Für die sekundären Endpunkte „Gesamtüberleben“, „Objektive Ansprechrate“, „Krankheitskontrollrate“, „Dauer objektiven Ansprechens“, „gesundheitsbezogene Lebensqualität“ und „unerwünschte Ereignisse“ waren keine prä-spezifizierten Subgruppenanalysen geplant oder durchgeführt, ebenso wenig für die explorativen Endpunkte. Prä-spezifizierte Subgruppenanalysen basieren auf erhobenen und verfügbaren Baseline-Informationen (Baseline-Charakteristika, Biomarkern oder Krankheitsstatus). Des Weiteren ist darauf hinzuweisen, dass die Resultate von Subgruppenanalysen, insbesondere für die bereits eingeschränkte genäherte Zielpopulation, mit besonderer Vorsicht beurteilt

werden sollten. Kleine Stichproben innerhalb der Subgruppen, ein Ungleichgewicht zwischen den Subgruppen und die steigende Anzahl statistischer Tests (Multiple Analysen) erhöhen prinzipiell die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse [89].

Um nicht durch ein *post-hoc* durchgeführtes multiples Testen die Wahrscheinlichkeit zufälliger statistischer signifikanter Ergebnisse zu erhöhen, wurden keine *a posteriori* spezifizierten Subgruppenanalysen für oben genannte sekundäre Endpunkte durchgeführt und im Modul 4 dargestellt. Eine numerische Auswertung der wichtigsten Patientenmerkmale (Geschlecht, Alter, Krankheitsstadium) ist für die einzelnen Endpunkte jedoch erfolgt und im Modul 5 hinterlegt.

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 möglich[7]. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“[8], „Multiple-Treatment-Meta-Analysen“[9] oder auch „Netzwerk-Meta-Analysen“[10], sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[11].

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen[12].

[7] *Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D‘Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.*
[8] *Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.*
[9] *Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331(7521): 897-900.*
[10] *Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks ofrandomized trials. Stat Methods Med Res 2008;17(3): 279-301.*
[11] *B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*
[12] *Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009; 338: b1147.*

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

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Es wurden keine indirekten Vergleiche zwischen Vandetanib und der zweckmäßigen Vergleichstherapie durchgeführt, da die Zulassung von Vandetanib auf einer doppel-blinden, randomisierten, placebo-kontrollierten Studie beruht. Der Vergleichsarm „Placebo“ erlaubt präzise und quantifizierbare Ergebnisse und stellt eine Vergleichstherapie im Sinne von *Best Supportive Care* (BSC) dar, das heißt, der bestmöglichen unterstützenden Fürsorge.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, 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“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D4200C00058	ja	laufend	Seit 23. November 2006 laufend, voraussichtlich bis Dezember 2016	Vandetanib Placebo
D4200C00097	nein	laufend	Seit Juni 2012 laufend, geplant bis Mai 2015	Vandetanib 300 mg Vandetanib 150 mg

Geben Sie an, welchen Stand die Information in Tabelle 4-2 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Es gibt nur eine kontrollierte klinische Studie, die Vandetanib in der Indikation MTC gegenüber der zweckmäßigen Vergleichstherapie untersucht (Vgl. Anhang 4-B). Der Stand der Information in Tabelle 4-2 ist vom 07.01.2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-2 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
D4200C00097	Studie ist nicht placebokontrolliert und befindet sich noch in der Rekrutierungsphase; es sind noch keine Resultate verfügbar

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. 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 Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.1 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

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

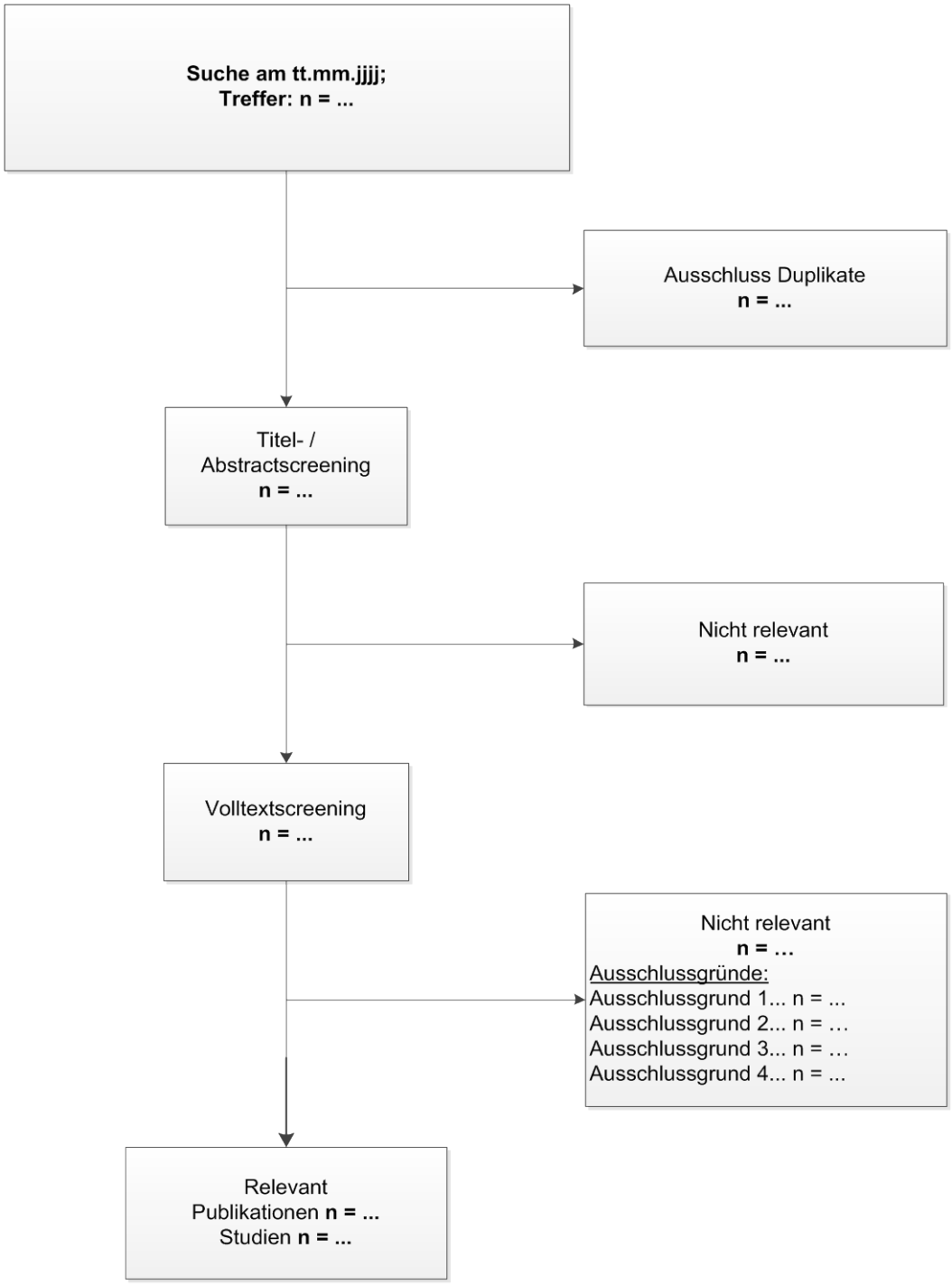


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Eine systematische bibliografische Literaturrecherche wurde am 01.10.2012 durchgeführt und am 11.01.2013 aktualisiert (siehe Anhang 4-A). Zur Identifizierung für die Nutzenbewertung von Vandetanib im Sinne der Fragestellung (siehe Abschnitt 4.2.1) relevanter Studien wurde

unter Bezugnahme auf die in Tabelle 4-1 gelisteten Kriterien die folgenden Ein- und Ausschlusskriterien festgelegt und für die Sichtung von Abstracts und Volltexten verwendet:

Tabelle 4-4: Ein-/Ausschlusskriterien für die bibliografische Literaturrecherche

E1	Patientenkollektiv der Studie umfasst erwachsene Patienten beiderlei Geschlechts mit aggressivem und symptomatischem medullärem Schilddrüsenkarzinom (MTC) in nicht resektablem lokal fortgeschrittenem oder metastasiertem Stadium	**A1***
E1a	Studie umfasst Patienten, die nach den Kriterien von G-BA und IQWiG als hinreichende Näherung des oben genannten Kollektivs dargestellt werden können (Patientenkollektiv mit progressivem und symptomatischem medullärem MTC in nicht resektablem lokal fortgeschrittenem oder metastasiertem Stadium)	
E2	Intervention: Vandetanib (Caprelsa / ZD6474) einmal täglich 300 mg	**A2**
E3	Erste-Hand-Daten aus randomisierter, kontrollierter Studie; keine Reviews	**A3**
E4	Kontrolle: Placebo (+ Best Supportive Care)	**A4**

E5	**Mindestens einer** der folgenden Endpunkte wird berichtet:
	• Mortalität:
	  o Gesamtüberleben
	• Morbidität
	  o progressionsfreies Überleben (*Progression-free Survival*, PFS)
	  o Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)
	  o Ansprechrate (*Overall Response Rate*, ORR)
	• gesundheitsbezogene Lebensqualität (*Health Related Quality of Life*, HRQoL)
	• Unerwünschte Ereignisse
		A5

***** Nichterfüllung eines Einschlusskriteriums bedeutet automatisch Ausschluss.

96 Treffer der Literaturrecherche wurden nach Eliminierung von Duplikaten nach diesen Kriterien zunächst anhand des Abstracts beurteilt, potentiell relevante Publikationen

anschließend im Volltext beschafft und evaluiert. Nachfolgendes Flussdiagramm illustriert den Verlauf der Recherche:

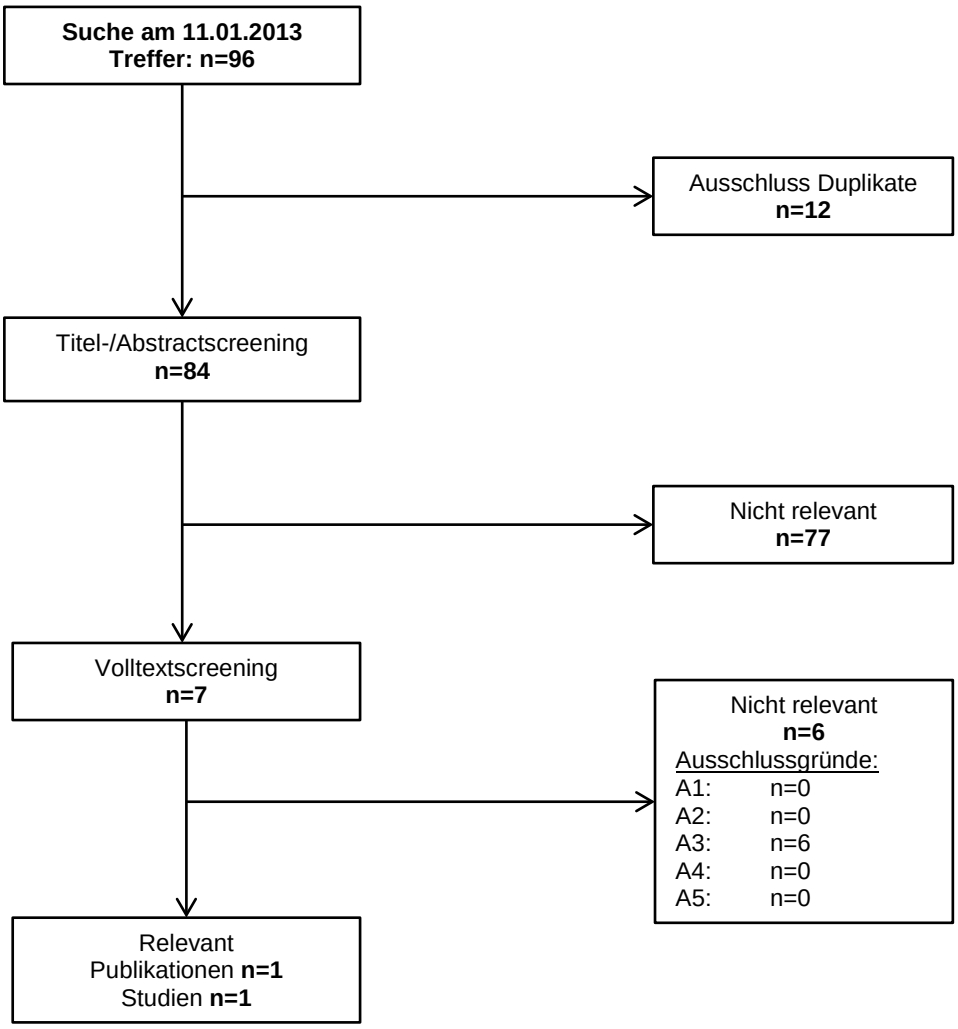


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit Vandetanib

Es konnte eine relevante Publikation zu einer randomisierten kontrollierten Studie mit Vandetanib im Anwendungsgebiet MTC identifiziert werden. Hierbei handelt es sich um die Zulassungsstudie zu Vandetanib, Studie D4200C00058. Weitere für die Beantwortung der unter 4.2.1 spezifizierten Fragestellung relevante Publikationen bzw. Studien wurden nicht gefunden.

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 4-2) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.1 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

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

Tabelle 4-5: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Identifikationsorte (Name der Studienregister und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.[b])**
D4200C00058	International Clinical Trials Registry Platform Search Portal [http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00410761] [90], clinicaltrials.gov [http://www.clinicaltrials.gov/ct2/show/NCT00410761] [91]	ja	ja

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle für die Nutzenbewertung berücksichtigten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. 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 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studienkategorie**			**verfügbare Datenquellen[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						
D4200C00058	ja	ja	nein	ja [22]	ja [91] (NCT00410761)	ja [92]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
--*						

a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.

b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

*: Die eingeschlossene RCT ist placebokontrolliert; der Placebo-Arm bildet jedoch aufgrund der zulässigen Begleitmedikation die zweckmäßige Vergleichstherapie BSC hinreichend ab.

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-F zu hinterlegen.

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

Studie	**Studiendesign** <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
D4200C00058	RCT, doppelblind, placebokontrolliert, 2:1 Randomisierung, Nach objektiv festgestellter Progression wurden die Patienten aus der randomisierten, verblindeten Studie entlassen und hatten die Option, in die offene Phase der Studie überzutreten (*Open-Label* Vandetanib)	Erwachsene Patienten mit messbarem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom. Die Teilpopulation mit symptomatischer und progressiver Erkrankung (=genäherte Zielpopulation) ist gesondert darstellbar.	Vandetanib (n=231) Placebo (n=100)	Start: 23. November 2006, noch andauernd, Behandlung bis zur objektiven Krankheitsprogres-sion, dann Option zu offener Behandlung, Nachbeobachtung für Gesamtüberleben	63 Studienzentren in 24 Ländern: Australien, Belgien, Brasilien, Dänemark, Deutschland, Frankreich, Indien, Italien, Kanada, Korea, Niederlande, Mexiko, Österreich, Polen, Portugal, Rumänien, Russland, Serbien, Spanien, Schweden, Schweiz, Tschechien, Ungarn und die USA; 11/2006 – 12/2016	Primärer Endpunkt: Progressionsfreies Überleben (PFS) Patientenrelevante sekundäre Endpunkte: • Gesamtüberleben (OS) • Zeit bis zur Schmerzprogression (TWP) • objektive Ansprechrate (ORR) • Krankheitskontrollrate (DCR) • Dauer objektiven Ansprechens (DOR) • gesundheitsbezogene Lebensqualität (HRQoL) anhand des FACT-G • unerwünschte Ereignisse (UE) Explorative Morbiditätsendpunkte: • Gewichtszunahme • Einfluss der Behandlung auf Diarrhöen

	• WHO-Performancestatus (WHO-PS) • Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline) • Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	**Vandetanib**	**Placebo**	***ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
D4200C00058	300 mg Vandetanib-Tablette einmal täglich	300 mg Vandetanib-Placebo-Tablette einmal täglich	Keine bestimmten Vorbehandlungen, keine Run-in-Phase

Tabelle 4-9: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 1)

Studie Gruppe	**N**	**Alter (Jahre)**	**Geschlecht w/m (%)**	**ethnische Zugehörigkeit (%)**				**Gewicht (kg)**	**Größe (cm)**
				S	**K**	**O**	**A**		
D4200C00058 (Gesamtpopulation)									
Vandetanib	231	50,7	42/58	0,4	94,4	3,5	1,7	70,4	170,8
Placebo	100	53,4	44/56	1,0	97,0	1,0	1,0	70,2	168,8
D4200C00058 (symptomatisch und progressiv)									
Vandetanib	126	53,1	37/63	n.d.	93,7	4,0	2,4	71,2	172,0
Placebo	60	53,9	35/65		98,3	0,0	1,7	68,9	169,9

S=Schwarz, K=Kaukasisch, O=Orientalisch, A=Andere, n.d.= nicht ermittelt

Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 2: Altersgruppen) (Alle Angaben in %)

Studie Gruppe	**Altersgruppen**			
	≥18 - <40	**≥40 - <65**	**≥65 - <75**	**≥75**
D4200C00058 (Gesamtpopulation)				
Vandetanib	21,6	57,1	18,2	3,0
Placebo	10,0	70,0	17,0	3,0
D4200C00058 (symptomatisch und progressiv)				
Vandetanib	16,7	57,1	23,8	2,4
Placebo	6,7	73,3	15,0	5,0

Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 3a: Bestes Ansprechen bei vorangegangener Chemotherapie gegen MTC) (Alle Angaben in % bezogen auf die Anzahl der Patienten im Studienarm)

Studie Gruppe	**Bestes Ansprechen bei vorangegangener Chemotherapie gegen Medulläres Schilddrüsenkarzinom (MTC)**							
	N	**Komplettes Ansprechen**	**Partielles Ansprechen**	**Stabile Erkrankung**	**Progression**	**Nicht auswertbar**	**Nicht anwendbar**	**Unbekannt**
D4200C00058 (Gesamtpopulation)								
Vandetanib	50	0,0	0,9	6,1	7,8	0,4	6,5	0,0
Placebo	18	1,0	0,0	11,0	1,0	1,0	3,0	1,0
D4200C00058 (symptomatisch und progressiv)								
Vandetanib	26	0,0	0,8	6,3	9,5	0,8	3,2	0,0
Placebo	11	1,7	0,0	10,0	0,0	0,0	5,0	1,7

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 3b: Systemische Vortherapie gegen MTC) (Alle Angaben in %)

Studie Gruppe	**Systemische Vortherapie gegen Medulläres Schilddrüsenkarzinom (MTC)**	
	Ja	**Nein**
D4200C00058 (Gesamtpopulation)		
Vandetanib	39,0	61,0
Placebo	42,0	58,0
D4200C00058 (symptomatisch und progressiv)		
Vandetanib	35,7	64,3
Placebo	48,3	51,7

Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 4: Primärtumor) (Alle Angaben in %)

Studie Gruppe	**Primärtumor**					
	T1	**T2**	**T3**	**T4a**	**T4b**	**TX**
D4200C00058 (Gesamtpopulation)						
Vandetanib	2,2	1,3	0,9	3,5	2,6	89,6
Placebo	1,0	0,0	5,0	5,0	1,0	88,0
D4200C00058 (symptomatisch und progressiv)						
Vandetanib	2,4	2,4	0,8	4,0	3,2	87,3
Placebo	0,0	0,0	6,7	6,7	1,7	85,0

Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 5: Regionäre Lymphknoten) (Alle Angaben in %)

Studie Gruppe	**Regionäre Lymphknoten**					
	N0	**N1a**	**N1b**	**N2**	**N3**	**NX**
D4200C00058 (Gesamtpopulation)						
Vandetanib	12,6	11,3	57,1	1,7	0,0	17,3
Placebo	13,0	10,0	59,0	3,0	1,0	14,0
D4200C00058 (symptomatisch und progressiv)						
Vandetanib	12,7	10,3	58,7	1,6	0,0	16,7
Placebo	15,0	6,7	56,7	3,3	0,0	18,3

Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 6: Fernmetastasen/ Stadienklassifikation) (Alle Angaben in %)

Studie Gruppe	**Fernmetastasen**			**Stadienklassifikation**			
	M0	**M1**	**MX**	**III**	**IVA**	**IVB**	**IVC**
D4200C00058 (Gesamtpopulation)							
Vandetanib	6,1	93,5	0,4	0,4	3,5	2,6	93,5
Placebo	3,0	97,0	0,0	2,0	0,0	1,0	97,0
D4200C00058 (symptomatisch und progressiv)							
Vandetanib	4,8	94,4	0,8	0,8	2,4	2,4	94,4
Placebo	1,7	98,3	0,0	1,7	0,0	0,0	98,3

Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 7: Patienten mit hereditärer Erkrankung) (Alle Angaben in %)

Studie Gruppe	**Familiengeschichte mit MTC**			**Begleitende endokrine Syndrome**				
	Ja	**Nein**	**Unbekannt**	**FMTC**	**MEN2a**	**MEN2b**	**Keins**	**Unbekannt**
D4200C00058 (Gesamtpopulation, n=33)								
Vandetanib (n=28)	42,9	42,9	14,3	14,3	50,0	25,0	7,1	3,6
Placebo (n=5)	80,0	20,0	0,0	20,0	60,0	0,0	20,0	0,0
D4200C00058 (symptomatisch und progressiv, n=13)								
Vandetanib (n= 11)	54,5	27,3	18,2	18,2	54,5	18,2	0,0	9,1
Placebo (n=2)	100	0,0	0,0	0,0	100	0,0	0,0	0,0

Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 8a: RET-Mutationsstatus positiv, sporadisch) (Alle Angaben in %) (Gesamtpopulation)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,3% (n=137); Placebo: 50,0% (n=50)) Sporadisch (Vandetanib: 45,0% (n=104); Placebo: 43,0% (n=43))**										
	Mutationsposition										
	618	**630**	**632**	**633**	**634**	**883**	**899**	**900**	**901**	**902**	**918**
D4200C00058											
Vandetanib	0,4	0,4	0,4	0,4	1,3	0,9	0,0	0,0	0,0	0,0	41,6
Placebo	0,0	0,0	0,0	0,0	2,0	0,0	2,0	2,0	2,0	2,0	39,0

Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 8b: RET-Mutationsstatus positiv, sporadisch) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,5% (n=75); Placebo: 50,0% (n=30)) Sporadisch (Vandetanib: 47,6% (n=60); Placebo: 43,3% (n=26))**										
	Mutationsposition										
	618	**630**	**632**	**633**	**634**	**883**	**899**	**900**	**901**	**902**	**918**
D4200C00058											
Vandetanib	0,8	0,8	0,0	0,0	2,4	1,6	0,0	0,0	0,0	0,0	42,1
Placebo	0,0	0,0	0,0	0,0	1,7	0,0	1,7	1,7	1,7	1,7	40,0

Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9a: RET-Mutationsstatus hereditär) (Alle Angaben in %) (Gesamtpopulation)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,3% (n=137); Placebo: 50,0% (n=50)) Hereditär (Vandetanib: 11,7% (n=27); Placebo: 5,0% (n=5))**									
	Mutationsposition									
	364	**609**	**611**	**618**	**620**	**634**	**781**	**791**	**804**	**918**
D4200C00058										
Vandetanib	0,4	0,4	0,4	0,9	0,9	4,8	0,0	0,4	0,4	3,5
Placebo	0,0	0,0	0,0	1,0	1,0	2,0	1,0	0,0	0,0	0,0

Tabelle 4-20: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9b: RET-Mutationsstatus hereditär) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,5% (n=75); Placebo: 50,0% (n=30))** **Hereditär (Vandetanib: 8,7% (n=11); Placebo: 3,3% (n=2))** ***Mutationsposition***									
	364	**609**	**611**	**618**	**620**	**634**	**781**	**791**	**804**	**918**
D4200C00058										
Vandetanib	0,0	0,8	0,8	1,6	0,8	3,2	0,0	0,0	0,0	1,6
Placebo	0,0	0,0	0,0	0,0	1,7	1,7	0,0	0,0	0,0	0,0

Tabelle 4-21: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9a: RET-Mutationsstatus unbekannt) (Alle Angaben in %) (Gesamt-population)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,3% (n=137); Placebo: 50,0% (n=50))** **Unbekannt (Vandetanib: 2,6% (n=6); Placebo: 2,0% (n=2))** ***Mutationsposition***									
	364	**609**	**611**	**618**	**620**	**634**	**781**	**791**	**804**	**918**
D4200C00058										
Vandetanib	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,0	2,2
Placebo	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,0

Tabelle 4-22: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9b: RET-Mutationsstatus unbekannt) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie Gruppe	**RET Mutationsstatus positiv (Vandetanib: 59,5% (n=75); Placebo: 50,0% (n=30))** **Unbekannt (Vandetanib: 3,2% (n=4); Placebo: 3,3% (n=2))** ***Mutationsposition***									
	364	**609**	**611**	**618**	**620**	**634**	**781**	**791**	**804**	**918**
D4200C00058										
Vandetanib	0,0	0,0	0,0	0,0	0,0	0,8	0,0	0,0	0,0	2,4
Placebo	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,3

Tabelle 4-23: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10a: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %) (Gesamtpopulation)

Studie Gruppe	**RET Mutationsstatus negativ (Vandetanib: 0,9% (n=2) , Placebo: 6,0% (n=6))**	**RET Mutationsstatus unbekannt (Vandetanib: 39,8% (n=92) , Placebo: 44,0% (n=44))**		
	Sporadisch	**Sporadisch**	**Hereditär**	**Unbekannt**
D4200C00058				
Vandetanib	0,9	38,5	0,4	0,9
Placebo	6,0	43,0	0,0	1,0

Tabelle 4-24: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10b: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie Gruppe	**RET Mutationsstatus negativ (Vandetanib: 0,8% (n=1) , Placebo: 10,0% (n=6))**	**RET Mutationsstatus unbekannt (Vandetanib: 39,7% (n=50) , Placebo: 40,0% (n=24))**		
	Sporadisch	**Sporadisch**	**Hereditär**	**Unbekannt**
D4200C00058				
Vandetanib	0,8	38,9	0,0	0,8
Placebo	10,0	40,0	0,0	0,0

Tabelle 4-25: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 11: WHO-PS bei Studieneintritt) (Alle Angaben in %)

Studie Gruppe	WHO *Performance Status* (WHO-PS)				
	(0) Normale Aktivität	**(1) Eingeschränkte Aktivität**	**(2) ≤50% der Zeit bettlägerig**	**(3) >50% der Zeit bettlägerig**	**(4) 100% bettlägerig**
D4200C00058 (Gesamtpopulation)					
Vandetanib	67,0	29,0	4,0	0,0	0,0
Placebo	58,0	38,0	4,0	0,0	0,0
D4200C00058 (symptomatisch und progressiv)					
Vandetanib	59,0	36,0	6,0	0,0	0,0
Placebo	55,0	42,0	3,0	0,0	0,0

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

Die Studie 58 ist die einzige identifizierte randomisierte Studie, die gemäß den Selektionskriterien für die Nutzenbewertung von Vandetanib zur Behandlung des medullären Schilddrüsenkarzinoms bei erwachsenen Patienten mit aggressivem und symptomatischem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem MTC relevant ist. In dieser multizentrischen, multinationalen, placebo-kontrollierten, doppelblinden Phase-III-Studie wurden erwachsene Patienten mit nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom eingeschlossen. Von 437 Patienten, welche eine Einverständniserklärung unterschrieben hatten, wurden 331 Patienten im Verhältnis 2:1 (Vandetanib:Placebo) auf die Studienarme randomisiert (siehe 4.3.1.2.2), davon ließen sich *post-hoc* 186 Patienten der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung zuordnen. Die Anzahl der Patienten mit progressiver und symptomatischer Erkrankung wurde auch in der IQWiG-Nutzenbewertung des ursprünglichen Dossiers zu Vandetanib [3] sowie in der mündlichen Anhörung thematisiert. Der EPAR wies in diesem Punkt eine numerische Unkorrektheit auf, wurde aber inzwischen von der EMA an die korrekte Patientenzahl von 186 Patienten angepasst [2]. 231 (126 in der genäherten Zielpopulation) Patienten wurden der Behandlung mit Vandetanib zugeteilt und 100 (60) Patienten der Placebo-Gruppe. Es erhielten 231 (126) Patienten die Behandlung mit Vandetanib und 99 (59) Patienten Placebo, nachdem ein Patient in der Placebo-Gruppe vor Therapiebeginn verstorben war. Beide Behandlungen wurden bis zur Krankheitsprogression fortgesetzt, sofern die Patienten nicht ihr Einverständnis zurückzogen oder aus anderen Gründen die Studienbehandlung nicht fortsetzen konnten. Gründe für den Studienabbruch waren Tod, Sicherheitsgründe, massive Einnahmeregelverletzungen, *Lost-to-follow-up*, Beendigung der Teilnahme auf Wunsch des Patienten oder andere Gründe.

Nach objektiv festgestellter Progression wurden die Patienten aus der randomisierten, verblindeten Behandlungsphase entlassen und hatten die Option, in die offene Phase der Studie zu wechseln und mit Vandetanib behandelt zu werden. Dies nahmen in der Gesamtpopulation 58% der Patienten aus dem Placebo-Arm sowie 19% der Patienten im Verum-Arm wahr. In der genäherten Zielpopulation erhielten 63,3% der Patienten aus der Placebo-Gruppe und 20,6% der Patienten aus der Vandetanib-Gruppe im Anschluss an die verblindete Behandlungsphase *Open-Label* Vandetanib. Alle Patienten werden für das Gesamtüberleben nachbeobachtet.

Hinsichtlich Alter, Geschlecht, Größe und Rasse sowie RET-Mutationsstatus gab es keine größeren Unterschiede zwischen der Vandetanib- und der Placebo-Gruppe. Mit 53,4 (53,9) Jahren in der Placebo-Gruppe und 50,7 (53,1) Jahren in der Vandetanib-Gruppe liegt das mittlere Alter der Patienten in der Placebo-Gruppe etwas über dem der Patienten in der Vandetanib-Gruppe. Betrachtet man die Kategorien bezüglich des Alters, siehe Tabelle 4-10, besteht trotz Randomisierung ein Unterschied zwischen den beiden Behandlungsgruppen. 70% (73,3%) der Placebo-Patienten entstammen der Altersklasse (≥40 – <65), deren Anteil im Behandlungsarm mit Vandetanib liegt hingegen bei nur 57,1% (gleicher Anteil in der Teilpopulation). In der Vandetanib-Gruppe sind 21,6% in der jüngsten Patientengruppe (≥18 - <40), hingegen nur 10% der Patienten in der vergleichbaren Placebo-Gruppe. Dieses leichte Ungleichverhältnis findet sich auch innerhalb der genäherten Zielpopulation mit

symptomatischer und progressiver Erkrankung (16,7% gegenüber 6,7%). Der Anteil älterer Patienten (≥65 – <75) ist in der Vandetanib-Gruppe in der genäherten Zielpopulation ebenfalls deutlich höher als in der Placebo-Gruppe (23,8% gegenüber 15%; in der Gesamtpopulation 18,2 gegenüber 17%). Der Anteil der kaukasischen Patienten mit 97% (98,3%) in der Placebo-Gruppe war etwas größer als in der Vandetanib-Gruppe mit 94,4% (93,7).

Hinsichtlich des besten Ansprechens bei vorheriger MTC Chemotherapie liegen für „Stabile Erkrankung“ und „Progression“ deutliche Unterschiede zwischen den Behandlungsarmen vor: Stabile Erkrankung Vandetanib vs. Placebo: 6,1% vs. 11,0% bzw. 6,3% vs. 10,0% in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung, Progression Vandetanib vs. Placebo 7,8% vs. 1,0% bzw. 9,5% vs. 0,0%. Auch die Aufteilung der Patienten auf die verschiedenen Tumorstadien unterscheidet sich zwischen den Behandlungsgruppen.

Um den möglichen Effekt von Unbalanciertheiten in *Baseline*-Charakteristika zwischen den Behandlungsgruppen auf den beobachteten Behandlungseffekt zu beurteilen, wurden entsprechende adjustierte Analysen durchgeführt. Es konnte in beiden betrachteten Patientenpopulationen kein möglicher Einfluss festgestellt werden.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-26: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver-zerrungspotenzial beeinflussende Punkte	Verzerrungs-potenzial auf Studienebene
			Patient	Behandler			
D4200C00058	ja	ja	ja	ja	nein	ja*	niedrig

* Möglichkeit der *Open-Label* Behandlung mit Vandetanib nach Progression; Abweichung von der geplanten Fallzahlplanung: Anstelle von 232 Patienten wurden 331 Patienten randomisiert

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die Studie 58 sind der Randomisierungsprozess und die verdeckte Zuteilung adäquat beschrieben und durchgeführt. Auch die Verblindungsmethodik wurde adäquat beschrieben bzw. durchgeführt.

Die Randomisierung der geeigneten Patienten erfolgte im Verhältnis 2:1 durch ein Randomisierungsschema, welches computerbasiert generiert wurde. Die Randomisierung erfolgte in 3-er Blöcken und wurde nach Zentrum stratifiziert.

Die Studienmedikation wurde mit einer eindeutigen MedID versehen, welche an das Randomisierungsschema gekoppelt war. Die gleichartige Verpackung und das Aussehen der Studienmedikation garantierte die Verblindung, sowohl der Patienten als auch der Behandler. Patienten und Behandler blieben bis zum Eintreten der Progression der Erkrankung verblindet. Anschließend erhielten die Patienten die Option, entweder mit Vandetanib 300 mg (oder mit einer reduzierten Dosis, wenn nötig) die Behandlung fortzuführen oder die Behandlung zu beenden und im *Follow-up* weiterhin hinsichtlich ihres Überlebens beobachtet zu werden.

Da Patienten gemäß *Good Clinical Practice* vor Beginn der Studie über zu erwartende Nebenwirkungen aufgeklärt werden müssen, ist die Gefahr einer Entblindung der Patienten und Behandler bei Auftreten charakteristischer Nebenwirkungen, insbesondere bei einem Vergleich gegen Placebo, gegeben. In der betrachteten Indikation ist eine Differenzierung zwischen krankheitsbedingten Beschwerden und therapieassoziierten Ereignissen aufgrund von Überlagerungen der Profile für den Patienten jedoch schwierig. Daher erscheint eine Entblindung aufgrund typischer Nebenwirkungen und damit eine mögliche Verzerrung durch die Kenntnis der Behandlungsgruppe unwahrscheinlich.

Die primären Analysen der Wirksamkeits-Endpunkte in der Studie basieren auf der ITT-Population, welche alle randomisierten Patienten einschließt. Die Auswertung der Wirksamkeits-Endpunkte erfolgte in einem zentralen, verblindeten Review-Verfahren nach RECIST. Somit liegt insbesondere für den primären Wirksamkeitsendpunkt Progressionsfreies Überleben (PFS), ebenso wie für die RECIST-basierten Endpunkte objektive Ansprechrate (ORR), Dauer des objektiven Ansprechens (DOR) und Krankheitskontrollrate (DCR), in Bezug auf eine Entblindung gegenüber der Studienmedikation ein niedriges Verzerrungspotential vor.

Gemäß Studienprotokoll hatten alle Patienten in beiden Armen die Möglichkeit, im Anschluss an die verblindete Behandlungsphase eine *Open-Label* Behandlung mit Vandetanib aufzunehmen, da zum Zeitpunkt der Studiendurchführung keine zugelassene Anschlusstherapie existierte. Die Umstellung auf *Open-Label* Vandetanib konnte erfolgen, sobald der behandelnde Studienarzt eine Progression festgestellt hatte. Eine Bestätigung der Progression durch den zentralen RECIST-Review war für den Therapiewechsel nicht erforderlich. Die Befundung im lokalen und zentralen Review stimmte in einigen Fällen nicht überein. Dies führte dazu, dass 13 Placebo-Patienten, die lokal als progredient befundet worden waren, Vandetanib erhielten, obwohl im zentralen Review kein Tumorprogress bestätigt wurde. In der genäherten Zielpopulation war dies für 10 Patienten der Fall. Dies legt eine mögliche Verzerrung der Studienergebnisse für die Endpunkte ORR und DCR zuungunsten von Vandetanib nahe, erhöht also die Wahrscheinlichkeit einer Unterschätzung des Effekts von Vandetanib. Da die Anzahl der betroffenen Patienten klein ist (13 bzw. 10

Patienten) wird der verzerrende Aspekt jedoch für diese Endpunkte als eher gering eingeschätzt.

Eine höhere Wahrscheinlichkeit für eine Verzerrung des Studienergebnisses muss aufgrund der *Open-Label* Option für den Endpunkt Gesamtüberleben angenommen werden. Die *Open-Label* Therapie mit Vandetanib wurde in der Gesamtpopulation der Studie im Placebo-Arm von 58% der Patienten wahrgenommen, im Verum-Arm von 19%. In der genäherten Zielpopulation erhielten 63,3% der Placebo-Patienten in der offenen Behandlungsphase Vandetanib, sowie 20,6% der Vandetanib-Patienten. Die Anschlussbehandlung mit *Open-Label* Vandetanib von ca. 60% der Placebo-Patienten, denen alternativ nur noch *Best Supportive Care* als zugelassene Behandlungsoption angeboten werden konnte, legt eine Verzerrung des Studienergebnisses zuungunsten von Vandetanib nahe, erhöht also die Wahrscheinlichkeit einer Unterschätzung des Effekts von Vandetanib.

Die Anzahl der randomisierten Patienten lag mit 331 Patienten über den in der Fallzahlplanung ursprünglich geplanten 232 Patienten. 231 Patienten wurden dem Behandlungsarm und 100 Patienten dem Placeboarm zugeteilt. Trotz regulär abgelaufener Randomisierung führten die bei Rekrutierungsende nicht aufgefüllten Randomisierungsblöcke zu einem Gesamtverhältnis von 2,31:1 anstelle von 2:1. Eine Erhöhung der Fallzahl bedeutet, dass die Anzahl der Ereignisse für den primären Endpunkt mit einer kürzeren *Follow-up* Periode erreicht werden könnte als für die ursprünglich 232 geplanten Patienten. Die Überrekrutierung lag darin begründet, dass vor dem 19. September 2007 297 Patienten in 310 Tagen (d. h. durchschnittlich 29 Patienten pro Monat) gescreent wurden. Zwischen dem 19. September 2007 und dem 19. Oktober 2007 wurden, nach erfolgter Benachrichtigung der Prüfzentren über die baldige Beendigung der Rekrutierungsphase, von diesen weitere 139 Patienten gescreent. Dieser rapide Anstieg war nicht vorhersehbar und endete letztlich in einer Überrekrutierung von 99 Patienten. Eine Sensitivitätsanalyse zeigte jedoch, dass auch mit den ersten 232 Patienten der primäre Endpunkt erreicht worden wäre. Eine entsprechend beantwortete Anfrage der EMA im Rahmen des Zulassungsprozesses zu diesem Sachverhalt wurde auch offiziell akzeptiert [93].

Sowohl aufgrund der zusätzlich durchgeführten Sensitivitätsanalyse und der angeführten Erläuterungen als auch wegen der sonstigen Beurteilung der einzelnen Aspekte wird das Verzerrungspotential insgesamt als „niedrig“ bewertet. Die Überrekrutierung wird nicht als verzerrender Aspekt betrachtet. Im Rahmen der *Post-hoc*-Auswertung der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung wurde auch das Verzerrungspotential, sowohl auf Studien- als auch auf Endpunktebene, neu beurteilt. Es ergab sich jedoch kein Hinweis darauf, dass die für die Gesamtpopulation getroffene Einschätzung bei Betrachtung der genäherten Zielpopulation geändert werden muss.

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 4-27: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	**PFS**	**TWP**	**ORR**	**DCR**	**DOR**	**OS**	**FACT-G**	**CTN/CEA***	**Unerwünschte Ereignisse**
D4200C00058	ja	ja	ja	ja	ja	ja	ja	ja	ja

*Surrogat

Tabelle 4-28: Matrix der explorativen Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	**Gewichtszunahme**	**Einfluss der Behandlung auf Diarrhöen**	**WHO-PS**	**TWP nach Opioid-Gebrauch zu Studienbeginn**	**Reduktion des Opioid-Gebrauchs gegenüber dem Anfangswert**
D4200C00058	ja	ja	ja	ja	ja

4.3.1.3.1 Endpunkte – RCT

In diesem Kapitel sind ergänzend zu den Ergebnissen für die gesamte Studienpopulation alle Ergebnisse auch für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung als Näherung für die Zielpopulation „aggressiv und symptomatisch“ dargestellt. Die im Studienprotokoll der für die Nutzenbewertung herangezogenen RCT geplanten Analysen waren allerdings nicht für diese genäherte Zielpopulation geplant; durch die Beschränkung der Analysen auf die genäherte Zielpopulation sinkt die statistische *Power*. Dies gilt insbesondere z. B. für die Durchführung von Subgruppenanalysen zur Identifikation von potentiellen Effektmodifikatoren (vgl. Abschnitt 4.3.1.3.2). Um zu überprüfen, ob die Ergebnisse der gesamten Studienpopulation im Prinzip auch auf die genäherte Zielpopulation übertragbar sind, wurden daher im Folgenden die Ergebnisse für die beiden Populationen gegenübergestellt, und (wenn statistisch auswertbar) die Ergebnisse von Interaktionstests für „progressiv und symptomatisch“ vs. „nicht progressiv und symptomatisch“ angegeben. Im Rahmen dieser Dossiererstellung wurden einige statistische Werte zusätzlich berechnet, die entsprechend weder im Studienbericht der Studie 58, noch in der *post-hoc* Auswertung für die genäherte Zielpopulation angegeben sind.

Tabelle 4-29: Interaktionstests für die Populationen „progressiv und symptomatisch“ und „nicht progressiv und symptomatisch“

Wirksamkeits-Endpunkte				
	Gesamte Studienpopulation (n=331)	**Teilpopulationen**		
		Progressiv und symptomatisch (n=186)	**Nicht progressiv und symptomatisch (n=139)**	**p-Wert Interaktion**
Progressionsfreies Überleben (PFS)	HR=0,46 95%- KI: [0,31;0,69]	HR=0,47 95%- KI: [0,29;0,77]	HR=0,45 95%- KI: [0,22;0,93]	0,9617
Zeit bis zur Schmerzprogression (TWP)	HR=0,61 95%- KI: [0,43;0,87]	HR=0,62 95%- KI: [0,39; 0,99]	HR=0,57 95%- KI: [0,33;0,99]	0,8314
Objektive Ansprechrate (ORR)	OR=5,48 95%- KI: [2,99;10,79]	OR=4,00 95%- KI: [1,93;9,01]	OR=10,15 95%- KI: [3,37;44,03]	0,1910
Krankheitskontrollrate (DCR)	OR=2,64 95%- KI: [1,48;4,69]	OR=2,46 95%- KI: [1,19;5,09]	OR=2,82 95%- KI: [1,03;7,68]	0,8238
Biochemisches Ansprechen (CTN)	OR=72,86 95%- KI: [26,22;303,2]	OR=60,12 95%- KI: [17,56;378,0]	OR=87,54 95%- KI: [17,68;1589]	0,7646
Biochemisches Ansprechen (CEA)	OR=52,03 95%- KI: [15,95;320,3]	OR=NC 95%- KI: NA	OR=24,28 95%- KI: [6,89;154,5]	NC
Sicherheits-Endpunkte				
	Gesamte Studienpopulation (n=330)	**Teilpopulationen**		
		Progressiv und symptomatisch (n=185)	**Nicht progressiv und symptomatisch (n=139)**	**p-Wert Interaktion**
Mindestens ein schweres, unerwünschtes Ereignis	OR=2,94 95%- KI: [1,54;5,60]	OR=2,28 95%- KI: [1,05;4,96]	OR=7,61 95%- KI: [1,72;33,63]	0,1445*
Studienabbruch wegen unerwünschten Ereignissen	OR=4,41 95%- KI: [1,31;14,80]	OR=7,84 95%- KI: [1,01;60,82]	OR=2,66 95%- KI: [0,57;12,38]	0,3949*
Mindestens ein unerwünschtes Ereignis	OR=23,00 95%- KI: [2,87;184,16]	OR=15,67 95%- KI: [0,80;308,48]	OR=18,75 95%- KI: [2,18;161,63]	0,5430*

*Breslow-Day-Test auf Homogenität der rohen *Odds Ratios*, *post-hoc* berechnet
NC: nicht berechnet (kein Ansprechen in der Placebogruppe progressiv und symptomatisch)
NA: nicht anwendbar
6 Patienten wurden auf Grund fehlender Werte vor der Randomisierung aus der Berechnung der Teilpopulationen ausgeschlossen.

Es fanden sich keine Hinweise auf qualitative Interaktionen. Es fand sich ein Hinweis auf eine quantitative Interaktion bezüglich der Endpunkte „Objektive Ansprechrate“ und „Mindestens ein schweres, unerwünschtes Ereignis“, deren p-Werte nur sehr knapp unter dem Signifikanzlevel 0,2 liegen und sich in den übrigen Wirksamkeits- oder Sicherheitsendpunkten nicht abbildeten. Insbesondere gab es keine Hinweise auf Interaktion bei den Endpunkten „progressionsfreies Überleben“ und „Zeit bis zur Schmerzprogression“.

Zusammenfassend kann daher von einer Übertragbarkeit der in der gesamten Studienpopulation beobachteten Ergebnisse auf die Teilpopulationen inklusive der nach Auffassung des IQWiG bewertungsrelevanten genäherten Zielpopulation „progressiv und symptomatisch“ ausgegangen werden.

4.3.1.3.1.1 Progressionsfreies Überleben (PFS)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-30: Operationalisierung von Progressionsfreies Überleben (PFS)

Studie	Operationalisierung
D4200C00058	Zeit bis zum Auftreten der ersten objektiven Krankheitsprogression nach RECIST modifiziert für MTC (zentraler Review) oder Tod (ungeachtet der zugrundeliegenden Ursache), relativ zum Zeitpunkt der Randomisierung

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-31: Bewertung des Verzerrungspotenzials für Progressionsfreies Überleben (PFS) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Progressionsfreies Überleben ist operationalisiert als Zeit bis zum Auftreten der ersten objektiven Krankheitsprogression nach modifiziertem RECIST v1.0 [22, 23](zentraler, unabhängiger Review) oder Tod (ungeachtet der zugrundeliegenden Ursache) relativ zum Zeitpunkt der Randomisierung. Patienten, die zum Zeitpunkt des Daten *Cut-off* für die statistische Analyse noch lebten oder progressionsfrei waren, wurden zum Zeitpunkt ihrer letzten objektiven Tumormessung (RECIST) zensiert.

Die primäre Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen. Der Endpunkt „Tod“ ist stets objektiv zu bewerten.

Für die primäre Analyse des Progressionsfreien Überlebens wurde auf Basis der ITT-Population ein *Log-Rank* Test durchgeführt. Unterschiede zwischen den Gruppen wurden mithilfe des *Hazard Ratio* und des dazugehörigen Intervallschätzers bewertet. Punktschätzer für das mediane progressionsfreie Überleben in jeder Behandlungsgruppe wurden errechnet und PFS wurde mittels einer Kaplan-Meier-Kurve grafisch dargestellt. Zusätzlich wurden die

Anteile der progressionsfreien Patienten nach 6, 12 und 24 Monaten nach Behandlungsbeginn in den Behandlungsgruppen gegenübergestellt. Desweiteren erfolgte eine Zusammenfassung des Progressionsstatus (Anzahl (%) der Patienten mit Progression, Art der Progression) zum Zeitpunkt der Analyse.

Unterstützend wurde eine Vielzahl an Sensitivitätsanalysen zur Überprüfung der Robustheit der Ergebnisse durchgeführt. Diese Sensitivitätsanalysen adressierten sowohl methodische Variationen als auch Variationen der Definition des PFS. So wurden unter anderem adjustierte Schätzer für das *Hazard Ratio* berechnet, um den möglichen Einfluss von Unbalanciertheiten bezüglich *Baseline*-Charakteristika zwischen den Behandlungsgruppen auf den beobachteten Behandlungseffekt zu beurteilen. Das adjustierte Cox-PH-Modell enthielt folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status (hereditär, sporadisch, unbekannt).

Für den beschriebenen Endpunkt erfolgte in der Studie 58 eine Analyse der ITT-Population sowohl innerhalb der Gesamtpopulation der Studie als auch innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung mittels adäquater statistischer Verfahren. Alle Ergebnisse wurden ausführlich dargestellt. Hinweise auf eine ergebnisgesteuerte Berichtserstattung wurden nicht gefunden. Eine mögliche Entblindung aufgrund des Nebenwirkungsprofils wird ausgeschlossen.

Unter Einbezug aller oben angeführten Aspekte, sowie auch aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips wurde das endpunktbezogene Verzerrungspotential als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

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.

Tabelle 4-32: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	73 (31,6)	100	51 (51,0)	0,46** [0,31;0,69]	0,0001*
					0,46*** [0,32;0,68]	0,0001

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse: Cox-PH-Modell mit Behandlung als Faktor und den Kovariaten RET Mutationsstatus, Zeitraum bis zur Verdopplung des Calcitoninwertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär, sporadisch, unbekannt).

Tabelle 4-33: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	46 (36,5)	60	35 (58,3)	0,47** [0,29;0,77]	0,0024*
					0,47*** [0,29;0,76]	0,0027

Anmerkung: Für 6 Patienten lagen keine Daten bezüglich der letzten Progression vor Randomisierung vor. „Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.
* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse: Cox-PH-Modell mit Behandlung als Faktor und den Kovariaten RET Mutationsstatus, Zeitraum bis zur Verdopplung des Calcitoninwertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär, sporadisch, unbekannt).

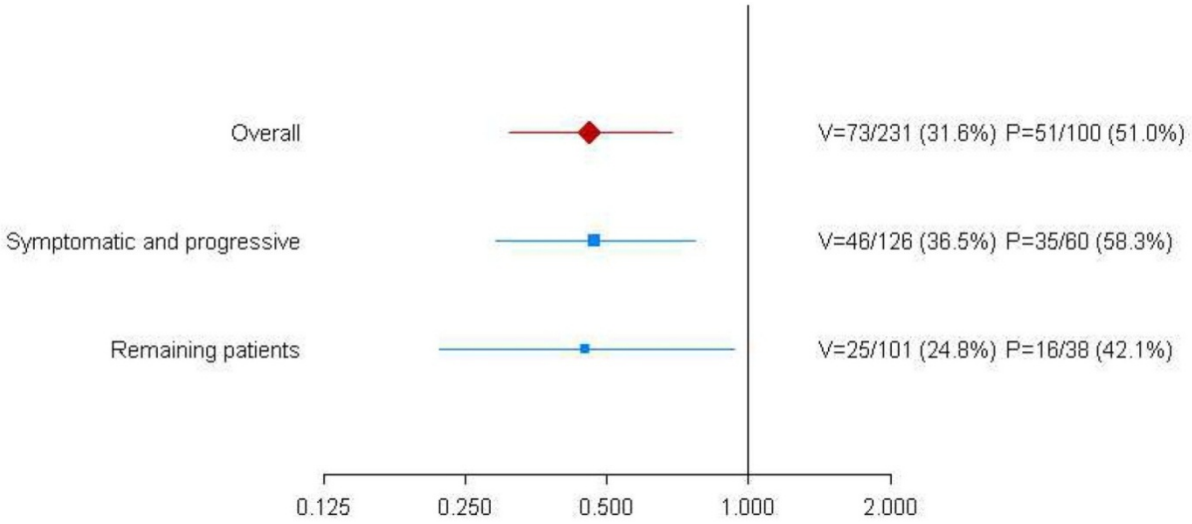


Abbildung 3: Forest Plot für Progressionsfreies Überleben (PFS) für Gesamtpopulation, genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung und Restpopulation

Tabelle 4-34: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib (N=231)** **Patienten (%) ohne Progression**	**Placebo (N=100)** **Patienten (%) ohne Progression**	**Zeit** **[Monate]**
D4200C00058	91,4	74,5	nach Monat 6
	83,3	62,7	nach Monat 12
	58,8	41,7	nach Monat 24

Patienten ohne Progression in Prozent nach Monat 6, 12 und 24 (Kaplan-Meier-Schätzer)

Tabelle 4-35: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib (N=126)** **Patienten (%) ohne Progression**	**Placebo (N=60)** **Patienten (%) ohne Progression**	**Zeit** **[Monate]**
D4200C00058	89,4	72,9	nach Monat 6
	80,8	57,1	nach Monat 12
	51,7	36,0	nach Monat 24

Patienten ohne Progression in Prozent nach Monat 6, 12 und 24 (Kaplan-Meier-Schätzer). „Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.

Tabelle 4-36: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode) (Gesamtpopulation)

Studie	**Vandetanib (N=231)**		**Placebo (N=100)**	
	Anzahl Ereignisse	**Median PFS (Monate)**	**Anzahl Ereignisse**	**Median PFS (Monate)**
D4200C00058	73	NC	51	19,3

NC = nicht schätzbar, wegen der geringen Anzahl an PFS-Ereignissen in der Vandetanib-Gruppe wurde das mediane PFS in der Vandetanib-Gruppe nicht erreicht; eine Extrapolation unter Verwendung eines Weibull-Modells ergab einen Schätzwert für das mediane PFS von 30,5 Monaten.

Tabelle 4-37: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib (N=126)**		**Placebo (N=60)**	
	Anzahl Ereignisse	**Median PFS (Monate)**	**Anzahl Ereignisse**	**Median PFS (Monate)**
D4200C00058	46	NC	35	16

„Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptome, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.
NC = nicht schätzbar, wegen der geringen Anzahl an PFS-Ereignissen in der Vandetanib-Gruppe wurde das mediane PFS in der Vandetanib-Gruppe nicht erreicht; eine Extrapolation unter Verwendung eines Weibull-Modells ergab einen Schätzwert für das mediane PFS von 28 Monaten (siehe auch Abbildung 6).

Tabelle 4-38: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel: Progressionsstatus zum Zeitpunkt des Daten *Cut-off* (Gesamtpopulation)

Studie			**Anzahl Patienten (%)**		
	Progressionsstatus	**Ereignistyp**	**Vandetanib (N=231)**	**Placebo (N=100)**	**Total (N=331)**
D4200C00058	Progression	Gesamt	73 (31,6)	51 (51,0)	124 (37,5)
		RECIST Progression	64 (27,7)	46 (46,0)	110 (33,2)
		Tod*	9 (3,9)	5 (5,0)	14 (4,2)
	Keine Progression	Gesamt	158 (68,4)	49 (49,0)	207 (62,5)
		Tod**	5 (2,2)	3 (3,0)	8 (2,4)
		Lebend	137 (59,3)	41 (41,0)	178 (53,8)
		Vor Progression ausgeschieden	16 (6,9)	5 (5,0)	21 (6,3)
		Sicherheitsgründe	1 (0,4)	0 (0,0)	1 (0,3)

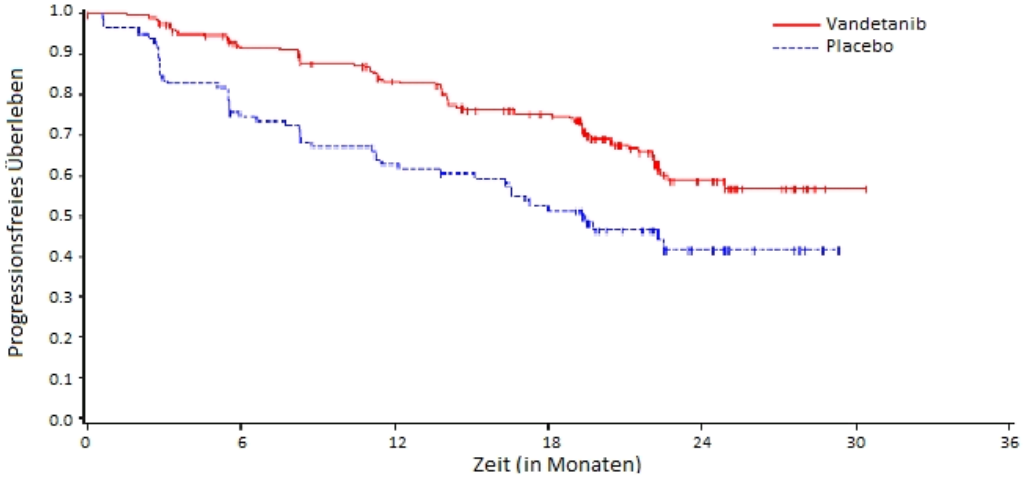
Studie			Anzahl Patienten (%)		
	Progressionsstatus	**Ereignistyp**	**Vandetanib (N=231)**	**Placebo (N=100)**	**Total (N=331)**
		Schwerwiegende Nichtbefolgung des Protokolls	1 (0,4)	1 (1,0)	2 (0,6)
		Beendigung auf Wunsch des Probanden	14 (6,1)	4 (4,0)	18 (5,4)

* Tod, wobei der Tod innerhalb 3 Monate nach der letzten evaluierbaren RECIST Bewertung auftritt
** Tod, wobei der Tod später als 3 Monate nach der letzten evaluierbaren RECIST Bewertung auftritt

Tabelle 4-39: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel: Progressionsstatus zum Zeitpunkt des Daten *Cut-off* (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

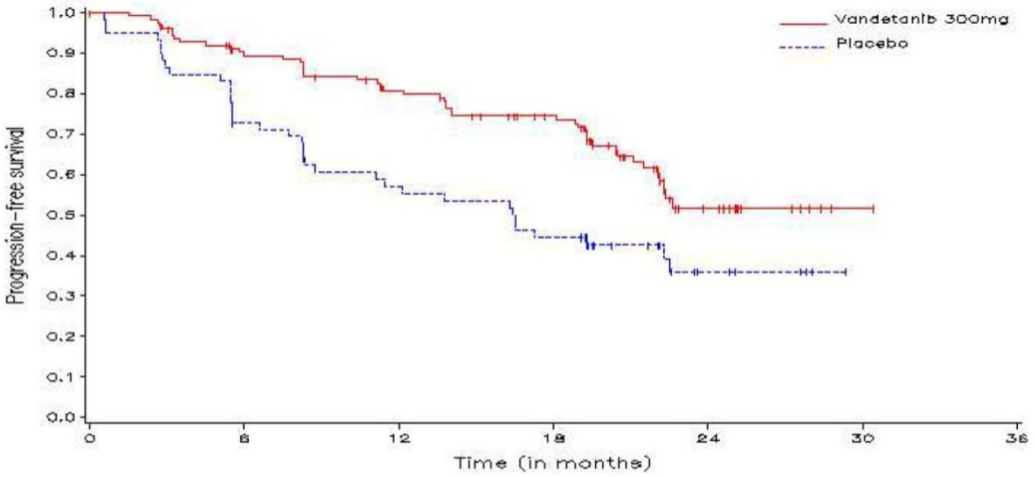
Studie			Anzahl Patienten (%)		
	Progressionsstatus	**Ereignistyp**	**Vandetanib (N=126)**	**Placebo (N=60)**	**Total (N=186)**
D4200C00058	Progression	Gesamt	46 (36,5)	35 (58,3)	81 (43,5)
		RECIST Progression	40 (31,7)	31 (51,7)	71 (38,2)
		Tod*	6 (4,8)	4 (6,7)	10 (5,4)
	Keine Progression	Gesamt	80 (63,5)	25 (41,7)	105 (56,5)
		Tod**	2 (1,6)	1 (1,7)	3 (1,6)
		Lebend	72 (57,1)	21 (35,0)	93 (50,0)
		Vor Progression ausgeschieden	6 (4,8)	3 (5,0)	9 (4,8)
		Sicherheitsgründe	1 (0,8)	0 (0,0)	1 (0,5)
		Schwerwiegende Nichtbefolgung des Protokolls	0 (0,0)	1 (1,7)	1 (0,5)
		Beendigung auf Wunsch des Probanden	5 (4,0)	2 (3,3)	7 (3,8)

Anmerkung: Für 6 Patienten lagen keine Daten bezüglich der letzten Progression vor Randomisierung vor. „Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.
* Tod, wobei der Tod innerhalb 3 Monate nach der letzten evaluierbaren RECIST Bewertung auftritt
** Tod, wobei der Tod später als 3 Monate nach der letzten evaluierbaren RECIST Bewertung auftritt



n in Monat	0	6	12	18	24	30	36
Vandetanib 300 mg	231	196	169	140	40	1	0
Placebo	100	71	57	45	13	0	0

Abbildung 4: Kaplan-Meier-Kurve für Progressionsfreies Überleben (Gesamtpopulation → ITT-Population)



n at months	0	6	12	18	24	30	36
Vandetanib 300mg	126	105	91	75	19	1	0
Placebo	60	42	32	25	8	0	0

Abbildung 5: Kaplan-Meier-Kurve für Progressionsfreies Überleben (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)

„Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptome, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.

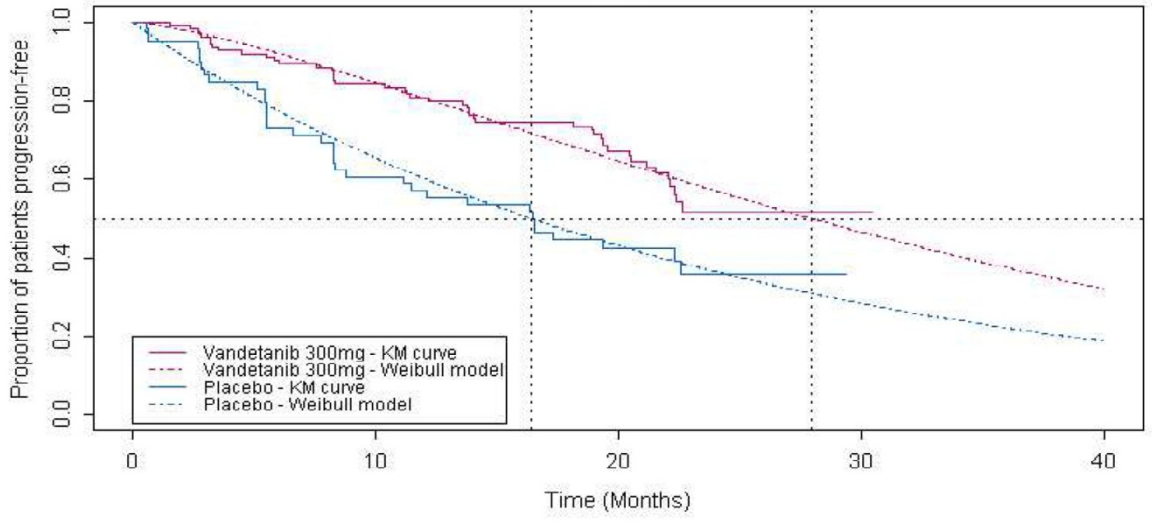


Abbildung 6: Grafische Darstellung des Weibull-Modells, angepasst an PFS mittels Kaplan-Meier-Kurve (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)

„Symptomatisch" zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv" zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.

Bezüglich des PFS (Tabelle 4-32) zeigt sich in der Gesamtpopulation ein statistisch signifikanter Unterschied zugunsten der Vandetanib-Gruppe. Mit einer *Hazard Ratio* von 0,46 (95%-KI: [0,31;0,69], p-Wert: 0,0001) liegt in der Vandetanib-Gruppe eine Reduktion des Risikos von 54% gegenüber der Placebo-Gruppe vor. Das mediane PFS wurde im Beobachtungszeitraum nur in der Placebo-Gruppe erreicht. In der Vandetanib-Gruppe wurde das mediane PFS aufgrund der geringen Anzahl an Ereignissen nicht erreicht und durch Anpassung eines Weibull-Modells geschätzt. Das geschätzte mediane PFS für den Vandetanib-Arm liegt bei 30,5 Monaten, was, verglichen mit 19,3 Monaten (Kaplan-Meier-Schätzer) in der Placebo-Gruppe, einer Verlängerung des medianen PFS um ca. 11 Monate entspricht (Tabelle 4-36). Die Ergebnisse der *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung unterscheiden sich von den für die Gesamtpopulation genannten nur marginal. Bezüglich des PFS zeigt sich in der genäherten Zielpopulation ein statistisch signifikanter Unterschied zugunsten der Vandetanib-Gruppe. Mit einer *Hazard Ratio* von 0,47 (95%-KI: [0,29;0,77], p-Wert: 0,0024), entsprechend einer Reduktion des Risikos gegenüber der Placebo-Gruppe um 53% (Tabelle 4-33). In der Vandetanib-Gruppe der genäherten Zielpopulation wurde das mediane PFS bis zur primären Auswertung der Studie laut Protokoll nicht erreicht (Abbildung 5) und stattdessen durch Anpassung eines Weibull-Modells geschätzt (Abbildung 6). Danach liegt das geschätzte mediane PFS für den Vandetanib-Arm innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung bei 28 Monaten, in der Placebo-Gruppe hingegen nur bei 16 Monaten, was eine Verlängerung des medianen PFS um rund ein Jahr bedeutet. In der genäherten Zielpopulation mit dem höheren therapeutischen Bedarf liegt

somit das progressionsfreie Überleben um einen Monat über dem der gesamten Studienpopulation.

In Tabelle 4-35 und Tabelle 4-37 zeigt sich, dass sowohl bei Betrachtung der Gesamtpopulation als auch der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung der Anteil an progressionsfreien Patienten (Kaplan-Meier-Schätzer) in der Vandetanib-Gruppe zu jedem der Zeitpunkte höher war als in der Vergleichsgruppe. Nach 6 Monaten betrug der Anteil der Patienten ohne Progression in der Vandetanib-Gruppe 91,4% (89,4%) gegenüber 74,5% (72,9%) in der Placebo-Gruppe. Nach Monat 12 lagen die Anteile der progressionsfreien Patienten bei 83,3% und 80,8% (Vandetanib), bzw. 62,7% und 57,1% (Placebo) und nach Monat 24 bei 58,5% und 51,7% bzw. 41,7% und 36,0%.

Abbildung 4 und Abbildung 5 unterstreichen die oben genannten Ergebnisse anhand der grafischen Darstellung der Kaplan-Meier-Kurven für das PFS, welche in beiden betrachteten Populationen übereinstimmend eine frühe und gleichbleibend deutliche Separierung zeigen.

In Tabelle 4-38 wird der PFS-Status in der Gesamtpopulation am Tag des Daten *Cut-off* angegeben. Auch hier liegt der Anteil der progressiven Patienten in der Vandetanib-Gruppe mit 31,6% um knapp 20% niedriger als in der Placebo-Gruppe, wo bereits etwas über die Hälfte der Patienten (51%) eine Krankheitsprogression erlitten haben. In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ist dieser Unterschied sogar noch geringfügig größer (36,5% vs. 58,3%). Die Aufteilung der Progressionsfälle auf RECIST-Progression und Tod ist in beiden Gruppen in beiden betrachteten Populationen sehr ähnlich.

Somit ist der Anteil der Patienten in der Gesamtpopulation, bei denen zum Zeitpunkt des Daten *Cut-off* keine Progression festgestellt worden ist, in der Vandetanib-Gruppe 68,4% und in der Placebo-Gruppe 49%. Gründe dafür, dass keine Progression berichtet wurde, waren das tatsächliche Ausbleiben einer Progression (59,3% vs. 41,0%; Vandetanib vs. Placebo), Tod (später als 3 Monate nach letzter RECIST Evaluierung; 2,2% vs. 3,0%; Vandetanib vs. Placebo) oder Ausscheiden, bevor eine Progression auftreten konnte (6,9% vs. 5,0%; Vandetanib vs. Placebo). Gründe für das vorzeitige Ausscheiden waren Sicherheitsgründe (0,4% vs. 0,0%; Vandetanib vs. Placebo), Protokollabweichung (0,4% vs. 1,0%; Vandetanib vs. Placebo) und Beendigung auf Wunsch des Patienten (6,1% vs. 4,0%; Vandetanib vs. Placebo). Ein ähnliches Bild für den Anteil der Patienten, bei denen zum Zeitpunkt des Daten *Cut-off* keine Progression festgestellt worden ist, ergibt sich für die genäherte Zielpopulation. In der Vandetanib-Gruppe liegt der Anteil der Patienten bei 63,5% im Gegensatz zur Placebo-Gruppe mit 41,7%. Gründe dafür, dass keine Progression berichtet wurde, waren das tatsächliche Ausbleiben einer Progression (57,1% vs. 35,0%; Vandetanib vs. Placebo), Tod (später als 3 Monate nach letzter RECIST Evaluierung; 1,6% vs. 1,7%; Vandetanib vs. Placebo) oder Ausscheiden, bevor eine Progression auftreten konnte (4,8% vs. 5,0%; Vandetanib vs. Placebo). Gründe für das vorzeitige Ausscheiden waren Sicherheitsgründe (0,8% vs. 0,0%; Vandetanib vs. Placebo), Protokollabweichung (0,0% vs. 1,7%; Vandetanib vs. Placebo) und Beendigung auf Wunsch des Patienten (4,0% vs. 3,3%; Vandetanib vs. Placebo).

Zusammenfassend zeigt sich bezüglich des Endpunktes „progressionsfreies Überleben“ eine statistisch signifikante 54%-ige Reduktion des Risikos unter Vandetanib verglichen mit Placebo (HR:0,46, 95%-KI: [0,31;0,69]) in der Gesamtpopulation bzw. eine 53%-ige Reduktion innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung (HR:0,47 (95%-KI: [0,29;0,77]), was sich in ein verlängertes progressionsfreies Überleben um median ca. 11 bzw. 12 Monate (Gesamt-, genährte Zielpopulation) übersetzt. Auch bei Adjustierung für mögliche prognostische Faktoren (RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien und MTC-Status) ergab sich in beiden betrachteten Populationen ein statistisch signifikanter Vorteil von Vandetanib gegenüber Placebo hinsichtlich des PFS (HR_{adj}:0,46, 95%-KI: [0,32;0,68] für die Gesamtpopulation bzw. HR_{adj}:0,47 (95%-KI: [0,29;0,76] für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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

4.3.1.3.1.2 Zeit bis zur Schmerzprogression (TWP) – 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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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 von TWP

Studie	Operationalisierung
D4200C00058	Kombinierter Endpunkt abgeleitet aus der Bewertung für den stärksten Schmerz basierend auf dem validierten Schmerzfragebogen BPI-SF (*Brief Pain Inventory (Short Form)*) sowie der Angabe des Patienten über die Anwendung von Opioidanalgetika.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-41: Bewertung des Verzerrungspotenzials für Verzögerung der TWP in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

TWP wird als kombinierter Endpunkt aus der Bewertung für den stärksten Schmerz basierend auf dem validierten Schmerzfragebogen BPI-SF sowie der Angabe des Patienten über die Anwendung von Opioidanalgetika abgeleitet. Der kombinierte Endpunkt für den TWP wurde erreicht, wenn der Patient eine Progression des „stärksten Schmerzes“ (≥2 Punkte im *Worst Pain Score* des BPI-SF) berichtete oder Opioidanalgetika verwenden bzw. in ihrer Dosis steigern musste, ohne dass innerhalb der folgenden 14 Tage eine Linderung der Schmerzen (ebenfalls gemessen am Schmerzmittelgebrauch und dem BPI-SF) eintrat. Tod und RECIST Progression wurden nicht als „stärkster Schmerz“ gewertet.

Patientenberichteter Schmerz wird durch den BPI-SF bewertet. Dieser spiegelt Information über Intensität (sensorische Dimension) des Schmerzes und den Grad, in welchem der Schmerz die Funktion beeinträchtigt (reaktive Dimension), wider. Der BPI-SF verwendet numerische Beurteilungsskalen von 0-10, wobei der Skalenwert 0 der Ausprägung „kein Schmerz“ und 10 der Einschätzung „stärkste vorstellbare Schmerzen“ entspricht. Patienten wurden gebeten, zum Zeitpunkt der Beantwortung des Fragebogens ihre Schmerzen über die letzten 24 Stunden zu bewerten, und zwar den stärksten, geringsten und durchschnittlichen Schmerz sowie die momentane Schmerzintensität. Mittels der gleichen Beurteilungsskala wird der Grad der Beeinträchtigung des Patienten hinsichtlich der allgemeinen Aktivität, der physischen und sozialen Aktivität, der Stimmung, des Gehvermögens, der normalen Arbeit, der Beziehungen zu anderen Menschen und des Schlafes bewertet. Der Gebrauch dieses Instruments bei Krebspatienten ist etabliert. Der BPI-SF wurde von den Patienten gemäß Studienplan einmal wöchentlich selbständig ausgefüllt.

Patientenberichteter Schmerzmittelgebrauch wurde durch ein ebenfalls wöchentlich geführtes Tagebuch bis zur Beendigung der randomisierten Behandlung erfasst. Patienten wurden gebeten, ihren Schmerzmittelgebrauch (aus jedem Grund) über die letzten 24 Stunden zu berichten. Der Schmerzmittelgebrauch wurde zu Vergleichszwecken in Morphinsulfat-Äquivalente in mg/Tag konvertiert.

Für die Bewertung für den stärksten Schmerz war „Progression“ definiert als eine Erhöhung um ≥2 Punkte auf der Beurteilungsskala gegenüber dem Ausgangswert. „Verbesserung“ war definiert als eine Reduzierung um ≥2 Punkte gegenüber dem Ausgangswert, „stärkster Schmerz“ ohne eine Erhöhung der Schmerzmittel gegenüber dem Ausgangswert um ≥10 mg/Tag Morphinsulfat-Äquivalent. „Sonstige“ war definiert als keine Veränderung.

Für den Schmerzmittelgebrauch war „Verschlechterung“ definiert als eine Erhöhung um ≥10 mg/Tag Morphinsulfat Äquivalent gegenüber dem Ausgangswert. „Verbesserung“ war definiert als Reduzierung um >50% des Gebrauchs von Schmerzmitteln gegenüber dem letzten Besuch, ohne Erhöhung der Bewertung für den stärksten Schmerz um ≥2 Punkte auf der Beurteilungsskala gegenüber dem Ausgangswert. „Sonstige“ war definiert als keine Veränderung.

Die primäre Analyse des TWP basierte auf einem Cox-PH-Modell mit Behandlung als Faktor. Die Behandlungsunterschiede wurden mithilfe des *Log-Rank* Tests getestet, zur Bewertung des Behandlungseffektes wurden der Schätzer für das *Hazard Ratio* und das entsprechende

Konfidenzintervall angegeben. Die Schätzung der medianen Zeit bis zur Schmerzprogression (TWP) und die grafische Darstellung der TWP erfolgten mittels Kaplan-Meier-Methodik.

Unterstützend wurden, neben der primären nicht-adjustierten Analyse, adjustierte Analysen als Sensitivitätsanalysen zur Überprüfung der Robustheit der Ergebnisse durchgeführt. Im multivariaten Cox-PH-Modell wurden neben der Behandlung die folgenden potenziell prognostischen Faktoren eingeschlossen: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien, MTC-Status (hereditär, sporadisch, unbekannt). Weitere Sensitivitätsanalysen umfassten Modifikationen der Definition der TWP: Erstens wurde die TWP allein anhand der Bewertung für den stärksten Schmerz basierend auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs hergeleitet, zweitens wurden zusätzlich fehlende Werte für den stärksten Schmerz zum Zeitpunkt der lokal festgestellten Progression extrapoliert, um zu bewerten, ob eine mögliche Verschlechterung nicht erfasst und somit ein verzerrter TWP-Wert für die Analyse verwendet worden wäre (Imputation).

Das endpunktbezogene Verzerrungspotential wurde aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Eine mögliche partielle Entblindung von Patienten in der Vandetanib-Gruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

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

Tabelle 4-42: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	114 (49,4)	100	57 (57,0)	0,61** [0,43;0,87]	0,0062*
					0,64*** [0,46;0,89]	0,0103

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse: Cox-PH-Modell mit Behandlung als Faktor und den Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien, MTC-Status (hereditär, sporadisch, unbekannt)
Schmerzprogression wird als „Verschlechterung“ betrachtet die sich innerhalb der nächsten 14 Tage nicht verbessert. Schmerzprogression wird abgeleitet aus dem BPI-SF und dem patientenberichteten Schmerzmittelgebrauch.

Tabelle 4-43: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	60 (47,6)	60	33 (55,0)	0,62** [0,39;0,99]	0,0446*
					0,65*** [0,40;1,04]	0,0789

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse: Cox-PH-Modell mit Behandlung als Faktor und den Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien, MTC-Status (hereditär, sporadisch, unbekannt)
Schmerzprogression wird als „Verschlechterung“ betrachtet die sich innerhalb der nächsten 14 Tage nicht verbessert. Schmerzprogression wird abgeleitet aus dem BPI-SF und dem patientenberichteten Schmerzmittelgebrauch.

Tabelle 4-44: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058						
BPI-SF*	231	105 (45,5)	100	48 (48,0)	0,76 [0,53;1,09]	0,1399
BPI-SF mit Ersetzen fehlender Werte**	231	109 (47,2)	100	53 (53,0)	0,73 [0,51;1,03]	0,0760

* *Log-Rank* Test allein anhand des „schwersten Schmerzes“ auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs
** *Log-Rank* Test allein anhand des „schwersten Schmerzes“ auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs. Fehlende Werte von Patienten, die aufgrund der im Prüfzentrum lokalen RECIST-Bewertung wegen Progression zensiert wurden, wurden mit Hilfe einer Regressionsanalyse extrapoliert.

Tabelle 4-45: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058						
BPI-SF*	126	52 (41,3)	60	26 (43,3)	0,79 [0,48;1,30]	0,3565
BPI-SF mit Ersetzen fehlender Werte**	126	55 (43,7)	60	29 (48,3)	0,76 [0,47;1,22]	0,2585

* *Log-Rank* Test allein anhand des „schwersten Schmerzes“ auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs
** *Log-Rank* Test allein anhand des „schwersten Schmerzes“ auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs. Fehlende Werte von Patienten, die aufgrund der im Prüfzentrum lokalen RECIST-Bewertungwegen Progression zensiert wurden, , wurden mit Hilfe einer Regressionsanalyse extrapoliert.

Tabelle 4-46: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**Mediane Zeit bis zur Schmerzprogression (Monate)**	**N**	**Mediane Zeit bis zur Schmerzprogression (Monate)**
D4200C00058	231	7,85	100	3,25

Tabelle 4-47: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**	
	N	**Mediane Zeit bis zur Schmerzprogression (Monate)**	**N**	**Mediane Zeit bis zur Schmerzprogression (Monate)**
D4200C00058	126	11,07	60	3,42

Anmerkung: Für 6 Patienten lagen keine Daten bezüglich der letzten Progression vor Randomisierung vor. „Symptomatisch“ zur *Baseline* waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymtomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei *Baseline*. „Progressiv“ zur *Baseline* waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.

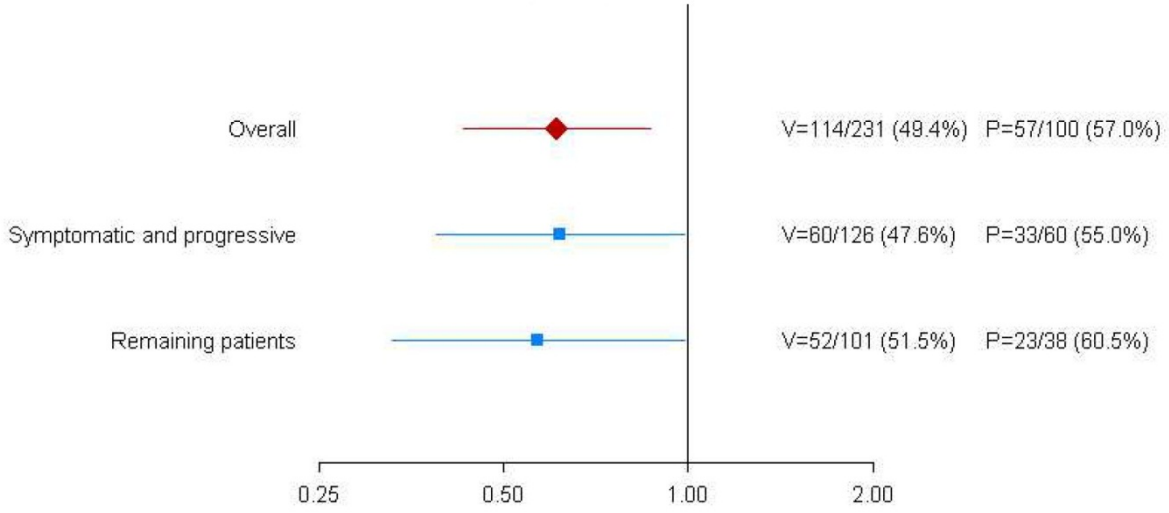


Abbildung 8: Forest Plot für TWP nach Symptomen und Progressionsstatus bei *Baseline* (Gesamtpopulation)

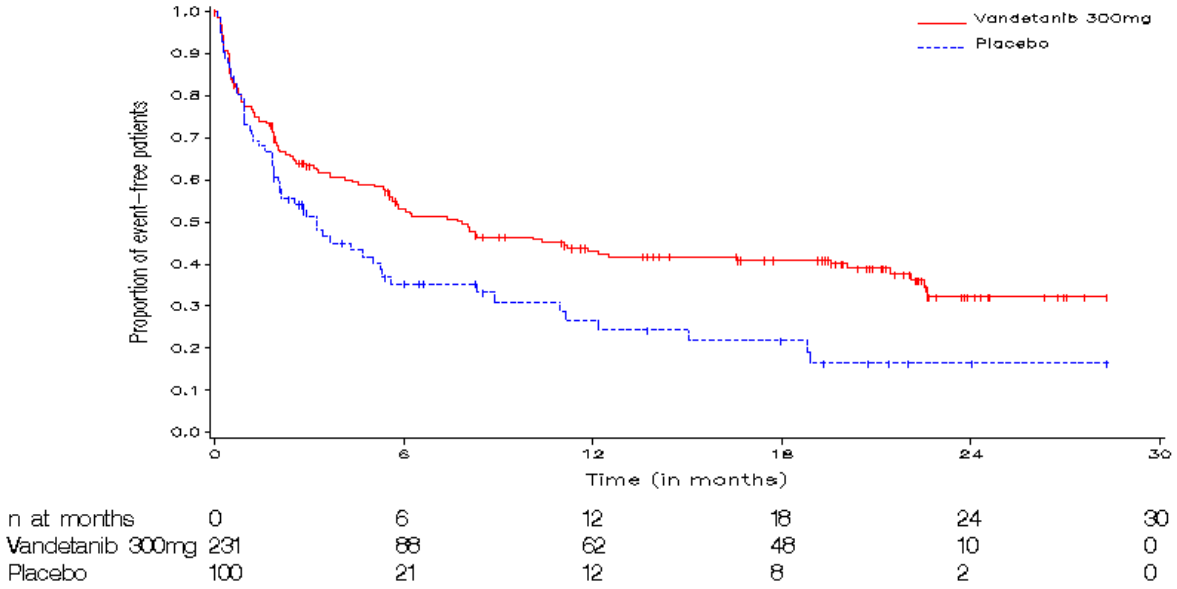


Abbildung 9: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population innerhalb der Gesamtpopulation)

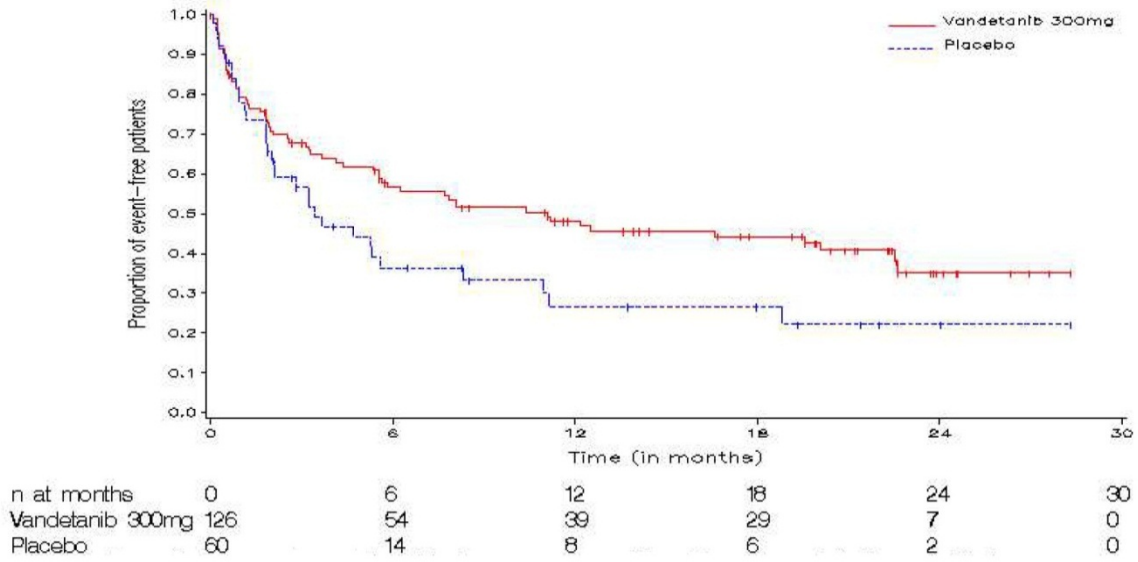


Abbildung 10: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)

Es zeigte sich eine statistisch signifikante Verlängerung der TWP unter Vandetanib im Vergleich zu Placebo mit einem *Hazard Ratio* von 0,61 (95%-KI: [0,43;0,87], p-Wert: 0,0062). In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ergab sich ein *Hazard Ratio* von 0,62 (95%-KI: [0,39;0,99], p-Wert: 0,0446). Die mediane Zeit (Tabelle 4-46) bis zur Schmerzprogression betrug in der Vandetanib-Gruppe 7,85 Monate und in der Placebo-Gruppe 3,25 Monate für die Gesamtpopulation, was einer Verbesserung um knapp 5 Monate entspricht. In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung betrug die Verbesserung sogar fast 8 Monate (11,07 vs. 3,42). Der *Forest Plot* (Abbildung 8) zeigt, dass beide betrachteten Patientenpopulationen vergleichbare Resultate zu diesem Endpunkt liefern. Abbildung 9 und Abbildung 10 unterstreichen die genannten Ergebnisse anhand der grafischen Darstellung der Kaplan-Meier-Kurven, welche die Zeit bis zur Schmerzprogression für beide Behandlungsgruppen zeigen. Die Kurven weisen übereinstimmend eine frühe und gleichbleibend deutliche Separierung auf. Aus den Abbildungen und den Ergebnissen wird ersichtlich, dass die Patienten der Vandetanib-Gruppe einen längeren Zeitraum erwarten konnten, bis sich Schmerzen entwickelten, eine Schmerzprogression eintrat oder eine Erhöhung der Opioiddosis nötig war, um die Schmerzen zu kontrollieren. Diese Ergebnisse wurden auch aus klinischer Sicht von der DGHO als „eindrucksvoll“ und als „echter Benefit“ für die Patienten bewertet [21].

Beide Sensitivitätsanalysen, die den Gebrauch von Schmerzmitteln für die Definition des Endpunktes nicht berücksichtigten, zeigten zwar ebenfalls einen numerischen Vorteil für Vandetanib, der Unterschied zwischen den Behandlungsgruppen war aber nicht mehr signifikant. Das *Hazard Ratio* lag bei 0,76 (95%-KI: [0,53;1,09]) bzw. bei 0,73 (95%-KI: [0,51;1,03]) nach Ersetzung fehlender Werte. Das Gleiche gilt für die genäherte

Zielpopulation mit symptomatischer und progressiver Erkrankung mit *Hazard Ratios* von 0,79 (95%-KI: [0,48;1,30]) bzw. bei 0,76 (95%-KI: [0,47;1,22]). Die Unterschiede zwischen den Resultaten der Sensitivitätsanalysen und dem Resultat der primären Analyse zeigen, dass eine valide Beurteilung der Schmerzprogression nur unter Berücksichtigung beider Komponenten, des Schmerzes (gemessen anhand validierter Skalen) und des Schmerzmittelgebrauchs, möglich ist. Eine isolierte Betrachtung der Schmerzskala erscheint somit nicht sinnvoll.

Ergebnisse zu weiteren explorativen Endpunkten für TWP bzw. Reduktion des Opioid-Gebrauchs sind in den Abschnitten 4.3.1.3.1.9 bzw. 4.3.1.3.1.10 dargestellt.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 11: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurden keine Meta-Analysen durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.3 Objektive Ansprechrate (ORR) – 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

– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-48: Operationalisierung der objektiven Ansprechrate (ORR)

Studie	Operationalisierung
D4200C00058	Anteil der Patienten, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression (BOR, *Best Objective Response*) einem kompletten Ansprechen (CR, *Complete Response*) oder partiellen Ansprechen (PR, *Partial Response*) entspricht.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-49: Bewertung des Verzerrungspotenzials für die objektive Ansprechrate (ORR) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Objektive Ansprechrate ist operationalisiert als Anteil der Patienten, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression (BOR, *Best Objective Response*), einem kompletten Ansprechen (CR, *Complete Response*) oder partiellen Ansprechen (PR, *Partial Response*) entspricht. Die Kriterien für CR oder PR entsprechen den

RECIST Kriterien und basieren auf standardisierten reproduzierbaren bildgebenden Verfahren.

Das beste objektive Ansprechen wurde zum Zeitpunkt der ersten eingetretenen RECIST Progression ermittelt, unter Berücksichtigung aller zentral verfügbaren RECIST Bewertungen. Patienten, die vor dem Ausscheiden aus dem jeweiligen randomisierten Behandlungsarm progressionsfrei waren, wurden unabhängig davon, ob die Bewertung während der offenen Behandlungsphase (*Open Label*) erhoben wurde, bis zum Zeitpunkt der Progression berücksichtigt. Progressionsfreie Patienten wurden mit ihrer letzten RECIST Bewertung zum Zeitpunkt der Analyse berücksichtigt. Für die primäre Analyse der objektiven Ansprechrate wurde auf Basis der ITT-Population eine logistische Regression mit dem Behandlungsarm als Faktor durchgeführt. Die Behandlungsunterschiede wurden anhand des *Odds Ratio* und des zugehörigen Konfidenzintervalls bewertet. Zusätzlich wurde der Anteil der Patienten mit objektivem Ansprechen gemäß allen verfügbaren zentralen RECIST Bewertungen dargestellt. Die Zusammenfassung beinhaltet ebenfalls Details über die Anzahl (%) der Patienten innerhalb jeder BOR Kategorie.

Im Zusammenhang mit der ORR wurde für Patienten mit partiellem oder komplettem Ansprechen anhand aller verfügbaren RECIST Bewertungen (zentraler Review) die Dauer des Ansprechens ermittelt. Erhoben wurde die Dauer des Ansprechens ab dem ersten bestätigten Ansprechen.

Unterstützend wurde eine Sensitivitätsanalyse zur Überprüfung der Robustheit der Ergebnisse nach Adjustierung für potenziell prognostische Faktoren durchgeführt. Das dafür verwendetete logistsche Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Der Patientenstatus bezüglich des Tumoransprechens (*Objective Response*) wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen. Einen das Verzerrungspotential beeinflussenden Punkt stellt der vorzeitige Wechsel (*Cross-Over*) von Placebo zu *Open-Label* Vandetanib dar. Ein Wechsel der Behandlungsgruppe von Placebo zu Vandetanib war bereits auf Grundlage eines positiven Reviews (Progression) durch den Prüfarzt erlaubt, welcher nicht immer durch den zentralen Review bestätigt werden konnte. Um den möglichen Effekt des *Cross-Over* auf den beobachteten Behandlungseffekt zu beurteilen, wurde eine nicht präspezifizierte Sensitivitätsanalyse – ohne Berücksichtigung der Daten aus der offenen Behandlungsphase – zur Überprüfung der Robustheit der Ergebnisse durchgeführt.

Sowohl unter Einbezug aller oben angeführten Aspekte als auch aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips wurde das endpunktbezogene Verzerrungspotential als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotential

beeinflussende Punkte gefunden werden. Die Richtung einer möglichen Verzerrung aufgrund einer *Open-Label* Therapie mit Vandetanib im Anschluss an die verblindete Behandlung wird eher zuungunsten von Vandetanib angenommen. Mögliche Abweichungen werden aufgrund der kleinen Patientenzahl (von 13 Patienten mit Ansprechen in der Placebo-Gruppe) als „gering“ eingestuft. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung (10 Patienten mit Ansprechen in der Placebo-Gruppe).

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.

Tabelle 4-50: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	104 (45,0)	100	13 (13,0)	5,48** [2,99;10,79]	<0,0001*
					5,64*** [3,00;11,33]	<0.0001

* Logistische Regression
** mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse (Adjustierung für Kovariaten): Das logistische Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Tabelle 4-51: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	56 (44,4)	60	10 (16,7)	4,00** [1,93;9,01]	0,0001*
					3,86*** [1,78;9,09]	0,0005

* Logistische Regression
** mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse (Adjustierung für Kovariaten): Das logistische Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Tabelle 4-52: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	101 (43,7)	100	1 (1,0)	76,91 [16,68;1366]	<0,0001

Tabelle 4-53: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	55 (43,7)	60	1 (1,7)	45,70 [9,58;819,7]	<0,0001

Tabelle 4-54: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

			Anzahl (%) der Patienten		
Studie	**Ansprech-Status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=231)**	**Placebo (N=100)**	**Total (N=331)**
D4200C00058	Ansprechen	Gesamt	104 (45,0)	13 (13,0)[a]	117 (35,3)
		Komplettes Ansprechen	0 (0,0)	0 (0,0)	0 (0,0)
		Partielles Ansprechen	104 (45,0)	13 (13,0)[a]	117 (35,3)
	Kein Ansprechen	Gesamt	127 (55,0)	87 (87,0)	214 (64,7)
		Stabile Erkrankung ≥8 Wochen	114 (49,4)	70 (70,0)	184 (55,6)
		Progressive Erkrankung	9 (3,9)	13 (13,0)	22 (6,6)
		Nicht erhebbar	4 (1,7)	4 (4,0)	8 (2,4)
		Keine Erkrankung	0 (0,0)	0 (0,0)	0 (0,0)

[a] Ansprechbeginn war in 12 von 13 Fällen während der offenen (*open label*) Behandlungsphase mit Vandetanib.

Tabelle 4-55: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

			Anzahl (%) der Patienten		
Studie	**Ansprech-Status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=126)**	**Placebo (N=60)**	**Total (N=186)**
D4200C00058	Ansprechen	Gesamt	56 (44,4)	10 (16,7)[a]	66 (35,5)
		Komplettes Ansprechen	0 (0,0)	0 (0.0)	0 (0,0)
		Partielles Ansprechen	56 (44,4)	10 (16,7)[a]	66 (35,5)
	Kein Ansprechen	Gesamt	70 (55,6)	50 (83,3)	120 (64,5)
		Stabile Erkrankung ≥8 Wochen	62 (49,2)	42 (70,0)	104 (55,9)
		Progressive Erkrankung	7 (5,6)	6 (10,0)	13 (7,0)
		Nicht erhebbar	1 (0,8)	2 (3,3)	3 (1,6)
		Keine Erkrankung	0 (0,0)	0 (0,0)	0 (0,0)

[a] Ansprechbeginn war in 9 von 10 Fällen während der offenen (*open label*) Behandlungsphase mit Vandetanib.

Tabelle 4-50 zeigt, dass es eine statistisch signifikante Verbesserung des ORR auf Vandetanib verglichen mit Placebo gibt. Das *Odds Ratio* liegt bei 5,48 (4,00 in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) mit einem 95%-Konfidenzintervall von [2,99;10,79] ([1,93;9,01]) und einem p-Wert von <0,0001 (0,0001). Dieses Ergebnis spricht deutlich für eine bessere ORR innerhalb der Vandetanib-Gruppe im Vergleich zu der Placebo-Gruppe, unabhängig von der betrachteten Patientenpopulation. Insgesamt betrug die ORR 45% in der Vandetanib-Gruppe und 13% in der Placebo-Gruppe (bzw. 44,4% vs. 16,7% in der genäherten Zielpopulation). Von den 13 Patienten mit Ansprechen in der Placebo-Gruppe (Gesamtpopulation) hatte lediglich ein Patient bereits unter Placebo ein beginnendes Ansprechen. Bei den verbleibenden 12 Patienten der Placebo-Gruppe setzte die Remission erst während der offenen Behandlungsphase unter Vandetanib ein (bzw. bei 9 von 10 Patienten in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung). Grundlage für den Wechsel auf *Open-Label* Vandetanib war ein durch den Prüfarzt festgestellter Progress, welcher nicht immer durch den zentralen Review bestätigt werden konnte (Tabelle 4-54). Die Bewertung der ORR anhand der Daten aus dem zentralen Review birgt einen Vorteil zugunsten von Placebo. Das Resultat der Sensitivitätsanalyse ohne Berücksichtigung der Daten nach C*ross-Over* auf *Open-Label* Vandetanib zeigt folglich einen noch deutlicheren Effekt für Vandetanib (Tabelle 4-52): Objektives Ansprechen erlebten 101/231 (55/126 in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) Patienten unter Vandetanib und 1/100 (1/60) Patienten unter Placebo (OR [95%-KI]: 76,91 [16,68;1366] bzw. 45,70 [9,58;819,7]).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 12: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.4 Krankheitskontrollrate (DCR)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-56: Operationalisierung von Krankheitskontrollrate (DCR)

Studie	Operationalisierung
D4200C00058	Anteil der Patienten, deren bestes objektives Ansprechen (BOR, *Best Objective Response*) einem kompletten Ansprechen (CR, *Complete Response*), einem partiellen Ansprechen (PR, *Partial Response*) oder einer Krankheitsstabilität für mindestens 24 Wochen seit Randomisierung entspricht.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-57: Bewertung des Verzerrungspotenzials für Krankheitskontrollrate (DCR) in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Hinweise auf ergebnisgesteuerte Berichterstattung**	**Sonstige das Verzerrungspotenzial beeinflussende Punkte**	**Verzerrungspotenzial Endpunkt**
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Krankheitskontrollrate (DCR) ist operationalisiert als der Anteil an Patienten, deren bestes objektives Ansprechen (BOR, *Best Objective Response*) einem kompletten Ansprechen (CR, *Complete Response*), einem partiellen Ansprechen (PR, *Partial Response*) oder einer Krankheitsstabilität (≥24 Wochen) seit Randomisierung entspricht. Die Kriterien für CR, PR oder Krankheitsstabilität entsprechen den RECIST Kriterien und basieren auf standardisierten reproduzierbaren bildgebenden Verfahren.

Alle verfügbaren zentralen RECIST Bewertungen wurden in ihrer ursprünglichen Form berücksichtigt, unabhängig davon, ob sich die Patienten zum Zeitpunkt der Bewertung in Therapie oder in der offenen Behandlungsphase befanden.

Für die primäre Analyse der Krankheitskontrollrate wurde auf Basis der ITT-Population eine logistische Regression mit dem Behandlungsarm als einzigem Faktor durchgeführt. Die Behandlungsunterschiede wurden anhand des *Odds Ratio* und des zugehörigen Konfidenzintervalls bewertet. Zusätzlich wurde der Anteil der Patienten mit der Krankheitskontrollrate gemäß allen verfügbaren zentralen RECIST Bewertungen dargestellt.

Unterstützend wurde eine Sensitivitätsanalyse zur Überprüfung der Robustheit der Ergebnisse nach Adjustierung für potenziell prognostische Faktoren durchgeführt. Das dafür verwendete logistische Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib wurde ausgeschlossen, da sich krankheitsbedingte Symptome und Therapie-assoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind. Einen das Verzerrungspotential beeinflussenden Punkt stellt der vorzeitige Wechsel von Placebo zu

Open-Label Vandetanib dar. Um den möglichen Effekt dieses *Cross-Over* auf den beobachteten Behandlungseffekt zu beurteilen, wurde eine nicht präspezifizierte Sensitivitätsanalyse – ohne Berücksichtigung der Daten aus der offenen Behandlungsphase – zur Überprüfung der Robustheit der Ergebnisse durchgeführt.

Sowohl unter Einbezug aller oben angeführten Aspekte als auch aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips wurde das endpunktbezogene Verzerrungspotential als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Die Richtung einer möglichen Verzerrung aufgrund eines *Cross-Over* auf *Open-Label* Vandetanib wird eher zuungunsten von Vandetanib angenommen. Patienten konnten nach lokal festgestellter Progression auf *Open-Label* Vandetanib umgestellt werden. Nicht in allen Fällen stimmte der lokale *read* mit der zentralen Bewertung überein.

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

Tabelle 4-58: Ergebnisse für die Krankheitskontrollrate (DCR) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Krankheitskontrolle (%)**	**N**	**Anzahl Patienten mit Krankheitskontrolle (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	200 (86,6)	100	71 (71,0)	2,64** [1,48;4,69]	0,0010*
					2,93*** [1,57; 5,49]	0,0007

* Logistische Regression
** mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse (Adjustierung für Kovariaten): Das logistische Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Tabelle 4-59: Ergebnisse für die Krankheitskontrollrate (DCR) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Krankheitskontrolle (%)**	**N**	**Anzahl Patienten mit Krankheitskontrolle (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	106 (84,1)	60	41 (68,3)	2,46** [1,19;5,09]	0,0156*
					3,03*** [1,34;7,02]	0,0080

* Logistische Regression
** mit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse (Adjustierung für Kovariaten): Das logistische Regressionsmodell enthielt neben dem Faktor Behandlungsarm folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Tabelle 4-60: Ergebnisse (explorativ) für die Krankheitskontrollrate (DCR) **ohne** offene Behandlungsphase aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Krankheitskontrolle (%)**	**N**	**Anzahl Patienten mit Krankheitskontrolle (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	196 (84,8)	100	61 (61,0)	3,58 [2,09;6,17]	<0,0001*

* Logistische Regression

Tabelle 4-61: Ergebnisse (explorativ) für die Krankheitskontrollrate (DCR) **ohne** offene Behandlungsphase aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Krankheitskontrolle (%)**	**N**	**Anzahl Patienten mit Krankheitskontrolle (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	104 (82,5)	60	34 (56,7)	3,61 [1,83;7,26]	0,0002

Die Krankheitskontrollrate war in der Gesamtpopulation mit 86,6% in der Vandetanib Gruppe im Vergleich zu 71,0% unter Placebo höher. In der genäherten Zielpopulation lag die Krankheitskonrollrate bei 84,1% gegenüber 68,3%. Die Tabelle 4-58 zeigt einen statistisch signifikanten Unterschied der Ergebnisse für die Krankheitskontrollrate zwischen Vandetanib

und Placebo. Das *Odds Ratio* liegt bei 2,64 (2,46 in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) mit einem 95%-Konfidenzintervall von [1,48;4,69] bzw. [1,19;5,09] und einem p-Wert von 0,010 (bzw. 0,0156). Zum Zeitpunkt der lokal festgestellten Progression konnten die Patienten auf *Open-Label* Vandetanib wechseln. Nicht bei allen Patienten wurde der lokal diagnostizierte RECIST-Progress in der zentralen Auswertung bestätigt. Der vorzeitige Wechsel von Placebo zu *Open-Label* Vandetanib birgt einen Vorteil zugunsten von Placebo. Das Resultat der Sensitivitätsanalyse ohne Berücksichtigung der Daten aus der offenen Behandlungsphase zeigt folglich einen noch deutlicheren Effekt für Vandetanib. Der Anteil der Patienten mit Krankheitskontrolle betrug in der Gesamtpopulation 84,8% gegenüber 61% (Vandetanib / Placebo-Gruppe), in der genäherten Zielpopulation war der Unterschied in der Krankheitskontrolle noch größer mit 82,5% gegenüber 56,7%. Es konnte ein statistisch signifikantes Ergebnis zugunsten der Vandetanib–Gruppe gezeigt werden (OR: 3,58; 95%-KI: [2,09;6,17] bzw. OR: 3,61; 95%-KI: [1,83;7,26]).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 13: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.5 Dauer des objektiven Ansprechens (DOR)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-62: Operationalisierung von Dauer des objektiven Ansprechens (DOR)

Studie	Operationalisierung
D4200C00058	Für Patienten, die als bestes Ansprechen ein komplettes oder partielles Ansprechen (CR/PR) aufwiesen, wurde die Dauer des objektiven Ansprechens als Zeitraum vom ersten dokumentierten Ansprechen bis zur Progression oder Tod ohne vorhergehende dokumentierte Progression berechnet.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-63: Bewertung des Verzerrungspotenzials für Dauer des objektiven Ansprechens (DOR) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für Patienten, die als bestes Ansprechen ein komplettes oder partielles Ansprechen (CR/PR) aufwiesen, wurde die Dauer des objektiven Ansprechens betrachtet. Die Dauer des objektiven Ansprechens (DOR) ist definiert als Zeitraum zwischen dem erstem objektiven Ansprechen und der ersten dokumentierten Progression oder dem Tod ohne eine vorhergehende Progression. Die Kriterien für CR oder PR entsprechen den RECIST Kriterien und basieren auf standardisierten reproduzierbaren bildgebenden Verfahren. Die Dauer des objektiven Ansprechens wurde zum Zeitpunkt der ersten eingetretenen RECIST Progression ermittelt, unter Berücksichtigung aller zentral verfügbaren RECIST Bewertungen. Für Patienten, die vor dem Ausscheiden aus dem jeweiligen randomisierten Behandlungsarm noch progressionsfrei gemäß RECIST waren, wurden die RECIST Bewertungen der offenen Behandlungsphase berücksichtigt.

Eine ursprünglich im Studienbericht vorgesehene weitere Operationalisierung der Dauer des objektiven Ansprechens als „Zeit zwischen Randomisierung und erster dokumentierter Progression oder dem Tod ohne vorhergehende Progression“ wurde im weiteren Verlauf der Studie nicht ausgewertet, da eine Darstellung von der Randomisierung bis zur ersten RECIST-Progression in der zentralen Bewertung zu einer erheblichen Verzerrung des Behandlungseffekts führen würde. Der Grund der erheblichen Verzerrung ist der Tatsache geschuldet, dass in dieser Operationalisierung gemäß ITT-Prinzip auch die Patienten berücksichtigt worden wären, deren lokal diagnostizierte Progression im zentralen *review* nicht bestätigt wurde und die erst in der offenen Behandlungsphase ein Ansprechen auf die Therapie zeigten. Der Vorteil (Ansprechen) ergab sich, bis auf einen Patienten in der Placebo-Gruppe, für die Placebo-Patienten erst in der offenen Behandlungsphase. Der Unterschied zwischen den Behandlungsarmen für den Beginn des objektiven Ansprechens war entsprechend deutlich erkennbar (Vgl. Tabellen Tabelle 4-66 und Tabelle 4-67)

Die Dauer des objektiven Ansprechens wurde, wenn möglich, als mediane Dauer des objektiven Ansprechens, gemäß der Definition „Zeit vom ersten dokumentierten Ansprechen bis zur Progression oder Tod ohne vorhergehende dokumentierte Progression“, zusammengefasst. Über die obige Operationalisierung hinaus wurde, wenn möglich, die mediane Zeit bis zum Ansprechen, definiert als die Anzahl der Tage von der Randomisierung bis zum Ansprechen, zusammengefasst.

Die Analysen wurden auf Basis der ITT-Population durchgeführt und mittels einer Kaplan-Meier-Kurve grafisch dargestellt. Eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib wurde ausgeschlossen, da sich krankheitsbedingte Symptome und Therapie-assoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind. Das endpunktbezogene Verzerrungspotential wird als „niedrig“ bewertet, da keine Anhaltspunkte für eine ergebnisgesteuerte Berichtserstattung vorliegen.

Stellen Sie die Ergebnisse für den Endpunkt "Dauer des objektiven Ansprechens" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-64: Ergebnisse für die geschätzte mediane Dauer des objektiven Ansprechens (DOR) aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**geschätze mediane DOR (Monate)**	**N**	**geschätzte mediane DOR (Monate)**
D4200C00058	104	22,2	13	16,3

Nur Patienten mit Ansprechen wurden in die Auswertung mit eingeschlossen, Zeit vom ersten Ansprechen bis zur Progression

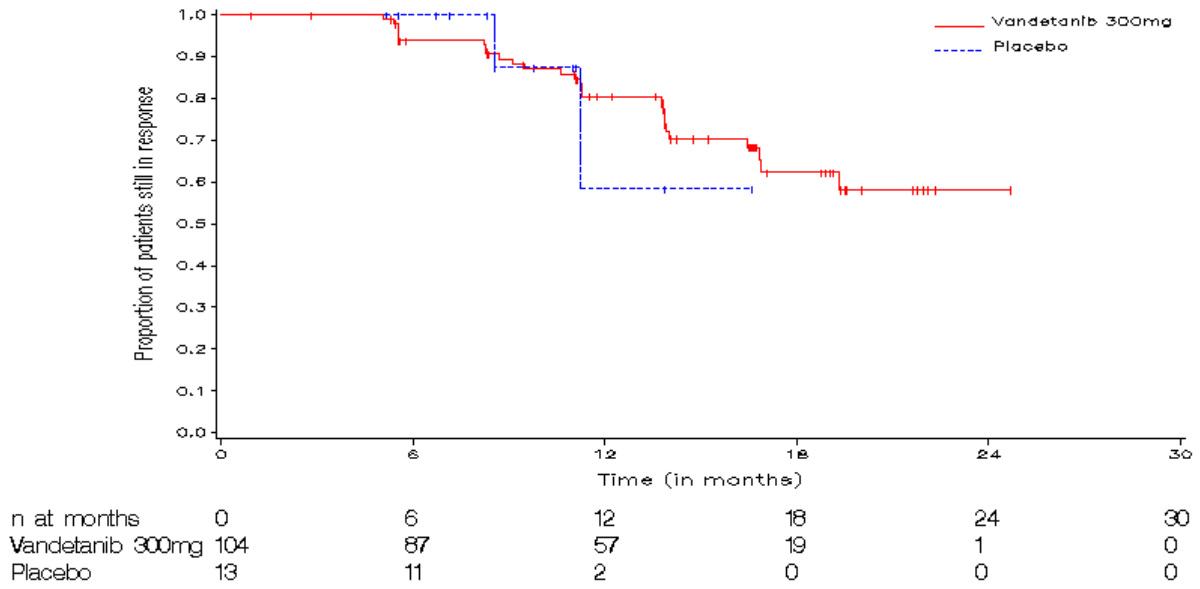


Abbildung 14: Kaplan-Meier-Kurve für die Dauer des objektiven Ansprechens (DOR) (Gesamtpopulation → ITT-Population), Zeit vom ersten Ansprechen bis zur Progression

Tabelle 4-65: Ergebnisse für die geschätzte mediane Dauer des objektiven Ansprechens (DOR) aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**	
	N	**geschätzte mediane DOR (Monaten)**	**N**	**geschätzte mediane DOR (Monate)**
D4200C00058	56	21,8	10	13,0

Nur Patienten mit Ansprechen wurden in die Auswertung mit eingeschlossen, Zeit vom ersten Ansprechen bis zur Progression.

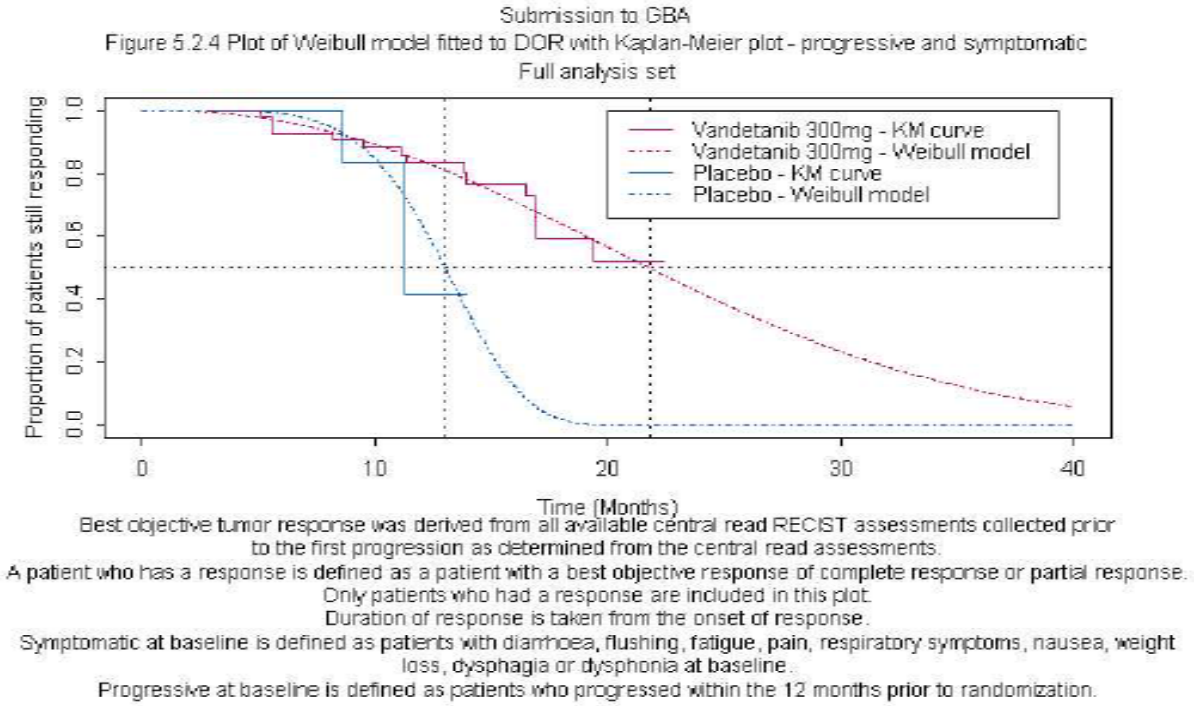


Abbildung 15: Kaplan-Meier-Kurve für Dauer des objektiven Ansprechens (DOR) (genäherte Zielpopulation → ITT-Population)

Tabelle 4-66: Ergebnisse für die mediane Zeit von der Randomisierung bis zum Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**mediane Zeit (Monate)**	**N**	**mediane Zeit (Monate)**
D4200C00058	104	5,8	13	13,7

Nur Patienten mit Ansprechen wurden in die Auswertung mit eingeschlossen

Tabelle 4-67: Ergebnisse für die mediane Zeit von der Randomisierung bis zum Ansprechen aus RCT mit dem zu bewertenden Arzneimittel (ITT-Population innerhalb der genäherten Zielpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**mediane Zeit (Monate)**	**N**	**mediane Zeit (Monate)**
D4200C00058	56	7,1	10	13,7

Nur Patienten mit Ansprechen wurden in die Auswertung mit eingeschlossen

Betrachtet man für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) die Abbildungen der Kaplan-Meier-Kurven der DOR für die Zeit vom ersten Ansprechen bis zur Progression, so zeigt sich für beide Behandlungsarme ein lang anhaltendes Ansprechen, wobei ein Ansprechen in der Placebo-Gruppe für 12 von 13 Patienten der Gesamtpopulation und 9 von 10 Patienten der genäherten Zielpopulation erst in der offenen Behandlungsphase unter Vandetanib einsetzte. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens liegen unter Vandetanib bei 22,2 Monaten in der Gesamtpopulation bzw. 21,8 Monaten in der genäherten Zielpopulation und unter Placebo bei 16,3 bzw. 13,0 Monaten.

Die mediane Zeit von der Randomisierung bis zum Ansprechen betrug in der Vandetanib-Gruppe 5,8 (Gesamtpopulation) bzw. 7,1 (genäherte Zielpopulation) Monate und unter Placebo für beide Populationen 13,7 Monate. In einer zusätzlichen Auswertung für die mediane Zeit von der Randomisierung bis zum Ansprechen unter Berücksichtigung der *Open-Label* Vandetanib-Therapie beträgt die mediane Zeit von der Randomisierung bis zum Ansprechen für den einen Patienten mit Ansprechen unter Placebo 12,2 Monate. Für die 12 verbleibenden Placebo-Patienten mit Ansprechen während der offenen Behandlungsphase (Gesamtpopulation) bzw. 9 Placebo-Patienten in der genäherten Zielpopulation beträgt die mediane Zeit von der ersten Dosis unter Vandetanib bis zum Ansprechen 6,9 Monate (Gesamtpopulation) bzw. 8,1 Monate (genäherte Zielpopulation).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 16: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.6 Explorativer Morbiditätsendpunkt Gewichtszunahme

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-68: Operationalisierung von Gewichtszunahme

Studie	Operationalisierung
D4200C00058	Rate der klinischen Gewichtszunahme, definiert als Anteil der randomisierten Patienten, auf welche das Kriterium der klinischen Gewichtszunahme zutraf. Klinische Gewichtszunahme ist definiert als Gewichtszunahme von >5% gegenüber dem Ausgangswert, i) nicht gefolgt von einer Gewichtsmessung innerhalb der nächsten 4 Wochen, die eine Gewichtsabnahme oder ein gleichbleibendes Gewicht bezogen auf die Ausgangsmessung anzeigt ii) nicht gefolgt von einer Gewichtsmessung, die eine Gewichtszunahme von 5% oder weniger anzeigt iii) nicht assoziiert mit einer Flüssigkeitsansammlung oder einem Ödem nach Einschätzung des Prüfarztes.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-69: Bewertung des Verzerrungspotenzials für Gewichtszunahme in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Klinische Gewichtszunahme ist definiert als Gewichtszunahme von >5% gegenüber dem Ausgangswert, i) nicht gefolgt von einer Gewichtsmessung innerhalb der nächsten 4 Wochen, die eine Gewichtsabnahme oder ein gleichbleibendes Gewicht bezogen auf die Ausgangsmessung anzeigt; ii) nicht gefolgt von einer Gewichtsmessung, die eine Gewichtszunahme von 5% oder weniger anzeigt; iii) nicht assoziiert ist mit einer Flüssigkeitsansammlung oder einem Ödem nach Einschätzung des Prüfarztes. Die Rate der klinischen Gewichtzunahme, war definiert als Anteil der randomisierten Patienten, welche obige Kriterien erfüllten. Der Endpunkt wurde rein exploratorisch analysiert.

Für die primäre Analyse der Gewichtszunahme wurde auf Basis der ITT-Population eine logistische Regression mit dem Behandlungsarm als einzigem Faktor durchgeführt. Die Behandlungsunterschiede wurden anhand der *Odds Ratio* und des zugehörigen Konfidenzintervalls bewertet.

Das endpunktbezogene Verzerrungspotential wurde aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Eine mögliche partielle Entblindung von Patienten in der Vandetanib-Gruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

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

Tabelle 4-70: Ergebnisse für die Rate der Gewichtszunahme aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Gewichtszunahme (%)**	**N**	**Anzahl Patienten mit Gewichtszunahme (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	66 (28,6)	100	11 (11,0)	3,24** [1,69;6,76]	0,0003*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-71: Ergebnisse für die Gewichtszunahme aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Gewichtszunahme (%)**	**N**	**Anzahl Patienten mit Gewichtszunahme (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	39 (31,0)	60	6 (10,0)	4,03** [1,71;11,17]	0,0010*

* Logistische Regression
** mit Behandlung als einzigem Faktor

In beiden Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigte sich ein deutliches Ergebnis zugunsten von Vandetanib (OR: 3,24 95%-KI: [1,69;6,76] bzw. OR: 4,03 (95%-KI: [1,71;11,17]). Der Anteil der Patienten mit klinischer Gewichtszunahme betrug in der Vandetanib-Gruppe 28,6% bzw. 31,0% verglichen zur Placebo-Gruppe mit 11,0% bzw. 10,0%. Zum arztberichteten unerwünschten Ereignis „Gewichtsverlust“ verweisen wir auf Abschnitt 4.3.1.3.1.14.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 17: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.7 Explorativer Morbiditätsendpunkt Einfluss der Behandlung auf Diarrhöen

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-72: Operationalisierung von Einfluss der Behandlung auf Diarrhöen

Studie	Operationalisierung
D4200C00058	Stuhlfrequenz: Durch den Patienten wöchentlich erhobene Werte anhand eines Tagebuches, bis zum Ausscheiden aus der randomisierten Behandlungsgruppe. Berichtet wurde die Stulfrequenz über die letzten 24 Stunden. Der Endpunkt wurde operationalisiert als Veränderung verglichen zum Ausgangswert.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-73: Bewertung des Verzerrungspotenzials für Einfluss der Behandlung auf Diarrhöen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Stuhlfrequenz wurde wöchentlich durch den Patienten anhand eines Tagebuches erhoben. Die Erhebung der Stuhlfrequenz erfolgte bis zum Ausscheiden aus der randomisierten Behandlungsgruppe. Berichtet wurde die Stuhlfrequenz über die letzten 24 Stunden. Der Endpunkt wurde operationalisiert als Veränderung verglichen zum Ausgangswert.

Dieser Endpunkt wurde im Rahmen der Studie rein explorativ ausgewertet. Es wurde keine formale statistische Analyse durchgeführt. Aufgrund der Tatsache, dass keine validen Vorher-Nachher-Vergleiche existieren, da die Mittelwerte zu den unterschiedlichen Zeitpunkten auf unterschiedlichen Stichprobenumfängen basieren und somit innerhalb jeder Gruppe keine „gepaarten“ Vergleiche möglich sind, sind die im Studienbericht dargestellten Ergebnisse nur schwer interpretierbar. Der Anteil der Patienten, die in der Berechnung deskriptiver Statistiken nicht berücksichtigt wurden, war zu allen Zeitpunkten hoch (bezogen auf die Gesamtpopulation in der Vandetanib- bzw. Placebo-Gruppe) und war zwischen den Behandlungsgruppen nicht balanciert.

Durch das Auftreten charakteristischer Nebenwirkungen besteht theoretisch die Möglichkeit der Entblindung von Patienten. Eine Entblindung kann sich insbesondere bei einer subjektiven Bewertung, wie der der Stuhlfrequenz durch die Patienten, auf das Resultat auswirken. Die Wahrscheinlichkeit einer solchen Verzerrung wurde, wie bereits in Abschnitt 4.3.1.2.2 dargestellt, als „gering“ angesehen. Diese Einschätzung ist dadurch begründet, dass die Patienten mit dieser symptomatischen Erkrankung wegen der Überlagerung von krankheitsbedingten Symptomen und therapie- sowie placeboassoziierten Nebenwirkungen schwer zwischen beiden differenzieren können.

Aufgrund der nach den Vorgaben der VerfO (siehe Anhang 4-G) formell nicht adäquaten Umsetzung des ITT-Prinzips wird das Verzerrungspotential für diesen Endpunkt als „hoch“ bewertet, auch wenn keine Anhaltspunkte für eine ergebnisgesteuerte Berichtserstattung vorliegen und die Resultate nur deskriptiv gegenübergestellt werden. Die Aussagekraft der Ergebnisse für den „Einfluss der Behandlung auf Diarrhöen“ ist gering. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

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

Tabelle 4-74: Ergebnisse für die Stuhlfrequenz aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	***Baseline***	212	3,32	2,94	2,3	0,0	18,3
	Woche 12	172	3,64	5,18	3,0	0,0	61,0
	Woche 24	152	3,07	2,26	3,0	0,0	12,0
	Woche 36	146	3,23	2,56	3,0	0,0	23,0
	Woche 48	135	3,37	2,88	3,0	0,0	22,0
	Woche 60	126	3,18	3,02	3,0	0,0	28,0
	Woche 72	111	2,88	1,72	3,0	0,0	8,0
	Woche 84	101	3,14	1,94	3,0	0,0	11,0
	Woche 96	60	3,40	3,05	3,0	0,0	21,0
	Woche 108	28	3,25	2,22	3,0	1,0	11,0
	Woche 120	14	3,57	2,82	3,0	1,0	10,0
	Woche 132	1	3,0	NC	3,0	3,0	3,0
Placebo	***Baseline***	92	2,95	2,32	2,3	0,0	11,0
	Woche 12	72	2,94	2,17	2,0	0,0	10,0
	Woche 24	61	3,10	2,34	3,0	0,0	10,0
	Woche 36	41	3,02	2,15	3,0	0,0	8,0
	Woche 48	43	3,35	2,38	3,0	1,0	10,0
	Woche 60	34	3,12	2,20	2,5	1,0	9,0
	Woche 72	28	3,29	2,34	2,5	0,0	8,0
	Woche 84	26	3,42	2,73	2,5	0,0	10,0
	Woche 96	11	3,27	3,04	2,0	0,0	9,0
	Woche 108	8	3,75	3,01	3,5	0,0	8,0
	Woche 120	3	3,33	2,52	3,0	1,0	6,0

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar
Zur besseren Übersicht werden die Werte nur nach Woche 12,24,36,48, 60,72,84,96,108,120 und 132 (sofern erhoben) dargestellt

Tabelle 4-75: Ergebnisse für die Stuhlfrequenz aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation)

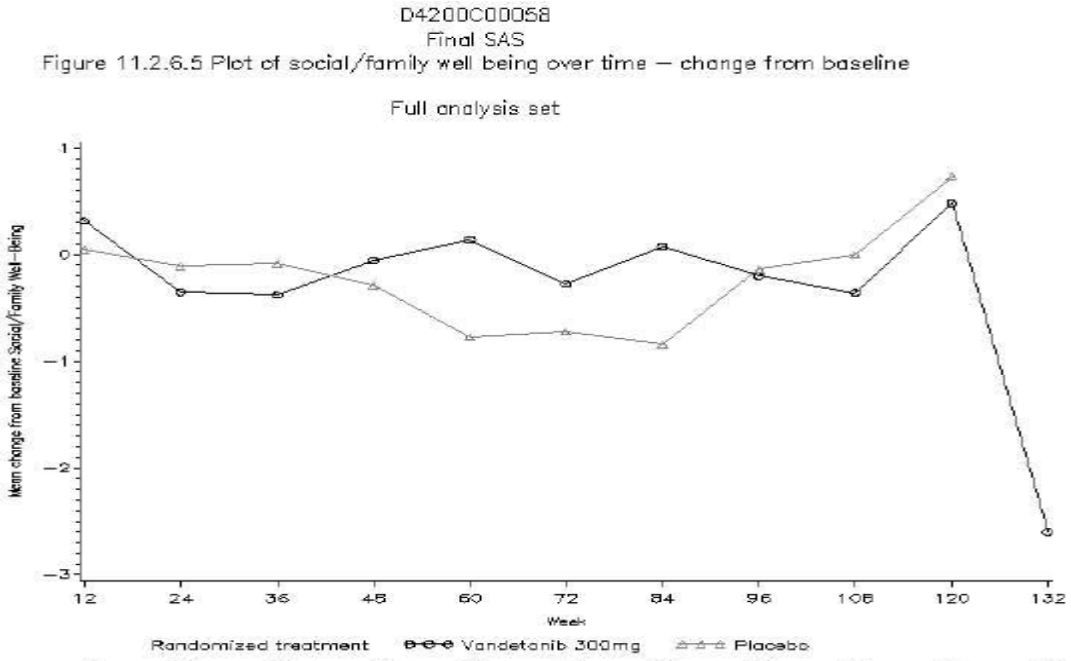
Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	***Baseline***	114	3,99	3,14	3,0	0,0	18,3
	Woche 12	91	3,88	6,40	3,0	0,0	61,0
	Woche 24	80	3,09	2,09	3,0	0,0	12,0
	Woche 36	83	3,58	2,96	3,0	0,0	23,0
	Woche 48	74	3,82	3,20	3,0	0,0	22,0
	Woche 60	65	3,18	1,72	3,0	0,0	8,0
	Woche 72	58	3,60	1,77	3,0	0,0	8,0
	Woche 84	56	3,70	2,15	3,0	1,0	11,0
	Woche 96	36	3,92	3,5	3,0	1,0	21,0
	Woche 108	16	3,06	1,88	3,0	1,0	7,0
	Woche 120	8	3,50	2,83	3,0	1,0	10,0
	Woche 132	1	3,0	NC	3,0	3,0	3,0
Placebo	***Baseline***	55	3,53	2,32	3,0	0,0	9,5
	Woche 12	43	3,44	2,40	3,0	0,0	10,0
	Woche 24	36	3,94	2,52	3,0	0,0	10,0
	Woche 36	27	3,44	2,28	3,0	0,0	8,0
	Woche 48	24	3,67	2,14	3,0	0,0	10,0
	Woche 60	19	3,37	2,61	2,0	1,0	9,0
	Woche 72	17	3,88	2,31	3,0	1,0	8,0
	Woche 84	16	4,06	2,54	3,5	1,0	9,0
	Woche 96	9	3,67	3,24	2,0	0,0	9,0
	Woche 108	6	4,17	3,31	4,5	0,0	8,0
	Woche 120	3	3,33	2,52	3,0	1,0	6,0

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar
Zur besseren Übersicht werden die Werte nur nach Woche 12,24,36,48, 60,72,84,96,108,120 und 132 (sofern erhoben) dargestellt

Tabelle 4-76: Veränderung für die Stuhlfrequenz zum Ausgangswert - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib							
	Woche 12	165	0,07	3,44	0,20	-15,3	16,5
	Woche 24	145	-0,16	3,17	0,0	-11,8	11,2
	Woche 36	140	-0,06	3,80	0,10	-18,3	18,5
	Woche 48	128	0,07	4,02	0,0	-12,3	20,0
	Woche 60	119	-0,12	3,75	0,0	-14,3	25,0
	Woche 72	104	-0,35	3,05	0,0	-14,3	6,7
	Woche 84	93	-0,18	3,20	0,0	-15,3	6,7
	Woche 96	52	-0,05	4,48	0,0	-13,3	19,0
	Woche 108	25	-0,64	3,65	0,0	-7,0	5,0
	Woche 120	12	-0,53	4,31	-0,8	-7,0	8,2
	Woche 132	1	-3,3	NC	-3-3	-3,3	-3,3
Placebo							
	Woche 12	71	0,01	1,75	0,0	-5,0	6,0
	Woche 24	60	-0,09	2,30	0,0	-8,0	9,0
	Woche 36	41	0,32	1,70	0,0	-3,8	5,2
	Woche 48	43	0,64	1,70	0,2	-2,8	6,2
	Woche 60	34	0,27	1,84	0,0	-3,8	5,0
	Woche 72	28	0,08	1,79	0,1	-3,8	3,0
	Woche 84	26	0,20	1,72	0,0	-3,8	4,0
	Woche 96	11	0,34	1,80	-0,3	-1,8	4,0
	Woche 108	8	0,60	3,49	-0,15	-4,8	6,0
	Woche 120	3	-1,60	2,95	-1,0	-4,8	1,0

Zur besseren Übersicht werden die Werte nur nach Woche 12,24,36,48, 60,72,84,96,108,120 und 132 (sofern erhoben) dargestellt; NC = nicht berechenbar



n at weeks	12	24	36	48	60	72	84	96	108	120	132
Vandetanib 300mg	169	169	149	143	137	113	112	58	28	14	1
Placebo	68	62	49	40	37	31	25	14	9	3	0

The baseline score is the average of up 4 of the most recent non-missing scores collected within 28 days prior to randomization.

/art/prod_2515/astrazeneca/oncology/d4200/d4200c00058/prog/reports/qs310.sas qs310e.png 07JUN2010:18:45 art_arv

Abbildung 18: Stuhlfrequenz über die Zeit – Veränderung der Stuhlfrequenz zum Ausgangswert (Gesamtpopulation)

Tabelle 4-77: Veränderung der Stuhlfrequenz gegenüber dem Ausgangswert - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib							
	Woche 12	85	-0,81	3,52	0,0	-15,3	7,2
	Woche 24	74	-0,79	3,49	-0,3	-11,8	10,7
	Woche 36	78	-0,48	4,67	-0,3	-18,3	18,5
	Woche 48	69	-0,41	0,79	0,0	-12,3	20,0
	Woche 60	60	-1,14	3,46	-0,8	-14,3	6,7
	Woche 72	53	-0,91	3,81	-0,3	-14,3	6,7
	Woche 84	51	-0,5	3,82	0,0	-15,3	6,7
	Woche 96	31	-0,04	5,19	0,0	-13,3	19,0
	Woche 108	14	-1,72	3,88	-1,4	-7,0	4,0
	Woche 120	6	-0,45	5,36	-1,8	-7,0	8,2
	Woche 132	1	-3,3	NC	-3,3	-3,3	-3,3
Placebo							
	Woche 12	42	-0,08	1,88	0,0	-4,0	6,0
	Woche 24	35	0,21	2,42	0,0	-3,8	9,0
	Woche 36	27	0,19	1,74	0,0	-3,8	4,0
	Woche 48	24	0,34	1,52	0,35	-2,8	4,0
	Woche 60	19	-0,17	1,97	-0,30	-3,8	5,0
	Woche 72	17	0,11	1,84	0,7	-3,8	2,2
	Woche 84	16	0,21	1,96	0,1	-3,8	4,0
	Woche 96	9	0,48	1,98	0,0	-1,8	4,0
	Woche 108	6	0,82	4,03	-0,15	-4,8	6,0
	Woche 120	3	-1,6	2,95	-1,0	-4,8	1,0

Zur besseren Übersicht werden die Werte nur nach Woche 12,24,36,48, 60,72,84,96,108,120 und 132 (sofern erhoben) dargestellt; NC = nicht berechenbar.

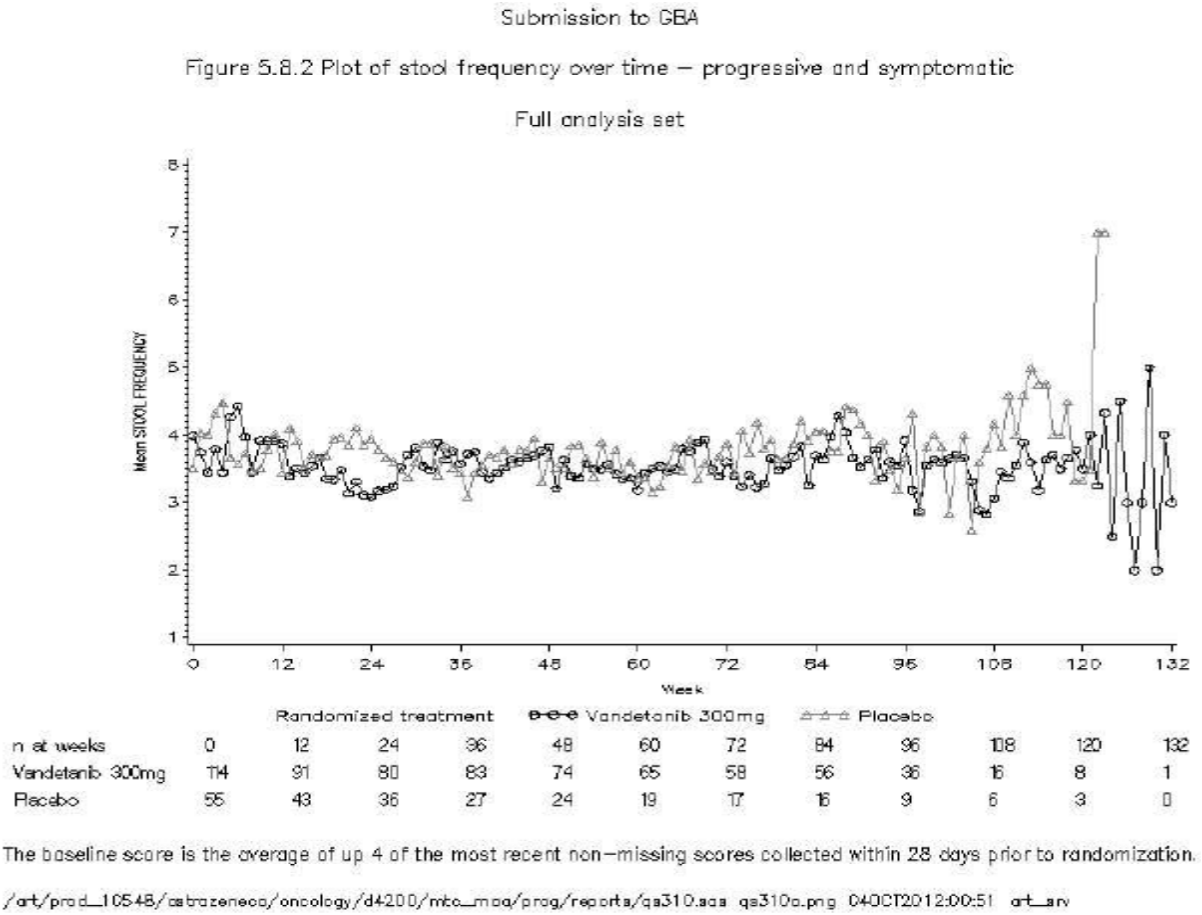


Abbildung 19: Stuhlfrequenz über die Zeit – Veränderung der Stuhlfrequenz gegenüber dem Ausgangswert (genäherte Zielpopulation)

Für diesen Endpunkt wurde keine formale statistische Analyse durchgeführt. Die Veränderung der Stuhlfrequenz ergab sich aus der Differenz der jeweiligen Beobachtung für die Wochen 1-132 gegenüber dem Ausgangswert (Baseline). Es wurden über den langen zeitlichen Verlauf in beiden Gruppen vergleichsweise stabile Gesamtscore-Werte erreicht.

Aufgrund der unterschiedlichen und im Zeitverlauf stark abnehmenden Patientenzahlen ist eine Interpretierbarkeit der Ergebnisse und auch der Veränderung der Stuhlfrequenz schwierig. Beim Vergleich der Werte für die Wochen 0-132 in Tabelle 4-74 und Tabelle 4-76 der beiden Behandlungsgruppen wird deutlich, dass in diesem patientenberichteten Endpunkt keine bedeutenden Unterschiede zwischen den Behandlungsgruppen bestehen. Die grafische Darstellung in Abbildung 18 verdeutlicht die Ergebnisse. Es ergab sich jedoch auch keine Evidenz dafür, dass die Stuhlfrequenz durch Vandetanib nachteilig beeinflusst wurde. Die Zahlen (Tabelle 4-75 und Tabelle 4-77) der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung spiegeln diese Einschätzung wider.

Zum arztberichteten unerwünschten Ereignis „Diarrhö“ verweisen wir auf Abschnitt 4.3.1.3.1.14.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 20: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.8 Explorativer Morbiditätsendpunkt WHO-Performancestatus (WHO-PS)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-78: Operationalisierung des WHO-Performancestatus (WHO-PS)

Studie	Operationalisierung
D4200C00058	Zeit bis zum ersten Nachweis der Verschlechterung des WHO-PS, relativ zum Zeitpunkt der Randomisierung. Verschlechterung des WHO-PS wurde definiert als Erhöhung um mindestens einen Punkt im Vergleich zum Ausgangswert

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-79: Bewertung des Verzerrungspotenzials für den WHO-Performancestatus (WHO-PS) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Daten für die Zeit bis zum ersten Nachweis der Verschlechterung des WHO-PS relativ zum Zeitpunkt der Randomisierung wurden rein explorativ ausgewertet. Eine Verschlechterung des WHO-PS wurde definiert als Erhöhung um mindestens einen Punkt im Vergleich zum Ausgangswert.

Die primäre Analyse für die Zeit bis zur Verschlechterung des WHO-Performancestatus erfolgte auf Basis der ITT-Population mithilfe eines *Log-Rank* Test mit dem Behandlungsarm

als einzigem Faktor. Unterschiede zwischen den Gruppen wurden mithilfe des *Hazard Ratio* und des dazugehörigen Intervallschätzers bewertet. Punktschätzer für die mediane Zeit bis zur Verschlechterung des WHO PS konnten aufgrund der geringen Ereignisrate für die Gesamtpopulation nicht errechnet werden. Die Zeit bis zur Verschlechterung des WHO PS wurde mittels einer Kaplan-Meier-Kurve grafisch dargestellt.

Das endpunktbezogene Verzerrungspotential wurde aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung als auch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Eine mögliche partielle Entblindung von Patienten in der Vandetanib-Gruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

Stellen Sie die Ergebnisse für den Endpunkt "Verschlechterung des WHO-Performancestatus" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-80: Ergebnisse für die Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	75 (32,5)	100	34 (34,0)	0,74** [0,48;1,14]	0,1654*
					-	

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
.

Tabelle 4-81: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib (N=231)**		**Placebo (N=100)**	
	Anzahl Ereignisse	**Mediane Zeit bis zur Verschlechterung des WHO-PS (Monate)**	**Anzahl Ereignisse**	**Mediane Zeit bis zur Verschlechterung des WHO-PS (Monate)**
D4200C00058	75	NC	34	24,9

NC = nicht berechenbar, wegen der geringen Anzahl an Ereignissen in der Vandetanib-Gruppe wurde die mediane Zeit bis zur Verschlechterung des WHO-PS in der Vandetanib-Gruppe nicht erreicht.

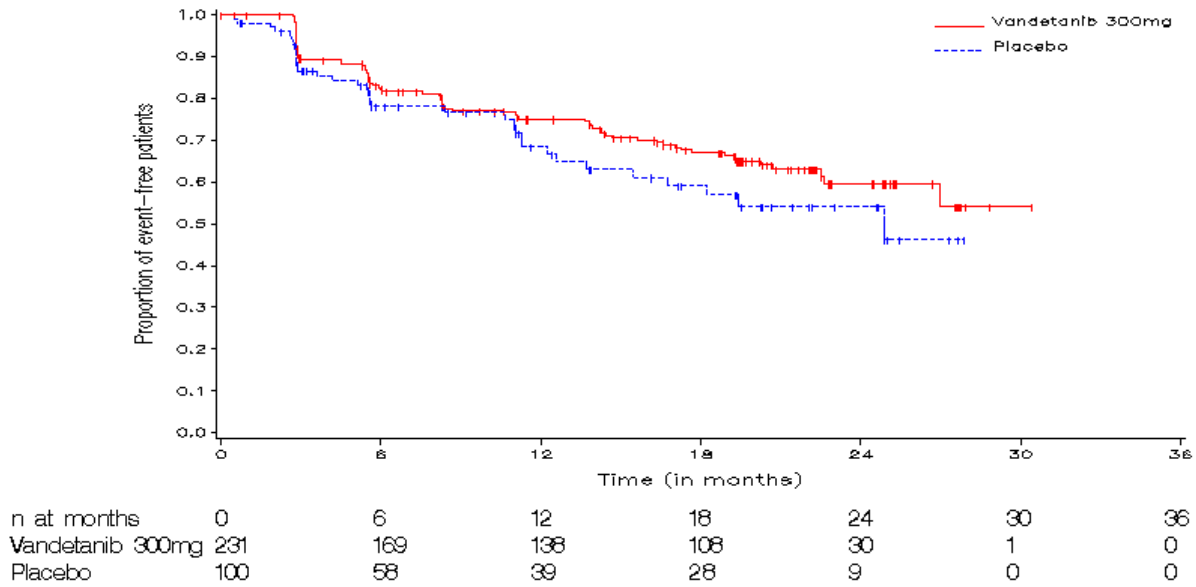


Abbildung 21: Kaplan-Meier-Kurve für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) (Gesamtpopulation → ITT-Population)

Tabelle 4-82: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	49 (38,9)	60	23 (38,3)	0,71** [0,42;1,21]	0,2115*

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor

Tabelle 4-83: Ergebnisse für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib (N=126)**		**Placebo (N=60)**	
	Anzahl Ereignisse	**Mediane Zeit bis zur Verschlechterung des WHO-PS (Monate)**	**Anzahl Ereignisse**	**Mediane Zeit bis zur Verschlechterung des WHO-PS (Monate)**
D4200C00058	49	22,6	23	15,4

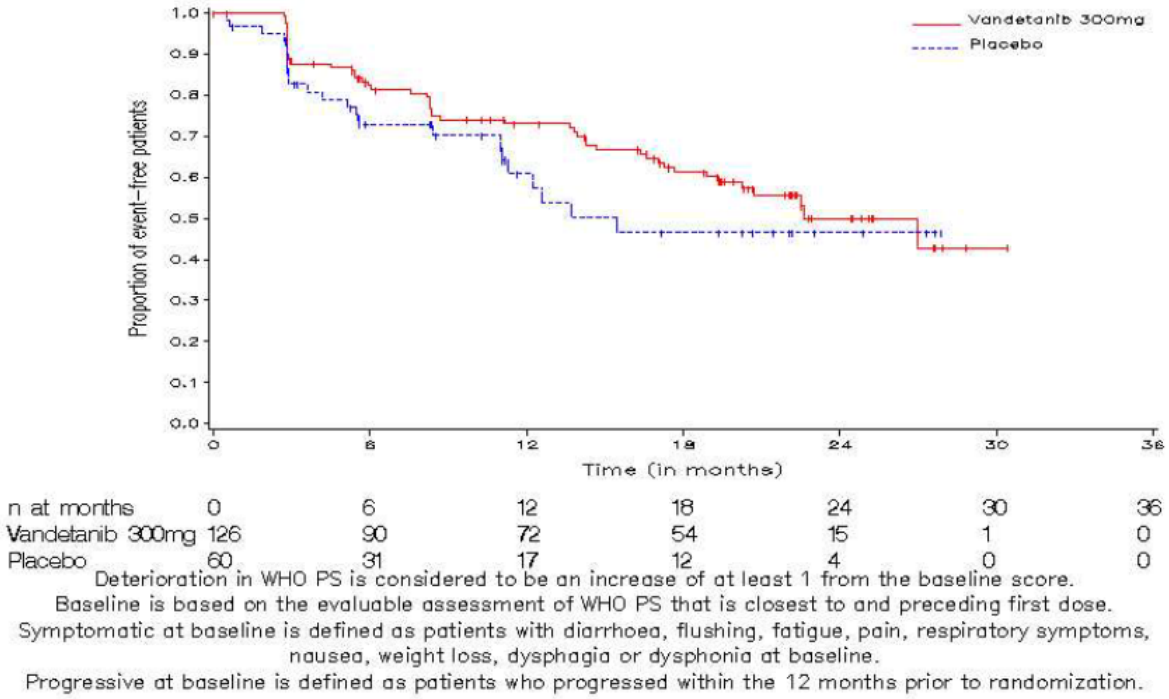


Abbildung 22: Kaplan-Meier-Kurve für die Zeit bis zur Verschlechterung des WHO-Performancestatus (WHO-PS) (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)

Für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigte sich kein deutlicher Unterschied in den Ergebnissen ((HR: 0,74 95%-KI: [0,48;1,14]) bzw. (HR: 0,71 95%-KI: [0,42; 1,21])). Die visuelle Prüfung der Kaplan-Meier-Kurven für die Zeit bis zur Verschlechterung des WHO-PS, sowie der deutlich unter 1 liegende Effektschätzer, deuten jedoch auf einen leichten Vorteil von Vandetanib gegenüber Placebo hin. Für die Gesamtpopulation konnte die mediane Zeit bis zur Verschlechterung des WHO-PS wegen der geringen Anzahl an Ereignissen in der Vandetanib-Gruppe bis zum Daten *Cut-off* nicht errechnet werden. Für die Placebo-Gruppe betrug die mediane Zeit bis zur Verschlechterung des WHO-PS 24,9 Monate. Die mediane Zeit bis zur Verschlechterung des WHO-PS für die

Vandetanib-Gruppe liegt in der genäherten Zielpopulation bei 22,6 Monaten, was, verglichen mit 15,4 Monaten in der Placebo-Gruppe, einer Verlängerung um ca. 7 Monate entspricht.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 23: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.9 Explorativer Morbiditätsendpunkt Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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 des TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn *(Baseline)*

Studie	Operationalisierung
D4200C00058	Patientenberichteter Schmerzmittelgebrauch durch ein wöchentlich geführtes Tagebuch bis zur Beendigung der randomisierten Behandlungphase. Patienten wurden gebeten, ihren Schmerzmittelgebrauch (aus jedem Grund) über die letzten 24 Stunden zu berichten. Der Schmerzmittelgebrauch wurde zu Vergleichszwecken in Morphinsulfat-Äquivalente in mg/Tag konvertiert. TWP in Abhängigkeit vom Opioid-Gebrauch wurde operationalisiert als: - Patienten ohne behandlungsbedürftigen Schmerz zu Studienbeginn mit einem Opioid-Gebrauch von <10 mg/Tag - Patienten mit behandlungsbedürftigem Schmerz zu Studienbeginn mit einem Opioid-Gebrauch von ≥ 10 mg/Tag - Patienten mit unbekanntem Opioid-Gebrauch

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-85: Bewertung des Verzerrungspotenzials für die Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline) in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Hinweise auf ergebnisgesteuerte Berichterstattung**	**Sonstige das Verzerrungspotenzial beeinflussende Punkte**	**Verzerrungspotenzial Endpunkt**
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der patientenberichtete Schmerzmittelgebrauch wurde durch ein wöchentlich geführtes Tagebuch bis zur Beendigung der randomisierten Behandlung erfasst. Patienten wurden gebeten, ihren Schmerzmittelgebrauch (aus jedem Grund) über die letzten 24 Stunden zu berichten. Der Schmerzmittelgebrauch wurde zu Vergleichszwecken in Morphinsulfat-Äquivalente in mg/Tag konvertiert. Der Endpunkt „Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn“ wurde rein explorativ ausgewertet und stellt in seiner Form eine Subgruppenanalyse zum TWP dar. TWP nach Opioid-Gebrauch wurde operationalisiert als i) Patienten ohne (behandlungsbedürftigen) Schmerz zu Studienbeginn mit einem Opioid-Gebrauch von <10 mg/Tag; ii) Patienten mit behandlungsbedürftigem Schmerz zu Studienbeginn mit einem Opioid-Gebrauch von ≥ 10 mg/Tag; iii) Patienten mit unbekanntem Opioid-Gebrauch.

Die Analyse für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn erfolgte auf Basis der ITT-Population mithilfe eines *Log-Rank* Test mit dem Behandlungsarm als einzigem Faktor. Schmerzprogression ohne Verbesserung innerhalb der nächsten 14 Tage wurde als „Verschlechterung“ definiert. Unterschiede zwischen den Gruppen wurden mithilfe des *Hazard Ratio* und des dazugehörigen Intervallschätzers bewertet.

Das endpunktbezogene Verzerrungspotential wurde aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

Stellen Sie die Ergebnisse für den Endpunkt "Zeit bis zur Schmerzprogression in Abhängigkeit vom Opioid-Gebrauch" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-86: Ergebnisse für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**
≥ 10 mg/Tag Opioidanalgetika Gebrauch	Vandetanib Placebo	24 16	13 (54,2) 14 (87,5)	0,36	[0,16;0,84]	0,0176
< 10 mg/Tag Opioidanalgetika Gebrauch	Vandetanib Placebo	168 64	89 (53,0) 37 (57,8)	0,68	[0,44;1,03]	0,0698
Opioidanalgetika Gebrauch unbekannt	Vandetanib Placebo	39 20	12 (30,8) 6 (30,0)	1,41	[0,53;3,74]	0,4936

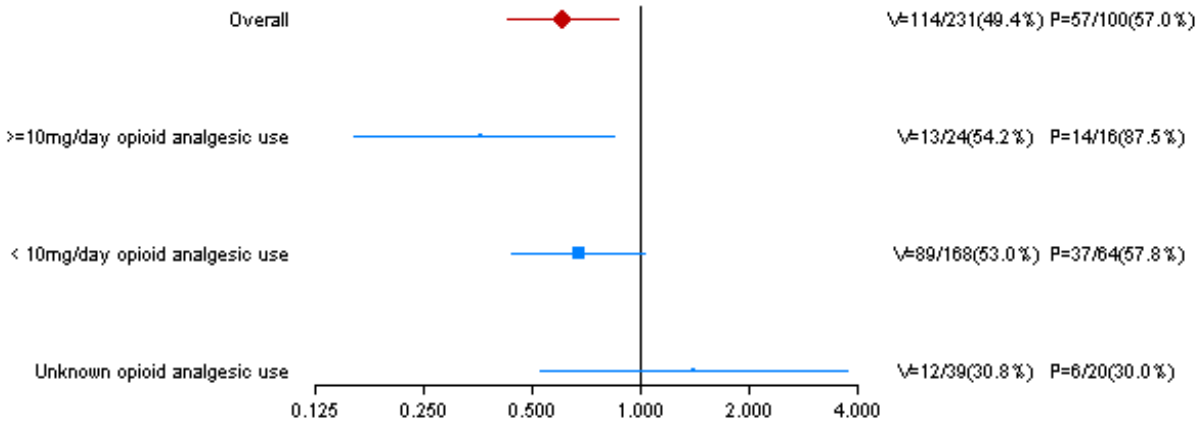
* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
Schmerzprogression wird als „Verschlechterung“ ohne Verbesserung innerhalb der nächsten 14 Tage erachtet; der Gebrauch von Opioidanalgetika zum Ausgangswert wurde zu Morphinsulfat- Äquivalenten in mg/Tag konvertiert

.

Tabelle 4-87: Ergebnisse für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

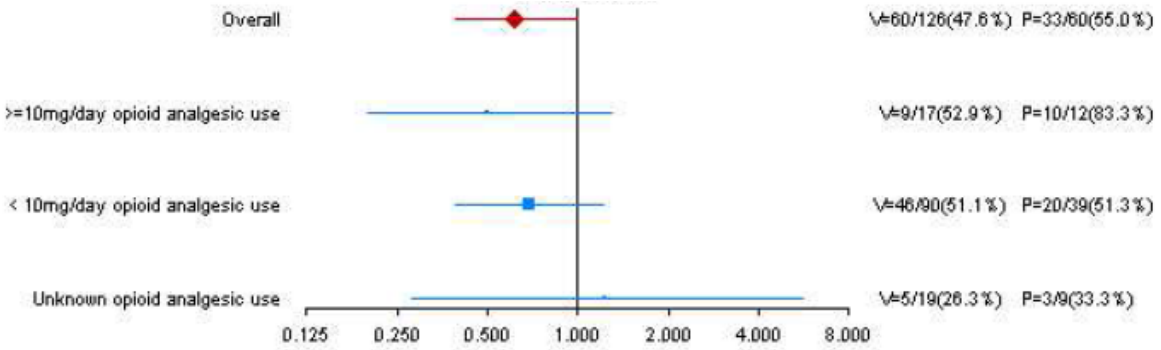
Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**
≥ 10 mg/Tag Opioidanalgetika-Gebrauch	Vandetanib Placebo	17 12	9 (52,9) 10 (83,3)	0,50	[0,20;1,30]	0,1573
< 10 mg/Tag Opioidanalgetika-Gebrauch	Vandetanib Placebo	90 39	46 (51,1) 20 (51,3)	0,69	[0,39;1,22]	0,1978
Opioidanalgetika-Gebrauch unbekannt	Vandetanib Placebo	19 9	5 (26,3) 3 (33,3)	1,24	[0,28;5,62]	0,7759

* *Log-Rank* Test
** Cox-PH-Modell mit Behandlung als einzigem Faktor
Schmerzprogression wird als „Verschlechterung“ ohne Verbesserung innerhalb der nächsten 14 Tage erachtet; der Gebrauch von Opioidanalgetika zum Ausgangswert wurde zu Morphinsulfat- Äquivalenten in mg/Tag konvertiert



V = Vandetanib, P = Placebo

Abbildung 24: *Forest-Plot* für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn – Gesamtpopulation



V = Vandetanib, P = Placebo

Abbildung 25: *Forest*-Plot für die TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Die visuelle Prüfung des *Forest-Plots* für die Gesamtpopulation spiegelt die Konsistenz des TWP über die zwei Gruppen mit bekanntem Opioidanalgetika Gebrauch hinweg wider. In der Subgruppe „≥10 mg/Tag Opioidanalgetika-Gebrauch“ zeigte sich mit einem *Hazard Ratio* von 0,36 (95%-KI: [0,16;0,84], p=0,0176) eine deutliche Überlegenheit der Vandetanib-Gruppe bezüglich des TWP. In der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika-Gebrauch“, liegt das *Hazard Ratio* bei 0,68 mit einem 95%-KI [0,44;1,03] und einem p-Wert von 0,0698.

Die visuelle Prüfung des *Forest-Plots* für die genäherte Zielpopulation spiegelt die Konsistenz des TWP über die zwei Gruppen mit bekanntem Opioidanalgetika-Gebrauch hinweg wider. Für die Gruppe „≥10 mg/Tag Opioidanalgetika-Gebrauch“ (Hazard Ratio 0,50 (95%-KI:

[0,20;1,30], p=0,1573)) und die komplementäre Gruppe „<10 mg/Tag Opioidanalgetika-Gebrauch“ (Hazard Ratio von 0,69 (95%-KI [0,39;1,22], p-Wert= 0,1978) wiesen die Effektschätzer zwar auf einen Vorteil von Vandetanib hin, statistische Signifikanz lag aber (vermutlich aufgrund der geringen Gruppengröße von nur 29 Patienten) nicht mehr vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 26: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.10 Explorativer Morbiditätsendpunkt Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-88: Operationalisierung der Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn *(Baseline)*

Studie	**Operationalisierung**
D4200C00058	Die Reduktion des Opioid-Gebrauches ist definiert als Reduzierung um >50% des Schmerzmittelgebrauchs in Relation zum Ausgangswert, wobei eine Reduktion zu jedem Zeitpunkt innerhalb der Studie erfolgen konnte. Nach der Reduzierung um >50% durfte innerhalb der darauf folgenden 14 Tage keines der genannten Ereignisse eintreten: - Anstieg des Schmerzmittelgebrauchs gegenüber dem Ausgangwert oder Schmerzmittelgebrauch analog zum Ausgangswert - Rückgang des Schmerzmittelgebrauchs um weniger als und bis zu 50% relativ zum Ausgangswert - Verschlechterung im patientenberichteten Endpunkt “stärkster Schmerz” des BPI Für die Reduktion des Opioid-Gebrauchs wurden nur Patienten mit einem Schmerzmittelgebrauch äquivalent zu ≥10 mg/Tag Morphinsulfat berücksichtigt. Für Patienten mit bestätigter Reduktion des Opioid-Gebrauchs wurde die Dauer der Reduktion des Opioid-Gebrauchs gemessen. Die Dauer der Reduktion des Opioid-Gebrauchs wurde operationalisiert als Zeit vom Datum der ersten bestätigten Schmerzmittelreduktion bis zum Zeitpunkt, an dem eines der folgenden Ereignisse eintrat: - Anstieg des Schmerzmittelgebrauchs gegenüber dem Ausgangwert oder Schmerzmittelgebrauch analog zum Ausgangswert - Rückgang des Schmerzmittelgebrauchs um weniger als und bis zu 50% relativ zum Ausgangswert - Verschlechterung im patientenberichteten Endpunkt “stärkster Schmerz” des BPI - Tod des Patienten Für Patienten, bei denen zum Zeitpunkt des Daten *Cut-Off* keines dieser vier Ereignisse eingetraten war, wurde die Dauer der Schmerzmittelreduktion als Zeit vom Datum der ersten bestätigten Schmerzmittelreduktion bis zum Zeitpunkt des letzten Eintrags zum Schmerzmittelgebrauch in das Patiententagebuch herangezogen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-89: Bewertung des Verzerrungspotenzials für Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (*Baseline*) in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Hinweise auf ergebnisgesteuerte Berichterstattung**	**Sonstige das Verzerrungspotenzial beeinflussende Punkte**	**Verzerrungspotenzial Endpunkt**
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Der explorative Endpunkt „Reduktion des Schmerzmittel-Gebrauchs zum Ausgangswert“ wurde als Reduzierung des Schmerzmittelgebrauchs um >50% in Relation zum Ausgangswert operationalisiert. Eine Reduktion konnte zu jedem Zeitpunkt innerhalb der Studie erfolgen. Der Reduktion durfte innerhalb der darauffolgenden 14 Tage keines der unten angeführten Ereignisse folgen: i) Anstieg des Schmerzmittelgebrauchs gegenüber dem Ausgangwert oder Schmerzmittelgebrauch analog zum Ausgangswert; ii) Rückgang des Schmerzmittelgebrauchs um weniger als und bis zu 50% relativ zum Ausgangswert; iii) Verschlechterung im patientenberichteten Endpunkt “stärkster Schmerz”. Für die Reduktion des Opioid-Gebrauchs wurde nur die Subgruppe der Patienten mit einem Schmerzmittelgebrauch von ≥10 mg/Tag Morphinsulfat-Äquivalenten berücksichtigt. Trat keines der drei oben genannten Ereignisse ein, lag eine bestätigte Reduktion des Opioid-Gebrauchs vor. Daten zur Reduktion des Opioid-Gebrauchs wurden ausschließlich während der randomisierten Behandlungsphase erhoben.

Für die Analyse der Reduktionsrate des Opioid-Gebrauchs gegenüber dem Ausgangswert wurde auf Basis der ITT-Population eine logistische Regression mit dem Behandlungsarm als einzigem Faktor durchgeführt. Die Behandlungsunterschiede wurden anhand des *Odds Ratio* und des zugehörigen Konfidenzintervalls bewertet.

Darüber hinaus wurde die mediane Dauer der Reduktion des Schmerzmittelgebrauchs dargestellt. Diese beinhaltet ausschließlich Patienten mit einer bestätigten Reduktion des Opioid-Gebrauchs. Die Dauer der Dauer der Reduktion des Opioid-Gebrauchs war dabei definiert als Zeit von der ersten bestätigten Reduktion bis zum Auftreten eines der drei (i)-iii)) oben genannten Ereignisse oder Tod. Für Patienten, bei denen zum Zeitpunkt des Daten *Cut-off* keines der oben genannten Ereignisse eingetreten war, wurden die Zeit vom Beginn der Reduktion bis zur letzten vervollständigten Opioidmessung in der statistischen Auswertung berücksichtigt.

Das endpunktbezogene Verzerrungspotential wurde aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips als „niedrig“ bewertet. Es konnten weder Hinweise

auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

Stellen Sie die Ergebnisse für den Endpunkt "Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-90: Ergebnisse für die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Opioid-Reduktion (%)**	**N**	**Anzahl Patienten mit Opioid-Reduktion (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	24	15 (62,5)	16	6 (37,5)	2,78** [0,77;10,80]	0,1193*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-91: Ergebnisse für die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Opioid-Reduktion (%)**	**N**	**Anzahl Patienten mit Opioid-Reduktion (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	17	9 (52,9)	12	4 (33,3)	2,25** [0,50;11,27]	0,2927*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-92: Ergebnisse für die mediane Dauer der Reduktion des Opioid-Gebrauchs - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**Mediane Dauer (Monate)**	**N**	**Mediane Dauer (Monate)**
D4200C00058	15	11,01	6	6,51

Tabelle 4-93: Ergebnisse für die mediane Dauer der Reduktion des Opioid-Gebrauchs - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation)

Studie	**Vandetanib**		**Placebo**	
	N	**Mediane Dauer (Monate)**	**N**	**Mediane Dauer (Monate)**
D4200C00058	9	11,56	4	6,51

Für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigte sich bei sehr kleinen Gruppengrößen (40 und 29 Patienten) ein Trend in der Reduktion des Schmerzmittelgebrauchs zugunsten von Vandetanib (Gesamtpopulation OR: 2,78; 95%-KI: [0,77;10,80] bzw. genäherte Zielpopulation OR: 2,25; 95%-KI: [0,50;11,27]). Die Reduktion des Schmerzmittelgebrauchs wurde in der Vandetanib-Gruppe bei 62,5% bzw. 52,9% Patienten gemessen verglichen mit 37,5% bzw. 33,3% Patienten in der Placebo-Gruppe. Die mediane Dauer der Reduktion des Opioid-Gebrauchs betrug für die Vandetanib-Gruppe 11,01 bzw. 11,56 Monate und für die Placebo-Gruppe in beiden Populationen 6,51 Monate. Alle Ergebnisse sind aufgrund der geringen Anzahl an Ereignissen mit Vorsicht zu interpretieren.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 27: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.11 Surrogatendpunkt: Biochemisches Ansprechen (CTN; CEA) – 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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-94: Operationalisierung des biochemischen Ansprechens (CTN; CEA)

Studie	**Operationalisierung**
D4200C00058	Die biochemische Ansprechrate für CTN bzw. CEA ist der Anteil der Patienten, deren bestes biochemisches Ansprechen einem kompletten Ansprechen (CR, *Complete Response*) oder partiellen Ansprechen (PR, *Partial Response*) entspricht. Es waren nur Erhebungen während der randomisierten Behandlung zu berücksichtigen. Die Analyse der Spiegel erfolgte durch ein Zentrallabor. CR: Vollständige Normalisierung des CTN- bzw. CEA-Spiegels (CTN: ≤10 pg/ml für Männer und ≤5 pg/ml für Frauen; CEA: ≤2,5 pg/ml), die durch eine Wiederholungsmessung >4Wochen später bestätigt sein muss PR: Eine Abnahme des CTN- bzw. CEA-Spiegels um mindestens 50% vom CTN- bzw. CEA-Ausgangswert (*Baseline*), die durch eine Wiederholungsmessung >4 Wochen später bestätigt sein muss

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-95: Bewertung des Verzerrungspotenzials für biochemisches Ansprechen (CTN; CEA) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die biochemische Ansprechrate ist operationalisiert als Anteil der Patienten, deren bestes biochemisches Ansprechen einem kompletten Ansprechen (CR, *Complete Response*) oder partiellen Ansprechen (PR, *Partial Response*) entspricht. CR ist definiert als vollständige Normalisierung des CTN- bzw. CEA-Spiegels (CTN: ≤10 pg/ml für Männer und ≤5 pg/ml für Frauen; CEA: ≤2,5 pg/ml), die durch eine Wiederholungsmessung >4 Wochen später bestätigt sein muss. PR ist definiert als Abnahme des CTN- bzw. CEA-Spiegels um mindestens 50% vom CTN- bzw. CEA-Ausgangswert (*Baseline*), die durch eine Wiederholungsmessung >4 Wochen später bestätigt sein muss.
Für die biochemische Ansprechrate wurden nur Spiegel herangezogen, die bei Visiten (außer in Woche 1 und 2) unter randomisierter Behandlung gewonnen wurden. Bei einem fehlenden Wert wurde das biochemische Ansprechen nicht ausgewertet. Die Analyse der Spiegel erfolgte durch ein Zentrallabor.

Zur Analyse der biochemischen Ansprechrate wurde auf Basis der ITT-Population eine logistische Regression mit dem Behandlungsarm als Faktor durchgeführt. Die Behandlungsunterschiede wurden anhand des *Odds Ratio* und des zugehörigen Konfidenzintervalls bewertet. Die Zusammenfassung beinhaltet ebenfalls Details über die Anzahl (%) der Patienten innerhalb jeder BOR Kategorie.

Die Analyse der CTN- bzw. CEA-Spiegel ist durch ein unabhängiges, verblindetes, Zentrallabor durchgeführt worden, wodurch eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen ist. Da nur Messungen unter Behandlung gemäß Randomisierung – und somit während der verblindeten Studienphase – zur Berechnung der biochemischen Ansprechrate herangezogen worden sind, liegt hier kein Verzerrungspotential auf Grund eines vorzeitigen Wechsels (*Cross-Over*) von Placebo auf *Open-Label* Vandetanib vor.

Unter Einbezug aller oben angeführten Aspekte sowie auch aufgrund der Ergebnisdarstellung und der adäquaten Umsetzung des ITT-Prinzips wurde das endpunktbezogene Verzerrungspotential als „niedrig“ bewertet. Es konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende

Punkte gefunden werden. Dies gilt auch für die *Post-hoc*-Analyse der Teilpopulation mit symptomatischer und progressiver Erkrankung.

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

Tabelle 4-96: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib 300 mg**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	160 (69,3)	100	3 (3,0)	72,86** [26,22;303,2]	<0,0001*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-97: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib 300 mg**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	85 (67,5)	60	2 (3,3)	60,12** [17,56;378,0]	<0,0001*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-98: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

			Anzahl (%) der Patienten		
Studie	**Ansprech-Status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=231)**	**Placebo (N=100)**	**Gesamt (N=331)**
D4200C00058	Ansprechen	Gesamt	160 (69,3)	3 (3,0)	163 (49,2)
		Komplettes Ansprechen	3 (1,3)	0 (0,0)	3 (0,9)
		Partielles Ansprechen	157 (68,0)	3 (3,0)	160 (48,3)
	Kein Ansprechen	Gesamt	71 (30,7)	97 (97,0)	168 (50,8)
		Stabile Erkrankung	58 (25,1)	76 (76,0)	134 (40,5)
		Progressive Erkrankung	4 (1,7)	8 (8,0)	12 (3,6)
		Nicht auswertbar	9 (3,9)	13 (13,0)	22 (6,6)

Tabelle 4-99: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

			Anzahl (%) der Patienten		
Studie	**Ansprech-Status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=126)**	**Placebo (N=60)**	**Gesamt (N=186)**
D4200C00058	Ansprechen	Gesamt	85 (67,5)	2 (3,3)	87 (46,8)
		Komplettes Ansprechen	1 (0,8)	0 (0,0)	1 (0,5)
		Partielles Ansprechen	84 (66,7)	2 (3,3)	86 (46,2)
	Kein Ansprechen	Gesamt	41 (32,5)	58 (96,7)	99 (53,2)
		Stabile Erkrankung	32 (25,4)	45 (75,0)	77 (41,4)
		Progressive Erkrankung	2 (1,6)	7 (11,7)	9 (4,8)
		Nicht auswertbar	7 (5,6)	6 (10,0)	13 (7,0)

Tabelle 4-100: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib 300 mg**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	119 (51,5)	100	2 (2,0)	52,03** [15,95;320,3]	<0,0001*

* Logistische Regression
** mit Behandlung als einzigem Faktor

Tabelle 4-101: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib 300 mg**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Patienten mit Ansprechen (%)**	**N**	**Anzahl Patienten mit Ansprechen (%)**	***Odds Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	58 (46,0)	60	0 (0,0)	NC	NC

NC= nicht berechnet (kein Ansprechen in der Placebogruppe der genäherten Zielpopulation)

Tabelle 4-102: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

			Anzahl (%) der Patienten		
Studie	**Ansprech-status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=231)**	**Placebo (N=100)**	**Gesamt (N=331)**
D4200C00058	Ansprechen	Gesamt	119 (51,5)	2 (2,0)	121 (36,6)
		Komplettes Ansprechen	7 (3,0)	2 (2,0)	9 (2,7)
		Partielles Ansprechen	112 (48,5)	0 (0,0)	112 (33,8)
	Kein Ansprechen	Gesamt	112 (48,5)	98 (98,0)	210 (63,4)
		Stabile Erkrankung	82 (35,5)	77 (77,0)	159 (48,0)
		Progressive Erkrankung	20 (8,7)	8 (8,0)	28 (8,5)
		Nicht auswertbar	10 (4,3)	13 (13,0)	23 (6,9)

Tabelle 4-103: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

			Anzahl (%) der Patienten		
Studie	**Ansprech-status**	**Bestes objektives Ansprechen (BOR)**	**Vandetanib 300 mg (N=126)**	**Placebo (N=60)**	**Gesamt (N=186)**
D4200C00058	Ansprechen	Gesamt	58 (46,0)	0 (0,0)	58 (31,2)
		Komplettes Ansprechen	2 (1,6)	0 (0,0)	2 (1,1)
		Partielles Ansprechen	56 (44,4)	0 (0,0)	56 (30,1)
	Kein Ansprechen	Gesamt	68 (54,0)	60 (100,0)	128 (68,8)
		Stabile Erkrankung	51 (40,5)	48 (80,0)	99 (53,2)
		Progressive Erkrankung	9 (7,1)	5 (8,3)	14 (7,5)
		Nicht auswertbar	8 (6,3)	7 (11,7)	15 (8,1)

Tabelle 4-96 bzw. Tabelle 4-97 und Tabelle 4-100 bzw. Tabelle 4-101 zeigen eine statistisch signifikante Verbesserung der biochemischen Ansprechrate für Vandetanib verglichen mit Placebo. Für CTN liegt das *Odds Ratio* bei 72,86 (60,12 in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung) mit einem 95%-Konfidenzintervall von [26,33;303,2] ([17,56;378,0]) und einem p-Wert von <0,0001 in beiden betrachteten Populationen. Die CTN-Ansprechraten betragen 69,3% (bzw. 67,5%) für Vandetanib und 3,0% (bzw. 3,3%) für Placebo. Für CEA liegt das *Odds Ratio* in der Gesamtpopulation bei 52,03 mit einem 95%igen Konfidenzintervall von [15,95;320,3] und einem p-Wert von <0,0001; für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung wurde aufgrund fehlenden Ansprechens in der Placebo-Gruppe kein *Odds Ratio* berechnet. Die CEA-Ansprechraten betragen 51,5% (bzw. 46,0%) für Vandetanib und 2,0% (bzw. 0,0%) für Placebo.

Zusammen betrachtet, unterstreichen die Resultate hinsichtlich des beobachteten Effekts einer Vandetanib-Behandlung auf diese beiden wichtigen biochemischen Parameter die Resultate der meisten Endpunkte zur Morbidität, insbesondere PFS und ORR, auf eindrucksvolle Weise.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 28: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

4.3.1.3.1.12 Gesamtüberleben (OS)

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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

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

Tabelle 4-104: Operationalisierung von Gesamtüberleben (OS)

Studie	Operationalisierung
D4200C00058	Gesamtüberleben (OS), definiert als Zeit zwischen Randomisierung und Tod unabhängig von der zugrundeliegenden Ursache.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-105: Bewertung des Verzerrungspotenzials für Gesamtüberleben (OS) in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja	nein	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Gesamtüberleben (OS) ist definiert als Zeit zwischen Randomisierung und Tod, unabhängig von der zugrundeliegenden Ursache. Die Analyse des OS erfolgt mithilfe einer Cox Regression und Kaplan-Meier-Methodik. Patienten, welche keine Studienmedikation mehr erhalten, werden bis zum Eintreten des Todes oder dem Abbruch der Studie alle 12 Wochen beobachtet. Patienten, welche zum Zeitpunkt der statistischen Analyse noch leben, werden zum letzten bekannten Überlebenszeitpunkt zensiert.

Für die primäre Analyse der Gesamtüberlebensrate wurde auf Basis der ITT-Population ein *Log-Rank* Test durchgeführt. Unterschiede zwischen den Gruppen wurden mithilfe des *Hazard Ratio* und des dazugehörigen Intervallschätzers bewertet. Punktschätzer für das mediane Gesamtüberleben konnten aufgrund der geringen Ereignisrate nicht errechnet werden. OS wurde mittels einer Kaplan-Meier-Kurve grafisch dargestellt.

Unterstützend wurden Sensitivitätsanalysen zur Überprüfung der Robustheit der Ergebnisse durchgeführt. Diese Sensitivitätsanalysen adressieren sowohl Variationen in der Analysepopulation (*Per Protocol* statt ITT) als auch methodische Variationen (Adjustierung für Kovariaten). So wurden unter anderem adjustierte Schätzer für das *Hazard Ratio* berechnet, um den möglichen Einfluss von Unbalanciertheiten bezüglich *Baseline-*

Charakteristika zwischen den Behandlungsgruppen auf den beobachteten Behandlungseffekt zu beurteilen. Das adjustierte Cox-PH-Model enthielt folgende Kovariaten: RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status (hereditär, sporadisch, unbekannt).

Der Endpunkt „Tod“ ist stets objektiv zu bewerten und ist durch eine mögliche Entblindung aufgrund des typischen Nebenwirkungsprofils nicht beeinflusst.

Die Aussagekraft der Ergebnisse zum Endpunkt „Gesamtüberleben“ wird allerdings durch die im Studienprotokoll vorgesehene Möglichkeit für Patienten aus der Placebo-Gruppe zum Wechsel auf *Open-Label* Vandetanib nach Krankheitsprogression stark eingeschränkt. Als Resultat dieses *Cross-Over* wird der Effekt in der Placebo-Gruppe durch die nachfolgenden Effekte von Vandetanib verschleiert. Dies legt eine Verzerrung des Studienergebnisses zuungunsten von Vandetanib nahe, erhöht also die Wahrscheinlichkeit einer Unterschätzung des Effekts von Vandetanib. Das endpunktbezogene Verzerrungspotential wird insgesamt als „hoch“ bewertet. Für eine gesonderte Betrachtung der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ändert sich an dieser Einschätzung nichts.

Die Gesamtüberlebensrate soll erneut untersucht werden, wenn 50% der Patienten in der Gesamtpopulation verstorben sind. Diese Ergebnisse müssen dann, unter Einbezug der zu *Open-Label* Vandetanib übergegangenen Placebo-Patienten, erneut interpretiert werden.

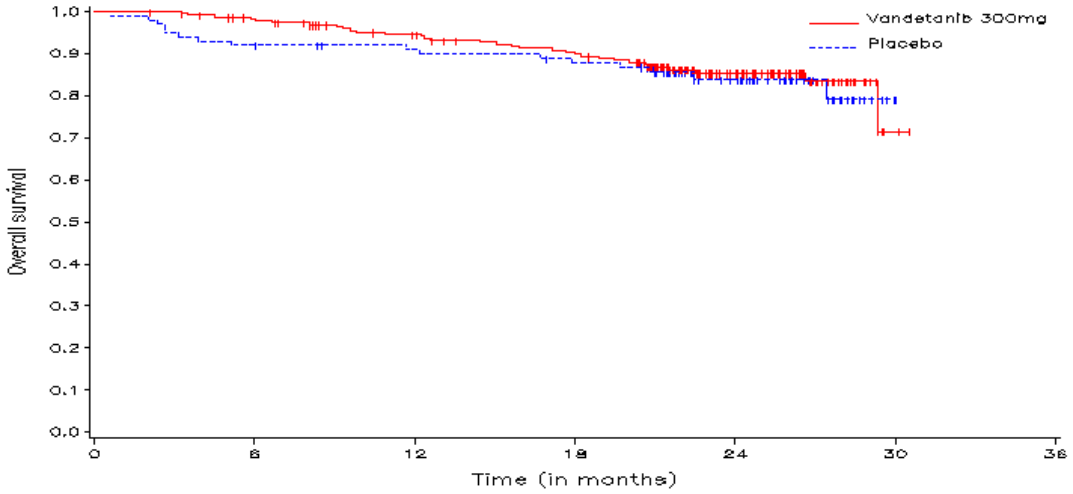
Unter Berücksichtigung aller oben genannten Aspekte und durch den Umstand, dass die Ergebnisse der Gesamtüberlebensrate durch die geringe Zahl an beobachteten Todesfällen sowie die Option der *Open-Label* Therapie mit Vandetanib nach Beendigung der verblindeten Behandlung verzerrt sein könnten, wird im Rahmen des Dossiers auf die weitere Bewertung dieses Endpunktes verzichtet.

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.

Tabelle 4-106: Ergebnisse für das Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	231	32 (13,9)	100	16 (16,0)	0,89* [0,28;2,85]	0,7115**
					0,91*** [0,28;2,93]	0,7693

* Cox-PH-Modell mit Behandlung als einzigem Faktor
** *Log-Rank* Test
*** Sensitivitätsanalyse: Cox-PH-Modellmit Behandlung als Faktor und den Kovariaten RET Mutationsstatus, Zeitraum bis zur Verdopplung des Calcitoninwertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär, sporadisch, unbekannt).



n at months	0	6	12	18	24	30	36
Vandetanib 300mg	231	221	202	189	81	2	0
Placebo	100	92	88	84	44	0	0

Abbildung 29: Kaplan-Meier-Kurve für das Gesamtüberleben (Gesamtpopulation → ITT-Population)

Tabelle 4-107: Ergebnisse für das Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib**		**Placebo**		**Behandlungseffekt (Vandetanib:Placebo)**	
	N	**Anzahl Ereignisse (%)**	**N**	**Anzahl Ereignisse (%)**	***Hazard Ratio* [95%-KI]**	**2-seitiger p-Wert**
D4200C00058	126	21 (16,7)	60	10 (16,7)	1,06** [0,26;4,38]	0,8791*
					1,15*** [0,25;5,22]	0,7355

* *Log-Rank* Test
** Cox-PH-Modellmit Behandlung als einzigem Faktor
*** Sensitivitätsanalyse: Cox-PH-Modellmit Behandlung als Faktor und den Kovariaten RET Mutationsstatus, Zeitraum bis zur Verdopplung des Calcitoninwertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär, sporadisch, unbekannt).

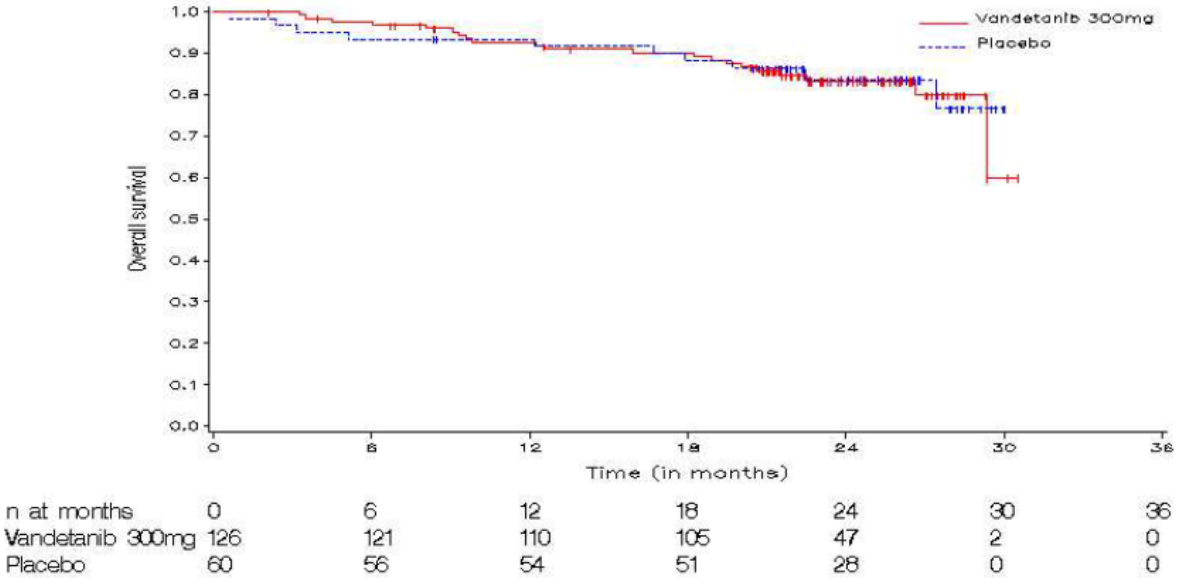


Abbildung 30: Kaplan-Meier-Kurve für das Gesamtüberleben (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung → ITT-Population)

Eine erste Analyse der Gesamtüberlebensrate (Tabelle 4-106) zeigte für die Gesamtpopulation ein *Hazard Ratio* von 0,89 zugunsten von Vandetanib. Der Unterschied zwischen beiden Therapiearmen war nicht statistisch signifikant (KI: [0,28;2,85], p=0,712). Diese Ergebnisse sind unter Vorbehalt zu interpretieren, da wie bereits erwähnt zum Zeitpunkt der Analyse (Daten *Cut-off*) erst 15% der Patienten verstorben waren. Zusätzlich – durch das Protokoll vorab definiert – bestand die Möglichkeit für alle Patienten, während der offenen Studienphase Vandetanib zu erhalten. Dies wurde im Placebo-Arm von 58% der Patienten wahrgenommen (19% im Verum-Arm). Infolgedessen sind die Ergebnisse der Placebo-

Gruppe systematisch verzerrt, der geschätzte Behandlungseffekt ist kaum interpretierbar. Abbildung 29 spiegelt obige Ergebnisse wider. Die finale Auswertung des Gesamtüberlebens wird nach Eintreten von ≥50% Ereignissen erfolgen. Dieser Punkt wird derzeit für 2014 erwartet, wobei die Aussagekraft dieses Endpunktes durch die Möglichkeit des *Cross-Overs* auf *Open-Label* Vandetanib eingeschränkt sein wird.

Eine erste Analyse der Gesamtüberlebensrate zeigt für die genäherte Zielpopulation ein *Hazard Ratio* von 1,06 (Tabelle 4-107). Der Unterschied war nicht statistisch signifikant (KI: [0,26; 4,38], p=0,8791). Allerdings sind die Ergebnisse der *Post-hoc*-Analyse zum Gesamtüberleben innerhalb der Patientengruppe mit symptomatischer und progressiver Erkrankung mit Vorsicht zu interpretieren, da zum einen, wie bereits oben ausgeführt, zum Zeitpunkt der Daten-Analyse (*Cut-off*) erst ein geringer Anteil an Patienten verstorben war und zum anderen die Ergebnisse durch das *Cross-Over* in die *Open-Label* Behandlung nur schwer interpretierbar sind. In der genäherten Zielpopulation erhielten 63,3% der Placebo-Patienten in der offenen Behandlungsphase Vandetanib, sowie 20,6% der Vandetanib-Patienten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 31: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurde keine Meta-Analyse durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.13 Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G – 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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-108: Operationalisierung des Endpunktes Gesundheitsbezogene Lebensqualität

Studie	**Operationalisierung**
D4200C00058	Der gesamte FACT-G Wert wird berechnet aus der Summe der individuellen Subskalen: Physisches Wohlbefinden (PWB), Soziales Wohlbefinden (SWB), Emotionales Wohlbefinden (EWB) und funktionelles Wohlbefinden (FWB). Der Endpunkt wurde operationalisiert als Veränderung verglichen mit dem Ausgangswert.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-109: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die gesundheitsbezogene Lebensqualität wurde mittels des Patientenfragebogens *Functional Assessment of Cancer Therapy General Scale* (FACT-G) [94] bewertet. Hierbei handelt es sich um ein validiertes, patientenrelevantes und weltweit eingesetztes Instrument zur Bewertung der gesundheitsbezogenen Lebensqualität, welches eine gute Korrelation zum klinischen Verlauf einer Erkrankung aufweist. Der Ausgangswert (*Baseline*) wurde durch Erfassung an vier Tagen während der letzten Woche der Screeningperiode ermittelt. Danach war der FACT-G Fragebogen von den Patienten alle 12 Wochen bis zur Beendigung der randomisierten Behandlung auszufüllen. Es wurde FACT-G Version 4 eingesetzt. Dieser setzt sich zusammen aus den vier Subskalen Physisches Wohlbefinden (PWB), Soziales Wohlbefinden (SWB), Emotionales Wohlbefinden (EWB) und Funktionelles Wohlbefinden (FWB). Anhand von 27 Aussagen, die auf einer Skala von 0 – 4 (trifft überhaupt nicht zu ... trifft sehr zu) von den Patienten bewertet werden, können die einzelnen Kategorien mit einem *Score* versehen werden. In den Einzelscores sind Maximalwerte zwischen 24 (EWB) und 28 (KWB, SWB, FWB) zu erreichen, im Summenscore maximal 108 Punkte. Niedrige FACT-Werte sprechen für eine niedrige Lebensqualität [95, 96].

Dieser Endpunkt wurde im Rahmen der Studie rein explorativ ausgewertet. Es wurde keine formale statistische Analyse durchgeführt. Aufgrund der Tatsache, dass keine validen Vorher-Nachher-Vergleiche existieren, da die Mittelwerte zu den unterschiedlichen Zeitpunkten auf unterschiedlichen Stichprobenumfängen basieren und somit innerhalb jeder Gruppe keine „gepaarten“ Vergleiche möglich sind, sind die im Studienbericht dargestellten Ergebnisse nur schwer interpretierbar. Der Anteil der Patienten, die in der Berechnung deskriptiver Statistiken nicht berücksichtigt wurden, war zu allen Zeitpunkten hoch (bezogen auf die Gesamtpopulation bereits in Woche 24 über 20% in der Vandetanib-Gruppe und bei fast 30% in der Placebo-Gruppe) und war zwischen den Behandlungsgruppen nicht balanciert.

Durch das Auftreten charakteristischer Nebenwirkungen besteht theoretisch die Möglichkeit der Entblindung von Patienten. Eine Entblindung kann sich insbesondere bei einer subjektiven Bewertung, wie der der Lebensqualität durch die Patienten, auf das Resultat auswirken. Die Wahrscheinlichkeit einer solchen Verzerrung wurde, wie bereits in Abschnitt 4.3.1.2.2 dargestellt, als „gering“ angesehen. Diese Einschätzung ist dadurch begründet, dass

die Patienten mit dieser symptomatischen Erkrankung wegen der Überlagerung von krankheitsbedingten Symptomen und therapie- sowie placeboassoziierten Nebenwirkungen schwer zwischen beiden differenzieren können.

Aufgrund der nach den Vorgaben der VerfO formell nicht adäquaten Umsetzung des ITT-Prinzips wird das Verzerrungspotential für diesen Endpunkt als „hoch“ bewertet, auch wenn keine Anhaltspunkte für eine ergebnisgesteuerte Berichtserstattung vorliegen und die Resultate nur deskriptiv gegenübergestellt werden. Die Aussagekraft der Ergebnisse zur „gesundheitsbezogenen Lebensqualität“ ist gering. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

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.

Tabelle 4-110: Ergebnisse für die gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	***Baseline***	218	79,11	16,716	81,45	34,1	106,0
	Woche 12	174	77,06	17,323	77,00	25,4	108,0
	Woche 24	177	77,22	17,410	79,50	37,2	108,0
	Woche 36	154	79,68	16,808	81,33	29,5	108,0
	Woche 48	149	78,70	17,114	80,00	30,3	108,0
	Woche 60	141	78,84	17,421	81,00	22,0	108,0
	Woche 72	117	78,93	16,871	79,00	31,5	108,0
	Woche 84	116	79,82	15,629	80,00	36,0	108,0
	Woche 96	62	79,51	15,843	81,00	44,0	108,0
	Woche 108	31	73,50	16,477	70,00	39,0	108,0
	Woche 120	15	72,53	17,230	70,00	42,7	95,0
	Woche 132	1	84,67	NC	84,67	84,7	84,7
Placebo	***Baseline***	92	77,53	14,180	76,40	42,7	105,0
	Woche 12	73	76,97	16,709	78,17	36,0	104,0
	Woche 24	63	78,89	15,478	81,00	32,8	104,0
	Woche 36	51	80,52	15,037	81,83	36,0	106,0
	Woche 48	40	77,82	15,940	78,58	36,0	101,0
	Woche 60	38	77,69	14,194	78,08	39,0	102,0
	Woche 72	32	77,81	17,026	75,00	32,0	105,0
	Woche 84	25	74,67	14,791	72,00	52,8	107,0
	Woche 96	14	75,21	12,085	74,42	49,2	95,0
	Woche 108	9	69,93	10,411	71,50	52,2	84,0
	Woche 120	3	67,22	19,687	60,00	52,2	89,5

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar

Tabelle 4-111: Ergebnisse für die gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	***Baseline***	118	77,40	16,288	78,90	42,0	105,3
	Woche 12	88	75,95	17,708	72,92	41,8	107,0
	Woche 24	100	77,07	18,042	80,57	37,2	108,0
	Woche 36	79	79,30	18,281	81,83	29,5	108,0
	Woche 48	82	78,58	17,329	79,00	30,3	105,0
	Woche 60	74	78,17	18,106	80,33	22,0	108,0
	Woche 72	62	77,60	18,246	79,00	31,5	108,0
	Woche 84	64	78,88	16,158	80,00	36,0	107,0
	Woche 96	37	79,66	14,587	82,00	49,0	108,0
	Woche 108	18	74,14	17,323	74,00	39,0	108,0
	Woche 120	9	74,56	16,224	78,33	42,7	95,0
	Woche 132	1	84,67	NC	84,67	84,7	84,7
Placebo	***Baseline***	56	77,67	14,087	76,40	51,4	105,0
	Woche 12	42	77,03	18,880	78,58	36,0	104,0
	Woche 24	35	81,62	15,594	85,00	52,0	104,0
	Woche 36	31	81,58	17,450	89,00	36,0	106,0
	Woche 48	22	80,12	17,325	83,58	36,0	101,0
	Woche 60	22	79,96	16,074	81,33	39,0	102,0
	Woche 72	18	81,19	19,743	81,75	32,0	105,0
	Woche 84	14	77,85	16,965	75,92	52,8	107,0
	Woche 96	10	75,65	13,954	72,33	49,2	95,0
	Woche 108	6	64,94	8,728	64,50	52,2	77,0
	Woche 120	3	67,22	19,687	60,00	52,2	89,5

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar

Tabelle 4-112: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) gegenüber dem Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	**Woche 12**	168	-2,01	11,171	-1,00	-38,4	24,7
	Woche 24	168	-3,58	11,678	-2,80	-34,7	31,9
	Woche 36	147	-2,32	11,782	-2,10	-34,8	29,0
	Woche 48	142	-2,25	12,743	-1,40	-38,6	40,0
	Woche 60	136	-2,23	13,316	-1,40	-35,0	54,0
	Woche 72	112	-2,07	13,490	-1,30	-44,1	38,0
	Woche 84	111	-1,74	12,693	-1,30	-40,0	44,0
	Woche 96	57	-2,28	12,089	-1,50	-24,8	26,2
	Woche 108	28	-3,27	13,499	-0,90	-46,8	19,0
	Woche 120	14	-6,09	15,155	-5,15	-43,1	20,0
	Woche 132	1	11,20	NC	11,20	11,2	11,2
Placebo	**Woche 12**	69	-1,87	9,107	-1,30	-30,5	16,4
	Woche 24	62	-0,13	9,167	0,20	-21,5	20,4
	Woche 36	49	-0,65	9,701	-0,60	-25,8	29,2
	Woche 48	39	-2,82	11,314	-1,10	-26,0	30,2
	Woche 60	37	-4,09	9,859	-3,00	-25,5	17,0
	Woche 72	31	-2,80	12,739	-3,00	-29,8	29,0
	Woche 84	25	-4,36	10,567	-5,00	-28,0	14,0
	Woche 96	14	-3,10	7,735	-3,45	-19,8	9,9
	Woche 108	9	-4,80	11,078	-3,00	-24,5	9,2
	Woche 120	3	-1,90	3,804	-2,10	-5,6	2,0

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar

Tabelle 4-113: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) gegenüber dem Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel (genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	Erhebungszeitpunkt	n[a]	Mittelwert	SD	Median	Min	Max
D4200C00058							
Vandetanib	**Woche 12**	85	-1,29	11,730	-0,50	-29,2	24,7
	Woche 24	93	-2,32	12,129	-2,00	-33,0	31,9
	Woche 36	75	-1,44	11,374	-2,00	-22,5	27,9
	Woche 48	78	-0,86	11,416	-0,25	-24,2	27,2
	Woche 60	72	-1,68	13,132	-1,15	-26,8	29,2
	Woche 72	59	-1,22	13,792	-0,50	-44,1	26,0
	Woche 84	60	-0,75	11,431	-1,30	-24,8	24,1
	Woche 96	34	-0,94	13,092	0,90	-23,5	26,2
	Woche 108	16	-2,38	14,875	1,45	-46,8	19,0
	Woche 120	8	-5,09	18,494	-2,90	-43,1	20,0
	Woche 132	1	11,20	NC	11,20	11,2	11,2
Placebo	**Woche 12**	40	-2,18	9,875	-1,15	-30,5	16,4
	Woche 24	35	0,37	9,139	0,00	-21,5	20,4
	Woche 36	30	-1,73	9,228	-2,00	-25,8	16,5
	Woche 48	22	-1,95	10,059	-0,95	-25,8	14,0
	Woche 60	21	-4,17	10,967	-4,50	-25,5	17,0
	Woche 72	17	-2,53	13,289	-3,00	-29,8	19,0
	Woche 84	14	-3,51	11,650	-3,25	-28,0	14,0
	Woche 96	10	-4,92	7,688	-5,30	-19,8	5,0
	Woche 108	6	-10,25	8,887	-7,30	-24,5	-2,1
	Woche 120	3	-1,90	3,804	-2,10	-5,6	2,0

[a] Anzahl an Patienten mit vorliegenden Daten
SD = Standardabweichung (*Standard Deviation*)
NC = nicht berechenbar

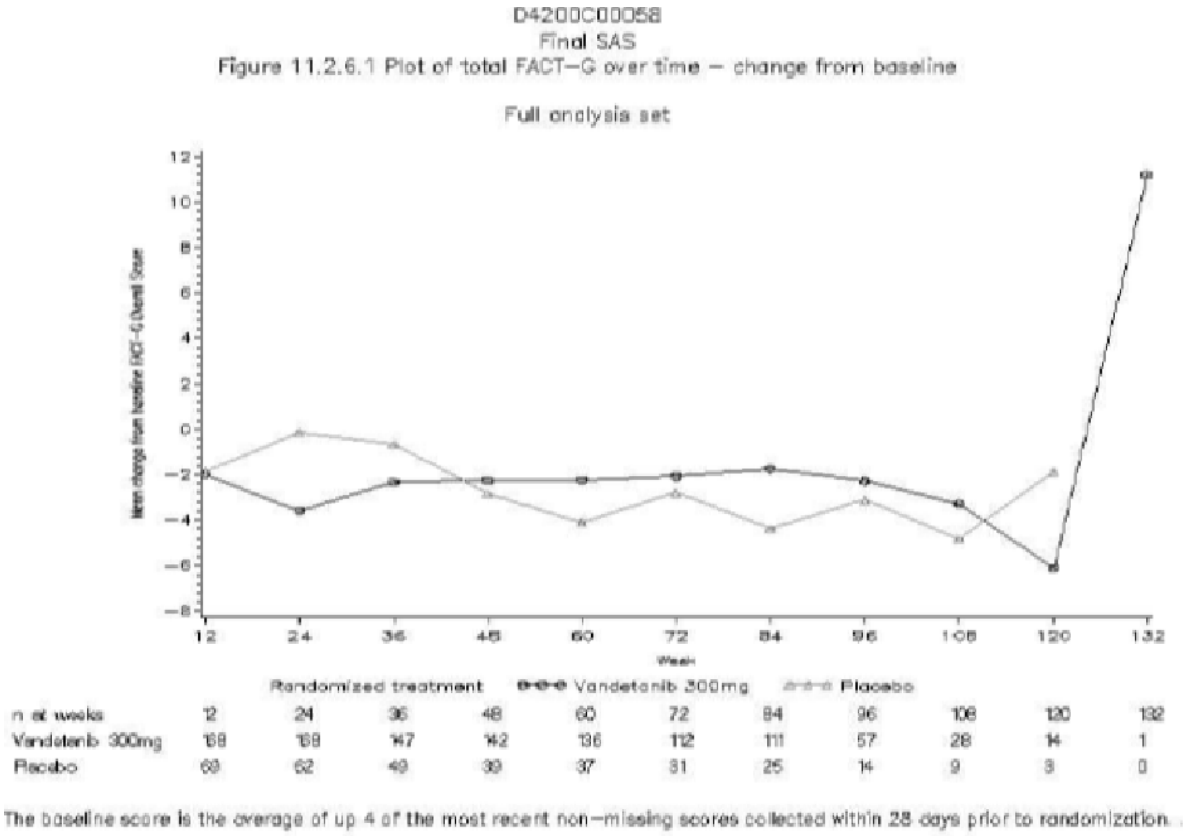


Abbildung 32: FACT-G über die Zeit – Veränderung der gesundheitsbezogenen Lebensqualität gegenüber dem Ausgangswert (Gesamtpopulation)

Für diesen Endpunkt wurde keine formale statistische Analyse durchgeführt. Die Veränderung der gesundheitsbezogenen Lebensqualität anhand des FACT-G ergab sich aus der Differenz der jeweiligen Beobachtung der Monate 12 bis 132 zum Ausgangswert (*Baseline*). Es wurden über den langen zeitlichen Verlauf in beiden Gruppen vergleichsweise stabile Gesamtscore-Werte erreicht, die bis auf ein Einzelereignis in der Vandetanib-Gruppe immer leicht unter dem *Baseline*-Ergebnis lagen.

Aufgrund der unterschiedlichen und im Zeitverlauf stark abnehmenden Patientenzahlen ist eine Interpretierbarkeit der Ergebnisse und auch der Veränderung der gesundheitsbezogenen Lebensqualität schwierig. Beim Vergleich der Werte für die Wochen 0-132 in Tabelle 4-110 und Tabelle 4-112 der beiden Behandlungsgruppen wird deutlich, dass keine bedeutenden Unterschiede zwischen den Behandlungsgruppen bestehen. Die grafische Darstellung (Abbildung 32) verdeutlicht die Ergebnisse. Es ergab sich jedoch auch keine Evidenz dafür, dass die gesundheitsbezogene Lebensqualität durch Vandetanib nachteilig beeinflusst wurde. Die Zahlen der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung spiegeln diese Einschätzung wider (Tabelle 4-111 und Tabelle 4-113).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 33: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurden keine Meta-Analysen durchgeführt (siehe Abschnitt 4.2.5.3).

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.

4.3.1.3.1.14 Unerwünschte Ereignisse (UE) – 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
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

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

Tabelle 4-114: Operationalisierung von Unerwünschte Ereignisse (UE)

Studie	Operationalisierung
D4200C00058	Unerwünschte Ereignisse Schwerwiegende unerwünschte Ereignisse Studienabbruch wegen unerwünschter Ereignisse - Jeweils operationalisiert als Anteil der Patienten mit mindestens einem Ereignis - Erhoben vom Tag der unterschriebenen Einverständniserklärung eines Patienten bis 60 Tage nach der letzten Einnahme der Studienmedikation (inkl. *Open-Label* Medikation). Alle schwerwiegenden unerwünschten Ereignisse und alle Vandetanib-zugeschriebenen Ereignisse wurden bis zur Wiederherstellung nachverfolgt, es sei denn, der Prüfarzt gab an, dass eine Wiederherstellung im Rahmen der Grunderkrankung des Patienten nicht wahrscheinlich sei - Die Beurteilung des Schweregrades erfolgte nach CTCAE Version3 Die Sicherheits- (*Safety*) Population umfasst alle Patienten, die mindestens eine Dosis der Studienmedikation eingenommen haben.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Die Bewertung des Verzerrungspotentials für unerwünschte Ereignisse erfolgte zusammenfassend für alle betrachteten Operationalisierungen. Dieses Vorgehen ist möglich, da die Meldung und Berichterstattung unerwünschter Ereignisse in der Studie 58 gemäß GCP Standards erfolgte und die in die Analyse eingeschlossenen Patienten für alle Operationalisierungen identisch ist.

Tabelle 4-115: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
D4200C00058	niedrig	ja	ja*	nein	nein	niedrig

* mittels der Sicherheits-Population, welche im weitesten Sinne eine modifizierte ITT-Population darstellt.

Begründen Sie für jede Studie die abschließende Einschätzung.

Der Beobachtungszeitraum war die Zeit zwischen der unterschriebenen Einverständniserklärung jedes Patienten und *Cut-off* für die Safety-Analyse (60 Tage nach der letzten Einnahme der Studienmedikation inklusive *Open-Label* Vandetanib). Die Analyse wurde anhand der Sicherheits (*Safety*) -Population durchgeführt, d. h. anhand aller Patienten, die mindestens eine Dosis der Studienmedikation eingenommen haben (1 Placebo-Patient ist bereits vor Beginn der verblindeten Behandlung verstorben). Trotz des doppelblinden Designs ist theoretisch die Möglichkeit einer Entblindung der Patienten und Behandler durch das Auftreten typischer Nebenwirkungen, insbesondere bei einem Vergleich Verum gegen Placebo, gegeben. Eine mögliche partielle Entblindung von Patienten in der Vandetanib-Gruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapie-assoziierte Nebenwirkungen überlagern können und vom Patienten schwer zu differenzieren sind.

Andere das Verzerrungspotential beeinflussende Punkte konnten nicht gefunden werden. Das endpunktbezogene Verzerrungspotential wurde insgesamt als „niedrig“ bewertet. Dies gilt auch für die *Post-hoc*-Analyse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung.

Tabelle 4-116 und Tabelle 4-117 fassen die Ergebnisse zu dem Anteil der Patienten mit unerwünschten Ereignissen, schwerwiegenden unerwünschten Ereignissen und die Studienabbruchraten aufgrund von unerwünschten Ereignissen für beide betrachtete Populationen zusammen.

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.

Tabelle 4-116: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation)

Studie	**Vandetanib (N=231)**	**Placebo (N=99)**	**Total (N=330)**
D4200C00058			
AE Kategorie	**Anzahl Patienten[1] (%)**	**Anzahl Patienten[1] (%)**	**Anzahl Patienten (%)**
Jegliche UE	230 (99,6)	90 (90,9)	320 (97,0)
Vandetanib zugeschriebene[2] UE	222 (96,1)	59 (59,6)	281 (85,2)
UE vom CTCAE Grad 3 oder höher	128 (55,4)	24 (24,2)	152 (46,1)
Jegliche SUE (einschließlich Ereignisse mit Ausgang = Tod)	71 (30,7)	13 (13,1)	84 (25,5)
Jegliche SUE mit Ausgang = Tod	5 (2,2)	2 (2,0)	7 (2,1)
Jegliche UE die zur Absetzung Vandetanibs führten	28 (12,1)	3 (3,0)	31 (9,4)
Alle anderen signifikanten UE[3]	0 (0,0)	0 (0,0)	0 (0,0)

[1] Patienten mit mehreren Ereignissen in derselben Kategorie werden nur einfach in dieser Kategorie gezählt
[2] Durch den Prüfarzt eingeschätzt
[3] Jegliche UE, die vom Prüfarzt als signifikant eingeschätzt werden.
SUE = Schwere unerwünschte Ereignisse

Tabelle 4-117: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib (N=126)**	**Placebo (N=59)**	**Gesamt (N=185)**
D4200C00058			
AE Kategorie	**Anzahl Patienten[1] (%)**	**Anzahl Patienten[1] (%)**	**Anzahl Patienten (%)**
Jegliche UE	126 (100,0)	56 (94,9)	182 (98,4)
Vandetanib zugeschriebene[2] UE	122 (96,8)	37 (62,7)	159 (85,9)
UE vom CTCAE Grad 3 oder höher	77 (61,1)	14 (23,7)	91 (49,2)
Jegliche SUE (einschließlich Ereignisse mit Ausgang = Tod)	40 (31,7)	10 (16,9)	50 (27,0)
Jegliche SUE mit Ausgang = Tod	3 (2,4)	1 (1,7)	4 (2,2)
Jegliche UE die zur Absetzung Vandetanibs führten	15 (11,9)	1 (1,7)	16 (8,6)
Alle anderen signifikanten UE[3]	0 (0,0)	0 (0,0)	0 (0,0)

[1] Patienten mit mehreren Ereignissen in derselben Kategorie werden nur einfach in dieser Kategorie gezählt
[2] Durch den Prüfarzt eingeschätzt
[3] Jegliche UE, die vom Prüfarzt als signifikant eingeschätzt werden.

Tabelle 4-118: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation)

Studie **Intervention**		**Pat. mit mindestens einem schwerwiegenden unerwünschten Ereignis** **n (%)**	**Studienabbruch wegen unerwünschten Ereignissen** **n (%)**	**Pat. mit mindestens einem unerwünschten Ereignis** **n (%)**
D4200C00058				
Vandetanib	231	71 (30,7)	28 (12,1)	230 (99,6)
Placebo	99	13 (13,1)	3 (3,0)	90 (90,9)
OR [95% -KI]		2,94 [1,54;5,60]	4,41 [1,31;14,88]	23,00 [2,87;184,16]
RR [95%-KI]		2,34 [1,36;4,03]	4,00 [1,24;12,85]	1,10 [1,03;1,17]

Tabelle 4-119: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie **Intervention**		**Pat. mit mindestens einem schwerwiegenden unerwünschten Ereignis** **n (%)**	**Studienabbruch wegen unerwünschten Ereignissen** **n (%)**	**Pat. mit mindestens einem unerwünschten Ereignis** **n (%)**
D4200C00058				
Vandetanib	126	40 (31,7)	15 (11,9)	126 (100,0)
Placebo	59	10 (16,9)	1 (1,7)	56 (94,9)
OR [95% -KI]		2,28 [1,05;4,96]	7,84 [1,01;60,82]	15,67 [0,80;308,48]
RR [95%-KI]		1,87 [1,01;3,48]	7,02 [0,95;51,93]	1,05 [0,99;1,12]

Tabelle 4-120: Ergebnisse für den Anteil an Patienten mit unerwünschten Ereignissen (nach MedDRA) mit einer Häufigkeit von ≥10% aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation und der genäherten Zielpopulation)

	Gesamtpopulation		**Genäherte Zielpopulation**	
Studie	**Vandetanib (N=231)**	**Placebo (N=99)**	**Vandetanib (N=126)**	**Placebo (N=59)**
D4200C00058				
Unerwünschte Ereignisse, N (%)	**230 (99,6)**	**90 (90,9)**	**126 (100,0)**	**56 (94,9)**
Diarrhö	130 (56,3)	26 (26,3)	66 (52,4)	13 (22,0)
Hautausschlag	104 (45,0)	11 (11,1)	62 (49,2)	8 (13,6)
Übelkeit	77 (33,3)	16 (16,2)	45 (35,7)	11 (18,6)
Bluthochdruck	73 (31,6)	5 (5,1)	39 (31,0)	5 (8,5)
Kopfschmerzen	59 (25,5)	9 (9,1)	33 (26,2)	6 (10,2)
Fatigue	55 (23,8)	23 (23,2)	37 (29,4)	16 (27,1)
Appetitverlust	49 (21,2)	12 (12,1)	35 (27,8)	8 (13,6)
Akne	46 (19,9)	5 (5,1)	21 (16,7)	4 (6,8)
Trockene Haut	35 (15,2)	5 (5,1)	21 (16,7)	2 (3,4)
Akneiforme Dermatitis	35 (15,2)	2 (2,0)	20 (15,9)	2 (3,4)
Asthenie	34 (14,7)	11 (11,1)	22 (17,5)	6 (10,2)
Erbrechen	34 (14,7)	7 (7,1)	19 (15,1)	5 (8,5)
Bauchschmerzen	33 (14,3)	5 (5,1)	15 (11,9)	4 (6,8)
QTc-Verlängerung	33 (14,3)	1 (1,0)	19 (15,1)	0 (0,0)
Photosensitivitätsreaktion	31 (13,4)	0 (0,0)	17 (13,5)	0 (0,0)
Schlaflosigkeit	30 (13,0)	10 (10,1)	22 (17,5)	9 (15,3)
Nasopharyngitis	26 (11,3)	9 (9,1)	12 (9,5)	5 (8,5)
Husten	25 (10,8)	10 (10,1)	17 (13,5)	6 (10,2)
Juckreiz	25 (10,8)	4 (4,0)	17 (13,5)	2 (3,4)
Dyspepsie	25 (10,8)	4 (4,0)	17 (13,5)	2 (3,4)
Hypokalzämie	25 (10,8)	3 (3,0)	18 (14,3)	2 (3,4)
Gewichtsverlust	24 (10,4)	9 (9,1)	20 (15,9)	6 (10,2)
Rückenschmerzen	21 (9,1)	20 (20,2)	14 (11,1)	13 (22,0)
Arthralgie	18 (7,8)	10 (10,1)	12 (9,5)	4 (6,8)
Gliederschmerzen	16 (6,9)	13 (13,1)	13 (10,3)	9 (15,3)
Geschmacksstörung	19 (8,2)	3 (3,0)	16 (12,7)	3 (5,1)
Erythem	23 (10,0)	3 (3,0)	16 (12,7)	3 (5,1)
Atemnot	18 (7,8)	9 (9,1)	13 (10,3)	7 (11,9)
Depression	22 (9,5)	3 (3,0)	13 (10,3)	2 (3,4)

	Gesamtpopulation		**Genäherte Zielpopulation**	
Studie	**Vandetanib (N=231)**	**Placebo (N=99)**	**Vandetanib (N=126)**	**Placebo (N=59)**
D4200C00058				
Verschwommenes Sehen	20 (8,7)	1 (1,0)	13 (10,3)	1 (1,7)
Proteinurie	23 (10,0)	2 (2,0)	13 (10,3)	1 (1,7)
Harnwegsinfektion	17 (7,4)	6 (6,1)	10 (7,9)	6 (10,2)

Tabelle 4-121: Ergebnisse für das Auftreten von QTc-bedingten Ereignissen und UE der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der Gesamtpopulation)

Studie	**Vandetanib 300 mg (N=231)**	**Placebo (N=99)**	**Gesamt (N=330)**
D4200C00058			
UE Kategorie	**Anzahl Patienten[1] (%)**	**Anzahl Patienten[1] (%)**	**Anzahl Patienten (%)**
Gesamtes Auftreten von QTc bedingten Ereignissen	36 (15,6)	4 (4,0)	40 (12,1)
Jegliche SUE mit Ausgang = Tod	0	0	0
Schwerwiegende UE	3 (1,3)	0	3 (0,9)
Jegliche UE die zur Absetzung Vandetanibs führten	2 (0,9)	1 (1,0)	3 (0,9)
UE vom CTCAE Grad 3 oder höher	20 (8,7)	3 (3,0)	23 (7,0)
Gesamtes Auftreten von UE der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA)	33 (14,3)	1 (1,0)	34 (10,3)
Jegliche SUE mit Ausgang = Tod	0	0	0
Schwerwiegende UE	1 (0,4)	0	1 (0,3)
Jegliche UE die zur Absetzung Vandetanibs führten	2 (0,9)	0	2 (0,6)
UE vom CTCAE Grad 3 oder höher	18 (7,8)	1 (1,0)	19 (5,8)

Tabelle 4-122: Ergebnisse für das Auftreten von QTc-bedingten Ereignissen und UEs der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung)

Studie	**Vandetanib 300 mg (N=126)**	**Placebo (N=59)**	**Gesamt (N=185)**
D4200C00058			
UE Kategorie	**Anzahl Patienten[1] (%)**	**Anzahl Patienten[1] (%)**	**Anzahl Patienten (%)**
Gesamtes Auftreten von QTc bedingten Ereignissen	20 (15,9)	1 (1,7)	21 (11,4)
Jegliche SUE mit Ausgang = Tod	0	0	0
Schwerwiegende UE	2 (1,6)	0	2 (1,1)
Jegliche UE die zur Absetzung Vandetanibs führten	1 (0,8)	0	1 (0,5)
UE vom CTCAE Grad 3 oder höher	10 (7,9)	0	10 (5,4)
Gesamtes Auftreten von UE der bevorzugten Bezeichnung *„electrocardiogram QT prolonged“* (nach MedDRA)	19 (15,1)	0	19 (10,3)
Jegliche SUE mit Ausgang = Tod	0	0	0
Schwerwiegende UE	1 (0,8)	0	1 (0,5)
Jegliche UE die zur Absetzung Vandetanibs führten	1 (0,8)	0	1 (0,5)
UE vom CTCAE Grad 3 oder höher	9 (7,1)	0	9 (4,9)

Wie in Tabelle 4-116 gezeigt, trat bei nahezu allen Patienten beider Gruppen mindestens ein unerwünschtes Ereignis auf (99,6% der Patienten in der Vandetanib-Gruppe und 90,9% der Placebo-Patienten). In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung lagen die entsprechenden Inzidenzen bei 100% bzw. 94,9% (Tabelle 4-117). Der Anteil der Patienten mit unerwünschten Ereignissen vom *Common Terminology Criteria for Adverse Events* (CTCAE) Grad 3 oder höher (55,4% vs. 24,2% bzw. 61,1% vs. 23,7%), mit SUE (30,7% vs. 13,1% bzw. 31,7% vs. 16,9%) und mit unerwünschten Ereignissen, die zum Absetzen der Studienbehandlung führten (12,1% vs. 3,0% bzw. 11,9% vs. 1,7%), war jeweils unter Vandetanib im Vergleich zu Placebo größer. Der Anteil der Todesfälle aufgrund unerwünschter Ereignisse war in beiden Gruppen ähnlich (2,2% vs. 2,0% bzw. 2,4% vs. 1,7%). Im Vergleich zu Placebo zeigte sich somit für unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse und Studienabbrüche wegen unerwünschter Ereignisse ein statistisch signifikanter Unterschied zuungunsten von Vandetanib. Daraus

ergibt sich in beiden für die Nutzenbewertung betrachteten Populationen ein Beleg für einen größeren Schaden von Vandetanib gegenüber Placebo für UE, SUE und Studienabbruch wegen unerwünschter Ereignisse (Tabelle 4-118 und Tabelle 4-119**)**.

Die mediane Therapiedauer auf verblindeter Therapie war in der Vandetanib-Gruppe erheblich länger als in der Placebo-Gruppe (90,1 vs. 39,9 Wochen in der Gesamtpopulation bzw. 88,6 vs. 37,1 Wochen in der genäherten Zielpopulation), wodurch sich für die Vandetanib-Gruppe eine um ca. 50 Wochen längere Expositionsdauer ergibt.

Das Nebenwirkungsprofil von Gesamt- und genäherter Zielpopulation war vergleichbar (siehe Tabelle 4-120). Die fünf am häufigsten berichteten unerwünschten Ereignisse in der Vandetanib-Gruppe vs. Placebo waren Diarrhö, Hautausschlag, Übelkeit, Bluthochdruck und Kopfschmerzen.

	Gesamtpopulation	genäherte Zielpopulation
• Diarrhö:	56,3% vs. 26,3%	52,4% vs. 22,0%
• Hautausschlag:	45,0% vs. 11,1%	49,2% vs. 13,6%
• Übelkeit:	33,3% vs. 16,2%	35,7% vs. 18,6%
• Bluthochdruck:	31,6% vs. 5,1%	31,0% vs. 8,5%
• Kopfschmerzen:	25,5% vs. 9,1%	26,2% vs. 10,2%

Die unerwünschten Ereignisse (Vandetanib vs. Placebo) Rückenschmerzen (9,1% vs. 20,2% bzw. 11,1% vs. 22,0%), und Gliederschmerzen (6,9% vs. 13,1% bzw. 10,3% vs. 15,3%) wurden weniger häufig in der Vandetanib-Gruppe berichtet als in der Placebo-Gruppe. Gelenkschmerzen waren in der Gesamtpopulation seltener unter Vandetanib gegenüber Placebo (7,8% vs. 10,1%), in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung allerdings häufiger (9,5% vs. 6,8%).

Im Rahmen des Nebenwirkungsmanagements waren Dosisreduktionen auf 200 mg bzw. 100 mg Vandetanib / Placebo und Dosisunterbrechungen im Studienprotokoll vorgesehen und wurden im Vandetanib-Arm deutlich häufiger durchgeführt als unter Placebo. Dosisreduktionen auf 200 mg wurden in der Gesamtpopulation bei 35,1% (Vandetanib) und 3% (Placebo) der Patienten durchgeführt, in der genäherten Zielpopulation bei 32,5% und 3,4%. Eine weitere Dosisreduktion auf 100 mg erhielten nur Patienten unter Vandetanib, in der Gesamtpopulation waren dies 13% und in der genäherten Zielpopulation 12,7%. Zwei Patienten waren ohne Zwischenschritt von 300 mg auf 100 mg Vandetanib umgestellt worden (0,9%). Eine Dosisunterbrechung wurde in der Gesamtpopulation bei 47,2% (Vandetanib) und 15,2% (Placebo) der Patienten durchgeführt, in der genäherten Zielpopulation bei 49,2% und 15,3%.

Aufgrund der Erfahrungen aus klinischen Studien in anderen Indikationen sowie mit anderen TKI und aufgrund des Gefährdungspotentials für die Patienten wurde ein besonderes Augenmerk auf QTc-Verlängerungen gerichtet (siehe insbesondere Modul 3 A, Abschnitt 3.4).

Im Vergleich zu Placebo zeigte sich in beiden betrachteten Populationen für QTc-bedingte unerwünschte Ereignisse ein Unterschied zuungunsten von Vandetanib, insbesondere für UE vom CTCAE Grad 3 oder höher. Im Falle einer Toxizität des Grades 3 oder höher nach CTCAE oder einer Verlängerung des QT-Intervalls im EKG soll die Gabe von Vandetanib zumindest zeitweilig ausgesetzt und in einer reduzierten Dosierung wieder aufgenommen werden, wenn keine Toxizität mehr vorliegt oder diese sich bis zum CTCAE-Grad 1 verbessert hat (Tabelle 4-121 und Tabelle 4-122). Die QTc-bedingten unerwünschten Ereignisse wurden im Risk-Management Plan für Vandetanib berücksichtigt (siehe Modul 3 A, Abschnitt 3.4).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

<Abbildung Meta-Analyse>

Abbildung 34: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Es wurden keine Meta-Analysen durchgeführt (siehe Abschnitt 4.2.5.3).

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

4.3.1.3.2 Subgruppenanalysen – RCT

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

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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.

Eine umfassende Subgruppenanalyse nach Geschlecht und Schwere der Krankheit wurde gemäß Studienprotokoll für den Endpunkt „Progressionsfreies Überleben (PFS)“ durchgeführt. Für diesen Endpunkt erfolgte außerdem noch eine Subgruppenanalyse nach Rasse und WHO *Performance Status* (WHO PS). Des weiteren wurden Ergebnisse getrennt nach RET-Mutationsstatus, CTN-Verdopplungszeit, CEA-Verdopplungszeit, Anzahl an vorherigen Therapien, Ansprechen auf die jüngste vorherige Therapie, MTC-Status, Ausgangswerte für p-VEGF, p-VEGFR2 und p-bFGF dargestellt. Die Subgruppen wurden vor Entblindung festgelegt. Subgruppenresultate (*Hazard Ratio* mit zugehörigem 95%-Konfidenzintervall und 2-seitigem p-Wert mittels *Log-Rank Test*) wurden mittels *Forest-Plots* dargestellt. Die Analyse getrennt nach Altersgruppen hat explorativen Charakter, da sie nicht vor der Entblindung festgelegt wurde (Abschnitt 4.3.1.3.2.1).

Für den Endpunkt „Zeit bis zur Schmerzprogression“ (TWP) wurde eine explorative Subgruppenanalyse, getrennt nach *Baseline*-Opioidanalgetika-Gebrauch, prospektiv geplant und durchgeführt und *post-hoc* auch für die genäherte Zielpopulation ausgewertet. Diese Ergebnisse sind in Abschnitt 4.3.1.3.1.9 als explorativer Morbiditätsendpunkt dargestellt.

Zusätzlich wurden aufgrund der hohen Patientenrelevanz von TWP in der Vorbereitung des vorgelegten Dossiers alle auch für den Endpunkt PFS im Protokoll vordefinierten Subgruppenanalysen sowohl für die Gesamtpopulation als auch für die genäherte Zielpopulation *post-hoc* re-analysiert (siehe Abschnitt 4.3.1.3.2.2). Die besondere Patientenrelevanz gerade dieses Endpunkts lässt sich damit begründen, dass für die Patienten eine möglichst langfristige Verzögerung der Schmerzprogression von großer Bedeutung ist, da Schmerzen die Lebensqualität stark beeinträchtigen und auch der Einsatz von Schmerzmitteln von entsprechenden Nebenwirkungen begleitet ist (siehe Abschnitt 4.2.5.2).

Die im Rahmen des EU-Zulassungsverfahrens *post-hoc* vorgenommenen Subgruppenanalysen zum RET-Mutationsstatus wurden in Vorbereitung dieses Nutzendossiers auch für die genäherte Zielpopulation analysiert (detaillierte Erläuterung siehe 4.3.1.3.2.3).

Für die sekundären Endpunkte „Gesamtüberleben“, „Objektive Ansprechrate“, „Krankheitskontrollrate“, „Dauer des objektiven Ansprechens“, „gesundheitsbezogene Lebensqualität“ und „unerwünschte Ereignisse“ waren keine präspezifizierten Subgruppenanalysen geplant oder durchgeführt, ebenso wenig für die explorativen Endpunkte. Präspezifizierte Subgruppenanalysen basieren auf erhobenen und verfügbaren Baseline-Informationen (Baseline-Charakteristika, Biomarkern oder Krankheitsstatus). Des Weiteren ist darauf hinzuweisen, dass die Resultate von Subgruppenanalysen, insbesondere für die bereits eingeschränkte genäherte Zielpopulation, mit besonderer Vorsicht beurteilt werden sollten. Kleine Stichproben innerhalb der Subgruppen, ein Ungleichgewicht zwischen

den Subgruppen und die steigende Anzahl statistischer Tests (Multiple Analysen) erhöhen prinzipiell die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse [89].

Um nicht durch ein *post-hoc* durchgeführtes multiples Testen die Wahrscheinlichkeit zufälliger statistischer signifikanter Ergebnisse zu erhöhen, wurden keine *a posteriori* spezifizierten Subgruppenanalysen für oben genannte sekundäre Endpunkte durchgeführt und im Modul 4 dargestellt. Eine numerische Auswertung der wichtigsten Patientenmerkmale (Geschlecht, Alter, Krankheitsstadium) ist für die einzelnen Endpunkte jedoch erfolgt und im Modul 5 hinterlegt.

Ein im Protokoll vorgesehener globaler Interaktionstest (zum Niveau 0,01), der die 6 als prognostisch eingestuften *Baseline*-Kovariaten (RET Mutationsstatus, Zeitraum bis zur Verdopplung des CTN-Wertes, Zeitraum bis zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien und der MTC-Status) im Cox-PH-Modell einschloss, ergab keine statistisch signifikante Interaktion (p=0,177).

Post-hoc wurden individuelle Interaktionstests für alle präspezifizierten Subgruppenfaktoren durchgeführt. Hinweise auf Effektmodifikationen wurden mithilfe der individuellen Interaktionstests abgeleitet. Ein p-Wert kleiner als 0,2 wurde gemäß IQWiG Methodik als Hinweis auf eine Effektmodifikation gewertet. Als Voraussetzung für einen Beleg unterschiedlicher Effekte galt ein statistisch signifikanter Interaktionstest ($p \leq 0,05$). Im Fall von gleichgerichteten Gesamtschätzern der Subgruppen liegen Interaktionen quantitativer Natur vor, ansonsten sind diese qualitativer Natur.

Die individuellen Interaktionstests zeigten für PFS in der Gesamtpopulation mit Ausnahme der Faktoren CEA Verdopplungszeit und p-VEGFR2 keinen Hinweis auf eine Effektmodifikation. Für die CEA Verdopplungszeit (p=0,0062) und p-VEGFR2 (p=0,0560) ergaben sich ein Beleg bzw. ein Hinweis auf eine Interaktion, die in beiden Fällen quantitativer Natur war (Tabelle 4-123). Für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ergaben die durchgeführten individuellen Interaktionstests (Tabelle 4-124) nur für die CEA-Verdopplungszeit (p=0,1807) einen Hinweis auf eine Effektmodifikation. Die Interaktion ist jedoch ebenfalls quantitativer Natur.

Die *post-hoc* gemäß den präspezifizierten Subgruppenfaktoren des PFS durchgeführten individuellen Interaktionstests zeigten für TWP in der Gesamtpopulation außer für die Faktoren CEA Verdopplungszeit (p=0,1687) und p-bFGF (p=0,0632) keine Hinweise auf eine Effektmodifikation. In beiden Fällen waren die Interaktionen quantitativer Natur (Tabelle 4-125). Für die genäherte Zielpopulation ergaben die *post-hoc* durchgeführten Interaktionstests außer für die Faktoren CTN-Verdopplungszeit (p=0,1632), CEA-Verdopplungszeit (p=0,0151) und Alter (p=0,1980) keinen Hinweis auf eine Effektmodifikation. Die Interaktionen waren in allen Fällen quantitativer Natur (Tabelle 4-126).

Aufgrund der beschränkten Aussagekraft von Subgruppenanalysen, insbesondere angesichts der Vielzahl der durchgeführten Analysen und Tests und der geringen Anzahl an Patienten in

den einzelnen Subgruppen, ist dieses Ergebnis mit Vorsicht zu interpretieren. Insgesamt ergibt sich für die genäherte Zielpopulation keine Änderung des aus der Bewertung der Gesamtpopulation abgeleiteten Fazits.

Tabelle 4-123: Individuelle Interaktionstests (*post-hoc*) – PFS - Gesamtpopulation

Subgruppenbildender Faktor (***Subgruppen Kategorien)***	**p-Wert Interaktionstest**	**Art der Interaktion**
RET Mutationsstatus *(positiv, negativ, unbekannt)*	0,8028	
Zeitraum bis zur Verdopplung des Calcitoninwertes *(≤24 Monate, >24 Monate, unbekannt)*	0,2784	
Zeitraum bis zur Verdopplung des CEA-Wertes *(≤24 Monate, >24 Monate, unbekannt)*	0,0062	quantitativ
Anzahl vorangegangener Therapien *(≥1 Vorbehandlung, keine Vorbehandlung)*	0,6806	
MTC-Status *(hereditär, sporadisch, unbekannt)*	0,8693	
Geschlecht *(Männer, Frauen)*	0,9974	
WHO Performance Status *(≥1, 0)*	0,9948	
p-VEGF (*hoch, niedrig, unbekann*t)	0,4544	
p-VEGFR2 *(hoch, niedrig, unbekannt)*	0,0560	quantitativ
p-bFGF *(hoch, niedrig, unbekannt)*	0,6365	
Alter *(jünger 65 Jahre, älter 65 Jahre)*	0,3147	

Tabelle 4-124: Individuelle Interaktionstests (*post-hoc*) – PFS – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Subgruppenbildender Faktor (***Subgruppen Kategorien)***	**p-Wert Interaktionstest**	**Art der Interaktion**
RET Mutationsstatus *(positiv, negativ, unbekannt)*	NC	
Zeitraum bis zur Verdopplung des Calcitoninwertes *(≤24 Monate, >24 Monate, unbekannt)*	0,2118	
Zeitraum bis zur Verdopplung des CEA-Wertes *(≤24 Monate, >24 Monate, unbekannt)*	0,1807	quantitativ
Anzahl vorangegangener Therapien *(≥1 Vorbehandlung, keine Vorbehandlung)*	0,6001	
MTC-Status *(hereditär, sporadisch, unbekannt)*	0,7255	
Geschlecht *(Männer, Frauen)*	0,8758	
WHO Performance Status *(≥1, 0)*	0,7730	
p-VEGF (*hoch, niedrig, unbekannt*)	0,2177	
p-VEGFR2 *(hoch, niedrig, unbekannt)*	0,4670	
p-bFGF *(hoch, niedrig, unbekannt)*	0,7917	
Alter *(jünger 65 Jahre, älter 65 Jahre)*	0,7090	

NC = nicht berechenbar

Tabelle 4-125: Individuelle Interaktionstests (*post-hoc*) – TWP- Gesamtpopulation

Subgruppenbildender Faktor (*Subgruppen Kategorien)*	**p-Wert Interaktionstest**	**Art der Interaktion**
RET Mutationsstatus *(positiv, negativ, unbekannt)*	NC	
Zeitraum bis zur Verdopplung des Calcitoninwertes *(≤24 Monate, >24 Monate, unbekannt)*	0,2186	
Zeitraum bis zur Verdopplung des CEA-Wertes *(≤24 Monate, >24 Monate, unbekannt)*	0,1687	quantitativ
Anzahl vorangegangener Therapien *(≥1 Vorbehandlung, keine Vorbehandlung)*	0,9099	
MTC-Status *(hereditär, sporadisch, unbekannt)*	0,8307	
Geschlecht *(Männer, Frauen)*	0,3763	
WHO Performance Status *(≥1, 0)*	0,3054	
p-VEGF (*hoch, niedrig, unbekannt*)	0,2094	
p-VEGFR2 *(hoch, niedrig, unbekannt)*	0,6140	
p-bFGF *(hoch, niedrig, unbekannt)*	0,0632	quantitativ
Alter *(jünger 65 Jahre, älter 65 Jahre)*	0,6278	

NC = Nicht berechenbar

Tabelle 4-126: Individuelle Interaktionstests (*post-hoc*) – TWP – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Subgruppenbildender Faktor (*Subgruppen Kategorien)*	**p-Wert Interaktionstest**	**Art der Interaktion**
RET Mutationsstatus *(positiv, negativ, unbekannt)*	NC	
Zeitraum bis zur Verdopplung des Calcitoninwertes *(≤24 Monate, >24 Monate, unbekannt)*	0,1632	quantitativ
Zeitraum bis zur Verdopplung des CEA-Wertes *(≤24 Monate, >24 Monate, unbekannt)*	0,0151	quantitativ
Anzahl vorangegangener Therapien *(≥1 Vorbehandlung, keine Vorbehandlung)*	0,4111	
MTC-Status *(hereditär, sporadisch, unbekannt)*	0,7237	
Geschlecht *(Männer, Frauen)*	0,9140	
WHO Performance Status *(≥1, 0)*	0,4201	
p-VEGF (*hoch, niedrig, unbekannt*)	0,2331	
p-VEGFR2 *(hoch, niedrig, unbekannt)*	0,7116	
p-bFGF *(hoch, niedrig, unbekannt)*	0,2685	
Alter *(jünger 65 Jahre, älter 65 Jahre)*	0,1980	quantitativ

NC = nicht berechenbar

4.3.1.3.2.1 Progressionsfreies Überleben (PFS)

Tabelle 4-127: PFS Subgruppenanalysen nach *Baseline*-Charakteristika - Gesamtpopulation

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Männlich	Vandetanib	134	47 (35,1)	0,46	[0,28;0,77]	0,0028	0,9974
	Placebo	56	32 (57,1)				
Weiblich	Vandetanib	97	26 (26,8)	0,44	[0,22;0,85]	0,0147	
	Placebo	44	19 (43,2)				
Kaukasisch	Vandetanib	218	70 (32,1)	0,46	[0,31;0,70]	0,0002	NC
	Placebo	97	50 (51,5)				
Andere Rasse	Vandetanib	13	3 (23,1)	NC	NC	NC	
	Placebo	3	1 (33,3)				
WHO PS ≥ 1	Vandetanib	77	28 (36,4)	0,49	[0,27;0,90]	0,0204	0,9948
	Placebo	42	23 (54,8)				
WHO PS = 0	Vandetanib	154	45 (29,2)	0,46	[0,27;0,79]	0,0051	
	Placebo	58	28 (48,3)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test
NC = nicht berechenbar

Tabelle 4-128: PFS Subgruppenanalysen nach *Baseline*-Charakteristika - genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Männlich	Vandetanib	79	31 (39,2)	0,50	[0,28;0,89]	0,0190	0,8758
	Placebo	39	24 (61,5)				
Weiblich	Vandetanib	47	15 (31,9)	0,44	[0,18;1,05]	0,0638	
	Placebo	21	11 (52,4)				
Kaukasisch	Vandetanib	118	43 (36,4)	0,48	[0,29;0,79]	0,0038	NC
	Placebo	59	34 (57,6)				
Andere Rasse	Vandetanib	8	3 (37,5)	NC	NC	NC	
	Placebo	1	1 (100,0)				
WHO PS ≥= 1	Vandetanib	52	19 (36,5)	0,53	[0,26;1,09]	0,0846	0,7730
	Placebo	27	16 (59,3)				
WHO PS = 0	Vandetanib	74	27 (36,5)	0,42	[0,22;0,82]	0,0114	
	Placebo	33	19 (57,6)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test
NC = nicht berechenbar

Tabelle 4-129: PFS Subgruppenanalysen nach Biomarkern - Gesamtpopulation

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
RET-Mutationsstatus positiv	Vandetanib Placebo	137 50	47 (34,3) 27 (54,0)	0,45	[0,26;0,78]	0,0041	
RET-Mutationsstatus negativ	Vandetanib Placebo	2 6	1 (50,0) 5 (83,3)	0,87	[0,10;7,66]	0,8967	0,8028
RET-Mutationsstatus unbekannt	Vandetanib Placebo	92 44	25 (27,2) 19 (43,2)	0,49	[0,26;0,95]	0,0341	
CTN Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	124 46	39 (31,5) 27 (58,7)	0,31	[0,17;0,56]	0,0001	
CTN Verdopplungszeit > 24 Monate	Vandetanib Placebo	83 43	23 (27,7) 19 (44,2)	0,56	[0,29;1,06]	0,0754	0,2784
CTN Verdopplungszeit unbekannt	Vandetanib Placebo	24 11	11 (45,8) 5 (45,5)	0,57	[0,17;1,88]	0,3548	
CEA Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	69 33	25 (36,2) 26 (78,8)	0,19	[0,10;0,36]	<0,0001	
CEA Verdopplungszeit > 24 Monate	Vandetanib Placebo	119 48	28 (23,5) 14 (29,2)	0,81	[0,42;1,58]	0,5349	0,0062
CEA Verdopplungszeit unbekannt	Vandetanib Placebo	43 19	20 (46,5) 11 (57,9)	0,55	[0,24;1,25]	0,1551	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

CTN/CEA Verdopplungszeit wurde abgeleitet aus der natürlichen Logarithmierung der einzelnen CTN/CEA Messungen (*pre-dose*). Zur Ermittlung der Steigung m wird an diese Daten wird eine lineare Regressionsgerade anpasst. Die CTN/CEA Verdopplungszeit ergibt sich als: m/ ln 2.

Negative Verdopplungszeiten werden der >24 Monate Gruppe zugewiesen

Tabelle 4-130: PFS Subgruppenanalysen nach Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
RET-Mutationsstatus positiv	Vandetanib Placebo	75 30	30 (40,0) 19 (63,3)	0,41	[0,21;0,80]	0,0090	
RET-Mutationsstatus negativ	Vandetanib Placebo	1 6	0 (0,0) 5 (83,3)	NC	NC	NC	NC
RET-Mutationsstatus unbekannt	Vandetanib Placebo	50 24	16 (32,0) 11 (45,8)	0,60	[0,27;1,38]	0,2306	
CTN Verdopplungszeit ≤= 24 Monate	Vandetanib Placebo	65 30	25 (38,5) 21 (70,0)	0,32	[0,16;0,62]	0,0009	
CTN Verdopplungszeit > 24 Monate	Vandetanib Placebo	45 27	14 (31,1) 12 (44,4)	0,67	[0,30;1,49]	0,3277	0,2118
CTN Verdopplungszeit unbekannt	Vandetanib Placebo	16 3	7 (43,8) 2 (66,7)	0,02	[0,00;0,51]	0,0166	
CEA Verdopplungszeit ≤= 24 Monate	Vandetanib Placebo	40 24	17 (42,5) 19 (79,2)	0,38	[0,19;0,76]	0,0067	
CEA Verdopplungszeit > 24 Monate	Vandetanib Placebo	59 27	16 (27,1) 9 (33,3)	0,81	[0,35;1,88]	0,6226	0,1807
CEA Verdopplungszeit unbekannt	Vandetanib Placebo	27 9	13 (48,1) 7 (77,8)	0,17	[0,05;0,61]	0,0065	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

CTN/CEA Verdopplungszeit wurde abgeleitet aus der natürlichen Logarithmierung der einzelnen CTN/CEA Messungen (*pre-dose*). Zur Ermittlung der Steigung m wird an diese Daten wird eine lineare Regressionsgerade anpasst. Die CTN/CEA Verdopplungszeit ergibt sich als: m/ ln 2.

Negative Verdopplungszeiten werden der >24 Monate Gruppe zugewiesen

NC = nicht berechenbar

Tabelle 4-131: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung) - Gesamtpopulation

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hereditäre oder Keimbahn-Mutation	Vandetanib Placebo	28 5	7 (25,0) 2 (40,0)	0,30	[0,04;2,50]	0,2657	0,8693
Sporadische oder unbekannte Mutation	Vandetanib Placebo	203 95	66 (32,5) 49 (51,6)	0,48	[0,32;0,72]	0,0004	
Metastatisch	Vandetanib Placebo	217 97	67 (30,9) 51 (52,6)	0,42	[0,28;0,63]	<0,0001	NC
Lokal fortgeschritten	Vandetanib Placebo	14 3	6 (42,9) 0 (0)	NC	NC	NC	
≥ 1 Vorbehandlung	Vandetanib Placebo	90 42	31 (34,4) 24 (57,1)	0,43	[0,24;0,78]	0,0056	0,6806
Keine Vorbehandlung	Vandetanib Placebo	141 58	42 (29,8) 27 (46,6)	0,49	[0,29;0,85]	0,0106	
Ansprechen nach Vorbehandlung	Vandetanib Placebo	5 2	1 (20,0) 0 (0,0)	NC	NC	NC	NC
Kein Ansprechen nach Vorbehandlung	Vandetanib Placebo	42 23	19 (45,2) 11 (47,8)	0,73	[0,33;1,60]	0,4291	
NC/ unbekannte BOR zu Vorbehandlung	Vandetanib Placebo	184 75	53 (28,8) 40 (53,3)	0,38	[0,23;0,60]	<0,0001	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank Test*

Vorbehandlung bezieht sich auf systemische Antikrebs-Therapie für medullären Schilddrüsenkrebs, die vor der ersten Dosis der randomisierten Behandlung eingenommen wurde.

BOR = bestes objektives Ansprechen *(Best Objective Response)*

NC = nicht berechenbar

Tabelle 4-132: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung) – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hereditäre oder Keimbahn-Mutation	Vandetanib Placebo	11 2	2 (18,2) 1 (50,0)	NC	NC	NC	0,7255
Sporadische oder unbekannte Mutation	Vandetanib Placebo	115 58	44 (38,3) 34 (58,6)	0,50	[0,31;0,82]	0,0055	
Metastatisch	Vandetanib Placebo	119 59	43 (36,1) 35 (59,3)	0,45	[0,28;0,74]	0,0015	NC
Lokal fortgeschritten	Vandetanib Placebo	7 1	3 (42,9) 0 (0,0)	NC	NC	NC	
=≥ 1 Vorbehandlung	Vandetanib Placebo	45 29	19 (42,2) 17 (58,6)	0,60	[0,30;1,20]	0,1471	0,6001
Keine Vorbehandlung	Vandetanib Placebo	81 31	27 (33,3) 18 (58,1)	0,42	[0,21;0,83]	0,0132	
Ansprechen nach Vorbehandlung	Vandetanib Placebo	3 1	1 (33,3) 0 (0,0)	NC	NC	NC	NC
Kein Ansprechen nach Vorbehandlung	Vandetanib Placebo	23 14	12 (52,2) 5 (35,7)	1,26	[0,46;3,44]	0,6486	
NC/ unbekannte BOR zu Vorbehandlung	Vandetanib Placebo	100 45	33 (33,0) 30 (66,7)	0,32	[0,18;0,57]	<0,0001	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

Vorbehandlung bezieht sich auf systemische Antikrebs-Therapie für medullären Schilddrüsenkrebs, die vor der ersten Dosis der randomisierten Behandlung eingenommen wurde.

BOR = bestes objektives Ansprechen (*Best Objective* R*esponse*)

NC = nicht berechenbar

Tabelle 4-133: PFS Subgruppenanalysen nach Plasma Biomarkern - Gesamtpopulation

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hoher *Baseline* p-VEGF	Vandetanib Placebo	115 51	41 (35,7) 25 (49,0)	0,56	[0,32;0,97]	0,0382	
Niedriger *Baseline* p-VEGF	Vandetanib Placebo	101 42	25 (24,8) 20 (47,6)	0,37	[0,19;0,73]	0,0039	0,4544
Unbekannter *Baseline* p-VEGF	Vandetanib Placebo	15 7	7 (46,7) 6 (85,7)	0,27	[0,07;1,01]	0,0517	
Hoher *Baseline* p-VEGFR2	Vandetanib Placebo	155 69	40 (25,8) 26 (37,7)	0,60	[0,35;1,03]	0,0626	
Niedriger *Baseline* p-VEGFR2	Vandetanib Placebo	61 24	26 (42,6) 19 (79,2)	0,20	[0,09;0,42]	<0,0001	0,0560
Unbekannter *Baseline* p-VEGFR2	Vandetanib Placebo	15 7	7 (46,7) 6 (85,7)	0,27	[0,07;1,01]	0,0517	
Hoher *Baseline* p-bFGF	Vandetanib Placebo	107 49	39 (36,4) 26 (53,1)	0,54	[0,32;0,94]	0,0278	
Niedriger *Baseline* p-bFGF	Vandetanib Placebo	108 43	27 (25,0) 19 (44,2)	0,41	[0,21;0,80]	0,0090	0,6365
Unbekannter *Baseline* p-bFGF	Vandetanib Placebo	16 8	7 (43,8) 6 (75,0)	0,38	[0,11;1,31]	0,1252	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.
Hoher p-VEGF >55,0 pg/ml bei *Baseline*. Niedrige p-VEGF ≤55,0 pg/ml bei *Baseline.*
Hoher p-VEGFR2 >9,88165 pg/ml bei *Baseline*. Niedrige p-VEGFR2 ≤9,88165 pg/ml bei *Baseline.*
Hoher p-bFGF >2,10 pg/ml bei *Baseline*. Niedrige p-bFGF ≤2,10 pg/ml bei *Baseline*.
Baseline ist definiert als der Wert am nächsten zu und vor Randomisierung

Tabelle 4-134: PFS Subgruppenanalysen nach Plasma Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hoher *Baseline* p-VEGF	Vandetanib Placebo	60 26	26 (43,3) 14 (53,8)	0,68	[0,34;1,37]	0,2863	
Niedriger *Baseline* p-VEGF	Vandetanib Placebo	55 27	14 (25,5) 15 (55,6)	0,31	[0,14;0,69]	0,0044	0,2177
Unbekannter *Baseline* p-VEGF	Vandetanib Placebo	11 7	6 (54,5) 6 (85,7)	0,41	[0,12;1,42]	0,1600	
Hoher *Baseline* p-VEGFR2	Vandetanib Placebo	82 38	24 (29,3) 17 (44,7)	0,55	[0,28;1,09]	0,0851	
Niedriger *Baseline* p-VEGFR2	Vandetanib Placebo	33 15	16 (48,5) 12 (80,0)	0,36	[0,15;0,85]	0,0206	0,4670
Unbekannter *Baseline* p-VEGFR2	Vandetanib Placebo	11 7	6 (54,5) 6 (85,7)	0,41	[0,12;1,42]	0,1600	
Hoher *Baseline* p-bFGF	Vandetanib Placebo	60 27	23 (38,3) 16 (59,3)	0,53	[0,27;1,08]	0,0792	
Niedriger *Baseline* p-bFGF	Vandetanib Placebo	55 26	17 (30,9) 13 (50,0)	0,44	[0,19;0,98]	0,0447	0,7917
Unbekannter *Baseline* p-bFGF	Vandetanib Placebo	11 7	6 (54,5) 6 (85,7)	0,41	[0,12;1,42]	0,1600	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.

Hoher p-VEGF >55,0 pg/ml bei *Baseline*. Niedrige p-VEGF ≤55,0 pg/ml bei *Baseline*.

Hoher p-VEGFR2 >9,88165 pg/ml bei *Baseline*. Niedrige p-VEGFR2 ≤9,88165 pg/ml bei *Baseline*.

Hoher p-bFGF >2,10 pg/ml bei *Baseline*. Niedrige p-bFGF ≤2,10 pg/ml bei *Baseline*.

Baseline ist definiert als der Wert am nächsten zu und vor Randomisierung

Tabelle 4-135: PFS Subgruppenanalyse nach Altersgruppe – Gesamtpopulation

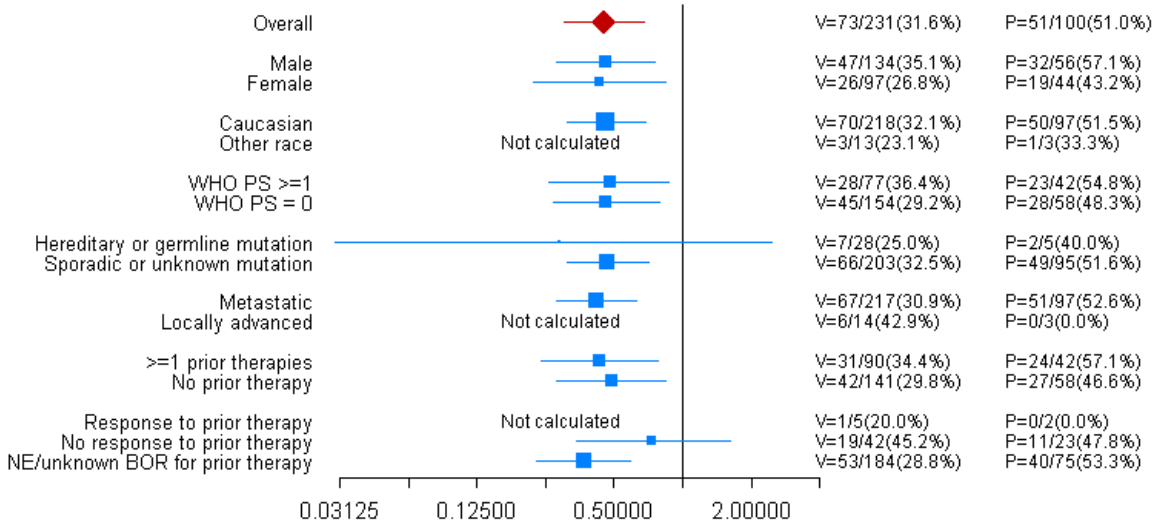
Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Jünger als 65 Jahre	Vandetanib	182	55 (30,2)	0,50	[0,32;0,80]	0,0036	0,3147
	Placebo	80	37 (46,3)				
65 Jahre und älter	Vandetanib	49	18 (36,7)	0,32	[0,14;0,74]	0,0071	
	Placebo	20	14 (70,0)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.

Tabelle 4-136: PFS Subgruppenanalyse nach Altersgruppe – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Jünger als 65 Jahre	Vandetanib	93	33 (35,5)	0,48	[0,27;0,84]	0,0101	0,7090
	Placebo	48	26 (54,2)				
65 Jahre und älter	Vandetanib	33	13 (39,4)	0,40	[0,15;1,08]	0,0696	
	Placebo	12	9 (75,0)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.



V = Vandetanib, P = Placebo

Abbildung 35: *Forest-Plot* der Subgruppenanalyse für PFS – *Baseline*- Charakteristiken und Krankheitsstatus – Gesamtpopulation

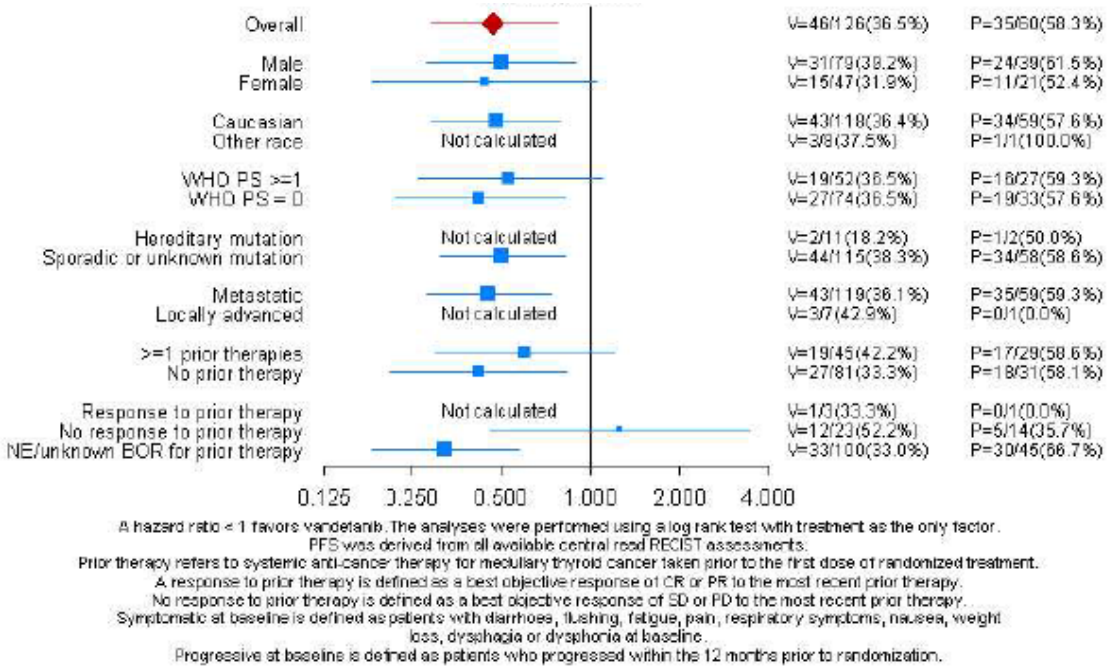
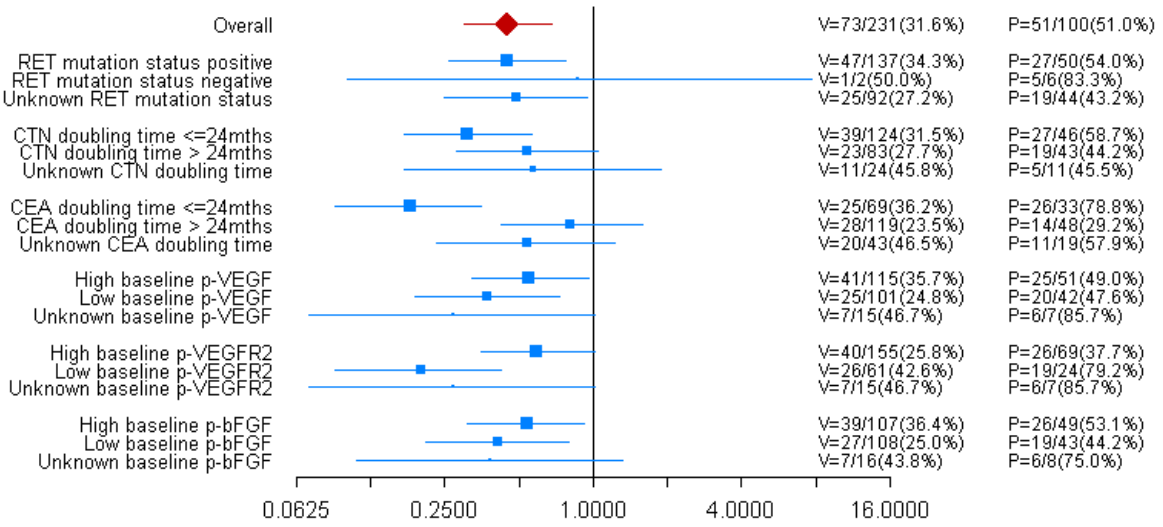
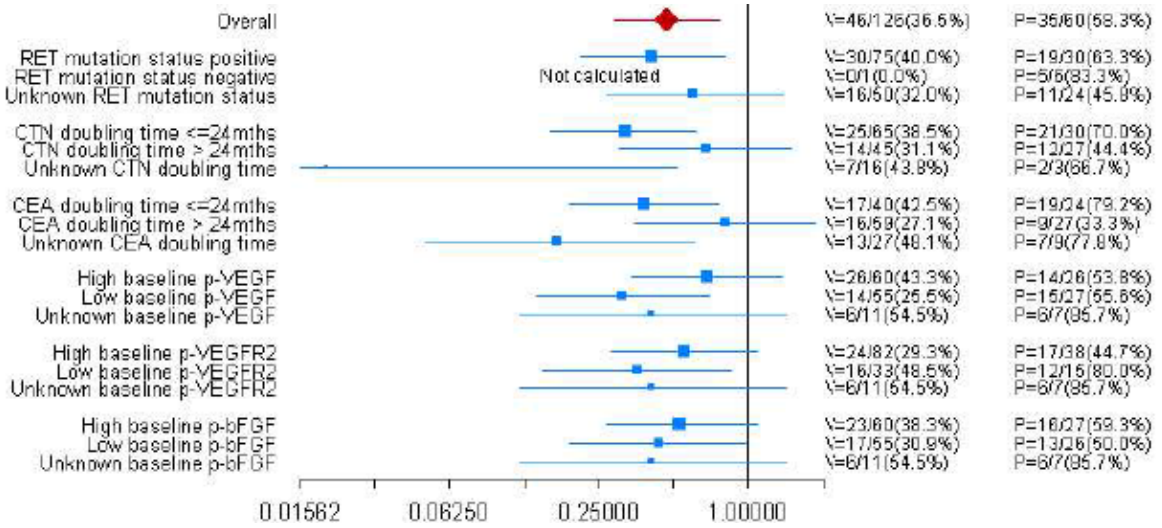


Abbildung 36: *Forest-Plot* der Subgruppenanalyse für PFS – *Baseline*- Charakteristiken und Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung



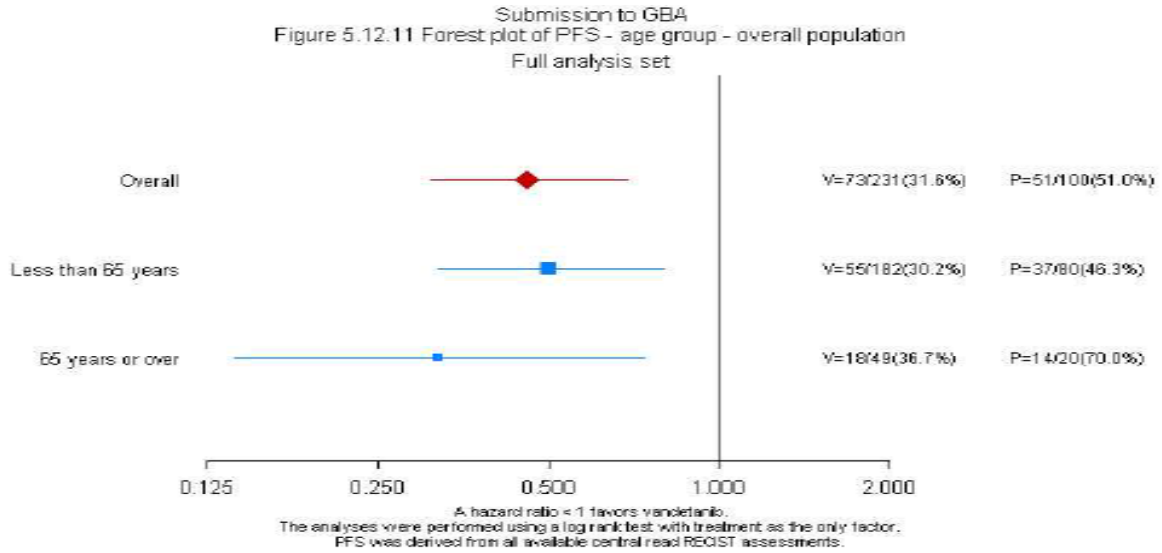
V = Vandetanib, P = Placebo

Abbildung 37: *Forest-Plot* der Subgruppenanalyse für PFS – Biomarker - Gesamtpopulation



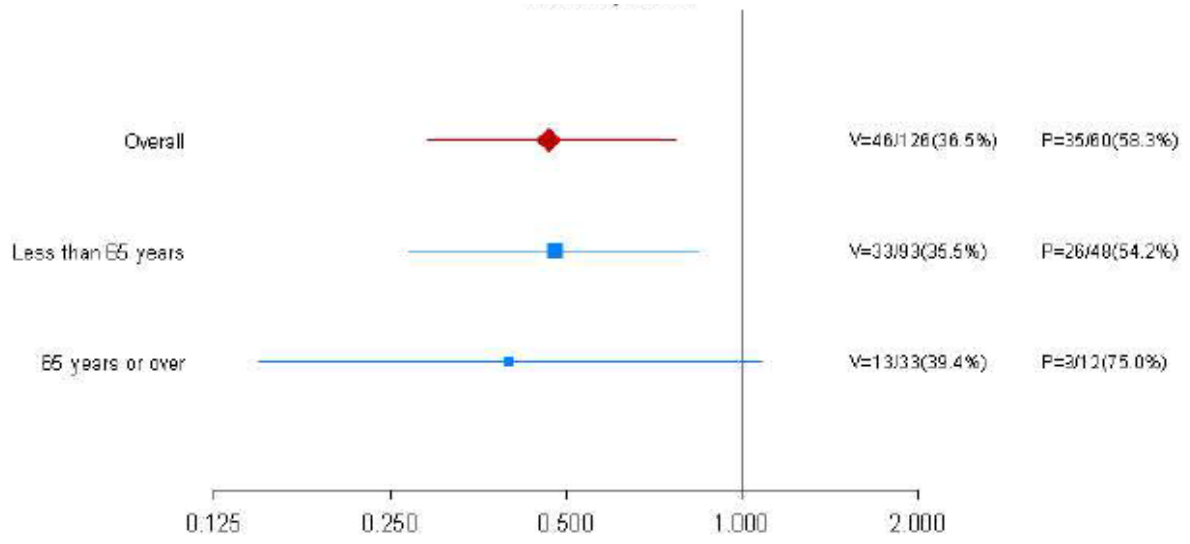
V = Vandetanib, P = Placebo

Abbildung 38: *Forest-Plot* der Subgruppenanalyse für PFS – Biomarker – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung



V = Vandetanib, P = Placebo

Abbildung 39: *Forest-Plot* der Subgruppenanalyse für PFS – Altersgruppen - Gesamtpopulation



V = Vandetanib, P = Placebo

Abbildung 40: *Forest-Plot* der Subgruppenanalyse für PFS – Altersgruppen – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Eine explorative Subgruppenanalyse nach Alter wurde für die Gesamtpopulation erst nach der Entblindung der Studiendaten durchgeführt. Mit einem *Hazard Ratio* von 0,50 (95%-KI: [0,32;0,80], p=0,0036) in der jüngeren Gruppe (<65 Jahre) und 0,32 (95%-KI: [0,14;0,74], p=0,0071) in der älteren Gruppe (>65 Jahre) zeichnete sich einheitlich ein statistisch signifikanter Unterschied des PFS zu Gunsten der Vandetanib-Gruppe ab. Für die genäherte Zielpopulation mit einem *Hazard Ratio* von 0,48 (95%-KI: [0,27;0,84], p=0,0101) in der jüngeren Gruppe (<65 Jahre) und 0,40 (95%-KI: [0,15;1,08], p=0,0696) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis nur für die jüngere Gruppe zu Gunsten der Vandetanib-Gruppe vor. Unter Einbezug der individuellen Interaktionstests liegt keine Effektmodifikation bzgl. des Alters vor. Die in beiden betrachteten Populationen beobachtete apparente Interaktion bezüglich des Parameters CEA Verdopplungszeit (Tabelle 4-129 und Tabelle 4-130) zeigte bei genauerer Betrachtung dennoch gleichgerichtete Effekte (Abbildung 37 und Abbildung 38) und ist somit nur quantitativer Natur.

Subgruppenanalysen – insbesondere bei nur einer vorliegenden Studie – sind allgemein kritisch zu sehen und vorsichtig zu interpretieren. Zusammenfassend existieren innerhalb der Subgruppen bzgl. *Baseline*-Charakteristika, Biomarkern, Krankheitsstatus und Plasma Biomarkern keine relevanten Unterschiede bezüglich des progressionsfreien Überlebens.

4.3.1.3.2.2 Zeit bis zur Schmerzprogression (TWP)

Für den Endpunkt TWP wurden in der Vorbereitung des vorgelegten Dossiers sowohl für die Gesamtpopulation als auch für die genäherte Zielpopulation alle für den Endpunkt PFS im Protokoll vordefinierten Subgruppenanalysen *post-hoc* re-analysiert. Die Analysen sind mit Vorsicht zu interpretieren, da die Patientenzahlen für die einzelnen Gruppen teilweise sehr

klein sind und somit zufällige Effekte entstehen können. Sie werden hier dennoch aufgeführt, um die Evidenz für Vandetanib unterstützend zu vervollständigen.

Tabelle 4-137: TWP Subgruppenanalysen nach *Baseline*-Charakteristika - Gesamtpopulation

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	p-Wert individueller Interaktionstest
Männlich	Vandetanib	134	59 (44,0)	0,51	[0,32;0,83]	0,0070	0,3763
	Placebo	56	33 (58,9)				
Weiblich	Vandetanib	97	55 (56,7)	0,74	[0,44;1,25]	0,2625	
	Placebo	44	24 (54,5)				
Kaukasisch	Vandetanib	218	109 (50,0)	0,60	[0,42;0,86]	0,0058	0,7454*
	Placebo	97	56 (57,7)				
Andere Rasse	Vandetanib	13	5 (38,5)	0,88	[0,09;8,67]	0,9158	
	Placebo	3	1 (33,3)				
WHO PS ≥ 1	Vandetanib	77	43 (55,8)	0,54	[0,33;0,89]	0,0164	0,3054
	Placebo	42	32 (76,2)				
WHO PS = 0	Vandetanib	154	71 (46,1)	0,77	[0,47;1,25]	0,2871	
	Placebo	58	25 (43,1)				

*In der ursprünglichen Auswertung wurde kein Test auf Interaktion durchgeführt, da die Anzahl von Patienten in der Gruppe „Andere Rasse“ für eine aussagekräftige Subgruppenanalyse als zu gering eingestuft wurde. Aus Gründen der Vollständigkeit wurde der hier dargestellte p-Wert *post-hoc* mit der Mantel-Haenszel Methode (Test auf Homogenität) berechnet.

Tabelle 4-138: TWP Subgruppenanalysen nach *Baseline*-Charakteristika - genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Männlich	Vandetanib	79	34 (43,0)	0,59	[0,33;1,07]	0,0815	0,9140
	Placebo	39	22 (56,4)				
Weiblich	Vandetanib	47	26 (55,3)	0,61	[0,27;1,35]	0,2203	
	Placebo	21	11 (52,4)				
Kaukasisch	Vandetanib	118	57 (48,3)	0,60	[0,37;0,96]	0,0335	NC
	Placebo	59	33 (55,9)				
Andere Rasse	Vandetanib	8	3 (37,5)	NC	NC	NC	
	Placebo	1	0 (0,0)				
WHO PS ≥= 1	Vandetanib	52	27 (51,9)	0,53	[0,28;0,99]	0,0475	0,4201
	Placebo	27	20 (74,1)				
WHO PS = 0	Vandetanib	74	33 (44,6)	0,81	[0,41;1,59]	0,5335	
	Placebo	33	13 (39,4)				

NC = Nicht berechenbar

Tabelle 4-139: TWP Subgruppenanalysen nach Biomarkern - Gesamtpopulation

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
RET-Mutationsstatus positiv	Vandetanib Placebo	137 50	66 (48,2) 29 (58,0)	0,47	[0,28;0,79]	0,0042	
RET-Mutationsstatus negativ	Vandetanib Placebo	2 6	0 (0,0) 4 (66,7)	NC	NC	NC	NC
RET-Mutationsstatus unbekannt	Vandetanib Placebo	92 44	48 (52,2) 24 (54,5)	0,85	[0,51;1,41]	0,5373	
CTN Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	124 46	57 (46,0) 23 (50,0)	0,75	[0,45;1,26]	0,2783	
CTN Verdopplungszeit > 24 Monate	Vandetanib Placebo	83 43	44 (53,0) 30 (69,8)	0,45	[0,27;0,77]	0,0032	0,2186
CTN Verdopplungszeit unbekannt	Vandetanib Placebo	24 11	13 (54,2) 4 (36,4)	0,99	[0,32;3,09]	0,9839	
CEA Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	69 33	38 (55,1) 16 (48,5)	0,86	[0,47;1,59]	0,6284	
CEA Verdopplungszeit > 24 Monate	Vandetanib Placebo	119 48	53 (44,5) 32 (66,7)	0,50	[0,30;0,82]	0,0057	0,1687
CEA Verdopplungszeit unbekannt	Vandetanib Placebo	43 19	23 (53,5) 9 (47,4)	0,54	[0,22;1,35]	0,1881	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

CTN/CEA Verdopplungszeit wurde abgeleitet aus der natürlichen Logarithmierung der einzelnen CTN/CEA Messungen (*pre-dose*). Zur Ermittlung der Steigung m wird an diese Daten wird eine lineare Regressionsgerade anpasst. Die CTN/CEA Verdopplungszeit ergibt sich als: m/ ln 2.

Negative Verdopplungszeiten werden der >24 Monate Gruppe zugewiesen

NC = Nicht berechenbar

Tabelle 4-140: TWP Subgruppenanalyse nach Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
RET-Mutationsstatus positiv	Vandetanib Placebo	75 30	35 (46,7) 17 (56,7)	0,43	[0,21;0,86]	0,0165	
RET-Mutationsstatus negativ	Vandetanib Placebo	1 6	0 (0,0) 4 (66,7)	NC	NC	NC	NC
RET-Mutationsstatus unbekannt	Vandetanib Placebo	50 24	25 (50,0) 12 (50,0)	1,02	[0,51;2,02]	0,9635	
CTN Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	65 30	27 (41,5) 15 (50,0)	0,87	[0,45;1,67]	0,6770	
CTN Verdopplungszeit > 24 Monate	Vandetanib Placebo	45 27	25 (55,6) 17 (63,0)	0,42	[0,21;0,85]	0,0166	0,1632
CTN Verdopplungszeit unbekannt	Vandetanib Placebo	16 3	8 (50,0) 1 (33,3)	0,00	[0,00;0,00]	0,0003	
CEA Verdopplungszeit ≤ 24 Monate	Vandetanib Placebo	40 24	23 (57,5) 11 (45,8)	1,32	[0,66;2,65]	0,4301	
CEA Verdopplungszeit > 24 Monate	Vandetanib Placebo	59 27	23 (39,0) 18 (66,7)	0,38	[0,19;0,77]	0,0075	0,0151
CEA Verdopplungszeit unbekannt	Vandetanib Placebo	27 9	14 (51,9) 4 (44,4)	0,30	[0,07;1,37]	0,1201	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

CTN/CEA Verdopplungszeit wurde abgeleitet aus der natürlichen Logarithmierung der einzelnen CTN/CEA Messungen (*pre-dose*). Zur Ermittlung der Steigung m wird an diese Daten wird eine lineare Regressionsgerade anpasst. Die CTN/CEA Verdopplungszeit ergibt sich als: m/ ln 2.

Negative Verdopplungszeiten werden der >24 Monate Gruppe zugewiesen

Tabelle 4-141: TWP Subgruppenanalyse nach Krankheitsstatus - Gesamtpopulation

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hereditäre oder Keimbahn-Mutation	Vandetanib	28	17 (60,7)	0,70	[0,17;2,83]	0,6177	0,8307
	Placebo	5	3 (60,0)				
Sporadische oder unbekannte Mutation	Vandetanib	203	97 (47,8)	0,60	[0,42;0,87]	0,0068	
	Placebo	95	54 (56,8)				
Metastatisch	Vandetanib	217	106 (48,8)	0,61	[0,43;0,88]	0,0077	0,6355*
	Placebo	97	55 (56,7)				
Lokal fortgeschritten	Vandetanib	14	8 (57,1)	0,91	[0,18;4,55]	0,9096	
	Placebo	3	2 (66,7)				
≥1 Vorbehandlung	Vandetanib	90	42 (46,7)	0,57	[0,33;0,97]	0,0371	0,9099
	Placebo	42	27 (64,3)				
Keine Vorbehandlung	Vandetanib	141	72 (51,1)	0,66	[0,41;1,05]	0,0797	
	Placebo	58	30 (51,7)				
Ansprechen nach Vorbehandlung	Vandetanib	5	4 (80,0)	2,07	[0,33;13,17]	0,4407	0,0583*
	Placebo	2	2 (100,0)				
Kein Ansprechen nach Vorbehandlung	Vandetanib	42	13 (31,0)	0,28	[0,12;0,62]	0,0019	
	Placebo	23	16 (69,6)				
NC/ unbekannte BOR zu Vorbehandlung	Vandetanib	184	97 (52,7)	0,71	[0,48;1,07]	0,1039	
	Placebo	75	39 (52,0)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

Vorbehandlung bezieht sich auf systemische Antikrebs-Therapie für medullären Schilddrüsenkrebs, die vor der ersten Dosis der randomisierten Behandlung eingenommen wurde.

BOR = bestes objektives Ansprechen (*Best Objective Response*)

NC = nicht berechenbar

*In der ursprünglichen Auswertung wurde kein Test auf Interaktion durchgeführt, da die Anzahl von Patienten in der Gruppe „lokal fortgeschritten“ bzw. „Ansprechen nach Vorbehandlung“ für eine aussagekräftige Subgruppenanalyse als zu gering eingestuft wurde. Aus Gründen der Vollständigkeit wurde der hier dargestellte p-Wert *post-hoc* mit der Mantel-Haenszel Methode (Test auf Homogenität) berechnet.

Tabelle 4-142: TWP Subgruppenanalyse nach Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hereditäre oder Keimbahn-Mutation	Vandetanib Placebo	11 2	7 (63,6) 1 (50,0)	0,38	[0,02;6,64]	0,5072	0,7237
Sporadische oder unbekannte Mutation	Vandetanib Placebo	115 58	53 (46,1) 32 (55,2)	0,60	[0,37;0,97]	0,0383	
Metastatisch	Vandetanib Placebo	119 59	57 (47,9) 33 (55,9)	0,64	[0,40;1,03]	0,0644	NC
Lokal fortgeschritten	Vandetanib Placebo	7 1	3 (42,9) 0 (0,0)	NC	NC	NC	
≥1 Vorbehandlung	Vandetanib Placebo	45 29	22 (48,9) 17 (58,6)	0,77	[0,40;1,49]	0,4424	0,4111
Keine Vorbehandlung	Vandetanib Placebo	81 31	38 (46,9) 16 (51,6)	0,56	[0,28;1,09]	0,0871	
Ansprechen nach Vorbehandlung	Vandetanib Placebo	3 1	2 (66,7) 1 (100,0)	NC	NC	NC	NC
Kein Ansprechen nach Vorbehandlung	Vandetanib Placebo	23 14	8 (34,8) 8 (57,1)	0,58	[0,21;1,64]	0,3068	
NB/ unbekannte BOR zu Vorbehandlung	Vandetanib Placebo	100 45	50 (50,0) 24 (53,3)	0,63	[0,37;1,08]	0,0945	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

Vorbehandlung bezieht sich auf systemische Antikrebs-Therapie für medullären Schilddrüsenkrebs, die vor der ersten Dosis der randomisierten Behandlung eingenommen wurde.

BOR = bestes objektives Ansprechen (*Best Objective* R*esponse*)

NC = nicht berechenbar

Tabelle 4-143: TWP Subgruppenanalysen nach Plasma Biomarkern - Gesamtpopulation

Studie D4200C00058					**Behandlungseffekt (Vandetanib:Placebo)**		
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hoher *Baseline* p-VEGF	Vandetanib	115	53 (46,1)	0,42	[0,24;0,71]	0,0012	
	Placebo	51	30 (58,8)				
Niedriger *Baseline* p-VEGF	Vandetanib	101	54 (53,5)	0,74	[0,43;1,27]	0,2747	0,2094
	Placebo	42	22 (52,4)				
Unbekannter *Baseline* p-VEGF	Vandetanib	15	7 (46,7)	1,30	[0,41;4,11]	0,6592	
	Placebo	7	5 (71,4)				
Hoher *Baseline* p-VEGFR2	Vandetanib	155	77 (49,7)	0,60	[0,39;0,93]	0,0208	
	Placebo	69	38 (55,1)				
Niedriger *Baseline* p-VEGFR2	Vandetanib	61	30 (49,2)	0,52	[0,25;1,09]	0,0844	0,6140
	Placebo	24	14 (58,3)				
Unbekannter *Baseline* p-VEGFR2	Vandetanib	15	7 (46,7)	1,30	[0,41;4,11]	0,6592	
	Placebo	7	5 (71,4)				
Hoher *Baseline* p-bFGF	Vandetanib	107	53 (49,5)	0,37	[0,22;0,63]	0,0003	
	Placebo	49	31 (63,3)				
Niedriger *Baseline* p-bFGF	Vandetanib	108	53 (49,1)	0,85	[0,50;1,46]	0,5663	0,0632
	Placebo	43	20 (46,5)				
Unbekannter *Baseline* p-bFGF	Vandetanib	16	8 (50,0)	1,14	[0,40;3,31]	0,8039	
	Placebo	8	6 (75,0)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.
Hoher p-VEGF >55,0 pg/ml bei *Baseline*. Niedrige p-VEGF ≤55,0 pg/ml bei *Baseline*.
Hoher p-VEGFR2 >9,88165 pg/ml bei *Baseline*. Niedrige p-VEGFR2 ≤9,88165 pg/ml bei *Baseline*.
Hoher p-bFGF >2,10 pg/ml bei *Baseline*. Niedrige p-bFGF ≤2,10 pg/ml bei *Baseline*.
Baseline ist definiert als der Wert am nächsten zu dem Wert vor Randomisierung.

Tabelle 4-144: TWP Subgruppenanalysen nach Plasma Biomarkern – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Hoher *Baseline* p-VEGF	Vandetanib Placebo	60 26	29 (48,3) 14 (53,8)	0,34	[0,15;0,77]	0,0091	
Niedriger *Baseline* p-VEGF	Vandetanib Placebo	55 27	26 (47,3) 14 (51,9)	0,75	[0,38;1,50]	0,4217	0,2331
Unbekannter *Baseline* p-VEGF	Vandetanib Placebo	11 7	5 (45,5) 5 (71,4)	1,38	[0,39;4,96]	0,6193	
Hoher *Baseline* p-VEGFR2	Vandetanib Placebo	82 38	40 (48,8) 19 (50,0)	0,61	[0,33;1,12]	0,1077	
Niedriger *Baseline* p-VEGFR2	Vandetanib Placebo	33 15	15 (45,5) 9 (60,0)	0,48	[0,18;1,27]	0,1421	0,7116
Unbekannter *Baseline* p-VEGFR2	Vandetanib Placebo	11 7	5 (45,5) 5 (71,4)	1,38	[0,39;4,96]	0,6193	
Hoher *Baseline* p-bFGF	Vandetanib Placebo	60 27	28 (46,7) 15 (55,6)	0,42	[0,20;0,89]	0,0230	
Niedriger *Baseline* p-bFGF	Vandetanib Placebo	55 26	27 (49,1) 13 (50,0)	0,76	[0,37;1,53]	0,4419	0,2685
Unbekannter *Baseline* p-bFGF	Vandetanib Placebo	11 7	5 (45,5) 5 (71,4)	1,38	[0,39;4,96]	0,6193	

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.

Hoher p-VEGF >55,0 pg/ml bei *Baseline*. Niedrige p-VEGF ≤55,0 pg/ml bei *Baseline*.

Hoher p-VEGFR2 >9,88165 pg/ml bei *Baseline*. Niedrige p-VEGFR2 ≤9,88165 pg/ml bei *Baseline*.

Hoher p-bFGF >2,10 pg/ml bei *Baseline*. Niedrige p-bFGF ≤2,10 pg/ml bei *Baseline*.

Baseline ist definiert als der Wert am nächsten zu dem Wert vor Randomisierung.

Tabelle 4-145: TWP Subgruppenanalyse nach Altersgruppe - Gesamtpopulation

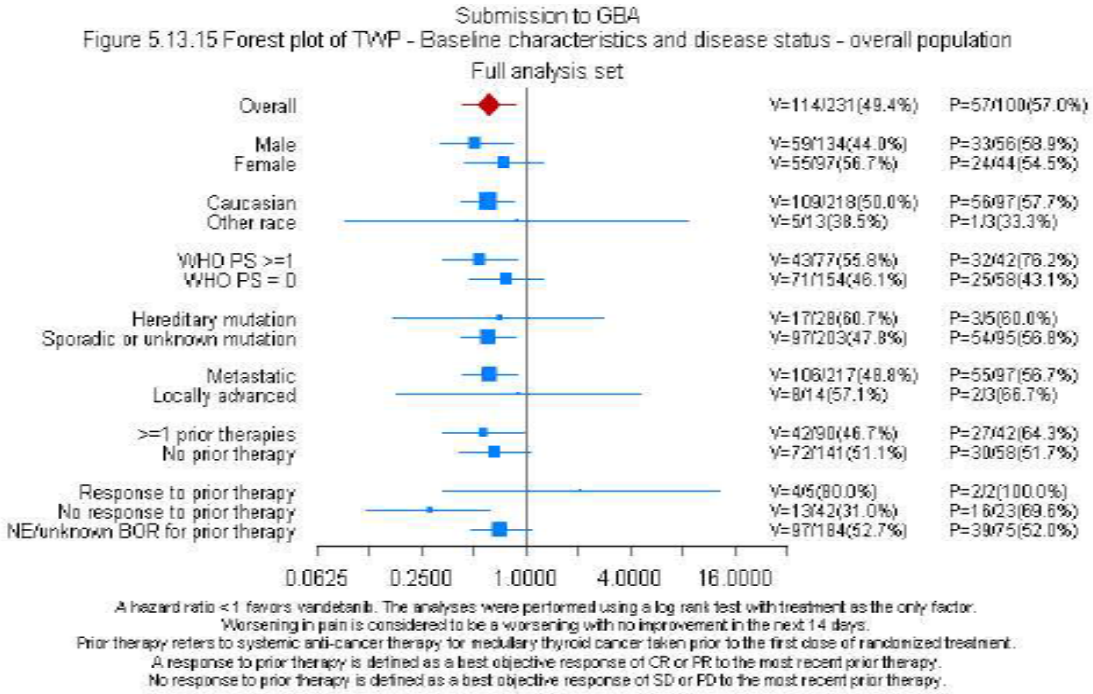
Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Jünger als 65 Jahre	Vandetanib	182	92 (50,5)	0,58	[0,40;0,86]	0,0058	0,6278
	Placebo	80	50 (62,5)				
65 Jahre und älter	Vandetanib	49	22 (44,9)	0,83	[0,33;2,05]	0,6829	
	Placebo	20	7 (35,0)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.

Tabelle 4-146: TWP Subgruppenanalyse nach Altersgruppe – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

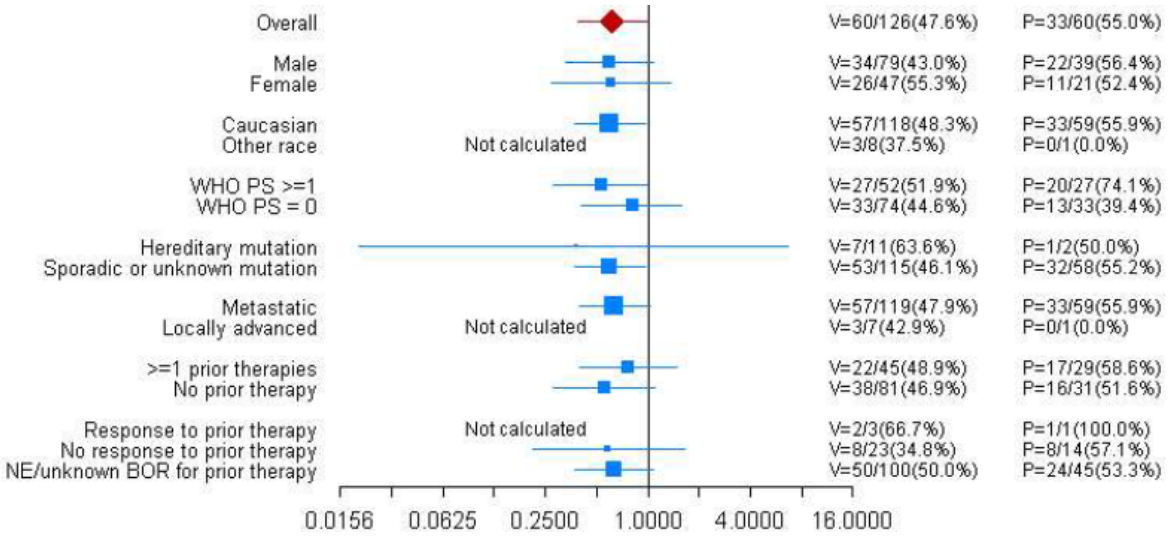
Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**			
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**p-Wert individueller Interaktionstest**
Jünger als 65 Jahre	Vandetanib	93	45 (48,4)	0,52	[0,31;0,88]	0,0142	0,1980
	Placebo	48	29 (60,4)				
65 Jahre und älter	Vandetanib	33	15 (45,5)	1,19	[0,41;3,49]	0,7472	
	Placebo	12	4 (33,3)				

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test.



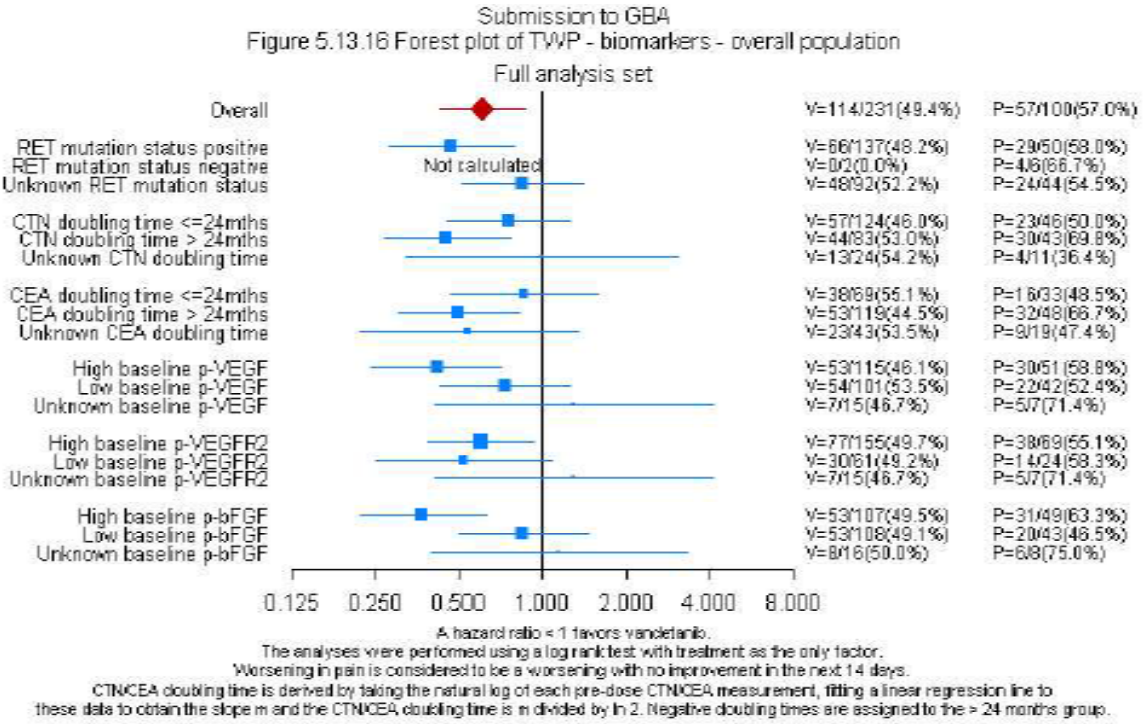
V = Vandetanib, P = Placebo

Abbildung 41: *Forest-Plot* der Subgruppenanalyse für TWP – *Baseline*- Charakteristiken und Krankheitsstatus – Gesamtpopulation



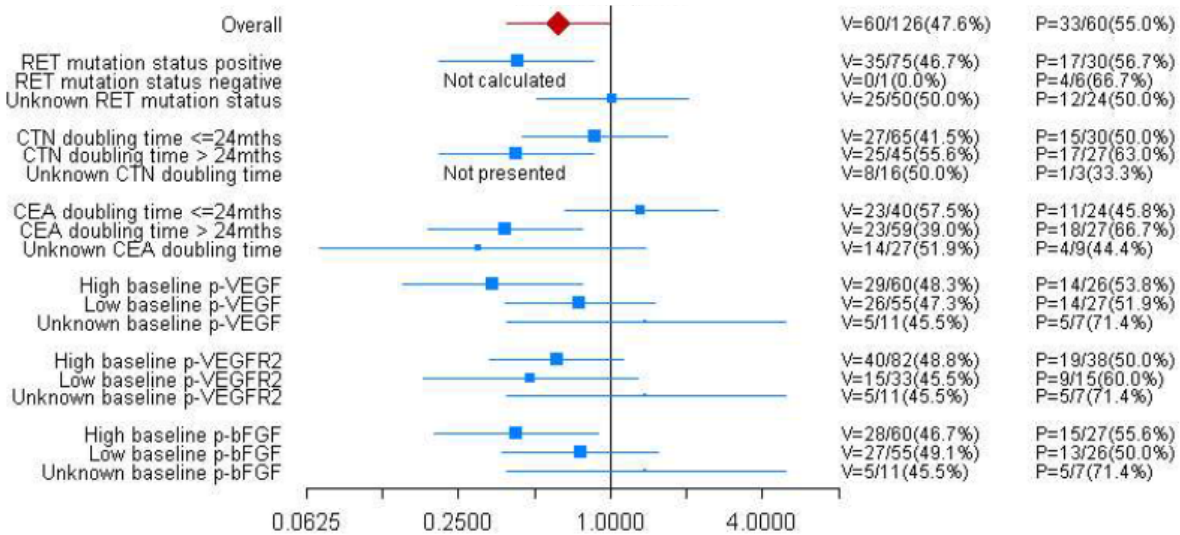
V = Vandetanib, P = Placebo

Abbildung 42: *Forest-Plot* der Subgruppenanalyse für TWP – *Baseline*- Charakteristiken und Krankheitsstatus – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung



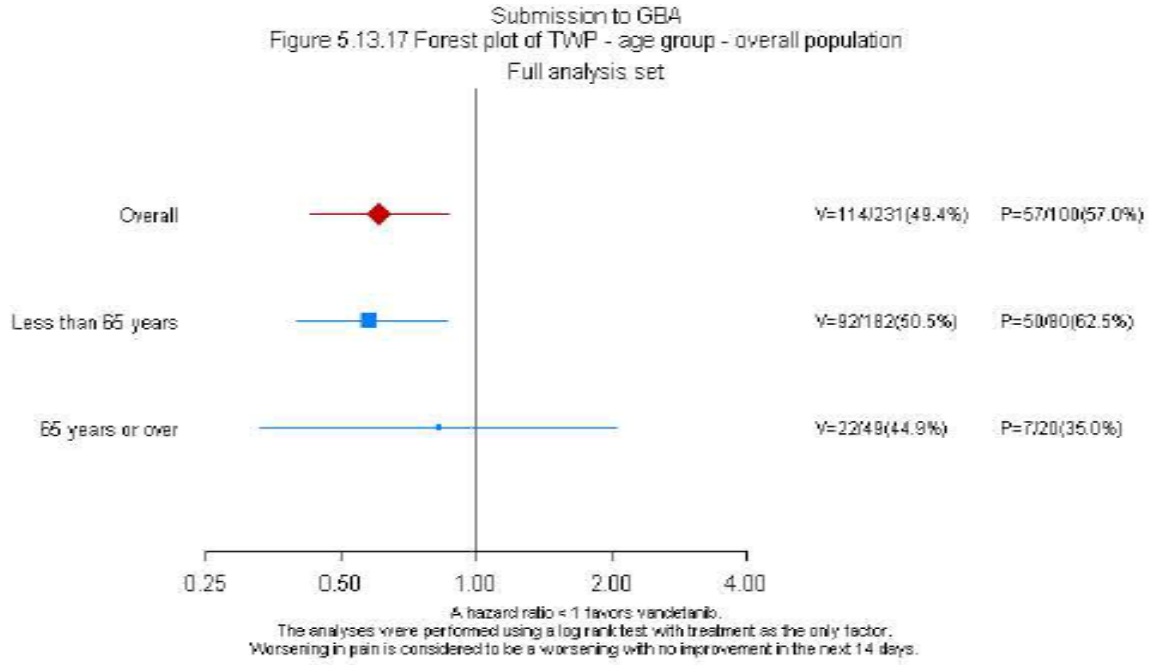
V = Vandetanib, P = Placebo

Abbildung 43: *Forest-Plot* der Subgruppenanalyse für TWP – Biomarker - Gesamtpopulation



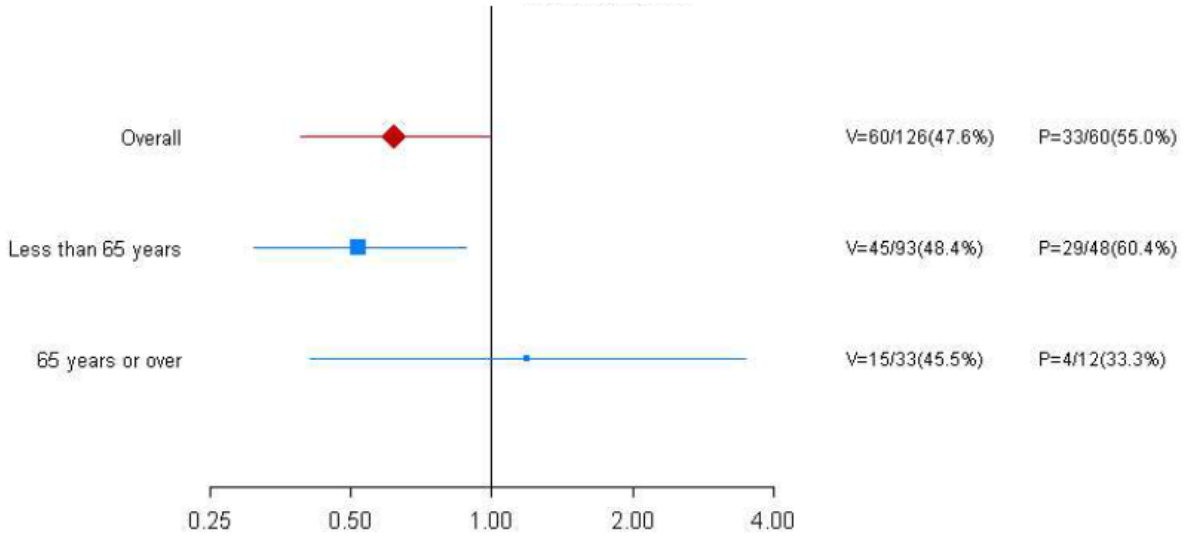
V = Vandetanib, P = Placebo

Abbildung 44: *Forest-Plot* der Subgruppenanalyse für TWP – Biomarker – genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung



V = Vandetanib, P = Placebo

Abbildung 45: *Forest-Plot* der Subgruppenanalyse für TWP – Altersgruppen – Gesamtpopulation



V = Vandetanib, P = Placebo

Abbildung 46: *Forest-Plot* der Subgruppenanalyse für TWP – Altersgruppen –genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung

Subgruppenanalysen – insbesondere bei nur einer vorliegenden Studie – sind allgemein kritisch zu sehen und vorsichtig zu interpretieren. Zusammenfassend existieren, unter Einbezug der individuellen Interaktionstests, innerhalb der Subgruppen bzgl. *Baseline*-Charakteristika, Biomarkern, Plasma Biomarkern und Krankheitsstatus keine relevanten Unterschiede bezüglich der Zeit bis zur Schmerzprogression.

Mit einem *Hazard Ratio* von 0,58 (95%-KI: [0,40;0,86], p=0,0058) in der jüngeren Gruppe (<65 Jahre) und 0,83 (95%-KI: [0,33;2,05], p=0,6829) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis zu Gunsten der Vandetanib-Gruppe in der Gesamtpopulation nur für die jüngere Gruppe vor. Hinsichtlich des TWP ergab sich für die Gesamtpopulation keine Interaktion durch das Alter.

Für die genäherte Zielpopulation mit einem *Hazard Ratio* von 0,52 (95%-KI: [0,31;0,88], p=0,0142) in der jüngeren Gruppe (<65 Jahre) und 1,19 (95%-KI: [0,41;3,49], p=0,7472) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis zu Gunsten der Vandetanib-Gruppe in der genäherten Zielpopulation nur für die jüngere Gruppe vor. Hinsichtlich des TWP ergab sich für die genäherte Zielpopulation knapp ein Hinweis auf eine signifikante Interaktion durch das Alter. Das Resultat ist jedoch auch aufgrund der unterschiedlichen Stichprobengrößen sowie der sehr geringen Anzahl an Patienten innerhalb der Gruppe der über 65-jährigen mit großer Vorsicht zu interpretieren.

4.3.1.3.2.3 RET-Mutations-Subgruppen

In Studie 58 wurde unter Verwendung des auf der Polymerase-Kettenreaktion (PCR) basierenden *Amplification Refractory Mutations System*-(ARMS)-Assays für die M918T-Mutation und der direkten Sequenzierung der DNA auf Mutationen in den Exons 10, 11, 13, 14, 15 und 16 (Stelle der M918T-Mutation) bei allen Patienten mit sporadischen MTC, bei denen DNA verfügbar war (297/298), ein Test auf RET-Mutation durchgeführt. Der RET-Status konnte allerdings bei 41,1% der Patienten nicht getestet werden (hauptsächlich aufgrund von nicht verfügbaren Ergebnissen der direkten Sequenzierung von DNA) [2, 42].

Im Rahmen des EU-Zulassungsverfahrens wurde eine *post-hoc* Subgruppen-Analyse des negativen RET-Status der pivotalen Studie 58 durchgeführt, die auf dem Fehlen der M918T-Mutation basierte. Eine RET-Mutation wurde angenommen, wenn bei einem Patienten entweder eine durch den ARMS-Assay nachgewiesene M918T-Mutation oder eine RET-Mutation in einem der sequenzierten Exons im Tumor vorhanden war. Die M918T-Mutation ist die am häufigsten beobachtete Mutation bei Patienten mit sporadischem MTC (siehe auch Modul 3 A, Abschnitt 3.2.1); es kann aber nicht ausgeschlossen werden, dass einige Patienten, die auf die M918T-Mutation RET-negativ getestet wurden, möglicherweise positiv für Mutationen in anderen Exons waren [42].

Die Ergebnisse für die Wirksamkeits-Endpunkte PFS, ORR und DOR gemäß RET-Status (positiv; unbekannt; RET-M918T-Mutations-negativ und andere Mutationen nicht getestet oder negativ) sind in den folgenden Tabellen dargestellt. Die entsprechenden Analysen für die

genäherte Zielpopulation der progressiven und symptomatischen Patienten wurden im Rahmen dieser Dossier-Erstellung vorgenommen.

Tabelle 4-147: PFS Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – Gesamtpopulation (ITT Population)

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**				**Test auf Inter-aktion**
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**geschätztes medianes PFS (Monate)**	**p-Wert**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	137 50	47 (34,3) 27 (54,0)	0,45	[0,26;0,78]	0,0041	29 18	
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	46 33	16 (34,8) 18 (54,5)	0,57	[0,29;1,13]	0,1087	28 18	0,854
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	48 17	10 (20,8) 6 (35,3)	0,44	[0,14;1,40]	0,1635	36 27	

** Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem Ergebnis)*

*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*

**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

Das mediane PFS wurde durch Anpassung eines Weibull-Modells in den beiden Behandlungsarmen geschätzt.

Der p-Wert für Interaktion wurde mit der Mantel-Haenszel Methode (Test auf Homogenität) berechnet.

Tabelle 4-148: PFS Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung (ITT Population)

Studie D4200C00058				**Behandlungseffekt (Vandetanib:Placebo)**				**Test auf Inter-aktion**
Subgruppe	**Behandlung**	**N**	**Anzahl Ereignisse (%)**	**Hazard Ratio**	**95%- KI**	**2-seitiger p-Wert**	**geschätztes medianes PFS (Monate)**	**p-Wert**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	75 30	30 (40,0) 19 (63,3)	0,41	[0,21;0,80]	0,0090	27 15	
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	31 20	11 (35,5) 14 (70,0)	0,47	[0,21;1,06]	0,0678	27 14	0,523
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	20 10	5 (25,0) 2 (20,0)	1,12	[0,23;5,59]	0,8871	34 82	

**Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem Ergebnis)*

*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*

**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*

Hazard Ratios basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem *Log-Rank* Test

Das mediane PFS wurde durch Anpassung eines Weibull-Modells in den beiden Behandlungsarmen geschätzt. „Symptomatisch“ zur Baseline waren definitionsgemäß Patienten mit Durchfall, Flush-Syndrom, Erschöpfungszuständen, Schmerzen, Atemwegssymptomen, Übelkeit, Gewichtsverlust, Appetitlosigkeit oder Dysphonie bei Baseline. „Progressiv“ zur Baseline waren definitionsgemäß Patienten, die innerhalb von 12 Monaten vor Randomisierung einen Progress erlitten hatten.

Der p-Wert für Interaktion wurde mit der Mantel-Haenszel Methode (Test auf Homogenität) berechnet.

Tabelle 4-149: ORR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus – Gesamtpopulation (ITT Population)

Studie D4200C00058					
Subgruppe	**Behandlung**	**N**	**Anzahl Patienten mit Ansprechen (%)**	**Odds ratio [95% KI]**	**Test auf Interaktion (p-Wert)**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	137 50	71 (51,8) 8 (16,0)	5,65 [2,47;12,91]	
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	46 33	16 (34,8) 4 (12,1)	3,87 [1,15;12,95]	0,773
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	48 17	17 (35,4) 1 (5,9)	8,77 [1,07;72,02]	

**Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem Ergebnis)*

*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*

**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*

Das beste objektive Ansprechen wurde zum Zeitpunkt der ersten eingetretenen RECIST Progression ermittelt, unter Berücksichtigung aller zentral verfügbaren RECIST Bewertungen. Die Objektive Ansprechrate entspricht dem Anteil der Patienten, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression (BOR, Best Objective Response), einem kompletten Ansprechen (CR, Complete Response) oder partiellen Ansprechen (PR, Partial Response) entspricht.

Der p-Wert für Interaktion wurde mit dem Breslow-Day-Test auf Homogenität der Odds-Ratios berechnet.

Tabelle 4-150: ORR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung (ITT Population)

Studie D4200C00058					
Subgruppe	**Behandlung**	**N**	**Anzahl Patienten mit Ansprechen (%)**	**Odds ratio [95% KI]**	**Test auf Interaktion (p-Wert)**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	75 30	38 (50,7) 6 (20,0)	4,11 [1,51;11,20]	
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	31 20	9 (29,0) 3 (15,0)	2,32 [0,54;9,90]	0,665
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	20 10	9 (45,0) 1 (10,0)	7,36 [0,78;69,58]	

**Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem Ergebnis)*

*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*

**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*

Das beste objektive Ansprechen wurde zum Zeitpunkt der ersten eingetretenen RECIST Progression ermittelt, unter Berücksichtigung aller zentral verfügbaren RECIST Bewertungen. Die Objektive Ansprechrate entspricht dem Anteil der Patienten, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression (BOR, Best Objective Response), einem kompletten Ansprechen (CR, Complete Response) oder partiellen Ansprechen (PR, Partial Response) entspricht.

Der p-Wert für Interaktion wurde mit dem Breslow-Day-Test auf Homogenität der Odds-Ratios berechnet.

Tabelle 4-151: DOR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - Gesamtpopulation

Studie D4200C00058			
Subgruppe	**Behandlung**	**N**	**Geschätzte mediane DOR (Monate)**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	71 8	22,6 12,6
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	16 4	18,3 16,3
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	17 1	31,6 NC

**Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem Ergebnis)*
*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*
**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*
NC = Nicht berechnet
Nur Patienten mit Ansprechen sind in der Tabelle enthalten. Die Dauer des Ansprechens wurde gemäß der Definition „Zeit vom ersten dokumentierten Ansprechen bis zur Progression oder Tod ohne vorhergehende dokumentierte Progression“, zusammengefasst. Die mediane Zeit bis zum Ansprechen, ist definiert als die Anzahl der Tage von der Randomisierung bis zum Ansprechen.

Tabelle 4-152: DOR Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus - genäherte Zielpopulation mit progressiver und symptomatischer Erkrankung

Studie D4200C00058			
Subgruppe	**Behandlung**	**N**	**Geschätzte mediane DOR (Monate)**
Patienten mit dokumentierter RET-Mutation*	Vandetanib Placebo	38 6	20,6 11,5
Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt***)**	Vandetanib Placebo	9 3	18,4 NC
Patienten mit unbekanntem RET-Status	Vandetanib Placebo	9 1	71,1 NC

**Patienten mit nachgewiesener Mutation (im ARMS-Assay oder mindestens eine direkte Sequenzierung mit positivem*
*** Bestätigter negativer RET-Mutationsstatus: Keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons sowie keine durch den ARMS-Assay nachgewiesene M918T-Mutation*
**** Unbestätigter negativer RET-Mutationsstatus:Keine durch den ARMS-Assay nachgewiesene M918T-Mutation sowie keine durch direkte Sequenzierung nachgewiesene Mutation der 6 untersuchten Exons, wobei nicht alle direkten Sequenzierungen erfolgreich durchgeführt werden konnten*
NC = Nicht berechnet
Nur Patienten mit Ansprechen sind in der Tabelle enthalten. Die Dauer des Ansprechens wurde gemäß der Definition „Zeit vom ersten dokumentierten Ansprechen bis zur Progression oder Tod ohne vorhergehende dokumentierte Progression“, zusammengefasst. Die mediane Zeit bis zum Ansprechen, ist definiert als die Anzahl der Tage von der Randomisierung bis zum Ansprechen.

Bei 79 Patienten der Gesamtpopulation wurde als die am häufigsten auftretende Mutation eine M918T-Mutation ausgeschlossen und auch keine RET-Mutation in einem der anderen 6 getesteten Exons gefunden (alle hatten sporadisches MTC; 71 dieser Patienten waren ursprünglich als „RET-Status unbekannt“ und 8 als „RET-negativ“ klassifiziert worden). Innerhalb der genäherten Zielpopulation waren dies 51 Patienten.

Bezüglich des PFS (Subgruppenanalyse nach *post-hoc* RET-Mutationsstatus) zeigt sich in der Subgruppe „Patienten mit dokumentierter RET-Mutation“ sowohl in der Gesamtpopulation als auch in der genäherten Zielpopulation ein statistisch signifikanter Unterschied zugunsten der Vandetanib-Gruppe mit 11 und 12 Monaten Verlängerung des PFS (HR: 0,45; 95%-KI: [0,26;0,78] bzw. HR: 0,41; 95%-KI: [0,21;0,80]).

Für die Subgruppe „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“ konnte für die Gesamtpopulation ein Trend für längeres progressionsfreies Überleben zugunsten von Vandetanib beobachtet werden, der jedoch nicht statistisch signifikant war (HR: 0,57; 95%-KI: [0,29;1,13]). Für die genäherte Zielpopulation wurde für die Subgruppe „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“ ein Vorteil zugunsten von Vandetanib beobachtet (13 Monate PFS-Vorteil), der ebenfalls statistisch nicht signifikant war (HR: 0,47; 95%-KI: [0,21;1,06]).

Das *Hazard Ratio* der Subgruppe „Patienten mit unbekanntem RET-Mutationsstatus“ beträgt für die Gesamtpopulation 0,44 (95%-KI: [0,14;1,40]) und liegt für die genäherte Zielpopulation über eins (HR: 1,12; 95%-KI: [0.23;5,59]), das Konfidenzintervall ist aufgrund der niedrigen Anzahl von Patienten (N=20 im Vandetanib-Arm und 10 Patienten im Placebo-Arm) aber sehr breit und lässt keine eindeutige Schlussfolgerung zu. Alle Auswertungen erfolgten nach dem ITT-Prinzip.

Das geschätzte mediane PFS für den Vandetanib-Arm entspricht einer Verlängerung in der Gesamtpopulation um ca. 10 Monate, in der genäherten Zielpopulation sogar um ca. 12 Monate für die Subgruppen „Patienten mit dokumentierter RET-Mutation“ und „Patienten mit negativem RET-Mutationsstatus“ (bestätigt oder unbestätigt).

Unabhängig von der betrachteten Patientenpopulation sprechen die Ergebnisse für alle drei Subgruppen deutlich für ein besseres ORR innerhalb der Vandetanib-Gruppe im Vergleich zur Placebo-Gruppe.

Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens liegen unter Vandetanib in den Subgruppen „Patienten mit dokumentierter RET-Mutation“ und „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt) für die Gesamtpopulation bei 22,6 bzw. 18,3 Monaten, verglichen mit Placebo mit 12,6 bzw. 16,3 Monaten. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens innerhalb der Subgruppe „Patienten mit unbekanntem RET-Mutationsstatus“ konnten für die Placebo-Gruppe nicht berechnet werden. Die Werte der Vandetanib-Gruppe liegen bei 31,6 Monaten.

Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens liegen unter Vandetanib in den Subgruppen „Patienten mit dokumentierter RET-Mutation“ für die genäherte Zielpopulation bei 20,6 Monaten, verglichen mit Placebo mit 11,5 Monaten. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens innerhalb der Subgruppen „Patienten mit negativem RET-Mutationsstatus“ (bestätigt oder unbestätigt) und „Patienten mit unbekanntem RET-Mutationsstatus“ konnten für die Placebo-Gruppe nicht berechnet werden. Die entsprechenden Werte der Vandetanib-Gruppe liegen bei 18,4 bzw. 71,1 Monaten.

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.

Vandetanib stellt eine klinisch bedeutsame Verbesserung bei der medizinischen Versorgung von Patienten mit MTC im zugelassenen Erkrankungsstadium dar. Dies konnte konsistent über eine große Anzahl von patientenrelevanten Endpunkten sowohl für die Gesamtpopulation als auch für die gesondert betrachtete genäherte Zielpopulation der Patienten mit progressiver und symptomatischer Erkrankung gezeigt werden.

Die Studie 58 zeigt für den primären Endpunkt „Progressionsfreies Überleben (PFS)“ mit einer Reduktion des Risikos um 54% eine statistisch signifikant bessere Wirkung von Vandetanib im Vergleich zu Placebo (HR: 0,46; 95%-KI: [0,31;0,69]). Dieses Resultat in der Gesamtpopulation der Studie ließ sich durch die Ergebnisse der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung bestätigen: hier lag die Risikoreduktion gegenüber Placebo bei 53% (HR: 0,47; 95%-KI: [0,29;0,77]). Die Wahrscheinlichkeit, nach 6, 12 und 24 Monaten progressionsfrei zu leben, liegt in beiden betrachteten Populationen in der Vandetanib-Gruppe im Schnitt um fast 20% über der in der Placebo-Gruppe. Am Daten *Cut-off* wies in der Gesamtpopulation ein um 20% kleinerer Anteil der Vandetanib-Gruppe eine Krankheitsprogression auf als in der Placebo-Gruppe; dieser Anteil lag in der Teilpopulation mit symptomatischer und progressiver Erkrankung sogar bei knapp 23%. Das geschätzte mediane progressionsfreie Überleben für die Vandetanib-Gruppe entspricht mit 30,5 (28) Monaten verglichen mit 19,3 (16) Monaten in der Placebo-Gruppe, einer Verlängerung des medianen PFS um ca. 11 Monate in der Gesamtpopulation und 12 Monate in der genäherten Zielpopulation.

Für die bezüglich des PFS durchgeführten Subgruppenanalysen existieren, unter Einbezug der individuellen Interaktionstests bzgl. *Baseline*-Charakteristika, Biomarkern, Krankheitsstatus, und Plasma Biomarkern, in beiden Populationen keine relevanten Unterschiede. Eine

explorative Subgruppenanalyse nach Alter wurde für die Gesamtpopulation erst nach der Entblindung der Studiendaten durchgeführt. Mit einem *Hazard Ratio* von 0,50 (95%-KI: [0,32;0,80], p=0,0036) in der jüngeren Gruppe (<65 Jahre) und 0,32 (95%-KI: [0,14;0,74], p=0,0071) in der älteren Gruppe (>65 Jahre) zeichnete sich einheitlich ein statistisch signifikanter Unterschied des PFS zu Gunsten der Vandetanib-Gruppe ab. Für die genäherte Zielpopulation mit einem *Hazard Ratio* von 0,48 (95%-KI: [0,27;0,84], p=0,0101) in der jüngeren Gruppe (<65 Jahre) und 0,40 (95%-KI: [0,15;1,08], p=0,0696) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis nur noch für die jüngere Gruppe zu Gunsten der Vandetanib-Gruppe vor. Unter Einbezug der individuellen Interaktionstests liegt keine Effektmodifikation bzgl. des Alters vor.

In der Subgruppenanalyse des PFS nach *post-hoc* RET-Mutationsstatus zeigt sich sowohl in der Gesamtpopulation als auch in der genäherten Zielpopulation ein statistisch signifikanter Unterschied zugunsten der Vandetanib-Gruppe in der Subgruppe „Patienten mit dokumentierter RET-Mutation“ (HR: 0,45; 95%-KI: [0,26;0,78] bzw. HR: 0,41; 95%-KI: [0,21;0,80]).

Für die Subgruppe „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“ konnte für die Gesamtpopulation ein Trend zugunsten von Vandetanib beobachtet werden, der jedoch nicht statistisch signifikant war (HR: 0,57; 95%-KI: [0,29;1,13]). Für die genäherte Zielpopulation wurde für diese Subgruppe sogar eine 53%ige Risikoreduktion zugunsten von Vandetanib beobachtet (13 Monate PFS-Vorteil), die statistisch aber ebenfalls nicht signifikant war (HR: 0,47; 95%-KI: [0,21;1,06]).

Die Subgruppe „Patienten mit unbekanntem RET-Mutationsstatus“ weist für die Gesamtpopulation ein *Hazard Ratio* von 0,44 (95%-KI: [0,14;1,40]) auf, dieses liegt für die genäherte Zielpopulation über eins (HR: 1,12; 95%-KI: [0.23;5,59]), das Konfidenzintervall ist aufgrund der niedrigen Anzahl von Patienten (N=20 im Vandetanib-Arm und 10 im Placebo-Arm) aber sehr breit und lässt keine eindeutige Schlussfolgerung zu. Das geschätzte mediane PFS für den Vandetanib-Arm entspricht einer Verlängerung in der Gesamtpopulation um ca. 10 Monate, in der genäherten Zielpopulation sogar um ca. 12 Monate für die Subgruppen „Patienten mit dokumentierter RET-Mutation“ und „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“.

Eines der bedeutsamsten Symptome für Patienten mit MTC ist, ebenso wie bei vielen anderen onkologischen Erkrankungen, der Schmerz. Der prädefinierte Endpunkt „Zeit bis zur Schmerzprogression (TWP)“ setzt sich zusammen aus den Angaben des Patienten zum Schmerzmittelgebrauch und den Antworten des BPI-SF. Es zeigte sich in der Gesamtpopulation eine statistisch signifikante Verbesserung der TWP unter Vandetanib im Vergleich zu Placebo mit einem *Hazard Ratio* von 0,61 (95%-KI: [0,43;0,87], p-Wert: 0,0062). In der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ergab sich ein *Hazard Ratio* von 0,62 (95%-KI: [0,39;0,99], p-Wert: 0,0446). Die mediane Zeit bis zur Schmerzprogression betrug innerhalb der Gesamtpopulation in der Vandetanib-Gruppe 7,85 Monate und in der Placebo-Gruppe 3,25 Monate, was einer Verbesserung von mehr als 4 Monaten entspricht. In der genäherten Zielpopulation mit symptomatischer und

progressiver Erkrankung betrug die Verbesserung sogar fast 8 Monate (11,07 vs. 3,42). Sowohl aus den dargestellten Kaplan-Meier-Kurven, welche die Zeit bis zur Schmerzprogression für beide Behandlungsgruppen zeigen, als auch aus den genannten Ergebnissen wurde ersichtlich, dass die Patienten der Vandetanib-Gruppe einen längeren Zeitraum erwarten konnten, bis sich Schmerzen entwickeln, eine Schmerzprogression eintritt oder eine Erhöhung der Opioiddosis nötig ist, um die Schmerzen zu kontrollieren.

Für die bezüglich der TWP *post-hoc* durchgeführten Subgruppenanalysen existieren, unter Einbezug der individuellen Interaktionstests, bzgl. *Baseline*-Charakteristika, Biomarkern, Plasma Biomarkern und Krankheitsstatus zwischen Gesamtpopulation und genäherter Zielpopulationen keine relevanten Unterschiede.

Mit einem *Hazard Ratio* von 0,58 (95%-KI: [0,40;0,86], p=0,0058) in der jüngeren Gruppe (<65 Jahre) und 0,83 (95%-KI: [0,33;2,05], p=0,6829) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis in der Gesamtpopulation zu Gunsten der Vandetanib-Gruppe nur für die jüngere Gruppe vor.

Für die genäherte Zielpopulation mit einem *Hazard Ratio* von 0,52 (95%-KI: [0,31;0,88], p=0,0142) in der jüngeren Gruppe (<65 Jahre) und 1,19 (95%-KI: [0,41;3,49], p=0,7472) in der älteren Gruppe (>65 Jahre) liegt ein statistisch signifikantes Ergebnis in der genäherten Zielpopulation zu Gunsten der Vandetanib-Gruppe nur für die jüngere Gruppe vor. Hinsichtlich des TWP ergab sich für die genäherte Zielpopulation knapp ein Hinweis auf eine signifikante Interaktion mit dem Alter.

Für die objektive Ansprechrate (ORR) konnte ebenfalls eine statistisch signifikante Verbesserung zugunsten von Vandetanib gegenüber Placebo gezeigt werden. In der Gesamtpopulation betrug die ORR 45% in der Vandetanib-Gruppe und 13% in der Placebo-Gruppe (OR: 5,48, 95%-KI: [2,99;10,79], p<0,0001). Die entsprechenden Werte in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung lagen bei 44,4% in der Vandetanib-Gruppe und 16,7% in der Placebo-Gruppe (OR: 4,00, 95%-KI: [1,93;9,01]). Der Effekt von Vandetanib bestätigte sich für 12 von 13 Patienten (bzw. 9 von 10 Patienten) der Placebo-Gruppe nach der Umstellung auf Vandetanib. In beiden betrachteten Populationen hatte nur ein Patient bereits unter Placebo ein beginnendes Ansprechen des Tumors. Für die verbleibenden 12 (9) Patienten der Placebo-Gruppe setzte das Ansprechen erst während der offenen Behandlungsphase unter Vandetanib ein.

Bezüglich der Subgruppenanalyse der ORR nach *post-hoc* RET-Mutationsstatus sprechen, unabhängig von der betrachteten Patientenpopulation, die Ergebnisse für alle drei Subgruppen deutlich für ein besseres ORR innerhalb der Vandetanib-Gruppe im Vergleich zur Placebo-Gruppe.

Die Krankheitskontrollrate war in beiden Populationen in der Vandetanib-Gruppe höher, in der Gesamtpopulation mit 86,6% unter Vandetanib im Vergleich zu 71,0% unter Placebo, in der genäherten Zielpopulation mit 84,1% gegenüber 68,3%. Es konnte eine statistisch signifikante Verbesserung mit einem *Odds Ratio* von 2,64 (2,46 in der genäherten

Zielpopulation) und einem 95%-Konfidenzintervall von [1,48;4,69] bzw. [1,19;5,09] (genäherte Zielpopulation) und einem p-Wert von 0,010 (0,0156 genäherte Zielpopulation) zugunsten von Vandetanib gegenüber Placebo gezeigt werden. Das Resultat der Sensitivitätsanalyse ohne Berücksichtigung der Daten nach Wechsel auf *Open-Label* Vandetanib zeigt einen noch deutlicheren Effekt für Vandetanib (OR: 3,58 95%-KI: [2,09;6,17] bzw. OR: 3,61 (95%-KI: [1,83;7,26]). Der Anteil der Patienten mit Krankheitskontrolle betrug hier in der Gesamtpopulation 84,8% gegenüber 61% (Vandetanib / Placebo), in der genäherten Zielpopulation war der Unterschied in der Krankheitskontrolle mit 82,5% gegenüber 56,7% noch größer.

Betrachtet man für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) die Kaplan-Meier-Kurven der Dauer des objektiven Ansprechens (DOR) für die Zeit vom ersten Ansprechen bis zur ersten Progression, so zeigt sich für beide Behandlungsarme ein lang anhaltendes Ansprechen, wobei ein Ansprechen für 12 bzw. 9 von 13 bzw. 10 Patienten der Placebo-Gruppe erst in der offenen Behandlungsphase unter Vandetanib einsetzte. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens liegen unter Vandetanib bei 22,2 (Gesamtpopulation) bzw. 21,8 (genäherten Zielpopulation) Monaten und unter Placebo bei 16,3 (Gesamtpopulation) bzw. 13,0 (genäherte Zielpopulation) Monaten. Die mediane Zeit von der Randomisierung bis zum Ansprechen betrug in der Vandetanib-Gruppe 5,8 (Gesamtpopulation) bzw. 7,1 (genäherte Zielpopulation) Monate und unter Placebo für beide Populationen 13,7 Monate. In einer zusätzlichen Auswertung für die mediane Zeit von der Randomisierung bis zum Ansprechen beträgt diese für den einen Patienten mit Ansprechen unter Placebo 12,2 Monate. Für die 12 verbleibenden Placebo-Patienten mit Ansprechen während der offenen Behandlungsphase (Gesamtpopulation) bzw. 9 Placebo-Patienten in der genäherten Zielpopulation beträgt die mediane Zeit von der ersten Dosis unter Vandetanib bis zum Ansprechen 6,9 Monate (Gesamtpopulation) bzw. 8,1 Monate (genäherte Zielpopulation).

Bezüglich der Subgruppenanalyse der DOR nach *post-hoc* RET-Mutationsstatus liegen die geschätzten medianen Werte für die Dauer des objektiven Ansprechens unter Vandetanib in den Subgruppen „Patienten mit dokumentierter RET-Mutation“ und „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“ für die Gesamtpopulation bei 22,6 bzw. 18,3 Monaten, verglichen mit 12,6 bzw. 16,3 Monaten unter Placebo. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens innerhalb der Subgruppe „Patienten mit unbekanntem RET-Mutationsstatus“ konnten für die Placebo-Gruppe nicht berechnet werden. Die Werte der Vandetanib-Gruppe liegen bei 31,6 Monaten.

Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens liegen unter Vandetanib in den Subgruppen „Patienten mit dokumentierter RET-Mutation“ für die genäherte Zielpopulation bei 20,6 Monaten, verglichen mit 11,5 Monaten unter Placebo. Die geschätzten medianen Werte für die Dauer des objektiven Ansprechens innerhalb der Subgruppen „Patienten mit negativem RET-Mutationsstatus (bestätigt oder unbestätigt)“ und „Patienten mit unbekanntem RET-Mutationsstatus“ konnten für die Placebo-Gruppe nicht

berechnet werden. Die entsprechenden Werte der Vandetanib-Gruppe liegen bei 18,4 bzw. 71,1 Monaten.

Für den Endpunkt „Gewichtszunahme“ zeigte sich in beiden Populationen (Gesamt- bzw. genäherte Zielpopulation) ein deutliches Ergebnis zugunsten von Vandetanib (OR: 3,24 95%-KI: [1,69;6,76] bzw. OR: 4,03 (95%-KI: [1,71;11,17]). Der Anteil der Patienten mit klinischer Gewichtszunahme betrug in der Vandetanib-Gruppe 28,6% bzw. 31,0% verglichen mit 11,0% bzw. 10,0% in der Placebo-Gruppe.

Für die Gesamtpopulation konnte die mediane Zeit bis zur Verschlechterung des WHO-PS wegen der geringen Anzahl an Ereignissen in der Vandetanib- Gruppe bis zum Daten C*ut-off* nicht errechnet werden. Für die Placebo-Gruppe betrug die mediane Zeit bis zur Verschlechterung des WHO-PS 24,9 Monate. Die mediane Zeit bis zur Verschlechterung des WHO-PS für die Vandetanib-Gruppe liegt in der genäherten Zielpopulation bei 22,6 Monaten, was, verglichen mit 15,4 Monaten in der Placebo-Gruppe, einer Verlängerung um ca. 7 Monate entspricht.

Für den explorativen Morbiditätsendpunkt „TWP in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn“ zeigt die visuelle Prüfung der *Forest-Plots* sowohl für die Gesamtpopulation, als auch für die genäherte Zielpopulation die Konsistenz des TWP über die zwei Behandlungsgruppen mit bekanntem Opioidanalgetika-Gebrauch „≥10 mg/Tag Opioidanalgetika-Gebrauch“ vs. „<10 mg/Tag Opioidanalgetika Gebrauch“ hinweg. Aufgrund sehr kleiner Populations- und Ereigniszahlen sind diese Ergebnisse allerdings in beiden Populationen mit Vorsicht zu interpretieren.

Für den explorativen Morbiditätsendpunkt „Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert“ zeigte sich für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) ein Trend in der Reduktion des Schmerzmittelgebrauchs zugunsten von Vandetanib (Gesamtpopulation OR: 2,78; 95%-KI: [0,77;10,80] bzw. genäherte Zielpopulation OR: 2,25; 95%-KI: [0,50;11,27]). Eine Reduktion des Schmerzmittelgebrauchs wurde in der Vandetanib-Gruppe von 62,5% bzw. 52,9% der Patienten erreicht verglichen mit 37,5% bzw. 33,3% in der Placebo-Gruppe. Die mediane Dauer der Reduktion des Schmerzmittelgebrauchs betrug für die Vandetanib-Gruppe 11,01 bzw. 11,56 Monate und für die Placebo-Gruppe in beiden Populationen 6,51 Monate. Alle Ergebnisse sind aufgrund der geringen Anzahl an Ereignissen mit Vorsicht zu interpretieren.

Auch für den Surrogatendpunkt „biochemisches Ansprechen (CTN und CEA)“, der zur Unterstützung der Ergebnisse zu PFS und ORR dargestellt wird, konnte unter Vandetanib eine signifikante Verbesserung gegenüber Placebo gezeigt werden: CTN-Ansprechen in der Gesamtpopulation: 69,3% versus 3,0% (OR: 72,86; 95%-KI: [26,22;303,2], p<0,0001) und in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung: 67,5% vs. 3,3% (OR: 60,12; 95%-KI: [17,56;378,0]), wobei hohe CTN-Werte klinisch mit sekretorischer Diarrhö verknüpft sind. Das CEA-Ansprechen lag in der Gesamtpopulation bei 51,5% versus 2,0% (OR: 52,03; 95%-KI: [15,95;320,3], p<0,0001). Für die genäherte

Zielpopulation mit symptomatischer und progressiver Erkrankung wurde aufgrund des fehlenden Ansprechens in der Placebo-Gruppe kein OR berechnet.

Die Resultate zum sekundären Wirksamkeitsendpunkt der Studie, Gesamtüberleben (OS), waren aufgrund der geringen Ereignisraten bis zum Zeitpunkt des Cut-off (15% der Patienten verstorben) und des möglichen Einflusses der erlaubten *Open-Label* Behandlung mit Vandetanib post Progression, wodurch eine Verzerrung der Ergebnisse wahrscheinlich ist, in beiden betrachteten Populationen nicht signifikant und werden im vorliegenden Dossier nicht für die Nutzenbewertung herangezogen.

Sowohl für den Endpunkt „gesundheitsbezogene Lebensqualität anhand des FACT-G“, als auch für den explorative Endpunkt „Einfluss der Behandlung auf Diarrhö“ wurden keine formalen statistischen Analysen durchgeführt. Es wurden über den langen zeitlichen Verlauf in beiden Behandlungsgruppen und Populationen (Gesamt-/ genäherte Zielpopulation) vergleichsweise stabile Gesamtscore-Werte erreicht. Aufgrund der unterschiedlichen und im Zeitverlauf stark abnehmenden Patientenzahlen ist eine Interpretierbarkeit der Ergebnisse und auch der Veränderung der gesundheitsbezogenen Lebensqualität bzw. der Stuhlfrequenz schwierig. Es ergab sich jedoch keine Evidenz dafür, dass die gesundheitsbezogene Lebensqualität bzw. die Stuhlfrequenz durch Vandetanib nachteilig beeinflusst wurde.

Bei nahezu allen Patienten beider Gruppen trat mindestens ein unerwünschtes Ereignis auf (Vandetanib vs. Placebo 99,6% vs. 90,9% in der Gesamtpopulation und 100% vs. 94,9% in der Teilpopulation mit symptomatischer und progressiver Erkrankung). Das Ergebnis bezüglich der Gesamtrate der unerwünschten Ereignisse zeigte für die Gesamtpopulation (OR: 23,00; 95%-KI: [2,87;184,16]), nicht aber für die Teilpopulation mit symptomatischer und progressiver Erkrankung (OR: 15,67; 95%-KI: [0,80;308,48]) statistisch signifikante Unterschiede zwischen den Behandlungsgruppen. Die Effektschätzer sind aufgrund der hohen absoluten Ereignishäufigkeiten und der Verwendung des Odds Ratios als Effektmaß mit Vorsicht zu interpretieren. Die Anteile der Patienten mit SUE (OR: 2,94; 95%-KI: [1,54;5,60] bzw. OR: 2,28; 95%-KI: [1,05;4,96]) und unerwünschten Ereignissen die zum Studienabbruch führten (OR: 4,41; 95%-KI: [1,31;14,80] bzw. RR: OR: 7,84; 95%-KI: [1,01;60,82]) waren unter Vandetanib im Vergleich zu Placebo statistisch signifikant größer. Die durch Vandetanib am häufigsten verursachten Nebenwirkungen können jedoch durch medizinische Interventionen sehr gut beherrscht werden. Die gemäß Studienprotokoll vorgesehene Möglichkeit der Dosisunterbrechung oder Dosisreduktion wurde in beiden Populationen von etwa der Hälfte der Vandetanib-Patienten gegenüber ca. 15% der Placebo-Patienten wahrgenommen. Der Anteil aufgrund unerwünschter Ereignisse verstorbener Patienten war in beiden Gruppen annähernd gleich (Gesamtpopulation 2,2% vs. 2,0% bzw. genäherte Zielpopulation 2,4 vs. 1,7%).

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. Benennen Sie

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

Es wurden keine indirekten Vergleiche durchgeführt.

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.

Es wurden keine indirekten Vergleiche durchgeführt.

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 4-153: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in

3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

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

Tabelle 4-154: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	**Referenzen Studien**	**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.

Es wurden keine indirekten Vergleiche durchgeführt.

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

Tabelle 4-155: 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-G. 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-G.

Tabelle 4-156: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Hinweise auf ergebnisgesteuerte Berichterstattung**	**Sonstige das Verzerrungspotenzial beeinflussende Punkte**	**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.

Es wurden keine indirekten Vergleiche durchgeführt.

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 4-157: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Es wurden keine indirekten Vergleiche durchgeführt.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

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

<< Angaben des pharmazeutischen Unternehmers >>

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 4.3.1.3.1.2.

Es wurden keine indirekten Vergleiche durchgeführt.

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. Benennen Sie

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

Es wurden keine nicht randomisierten vergleichenden Studien als Nachweis für einen Zusatznutzen herangezogen.

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-G. 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-G.

Tabelle 4-158: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	**Zeitliche Parallelität der Gruppen**	**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**	**Verblindung**		**Hinweise auf ergebnisgesteuerte Berichterstattung**	**Andere Aspekte, die Verzerrungen verursachen können**
			Patient	**Behandler**		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

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 4-159: 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-G. 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-G.

Tabelle 4-160: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

1) *Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar.*

Es wurden keine nichtrandomisierten vergleichenden Studien zur Bestimmung des Zusatznutzens herangezogen.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

Es wurden keine nichtrandomisierten vergleichenden Studien zur Bestimmung des Zusatznutzens herangezogen.

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. Benennen Sie

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

Über die Informationen aus RCT (Abschnitt 4.3.1.3) hinausgehend werden in den folgenden Abschnitten unterstützende Informationen aus nicht randomisierten Studien dargestellt. Diese Studien sind gemäß den in Abschnitt 4.2.2 angegeben Kriterien nicht geeignet, die Fragestellung gemäß 4.2.1 zu beantworten. Sie werden hier dennoch aufgeführt, um die Evidenz zu Vandetanib unterstützend zu vervollständigen. Ihre Resultate werden allerdings nicht für die Bewertung des medizinischen Nutzens oder die Quantifizierung des Zusatznutzens berücksichtigt.

Die Studien wurden in der Registersuche zu Vandetanib als weitere Studien im Indikationsgebiet identifiziert. Beide Studien sind Studien des pharmazeutischen Unternehmers.

Nachfolgend werden als weitere Unterlagen Daten aus den oben erwähnten nicht-randomisierten Studien im betrachteten Indikationsgebiet mit dem zu bewertenden Arzneimittel dargestellt und bewertet.

4.3.2.3.1.1 Studien des pharmazeutischen Unternehmers

Tabelle 4-161: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
D4200C00008	ja	abgeschlossen	Start: 11/2004 *Cut-off* für primäre Analyse: 2/2008	Vandetanib
D4200C00068	ja	abgeschlossen	Start: 8/2006 *Cut-off* für primäre Analyse: 1/2008	Vandetanib

Geben Sie an, welchen Stand die Information in Tabelle 4-161 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Tabelle 4-161 gibt einen Überblick über die Studien des pharmazeutischen Unternehmers im beschriebenen Anwendungsgebiet, die nicht bereits in Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Alle Studien sind nichtrandomisierte, nichtkontrollierte Studien mit dem zu

bewertenden Arzneimittel in dem in diesem Dossier beschriebenen Anwendungsgebiet. Der Stand der Information in der Tabelle ist zum 07.01.2013.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-161 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-162: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – weitere Untersuchungen

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
D4200C00008	Nicht randomisierte Studien sind laut Fragestellung nicht Bestandteil der relevanten Studien.
D4200C00068	Nicht randomisierte Studien sind laut Fragestellung nicht Bestandteil der relevanten Studien.

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

4.3.2.3.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. 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 Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[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 der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

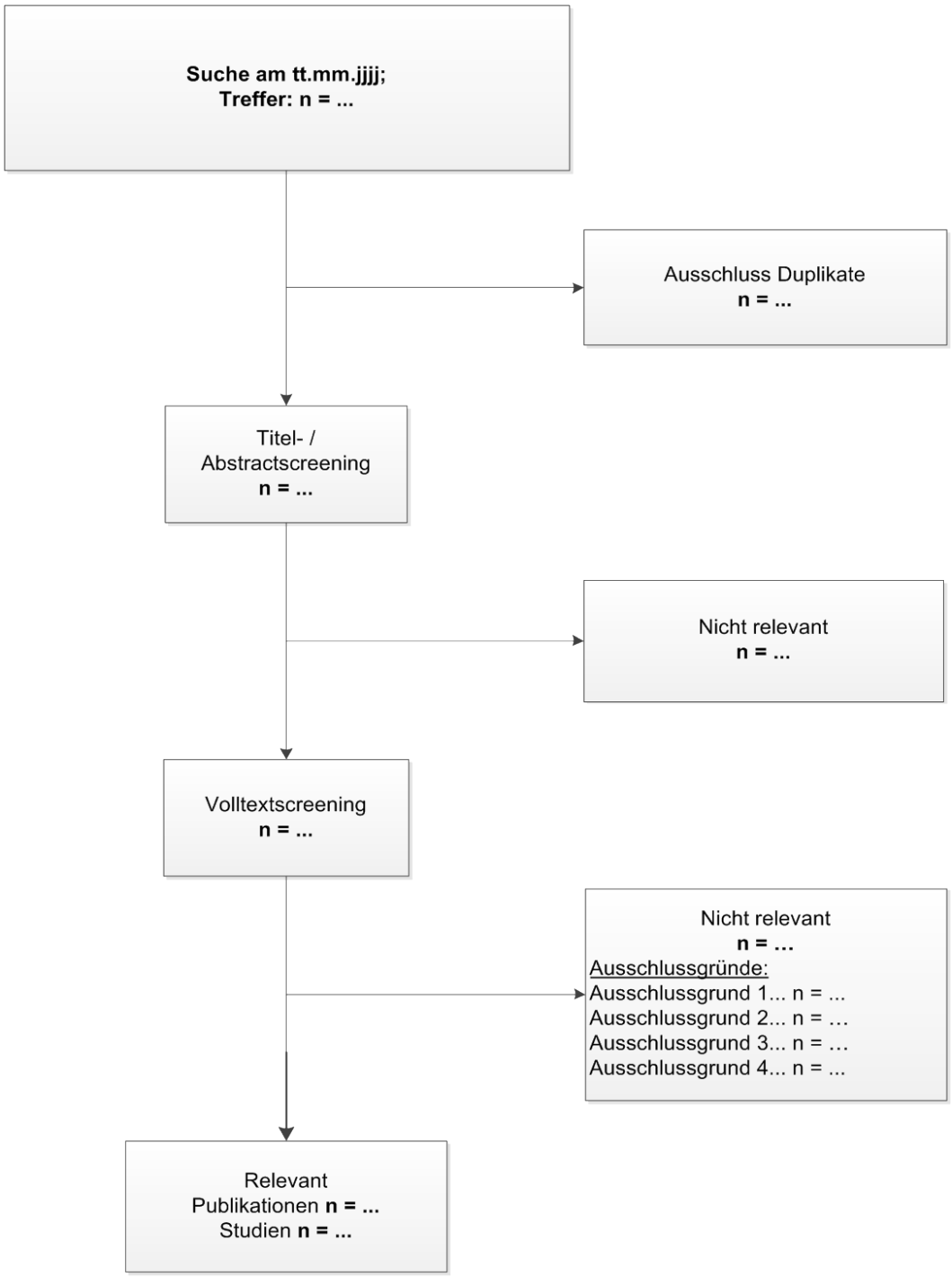


Abbildung 47: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen

Die im Abschnitt 4.3.2.3 dargestellten Ergebnisse dienen nur zur Illustration der in Abschnitt 4.3.1.3 präsentierten Daten und werden für die eigentliche Beantwortung der in 4.2.1 spezifizierten Fragestellung und somit zur Quantifizierung des Zusatznutzens von Vandetanib gegenüber BSC nicht herangezogen. Aufgrund dessen wurde keine bibliographische Literaturrecherche durchgeführt.

4.3.2.3.1.3 Studien aus der Suche in Studienregistern

Tabelle 4-163: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate[a])	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.[b])
D4200C00008	Clinicaltrials.gov: NCT00098345 [97]	ja	nein
D4200C00068	Clinicaltrials.gov: NCT00358956 [98]	ja	nein
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.2.3.1.4 Resultierender Studienpool: weitere Untersuchungen

Siehe Abschnitt 4.3.2.3.1.1.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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.

Zu weiteren Unterlagen wurden auf der Basis der Studienregistersuche zwei nicht-randomisierte, nichtkontrollierte Studien identifiziert: D4200C00008 und D4200C00068.

4.3.2.3.2.1 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-F zu hinterlegen.

Tabelle 4-164: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen

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**
D4200C00008	Phase-II, offen, 2-stufig, multizentrisch, einarmig	Patienten mit lokal fortgeschrittenem oder metastasierendem hereditärem MTC	n=30 (15 pro Stufe), Vandetanib 300 mg Tablette, oral, täglich bis Progression	Bis Progression, RECIST Messungen alle 4 Wochen (nach Amendment 7 alle 12 Wochen)	USA und Frankreich; 11/2004 – 8/2006	Primärer Endpunkt: • Objektive Ansprechrate (ORR) Patientenrelevante sekundäre Endpunkte: • Progressionsfreies Überleben (PFS) • WHO *Performance Status* (WHO PS) • Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACIT-D • Unerwünschte Ereignisse (UE)
D4200C00068	Phase-II, offen, multizentrisch, einarmig	Patienten mit lokal fortgeschrittenem oder metastasierendem hereditärem MTC für die keine Standardtherapie mehr verfügbar war	n= 19 Vandetanib 100 mg Tablette, oral, täglich bis Progression	Bis Progression, RECIST Messungen alle 12 Wochen	Australien, Kanada, Italien, Niederlande, Rumänien, Spanien, Schweiz und USA; 08/2006 – 01/2008	Primärer Endpunkt: • Objektive Ansprechrate (ORR) Patientenrelevante sekundäre Endpunkte: • Progressionsfreies Überleben (PFS) • WHO *Performance Status* (WHO PS) • Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G • Unerwünschte Ereignisse (UE)

* Bei beiden einarmigen Studien fand keine Randomisierung statt. Die Zahl der behandelten Patienten ist dargestellt.
FACT-G: *Functional Assessment of Cancer Therapy General Scale*; FACIT-D: *Functional Assessment of Chronic Illness Therapy – Diarrhea*; MTC: Medulläres Schilddrüsenkarzinom (*Medullary Thyroid Cancer*); ORR: Objektive Ansprechrate (*Objective Response Rate*); PFS: Progressionsfreies Überleben (*Progression-free Survival*); QoL: Lebensqualität (*Quality of Life*); RECIST: *Response Evaluation Criteria for Solid Tumors*; UE: Unerwünschte Ereignisse; WHO PS: *WHO Performance Status*.

Tabelle 4-165: Charakterisierung der Intervention – weitere Untersuchungen

Studie	**<Gruppe 1>**	**<Gruppe 2>**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
D4200C00008	Vandetanib 300 mg Tablette, oral, täglich bis Progression	-	Keine bestimmten Vorbehandlungen, keine Run-in-Phase
D4200C00068	Vandetanib 100 mg Tablette, oral, täglich bis Progression	--	Keine bestimmten Vorbehandlungen, keine Run-in-Phase

Tabelle 4-166: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 1)

Studie	**N**	**Alter (Jahre) Mittelwert (SD) [Spannweite]**	**Geschlecht w /m (%)**	**ethnische Zugehörigkeit (%)**		**Gewicht (kg) Mittelwert (SD) [Spannweite]**	**Größe (cm) Mittelwert (SD) [Spannweite]**
				K	**A**		
D4200C00008	30	48,7 (14,43) [20-77]	70,0 / 30,0	96,7	3,3	65,9 (20,71) [45,4-135,9]	165,5 (8,89) [149-188]
D4200C00068	19	44,7 (14,1) [22-79]	31,6 / 68,4	94,7	5,3	75,11 (12,107) [57,0-98,0]	174,07 (9,946) [155,0-195,0]

K=Kaukasisch, A=Andere

Tabelle 4-167: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 2)

Studie	**N**	**Zeit seit Diagnose (Jahre) Mittelwert (SD) [Spannweite]**	**Vorangegangene Theraphien für das MTC (%)**				***WHO Performance Status* (%)**		
			A	**C**	**R**	**O**	**0**	**1**	**2**
D4200C00008	30	16 (11) [1-43]	26,7	$	36,7	100	50,0	46,7	3,3
D4200C00068	19	13 (10) [5-33]	31,6	10,5	21,0	100*	84,2	5,3	10,5

A=Krebstherapie, C=Chemotherapie, R=Radiotherapie, S=OP
$ wurde nicht abgefragt und unter dem Stichpunkt „Andere“ dokumentiert
* OP oder unvollständig verheilter Eingriff

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

Beide Studien waren einarmige Phase-II-Studien mit Patienten mit lokal fortgeschrittenem oder metastasierendem hereditärem MTC. Primärer Endpunkt war die „objektive Ansprechrate“. Weitere Endpunkte waren unter anderem „progressionsfreies Überleben“, „gesundheitsbezogene Lebensqualität“ und „unerwünschte Ereignisse“.

Die Studie D4200C00008 hatte ein zweistufiges Studiendesign und die Patienten bekamen 300 mg Vandetanib täglich bis zur Progression. Die Studie D4200C00068 hatte ein einstufiges Studiendesign und die Patienten bekamen 100 mg Vandetanib täglich bis zur Progression.

Studie D4200C00008 wurde beendet, als alle Patienten eine Beobachtungszeit von einem Jahr aufwiesen.

4.3.2.3.2.2 Verzerrungspotenzial auf Studienebene

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

Tabelle 4-168: Verzerrungspotenzial auf Studienebene – Weitere Untersuchungen mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver-zerrungspotenzial beeinflussende Punkte	Verzerrungs-potenzial auf Studienebene
			Patient	Behandler			
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
D4200C00008	nein*	nein*	nein*	nein*	nein	nein	hoch**
D4200C00068	nein*	nein*	nein*	nein*	nein	nein	hoch**

* nicht anwendbar, da die Studie ein einarmiges Design hatte
** Gilt aufgrund des Studiendesigns generell als potenziell hoch verzerrt

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des Studiendesigns (offen, nicht vergleichend) der beiden Studien existiert keine Vergleichsgruppe. Folglich waren in beiden Studien sowohl die Patienten als auch die Prüfärzte unverblindet. Die Aussagekraft nicht vergleichender Studien ist wegen des Fehlens einer Vergleichsgruppe, die etwa hinsichtlich der natürlichen Krankheitsentwicklung aber auch hinsichtlich sogenannter Placeboeffekte adjustiert, gering einzustufen.

Im Hinblick auf den Zugewinn an Informationen, z. B. die Nachhaltigkeit des Behandlungseffekts oder Nutzenaussagen für begrenzte Populationen scheint jedoch die Vernachlässigung solcher Informationen in der Zusatznutzenbewertung kritisch.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Alle dargestellten Endpunkte der beiden Studien wurden auf der Basis aller Patienten, die mindestens eine Vandetanib-Dosis erhalten hatten, ausgewertet.

4.3.2.3.3.1 Objektive Ansprechrate (ORR) – 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 4-169: Operationalisierung der objektiven Ansprechrate

Studie	Operationalisierung
D4200C00008	Die objektive Ansprechrate ist definiert als der Anteil der Patienten, die ein bestes objektives Ansprechen von CR (komplettes Ansprechen) oder PR (partielles Ansprechen) erreichten. Das beste objektive Ansprechen wurde entsprechend den modifizierten RECIST Kriterien bestimmt. Hierfür wurden alle 4 Wochen (nach Amendment 7 alle 12 Wochen) radiologische Untersuchungen durchgeführt.
D4200C00068	Die objektive Ansprechrate ist definiert als der Prozentsatz von Patienten, die ein bestes objektives Ansprechen von CR (komplettes Ansprechen) oder PR (partielles Ansprechen) erreichten. Das beste objektive Ansprechen wurde entsprechend den modifizierten RECIST Kriterien bestimmt. Hierfür wurden alle 12 Wochen radiologische Untersuchungen durchgeführt.

Bewerten Sie das Verzerrungspotenzial 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.

Da die Einstufung des endpunktübergreifenden Verzerrungspotentials für nichtkontrollierte Studien generell als „hoch“ erfolgt, wird das Verzerrungspotenzial für die betrachteten Endpunkte auch als „hoch“ bewertet. Die CT/MRT-Aufnahmen wurden sowohl durch das Prüfzentrum als auch durch einen zentralen unabhängigen Review bewertet. Die zur Bewertung verwendeten RECIST Kriterien stellen ein etabliertes Bewertungsschema dar. Die primäre Analyse basierte auf dem im Prüfzentrum durchgeführten Review.

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.

Der primäre Endpunkt „objektive Ansprechrate“ ist für beide Studien in Tabelle 4-170 dargestellt.

Tabelle 4-170: Ergebnisse für die objektive Ansprechrate (ORR) – Weitere Untersuchungen

Studie	N	ORR (%) Wert [95%- KI]	Ansprechen n (%)				
			CR	PR	SD	PD	NA
D4200C00008	30	20 [7,7;38,6]	0 (0,0)	6 (20,0)	22 (73,3)	1 (3,3)	1 (3,3)
D4200C00068	19	15,8 [3,4;39,6]	0 (0,0)	3 (15,8)	12 (63,2)	3 (15,8)	1 (5,3)

CR=komplette Remission, PR= Partielles Ansprechen, SD=stabiler Krankheitsverlauf, PD= Progressive Erkrankung, NA=nicht auswertbar

Der primäre Endpunkt wurde in beiden Studien erreicht. In der Studie D4200C00008 wurde ein ORR von mindestens 20% prospektiv als Erfolg definiert. In der Studie D4200C00068 wurde ein ORR größer 0% prospektiv als Erfolg definiert.

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

4.3.2.3.3.2 Progressionsfreies Überleben (PFS) – weitere Untersuchungen

Tabelle 4-171: Operationalisierung des progressionsfreien Überleben

Studie	Operationalisierung
D4200C00008	Der Zeitraum zwischen erster Vandetanibgabe und einer objektiven Progression oder Tod. Eine Progression wurde entsprechend den modifizierten RECIST Kriterien bestimmt. Hierfür wurden alle 4 Wochen (nach Amendment 7 alle 12 Wochen) radiologische Untersuchungen durchgeführt. Als Zensierungsdatum wurde das Datum der letzten radiologischen Untersuchung für Patienten ohne Ereignis verwendet.
D4200C00068	Der Zeitraum zwischen erster Vandetanibgabe und einer objektiven Progression oder Tod. Eine Progression wurde entsprechend den modifizierten RECIST Kriterien bestimmt. Hierfür wurden alle 12 Wochen radiologische Untersuchungen durchgeführt. Als Zensierungsdatum wurde das Datum der letzten radiologischen Untersuchung für Patienten ohne Ereignis verwendet.

Da die Einstufung des endpunktübergreifenden Verzerrungspotentials für nichtkontrollierte Studien generell als „hoch“ erfolgt, wird das Verzerrungspotenzial für die betrachteten Endpunkte auch als „hoch“ bewertet. Die CT/MRT-Aufnahmen wurden sowohl durch das Prüfzentrum als auch durch einen zentralen unabhängigen Review bewertet. Die primäre Analyse basierte auf der Auswertung am Prüfzentrum. Die zur Bewertung verwendeten RECIST Kriterien stellen ein etabliertes Bewertungsschema dar.

Tabelle 4-172: Ergebnisse für das Progressionsfreies Überleben (PFS) – Weitere Untersuchungen

Studie	N	PFS (Monate) Wert [95% -KI]	Status* n (%)		
			Progression	Tod	Ereignisfrei
D4200C00008	30	27,9 [19,4;NC]	8 (26,7)	0 (0,0)	22 (73,3)
D4200C00068	19	15,8 [3,4;39,6]	4 (21,1)	1 (5,3)	14 (73,7)

NC = Nicht berechenbar
* Daten bis einschließlich dem 22. Februar 2008 (D4200C00008) bzw. dem 31. Januar 2008 (D4200C00068)

Die Ergebnisse für das Progressionsfreie Überleben sind für beide Studien in Tabelle 4-172 dargestellt.

4.3.2.3.3.3 WHO *Performance Status* (WHO PS) – 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 4-173: Operationalisierung des WHO *Performance Status*

Studie	Operationalisierung
D4200C00008	Der WHO *Performance Status* wurde alle 4 Wochen (nach Amendment 7 alle 12 Wochen) erhoben. Die Operationalisierung des Endpunkts erfolgte anhand der Veränderung des WHO *Performance Status* zu Woche 24 verglichen zum Ausgangswert
D4200C00068	Der WHO *Performance Status* wurde alle 12 Wochen erhoben. Die Operationalisierung des Endpunkts erfolgte anhand der Veränderung des WHO *Performance Status* zu Woche 24 verglichen zum Ausgangswert

Wegen des endpunktübergreifend als „hoch“ bewerteten Verzerrungspotentials wird auch das Verzerrungspotential für den Endpunkt „WHO *Performance Status*“ als „hoch“ eingeschätzt.

Tabelle 4-174: Ergebnisse des WHO *Performance Status* (WHO PS) – Weitere Untersuchungen

Studie	N	Veränderung zwischen dem Ausgangswert (*Baseline*) und der 24 Wochen (Follow-up) n (%)				Kein Wert in Woche 24
		Verbesserung	Gleichbleibend	Verschlechterung		
		-1	0	+1	+2	
D4200C00008	30	2 (6,7)	14 (46,7)	3 (10,0)	1 (3,3)	10 (33,3)
D4200C00068	19	0 (0,0)	11 (57,9)	2 (10,5)	0 (0,0)	6 (31,6)

Die Auswertung des WHO *Performance Status* ist für beide Studien in Tabelle 4-174 dargestellt

4.3.2.3.3.4 Gesundheitsbezogene Lebensqualität (HRQoL) – Weitere Untersuchungen

Tabelle 4-175: Operationalisierung der gesundheitsbezogenen Lebensqualität

Studie	Operationalisierung
D4200C00008	Der FACIT-D (Functional Assessment of Chronic Illness Therapy – Diarrhea) wurde alle 4 Wochen (nach Amendment 7 alle 12 Wochen) erhoben. Die Erhebung fand zu den jeweiligen Besuchszeitpunkten vor Beginn aller weiteren Untersuchungen statt. Der FACIT-D besteht aus 27 Aussagen bzgl. des physischen, sozialen, emotionalen und funktionellen Wohlbefindens. Höhere FACIT-D Werte sprechen für eine bessere Lebensqualität.
D4200C00068	Der FACT-G (Functional Assessment of Cancer Therapy General Scale) wurde alle 12 Wochen erhoben. Die Erhebung fand zu den jeweiligen Besuchszeitpunkten vor Beginn aller weiteren Untersuchungen statt. Der FACT-G besteht aus 27 Aussagen bzgl. des physischen, sozialen, emotionalen und funktionellen Wohlbefindens. Höhere FACT-G Werte sprechen für eine bessere Lebensqualität.

Wegen des endpunktübergreifend als „hoch“ bewerteten Verzerrungspotentials wird auch das Verzerrungspotential für den Endpunkt „gesundheitsbezogene Lebensqualität“ als „hoch“ eingeschätzt. Hinzuzufügen ist allerdings, dass beide Fragebögen etabliert und validiert sind.

Tabelle 4-176: Ergebnisse zur Gesundheitsbezogenen Lebensqualität – Weitere Untersuchungen

Studie	N*	Gesamtscore Veränderung verglichen zum Ausgangswert			
		Mittelwert	Standardabweichung	Minimum	Maximum
D4200C00008 (FACIT-D)	20	-12,1	12,89	-48,0	9,0
D4200C00068 (FACT-G)	10	-2,48	8,238	-21,2	8,7

* Nur Patienten mit Messungen in Woche 24 (*Follow-up*) wurden ausgewertet

Die Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität“ sind für beide Studien in Tabelle 4-176 dargestellt.

4.3.2.3.3.5 Unerwünschte Ereignisse (UE) – weitere Untersuchungen

Tabelle 4-177: Operationalisierung der unerwünschten Ereignisse

Studie	Operationalisierung
D4200C00008	Gesamtrate unerwünschter Ereignisse (UE) Gesamtrate schwerwiegender UE Gesamtrate zum Tode führender UE Gesamtrate der UE, die zum Abbruch der Studienmedikation führten Gesamtrate der medikamentenbedingten* UE Gesamtrate der UE mit CTCAE Grad ≥3
D4200C00068	Gesamtrate unerwünschter Ereignisse (UE) Gesamtrate schwerwiegender UE Gesamtrate zum Tode führender UE Gesamtrate der UE, die zum Abbruch der Studienmedikation führten Gesamtrate der medikamentenbedingten* UE Gesamtrate der UE mit CTCAE Grad ≥3

CTCAE = *Common Terminology Criteria for Adverse Events*
* Ein UE wurde als medikamentenbedingt eingestuft, falls der Arzt zumindest eine begründeten Verdacht hatte, dass die Medikmentengabe das Auftreten des UE bedingt hat.

Wegen des endpunktübergreifend als „hoch“ bewerteten Verzerrungspotentials wird auch das Verzerrungspotential für den Endpunkt „unerwünschte Ereignisse“ als „hoch“ eingeschätzt.

Tabelle 4-178: Unerwünschte Ereignisse – Weitere Untersuchungen

Studie	N	Gesamtrate					
		UE	Schwerwieg-endes UE	zum Tode führender UE	UE, die zum Abbruch der Studienmedi-kation führten	Medikamenten-bedingter* UE	UE mit CTCAE Grad ≥ 3 #
D4200C00008	30	30 (100,0)	11 (36,7)	2 (6,7)	7 (23,3)	30 (100,0)	24 (80,0)
D4200C00068	19	18 (94,7)	4 (21,1)	1 (5,3)	3 (15,8)	17 (89,5)	6 (31,6)

In Studie D4200C00008 wurden nur UE mit CTCAE Grad 3-4 tabelliert, in Studie D4200C00068 wurde CTCAE Grad 3-5 tabelliert.
* Ein UE wurde als medikamentenbedingt eingestuft, falls der Arzt zumindest einen begründeten Verdacht hatte, dass die Medikamentengabe das Auftreten des UE bedingt hat.

Die Ergebnisse bezüglich des Endpunktes „unerwünschte Ereignisse“ sind für beide Studien in Tabelle 4-178 dargestellt.

Die häufigsten UEs in den Studien D4200C00008 und D4200C00068 waren Diarrhö, Hautausschlag, Müdigkeit und Übelkeit (nur in Studie D4200C00008).

4.3.2.3.3.6 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.2.

In den unter weitere Untersuchungen aufgeführten Studien waren aufgrund der geringen Patientenzahlen keine Subgruppenanalysen geplant oder beschrieben.

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.

Die unter weiteren Unterlagen beschriebenen Studien, D4200C00008 und D4200C00068, sind in der Studienregistersuche identifiziert worden und sind Teil der Zulassungsdokumentation, die bei den Zulassungsbehörden eingereicht worden ist. Beide Studien sind nichtkontrollierte Phase-II-Studien, haben ein offenes, einarmiges, multizentrisches Design und schlossen Patienten mit lokal fortgeschrittenem oder metastasierendem hereditärem MTC (30 behandelte Patienten in D4200C00008 und 19 behandelte Patienten in D4200C00068) ein.

Die Studien unterschieden sich in der Dosierung: Während die Pilot-Studie D4200C00008 die Dosis von täglich 300 mg Vandetanib untersuchte, prüfte die spätere Studie (D4200C00068), die parallel zur Phase-III-Studie gestartet wurde, die niedrigere Dosis von täglich 100 mg.

In beiden Studien wurde die objektive Ansprechrate nach RECIST als primärer Endpunkt betrachtet. In Studie D4200C00008 lag die Ansprechrate bei 20%. In Studie D4200C00068, in der die Patienten die niedrigere Dosis Vandetanib erhielten, wurde eine Ansprechrate von 15,8% erreicht. In beiden Fällen lag die objektive Ansprechrate unter der in der Phase-III-Studie beobachteten Ansprechrate von 45%.

Die Dauer des Ansprechens war 10,2 Monate (D4200C00008) bzw. 5,5 Monate (D4200C00068). Das mediane PFS in Studie D4200C00008 war mit 28 Monaten vergleichbar mit dem in der Phase-III-Studie geschätzten medianen PFS (30 Monate). Das mediane progressionsfreie Überleben unter der niedrigeren Dosierung (100 mg) in Studie D4200C00068 betrug 16 Monate.

In Studie D4200C00008 führten zwei schwerwiegende unerwünschte Ereignisse zum Tod. Die anderen acht Patienten mit schwerwiegenden unerwünschten Ereignissen konnten noch während des Beobachtungszeitraums wiederhergestellt werden. In Studie D4200C00068

führte ein schwerwiegendes unerwünschtes Ereignis zum Tod. Drei weitere schwerwiegende unerwünschte Ereignisse traten während des Beobachtunsgzeitraums auf.

Zusammenfassend ist festzustellen, dass in beiden Studien ein Ansprechen auf die Vandetanib Therapie beobachtet wurde, wobei die bessere Wirksamkeit unter der höheren, der Zulassung entsprechendenen Dosis (300 mg) beobachtet werden konnte (Studie D4200C00008).

Das Verzerrungspotential in beiden Studien ist, bedingt durch das einarmige Design, generell als „hoch“ einzuschätzen.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Es gibt eine generelle Forderung nach Replizierbarkeit wissenschaftlicher Resultate und nach der Konsistenz der Ergebnisse.

Im Rahmen einer Nutzenbewertung gilt in der Regel für Aussagen eines Belegs die Mindestanforderung, dass zwei voneinander unabhängig durchgeführte Studien mit endpunktbezogen hoher Ergebnissicherheit und entsprechendem statistisch signifikantem Effekt vorliegen, deren Ergebnis nicht durch weitere vergleichbare, endpunktbezogene ergebnissichere Studien infrage gestellt wird [89]. In Ausnahmefällen kann jedoch auch auf der Basis einer einzelnen Studie der Beleg eines Nutzens oder Zusatznutzens abgeleitet werden. Die Erfordernisse hierfür beruhen auf den *CPMP Points to consider on application with: 1. meta-analyses; 2. one pivotal study* [99] und können im Fall von Vandetanib als erfüllt betrachtet werden. Im Folgenden wird dargelegt, warum diese Kriterien für die Studie Studie 58 mit Vandetanib erfüllt sind.

Die in die Nutzenbewertung eingeschlossene Studie ist eine RCT der Phase III und entspricht der Evidenzstufe Ib. Die Studie weist eine endpunktübergreifend hohe Ergebnissicherheit auf.

Sowohl für den primären Endpunkt „Progressionsfreies Überleben“, als auch für die sekundären Endpunkte „Zeit bis zur Schmerzprogression“, „objektive Ansprechrate“, „Krankheitskontrollrate“ und „unerwünschte Ereignisse“ sowie die explorativen Endpunkte „Gewichtszunahme“, „WHO-Performancestatus“, „Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn“ und „Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert“ konnten weder Hinweise auf eine ergebnisgesteuerte Berichtserstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden. Insbesondere wurde für den primären Endpunkt „Progressionsfreies Überleben“ sowie die Endpunkte „objektive Ansprechrate“ und „Krankheitskontrollrate“ ein

zentraler Review der CT/MRI *scans* durchgeführt, um einen Zentrumsbias zu vermeiden. Das endpunktbezogene Verzerrungspotential wurde aufgrund dessen als „niedrig“ bewertet.

Das endpunktspezifische Verzerrungspotential ist für die Endpunkte „Gesamtüberleben“, „gesundheitsbezogene Lebensqualität“, „Dauer des objektiven Ansprechens“ und „Einfluss der Behandlung auf Diarrhö“ als „hoch“ zu bewerten.

In Anbetracht der Seltenheit der Erkrankung und der damit verbundenen Schwierigkeit, die dennoch gültigen allgemeinen Bestimmungen hinsichtlich Einhaltung ethischer Vorgaben, Effizienz- und Sicherheitskriterien zu erfüllen – ein Problem, das sich bei vielen klinischen Untersuchungen zu seltenen Erkrankungen ergibt [100] – wurde ein vergleichsweise umfassendes Patientenkollektiv eingeschlossen. Studie 58 war folglich auch die erste Phase-III-Studie weltweit, die zur Behandlung des MTC durchgeführt worden ist. Die externe Validität ist aufgrund der gegenüber der in der Fachinformation spezifizierten Zielpopulation weniger restriktiven Studienpopulation gegeben. Die beobachteten Unterschiede zwischen Vandetanib und der Vergleichstherapie sind statistisch signifikant, die entsprechenden Effektschätzer sind präzise und klinisch relevant. Die beobachteten Effekte gelten innerhalb der Gesamtpopulation über prä-spezifizierte Subgruppen hinweg. Damit sind die wesentlichen Kriterien hinsichtlich interner und externer Validität, statistischer und klinischer Relevanz der Ergebnisse, sowie Konsistenz des Effekts über Subgruppen hinweg erfüllt. Auch die Vergleichbarkeit der Resultate der Gesamtpopulation mit denen der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung als Näherung für die Zielpopulation laut Indikation, „aggressiv und symptomatisch“, ließ sich nachvollziehbar mit Daten und anhand von Interaktionstests demonstrieren.

Die RCT mit Vandetanib beinhaltet einen direkten Vergleich mit Placebo. Die Kriterien hinsichtlich Population und Behandlung sind durch die mit der Zulassung festgelegten Zielpopulation des zu bewertenden Arzneimittels Vandetanib (siehe Fachinformation [42]) definiert, wobei der Beschränkung auf Patienten mit aggressiver und symptomatischer Erkrankung durch die vergleichende Betrachtung sowohl der Gesamtpopulation der Studie, als auch der abgrenzbaren Zielpopulation mit symptomatischer und progressiver Erkrankung als Näherung der Zulassungspopulation Genüge getan wird. Die Wahl der Kontrollgruppe erfolgte gemäß Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA am 29.09.2011 [1] und spiegelt zugleich die zum Zeitpunkt der Studieninitiierung aktuelle medizinische Praxis wider, da neben der symptomatischen und palliativen Symptomkontrolle (*Best Supportive Care*) bis zur Zulassung von Vandetanib keine geeigneten Therapien im Indikationsgebiet zur Verfügung standen. Erst mit der Zulassung von Vandetanib steht den Patienten eine Therapiealternative zu BSC zur Verfügung. Die Placebo-Gruppe der Zulassungsstudie korrespondiert aufgrund der individuell erforderlichen symptomatischen Behandlung mit „*Best Supportive Care*“ in beiden Behandlungsarmen mit der zweckmäßigen Vergleichstherapie. Basierend auf einer adäquaten Randomisierung kann von einer Strukturgleichheit der zu vergleichenden Patientenkollektive ausgegangen werden. Die Verwendung von Placebo als Kontrollgruppe erlaubt ein verblindetes Studiendesign und einen identischen Behandlungsprozess der beiden Gruppen. Sie ist somit die Basis für die

Vermeidung systematischer Verzerrungen und insofern die Voraussetzung für valide Resultate. Das Verzerrungspotenzial auf Studienebene ist demzufolge als „niedrig“ zu bewerten (Abschnitt 4.3.1.2.2); des Weiteren finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Faktoren. In diesem Dossier wurde der medizinische Nutzen von Vandetanib gegenüber Placebo in beiden betrachteten Populationen basierend auf dem direkten Vergleich dargelegt und damit die positive Risiko-Nutzen-Bewertung der Zulassungsbehörden untermauert.

Die in der Studie untersuchten Endpunkte sind etabliert und patientenrelevant.

Das progressionsfreie Überleben (PFS) ist, wie in Abschnitt 4.2.5.2 diskutiert, vor dem Hintergrund einer Erkrankung an fortgeschrittenem MTC ein patientenrelevanter Endpunkt, der sich direkt auf das Hinauszögern einer zunehmenden progressionsbedingten Morbidität überträgt, den Patienten über die Wirksamkeit der Therapie informiert und ein Frühindikator für den patientenrelevanten Endpunkt „Schmerz“ ist. Als primärer Endpunkt in der RCT präspezifiziert, wurde PFS durch ein unabhängiges, verblindetes, zentrales Reviewverfahren nach RECIST, modifiziert für MTC (standardisiertes Kriterium) erhoben [22, 23]. Die objektive Ansprechrate (ORR) misst die Verringerung der Tumorlast und umfasst sowohl komplettes (CR) als auch partielles Ansprechen (PR). Die Krankheitskontrollrate (DCR) ist ein Maß für die Krankheitsstabilisierung und berücksichtigt über die Patienten mit partieller (PR) oder kompletter (CR) Remission hinaus auch diejenigen Patienten, die durch die Therapie über einen längeren Zeitraum eine Stabilisierung ihrer Erkrankung erleben dürfen. Basierend auf einer kompletten (CR) oder partiellen (PR) Remission, misst die Dauer des objektiven Ansprechens (DOR) den Zeitraum, für den eine erreichte Tumorantwort bestehen bleibt. Die Kriterien für CR oder PR entsprechen den RECIST Kriterien und basieren auf standardisierten reproduzierbaren bildgebenden Verfahren. Die Validität der Endpunkte „progressionsfreies Überleben“, „objektive Ansprechrate“, „Krankheitskontrollrate“ und „Dauer des objektiven Ansprechens“ ist somit gegeben.

Als Instrument zur Erhebung einer Veränderung der Zeit bis zur Schmerzprogression (TWP) unter Studienmedikation im Vergleich zur Kontrollgruppe wurde in der Studie der BPI-SF (*Brief Pain Inventory (Short Form)*)-Fragebogen eingesetzt. Dabei handelt es sich um einen vom Patienten auszufüllenden Fragebogen. Der BPI-SF wurde 1983 von Daut und Mitarbeitern entwickelt und dient zur Messung des Schmerzes und der Schmerzfolgen. Sensorische Schmerzkomponenten und schmerzbedingte Beeinträchtigung bilden den Schwerpunkt des Fragebogens. Am häufigsten im angloamerikanischen Raum eingesetzt, wurde der BPI-SF von Radbruch und Mitarbeitern 1999 in die deutsche Sprache übersetzt und validiert [61]. Er wird von der *European Association for Palliative Care* als Messinstrument für klinische Studien der Tumorschmerztherapie bei Tumorpatienten empfohlen, kann aber auch für andere schmerzbehaftete Erkrankungen verwendet werden [62].

Der 1982 durch Onken et al. veröffentlichte *WHO Performance-Status* [80] ist neben dem 1948 von David Karnofsky publizierten [78] und 1984 durch Mor [79] validierten Index ein etabliertes, weltweit eingesetztes, valides Instrument zur Bewertung des Allgemeinzustand

und der Leistungsfähigkeit eines Patienten im täglichen Leben. Eine Studie von Buccheri et al. (1996) bestätigt die Gleichwertigkeit der Instrumente [81].

Die aufgezählten explorativen Morbiditätsendpunkte sind vor dem Hintergrund der aktuellen Empfehlungen der EMA als valide Messinstrumente akzeptabel [18]. Das biochemische Ansprechen stellt ebenfalls einen Endpunkt dar, der als Surrogat für die krankheitsbedingten Symptome die Ergebnisse zu PFS und ORR unterstreicht und ergänzt, da die Bewertung des CTN/CEA Status auf harten und objektiv messbaren Kriterien (Laborwerten; in der ausgewerteten Studie durch ein zentrales Labor erhoben) beruht.

Der *Functional Assessment of Cancer Therapy General Scale* (FACT-G) Fragebogen ist ein etabliertes, weltweit eingesetztes patientenbasiertes, valides Instrument zur Bewertung der gesundheitsbezogenen Lebensqualität und ermöglicht Rückschlüsse auf den Krankheitsverlauf [94, 95].

Die Meldung und Berichterstattung unerwünschter Ereignisse erfolgte in der RCT mit Vandetanib gemäß den Standards der *Good Clinical Practice* (GCP).

„Gesamtüberleben“ (OS) zählt aufgrund seiner eindeutigen Erhebung und klaren Definition (nur Tod zählt als Ereignis) grundsätzlich zu den harten und damit validen Endpunkten. Für die Bewertung von Vandetanib ist dieser Endpunkt jedoch von nachrangiger Wertigkeit, da die Ergebnisse der Gesamtüberlebensrate durch die geringe Zahl an beobachteten Todesfällen bis zum *Cut-off* sowie das *Cross-Over* auf *Open-Label* Vandetanib mit hoher Wahrscheinlichkeit verzerrt sein dürften. Insbesondere den Endpunkten PFS und TWP ist in dieser Indikation die massgebliche Bedeutung für die Bewertung des Zusatznutzens zuzuschreiben. Daher wurde, wie bereits an anderer Stelle im Dossier diskutiert, auf eine weitere Darstellung des Endpunktes Gesamtüberleben verzichtet (siehe Abschnitt 4.2.5.2).

Im Hinblick auf Evidenzstufe, Studienqualität und Validität der Endpunkte besitzen die erbrachten Nachweise zum Nutzen von Vandetanib anhand der doppel-blinden, randomisierten, placebo-kontrollierten Studie hohe Validität und Aussagekraft.

Damit ist der der Nutzenbewertung zugrundeliegende Nachweis geeignet, um im Falle statistisch signifikanter Ergebnisse **Belege** für einen Zusatznutzen abzuleiten.

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. 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.

Erheblicher Zusatznutzen von Vandetanib gegenüber *Best Supportive Care* bezüglich PFS, TWP, ORR und DCR, unterstützt durch die Resultate für DOR, biochemisches Ansprechen und weitere explorative Endpunkte:

Im Rahmen der randomisierten, kontrollierten Studie 58 konnte in beiden betrachteten Populationen, sowohl in der Gesamtpopulation der Studie als auch in der *post-hoc* untersuchten genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung, für die patientenrelevanten morbiditätsbezogenen Endpunkte „Progressionsfreies Überleben (*Progression Free Survival*, PFS)“, „Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)“, „Objektive Ansprechrate (*Objective Response Rate*, ORR)“ und „Krankheitskontrollrate (*Disease Control Rate*, DCR)“ ein statistisch signifikanter Vorteil von Vandetanib gegenüber der mit Placebo behandelten Kontrollgruppe, die ansonsten als aktive Therapie nur Maßnahmen der *Best Supportive Care* (BSC) erhielt, gezeigt werden.

Untermauert wurde der statistisch signifikante Vorteil von Vandetanib gegenüber Placebo durch die Ergebnisse des Endpunkts „Dauer des objektiven Ansprechens (*Duration of Objective Response*, DOR)“, die bezüglich Progression und Morbidität aussagekräftigen Surrogatparameter „biochemisches Ansprechen (CTN; CEA)“ sowie durch die Ergebnisse weiterer explorativer Morbiditätsendpunkte.

Im Folgenden wird das Ausmaß des Zusatznutzens auf Endpunktebene dargestellt:

Verbesserung des Progressionsfreien Überlebens (PFS)

In der palliativen onkologischen Therapie stellt, wie oben bereits detailliert begründet, ein Hinauszögern des Fortschreitens der Erkrankung und der damit verbundenen Symptomatik eines der wichtigsten Therapieziele dar [14]; dies gilt insbesondere auch bei einem zu erwartenden langen Gesamtüberleben [15-17]. Deshalb zählt das progressionsfreie Überleben (*Progression-Free Survival*, PFS) in der palliativen Situation als Funktion der Wachstumsrate des Tumors und der Wirksamkeit des untersuchten Präparates zu den wichtigsten patientenrelevanten Endpunkten [18], um die Effektivität einer eingesetzten Behandlung während des beobachteten Therapiezeitraums zu bewerten. Die Progression der Erkrankung geht mit einer zunehmenden Morbidität des Patienten einher.

Für das PFS wurde in der Vandetanib-Gruppe bezogen auf die Gesamtpopulation ein statistisch signifikanter Vorteil gegenüber Placebo beobachtet (HR: 0,46 95%-KI: [0,31;0,69]). Die entsprechende *post-hoc* Auswertung der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung ergab ebenfalls einen statistisch signifikanten Vorteil von Vandetanib gegenüber Placebo (HR: 0,47, 95%-KI: [0,29;0,77]). Das bedeutet für die Patienten eine Verzögerung ihrer Progression und damit fortschreitender Krankheitslast um 11 Monate in der Gesamt- bzw. um 12 Monate in der genäherten Zielpopulation. Grundsätzlich sind für diesen Endpunkt die Ergebnisse der Gesamtpopulation auf die genäherte Zielpopulation übertragbar, wie der Interaktionstest (Tab. 4-29) klar zeigt (p=0,9617).

Die Überlegenheit von Vandetanib bestätigt sich auch im zeitlichen Verlauf. Die Wahrscheinlichkeit für ein PFS (Kaplan-Meier Schätzer) war in der Vandetanib-Gruppe in der Gesamtpopulation zu allen Zeitpunkten um ca. 20% höher als in der Vergleichsgruppe: nach Monat 6 mit 91,4% versus 74,5%, Monat 12 mit 83,3% versus 62,7% und nach Monat 24 mit 58,8% versus 41,7%. In der genäherten Zielpopulation zeigte sich dies genauso: nach Monat 6 mit 89,4% versus 72,9%, Monat 12 mit 80,8% versus 57,1% und nach Monat 24 mit 51,7% versus 36,0%. Auch von der DGHO wird die durch Vandetanib erzielte Verlängerung des PFS als klinisch relevanter Zeitraum angesehen [101].

Angesichts des in Abschnitt 4.3.1.3.1.1 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie lässt sich die Evidenzsicherheit des Zusatznutzens von Vandetanib als Beleg ableiten.

Aus der erheblichen Verlängerung des progessionsfreien Überlebens und der besonderen Relevanz dieses Endpunktes im Hinblick auf die Morbidität und Lebensqualität der Patienten lässt sich für diesen patientenrelevanten Endpunkt gemäß der Klassifizierung des Zusatznutzens nach §5 Punkt 7 der AM-NutzenV [20] der Beleg für einen **erheblichen** Zusatznutzen ableiten.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.1.

Zeit bis zur Schmerzprogression (TWP)

Schmerz stellt eine bedeutsame Komponente der Symptomatik des aggressiven, nicht resektablen, lokal fortgeschrittenen oder metastasierten MTC dar. Dies bestätigt aus klinischer Sicht die DGHO, nach deren Aussage ein Großteil der Patienten extreme Schmerzen hat [21]. Für die Patienten ist eine möglichst langfristige Verzögerung der Schmerzprogression von großer Bedeutung, da Schmerzen nicht nur fortschreitende Morbidität signalisieren sondern auch die Lebensqualität stark beeinträchtigen und auch der Einsatz von Schmerzmitteln wieder von entsprechenden Nebenwirkungen begleitet ist. Die Vermeidung bzw. Hinauszögerung von Tumorschmerzen muss daher in der Palliativsituation als herausragendes Therapieziel bei der Behandlung dieser Patienten gelten.

Die Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP), abgeleitet aus der Bewertung für den stärksten Schmerz basierend auf validierten Schmerzfragebögen in Verbindung mit dem von den Patienten angegebenen Gebrauch von opioidhaltigen Schmerzmitteln, war unter Vandetanib im Median mehr als doppelt so lang wie in der Vergleichsgruppe, sowohl für die Gesamtpopulation (7,85 vs. 3,25 Monate), als auch in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung (11,1 vs. 3,4 Monate). Der Unterschied zwischen den Behandlungsgruppen war sowohl innerhalb der Gesamtpopulation (HR: 0,61, 95%-KI: [0,43;0,87]) als auch innerhalb der genäherten Zielpopulation (HR: 0,62 [0,39;0,99]) statistisch signifikant. Die Verlängerung der Zeit bis zur Schmerzprogression wurde auch aus klinischer Sicht von der DGHO als „eindrucksvoll“ und als „echter Benefit“ für die Patienten bewertet [21].

Angesichts des im Abschnitt 4.3.1.3.1.2 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie läßt sich das Ausmaß des Zusatznutzens von Vandetanib auch hier als Beleg ableiten.

Ebenso wie für den Endpunkt PFS zeigt sich für TWP eine nachhaltige und bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens [20]. Schmerz stellt ein besonders schwerwiegendes Symptom in der Palliativsitation dar. Aus der langfristigen Freiheit von schwerwiegenden Symptomen – hier von progredienten Schmerzen – und der besonderen Relevanz dieses Endpunktes ergibt sich gemäß der Klassifizierung des Zusatznutzens nach §5 Punkt 7 der AM-NutzenV [20] der Beleg für einen **erheblichen** Zusatznutzen.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.2.

Objektive Ansprechrate (ORR)

Bei der Behandlung onkologischer Erkrankungen in der metastasierten Situation stellt die Objektive Ansprechrate (*Objective Response Rate*, ORR) als Maß für die Kontrolle des Tumorwachstums und des Ansprechens der Therapie sowohl ein valides Messinstrument als auch ein bedeutsames Therapieziel dar. Die ORR misst dabei die Verringerung der Tumorlast

und umfasst sowohl komplettes als auch partielles Ansprechen, allerdings nicht den Stillstand der Erkrankung [24-26] und ist damit ein Parameter für reduzierte Morbidität.

Das Objektive Ansprechen (*Objective Response Rate*, ORR), in Form einer partiellen Remission gemäß RECIST Kriterien, wurde von 45% der Patienten in der Vandetanib-Gruppe der Gesamtpopulation und von 44,4% der Vandetanib-Gruppe innerhalb der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung erreicht. Im Vergleich erreichten nur 13% bzw. 16,7% der Patienten unter BSC/Placebo eine partielle Remission. Keiner der Patienten in den beiden Behandlungsgruppen und Populationen erreichte eine komplette Remission. Die Ergebnisse für die partielle Remission waren statistisch signifikant zugunsten von Vandetanib: OR: 5,48, 95%-KI: [2,99;10,79] (Gesamtpopulation) bzw. 4,00 [1,93;9,01] (genäherte Zielpopulation),

Angesichts des in Abschnitt 4.3.1.3.1.3 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie läßt sich das Ausmaß des Zusatznutzens von Vandetanib als Beleg ableiten.

Da das Wissen um ein Tumoransprechen und die Metastasenreduktion eine psychische Entlastung und damit einen direkten therapierelevanten Nutzen für den Patienten darstellen, ergibt sich mit den hier dargestellten signifikanten Ergebnissen für ORR gemäß der Klassifizierung des Zusatznutzens nach §5 Punkt 7 der AM-NutzenV [20] der Beleg für einen **erheblichen** Zusatznutzen.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.3.

Krankheitskontrollrate (DCR)

Die Krankheitskontrollrate (*Disease Control Rate*, DCR) ist ein Maß für die Krankheitsstabilisierung und berücksichtigt über die Patienten mit partieller oder kompletter Remission hinaus auch diejenigen Patienten, die durch die Therapie über einen längeren Zeitraum eine Stabilisierung ihrer Erkrankung erleben dürfen. Damit stellt die Krankheitskontrollrate ein weiteres Maß für den klinischen Nutzen einer Therapie dar, indem sie all jene Patienten einbezieht, bei denen über mindestens ein halbes Jahr kein Progress vorliegt und die damit für einen gewissen Zeitraum von einer Krankheitskontrolle durch die Therapie profitieren.

Die Krankheitskontrollrate beinhaltet den Anteil der Patienten, deren bestes objektives Ansprechen (B*est Objective Response*, BOR) einem kompletten Ansprechen (*Complete Response*, CR), einem partiellen Ansprechen (*Partial Response*, PR) oder einer Krankheitsstabilität für mindestens 24 Wochen gemäß RECIST Kriterien entspricht. Die Anzahl der Patienten mit Krankheitskontrolle war in beiden evaluierten Populationen (Gesamt- bzw. genäherte Zielpopulation) in der Vandetanib Gruppe (86,6% bzw. 84,1%) im Vergleich zu Placebo (71,0% bzw. 68,3%) höher. Bezüglich der Krankheitskontrollrate wurde in der Vandetanib-Gruppe sowohl der Gesamtpopulation als auch der genäherten

Zielpopulation mit symptomatischer und progressiver Erkrankung ein statistisch signifikant besseres Ergebnis als mit Placebo beobachtet (OR: 2,64 95%-KI: [1,48;4,69] bzw. OR: 2,46 (95%-KI: [1,19;5,09]). Unter Ausschluss der offenen Behandlungsphase – nicht bei allen Patienten wurde der lokal diagnostizierte Progress nach RECIST in der zentralen Auswertung bestätigt, wodurch für 12 von 13 Patienten (Gesamtpopulation) bzw. 9 von 10 Patienten (genäherte Zielpopulation) der Placebo-Gruppe noch ein beginnendes Ansprechen in der offenen Behandlungsphase erreicht wurde – betrug der Anteil der Patienten mit Krankheitskontrolle in der Vandetanib-Gruppe 84,8% bzw 82,5% gegenüber 61% bzw. 56,7% in der Placebo-Gruppe. Es konnte eine statistisch signifikante Verbesserung der Krankheitskontrollrate für die Vandetanib-Gruppe gezeigt werden (OR: 3,58; 95%-KI: [2,09;6,17] bzw. OR: 3,61; 95%-KI: [1,83;7,26]).

Angesichts des in Abschnitt 4.3.1.3.1.4 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie läßt sich das Ausmaß des Zusatznutzens von Vandetanib als Beleg ableiten.

Die Ergebnisse für die DCR belegen, dass ein statistisch signifikanter Anteil an Patienten über einen Zeitraum von mindestens 24 Wochen von der Therapie durch Vandetanib profitieren, was im Kontext dieser Pallitativsituation als langfristig anzusehen ist. Dies bedeutet für die betroffenen Patienten eine erhebliche psychische Entlastung und Arzt und Patient werden nicht vor die Situation gestellt, den nächstfolgenden Behandlungsschritt entscheiden zu müssen. Es ergibt sich daher gemäß der Klassifizierung des Zusatznutzens nach §5 Punkt 7 der AM-NutzenV [20] der Beleg für einen **erheblichen** Zusatznutzen.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.4.

Dauer des objektiven Ansprechens (DOR)

Die Dauer des objektiven Ansprechens (*Duration of Response*, DOR) misst den Zeitraum, für den eine erreichte Tumorantwort (komplette oder partielle Remission) bestehen bleibt. Die Patientenrelevanz dieses Endpunktes ergibt sich unmittelbar aus der Patientenrelevanz des Ansprechens selbst (siehe oben, ORR): für die Betroffen spielt es selbstverständlich eine Rolle, dass ein einmal erreichtes Ansprechen ihrer Erkrankung auch möglichst lange aufrecht erhalten werden kann.

Die Dauer des objektiven Ansprechens (DOR) wurde basierend auf dem zentralen, unabhängigen Review der Bildgebung nach den RECIST Kriterien bewertet. Die Abbildungen der Kaplan-Meier-Kurven beider Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigten für die DOR für die Zeit vom Ansprechen bis zur ersten Progression in beiden Behandlungsarmen ein lang anhaltendes Ansprechen, wobei ein Ansprechen für 12 bzw. 9 von 13 bzw. 10 Patienten der Placebo-Gruppe erst in der offenen Behandlungsphase unter Vandetanib einsetzte.

Die geschätzen medianen Werte für die Dauer des objektiven Ansprechens lagen unter Vandetanib in der Gesamtpopulation bei 22,2 Monaten und unter Placebo bei 16,3 Monaten.

Für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung ließ sich der gleiche Trend beobachten. So lagen die geschätzen medianen Werte für die Dauer des objektiven Ansprechens in der Vandetanib-Gruppe bei 21,8 Monaten und bei 13,0 Monaten in der Placebo-Gruppe.

Die Dauer des objektiven Ansprechens ist für Patienten, die auf Vandetanib ansprachen, im Trend damit länger als bei Patienten unter Placebo/BSC.

Die Ergebnisse für die Dauer des objektiven Ansprechens (DOR) untermauern die Belege für einen erheblichen Zusatznutzen der bereits dargestellten Endpunkte.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.5.

Weitere explorative patientenrelevante Morbiditätsendpunkte:

Die Ergebnisse für die nachfolgend aufgeführten explorativen Morbiditätsendpunkte untermauern die Ergebnisse der bereits dargestellten Morbiditätsendpunkte, welche eine Überlegenheit von Vandetanib zeigen und den sich daraus resultierenden erheblichen Zusatznutzen belegen:

Gewichtszunahme

Für die klinische Gewichtszunahme zeigte sich in beiden Populationen (Gesamt- bzw. genäherte Zielpopulation) ein deutliches Ergebnis zugunsten von Vandetanib (OR: 3,24; 95%-KI: [1,69;6,76] bzw. OR: 4,03; 95%-KI: [1,71;11,17]). Der Anteil der Patienten mit klinischer Gewichtszunahme betrug in der Vandetanib-Gruppe 28,6% bzw. 31,0% verglichen zur Placebo-Gruppe mit lediglich 11,0% bzw. 10,0%. Zum arztberichteten unerwünschten Ereignis „Gewichtsverlust“ verweisen wir auf Abschnitt 4.3.1.3.1.14. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.6.

Einfluss der Behandlung auf Diarrhöen

Der Einfluss der Behandlung auf krankheitsbedingte Diarrhöen wurde anhand der Stuhlfrequenz bestimmt. Die Angaben der Patienten bezüglich ihrer Stuhlfrequenzen im Hinblick auf den Einfluss der Behandlung auf Diarrhö zeigten in der Gesamtpopulation und auch der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung zwischen den untersuchten Behandlungsgruppen Vandetanib und Placebo keine auffälligen Unterschiede. Ein Anhaltspunkt für einen Einfluss der Behandlung mit Vandetanib auf Diarrhö ergab sich aus der exploratorischen Analyse in beiden betrachteten Populationen nicht. Zum arztberichteten unerwünschten Ereignis „Diarrhö“ verweisen wir auf Abschnitt 4.3.1.3.1.14. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.7.

WHO Perfomancestatus

Die Daten für die Zeit bis zum ersten Nachweis einerVerschlechterung des WHO-PS wurden relativ zum Zeitpunkt der Randomisierung ausgewertet. Eine Verschlechterung des WHO-PS wurde definiert als Erhöhung um mindestens einen Punkt im Vergleich zum Ausgangswert. Für beide Populationen (Gesamt- bzw. genäherte Zielpopulation) zeigte sich kein deutlicher Unterschied in den Ergebnissen (HR: 0,74; 95%-KI: [0,48;1,14] bzw. HR: 0,71; 95%-KI: [0,42;1,21]). Für die Gesamtpopulation konnte die mediane Zeit bis Verschlechterung des WHO-PS wegen der geringen Anzahl an Ereignissen in der Vandetanib- Gruppe zum Daten *Cut-off* nicht errechnet werden. Für die Placebo-Gruppe betrug die mediane Zeit bis zur Verschlechterung des WHO-PS 24,9 Monate. Die mediane Zeit bis zur Verschlechterung des WHO-PS für die Vandetanib-Gruppe lag in der genäherten Zielpopulation bei 22,6 Monaten, was, verglichen mit 15,4 Monaten in der Placebo-Gruppe, einer Verlängerung um ca. 7 Monate entspricht. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.8.

Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline)

Der Endpunkt „Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn“ stellt in seiner Form eine Subgruppenanalyse zum TWP dar. Sowohl für die Gesamt-, als auch für die genäherte Zielpopulation spiegelt die visuelle Prüfung des *Forest-Plots* die Konsistenz des TWP über die zwei Behandlungsgruppen mit bekanntem Opioidanalgetika Gebrauch hinweg wider. Für die Gesamtpopulation liegt in der Gruppe „≥10 mg/Tag Opioidanalgetika-Gebrauch“ das *Hazard Ratio* bei 0,36 (95%-KI: [0,16;0,84], p=0,0176). In der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika-Gebrauch“, liegt das *Hazard Ratio* bei 0,68 mit einem 95%-KI [0,44;1,03] und einem p-Wert von 0,0698. In der genäherten Zielpopulation ließen sich sowohl für die Gruppe „≥10 mg/Tag Opioidanalgetika-Gebrauch“ mit einem *Hazard Ratio* von 0,50 (95%-KI: [0,20;1,30], p=0,1573), als auch in der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika-Gebrauch“ mit einem *Hazard Ratio* von 0,69 (95%-KI [0,39;1,22], p=0,1978) keine deutlichen Unterschiede zeigen. Aufgrund kleiner Populations- und Ereigniszahlen sind diese Ergebnisse allerdings in beiden Populationen mit Vorsicht zu interpretieren. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.9.

Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

Die Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn stellt in seiner Form eine Subgruppenanalyse zum TWP dar. Für die Reduktion des Opioid-

Gebrauchs wurden nur Patienten mit einem Schmerzmittelgebrauch von ≥10 mg/Tag Morphinsulfat-Äquivalenten berücksichtigt. Sowohl in der Gesamt- als auch in der genäherten Zielpopulation zeigte sich ein Trend in der Reduktion des Schmerzmittelgebrauchs zugunsten von Vandetanib (Gesamtpopulation OR: 2,78; 95%-KI: [0,77;10,80] bzw. genäherte Zielpopulation OR: 2,25; 95%-KI: [0,50;11,27]). Eine Reduktion des Schmerzmittelgebrauchs wurde in der Vandetanib-Gruppe von 62,5% bzw. 52,9% der Patienten erreicht verglichen mit 37,5% bzw. 33,3% in der Placebo-Gruppe. Die mediane Dauer der Reduktion des Schmerzmittelgebrauchs betrug für die Vandetanib-Gruppe 11,01 bzw. 11,56 Monate und für die Placebo-Gruppe in beiden Populationen 6,51 Monate. Alle Ergebnisse sind aufgrund der geringen Anzahl an Ereignissen mit Vorsicht zu interpretieren. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.10.

Surrogatendpunkt: Biochemisches Ansprechen (CTN; CEA)

Das biochemische Ansprechen, gemessen an den Serumspiegeln von CTN und CEA, stellt gerade im Falle des MTC gut geeignete Marker für den klinischen Endpunkt Progression zur Verfügung und dient auf Grund seiner Signalwirkung für ein Ansprechen auf die Therapie wie auch für die Progredienz der Erkrankung als Surrogat für patientenrelevante Endpunkte des Bereichs Morbidität (krankheitsbedingte Symptome). Durchfall und Flush-Syndrom treten häufig infolge humoraler Faktoren auf (krankheitsbedingt hohe CTN-Werte). Das biochemische Ansprechen der CTN- und CEA-Werte zeigt sich in beiden betrachteten Populationen. Für CTN ergibt sich eine Spiegelreduzierung für 69,3% / 67,5% (Gesamtpopulation / genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung) der mit Vandetanib behandelten Patienten im Vergleich zu 3,0% / 3,3% bei Behandlung gemäß BSC (OR für CTN-Ansprechen: 72,86; 95%-KI: [26,22;303,2] / OR für CTN-Nicht-Ansprechen: 0,01; 95%-KI: [0,00;0,04] in der Gesamtpopulation bzw. OR für CTN-Ansprechen 60,12; 95%-KI: [17,56;378,0] / OR für CTN-Nicht-Ansprechen: 0,02; 95%-KI: [0,00;0,06] in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung).

Für CEA ergibt sich eine Spiegelreduzierung für 51,5% / 46,0% der Patienten in der Vandetanib-Gruppe im Vergleich zu 2,0% / 0,0% in der Placebo-Gruppe unter einer Behandlung gemäß BSC (OR für CEA-Ansprechen: 52,03; 95%-KI: [15,95;320,3] / OR für CEA-Nicht-Ansprechen: 0,02; 95%-KI: [0,00;0,08] in der Gesamtpopulation. Für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung wurde aufgrund des fehlenden Ansprechens in der Placebo-Gruppe kein OR berechnet.

Zusammen betrachtet unterstreichen die Resultate hinsichtlich des beobachteten Effekts einer Vandetanib-Behandlung auf diese beiden wichtigen biochemischen Parameter die Resultate der meisten Endpunkte zur Morbidität, insbesondere PFS und ORR, auf eindrucksvolle Weise.

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.11.

Gesamtüberleben (OS)

Aufgrund der geringen Ereignisraten bis zum Zeitpunkt des C*ut-off* und des möglichen Einflusses der erlaubten *Open-Label* Behandlung mit Vandetanib post Progression, wodurch eine Verzerrung der Ergebnisse wahrscheinlich ist, wurde die statistische Auswertung des Gesamtüberlebens nicht für die Nutzenbewertung herangezogen. Für die Gesamtpopulation zeigte eine erste Analyse der Gesamtüberlebensrate bei 48 eingetretenen Ereignissen ein *Hazard Ratio* von 0,89 zugunsten von Vandetanib. Der Unterschied war nicht statistisch signifikant (KI: [0,28; 2,85], p=0,713). Für die genäherte Zielpopulation zeigte eine erste Analyse der Gesamtüberlebensrate bei 31 eingetretenen Ereignissen ein *Hazard Ratio* von 1,06. Der Unterschied war auch hier nicht statistisch signifikant (KI: [0,26; 4,38], p=0,8791).

Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.12.

Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G

Daten zur gesundheitsbezogenen Lebensqualität wurden rein exploratorisch ausgewertet. Die Angaben der Patienten bezüglich ihrer Lebensqualität im Hinblick auf ihre Gesundheit in den FACT-G Fragebögen zeigten jedoch in der Gesamtpopulation und auch der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung zwischen den untersuchten Behandlungsgruppen Vandetanib und BSC keine auffälligen Unterschiede. Es ist anzunehmen, dass das objektive Ansprechen auf die Behandlung in Verbindung mit einer Schmerzverminderung einigen der Patienten in der palliativen Situation ein gewisses Maß an Zuversicht und Lebensmut gibt. Ein Anhaltspunkt für eine Verminderung der gesundheitsbezogenen Lebensqualität durch die Therapie mit Vandetanib ergab sich aus der exploratorischen Auswertung in beiden betrachteten Populationen nicht. Ein Zusatznutzen bezüglich des Endpunkts „gesundheitsbezogene Lebensqualität“ ist für Vandetanib nicht belegt. Zur detaillierten Diskussion des Endpunktes siehe auch Abschnitt 4.3.1.3.1.13.

Unerwünschte Ereignisse

Bei nahezu allen Patienten beider Gruppen und beider Populationen (Gesamtpopulation bzw. genäherte Zielpopulation) trat mindestens ein unerwünschtes Ereignis auf. Der Anteil der Patienten mit unerwünschten Ereignissenin der Vandetanib- bzw. Placebo-Gruppe betrug in der Gesamtpopulation 99,6% bzw. 90,9%, in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung 100% bzw. 94,9%. Für von schwerwiegenden unerwünschten Ereignissen, Abbrüchen aufgrund unerwünschter Ereignisse und von unerwünschten Ereignissen vom Schweregrad CTCAE Grad 3 oder höher betroffene Patienten ergaben sich in der Vandetanib-Gruppe signifikant höhere Prozentsätze als für die Placebo-Gruppe (CTCAE Grad 3 oder höher 55,4% vs. 24,2% bzw. 61,1% vs. 23,7%; SUE

30,7% vs. 13,1% bzw. 31,7% vs. 16,9%). Der Anteil der Patienten mit unerwünschten Ereignissen, die zum Absetzen der Studienbehandlung führten, war unter Vandetanib im Vergleich zu Placebo größer (Gesamtpopulation 12,1% vs. 3,0%, OR: 4,41; 95%-KI: [1,31;14,8], genäherte Zielpopulation 11,9% vs. 1,7%, OR: 7,84; 95%-KI: [1,01;60,82]). Die Anzahl der Ereignisse muss allerdings immer in Relation zur Expositionsdauer gesehen werden: Die mediane Therapiedauer auf verblindeter Therapie war in der Vandetanib-Gruppe erheblich länger als in der Placebo-Gruppe (90,1 vs. 39,9 Wochen in der Gesamtpopulation bzw. 88,6 vs. 37,1 Wochen in der genäherten Zielpopulation), wodurch sich für die Vandetanib-Gruppe eine um ca. 50 Wochen längere Expositionsdauer ergibt.

Die durch Vandetanib am häufigsten verursachten Nebenwirkungen Diarrhö, Hautausschlag, Übelkeit, Bluthochdruck und Kopfschmerzen ließen sich durch medizinische Interventionen gut beherrschen. Das potenzielle Risiko einer QTc-Verlängerung sowie von Torsade de Pointes ist sowohl durch Warnhinweise in der Fachinformation als auch durch Maßnahmen im Risk-Management-Plan für Vandetanib angemessen berücksichtigt worden (siehe Modul 3 A, Abschnitt 3.4.3). Auch nach Einschätzung der DHGO stellen die kardialen Nebenwirkungen die bedrohlichsten Nebenwirkungen dar. Es sei allerdings erfahrenen Klinikern sehr wohl möglich, Patienten mit kardialen Vorerkrankungen herauszufiltern. Darüber hinaus seien die kardialen Nebenwirkungen durch ein engmaschiges Monitoring gut beherrschbar [21].

Die *post-hoc* Auswertungen für die genäherte Zielpopulation mit symptomatischer und progressiver Erkrankung bestätigen alle Erkenntnisse aus der Gesamtpopulation. Zusammenfassend ist ein größerer Schaden durch Vandetanib somit belegt; das Ausmaß muss jedoch im Kontext der Gesamtraten unerwünschter Ereignisse in beiden Behandlungsgruppen, der guten Beherrschbarkeit der meisten unerwünschten Ereignisse, der langen Therapiedauer unter Vandetanib sowie unter Berücksichtigung des Schweregrades der Erkrankung betrachtet werden (siehe auch die Zusammenfassung am Ende von Abschnitt 4.3.1.3.1.14).

Gesamtschau: Erheblicher Zusatznutzen

Bei Patienten mit aggressivem und symptomatischem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem MTC handelt es sich um einen kleinen Teil der Betroffenen einer seltenen Erkrankung (*Orphan Indication*). Von den geschätzt über 6.000 (prognostizierte Zahl für 2012: 6.100) Schilddrüsenkarzinom-Neuerkrankungen in Deutschland im Jahr 2012 [102] haben nur ungefähr 3-10% MTC (also ca. 180 bis 610 Patienten [36, 103, 104]), wovon sich grob geschätzt gut 32% in einem fortgeschrittenen Stadium gemäß Indikation von Vandetanib befinden (vgl. Modul 3 A, Abschnitt 3.2.3).

Zum gegenwärtigen Zeitpunkt ist neben Vandetanib keine effektive oder anerkannte Behandlungsoption für fortgeschrittenes MTC verfügbar. Die Patienten leiden körperlich (z. B. Schmerzen, Durchfälle, Kachexie) und psychisch (z. B. Angst, Depression, Verlust an Selbstbestimmung); das mediane Gesamtüberleben beträgt nur noch 2 bis 3 Jahre. Die

überwiegende Mehrheit der betroffenen Patienten verstirbt letztendlich an MTC. Das langfristige Hinauszögern einer Tumorprogression mit Erhalt einer akzeptablen Lebensqualität steht für diese Patienten an erster Stelle, da damit auch die Zunahme krankheitsbedingter Symptome hinausgezögert wird. Da sich mit hoher Wahrscheinlichkeit früher oder später Tumor- und/oder CTN-bedingte Symptome (siehe Modul 3A, Abschnitt 3.2.1) manifestieren und spätestens bei Auftreten von Knochenmetastasen Schmerzen entstehen, deren Intensität selbst durch den Einsatz von Opiaten oft nur noch unzureichend begegnet werden kann, gibt es einen hohen unbefriedigten medizinischen Bedarf an einer Behandlungsoption. Eine neue Therapie bedeutet folglich dann für die Patienten einen erheblichen Zusatznutzen, wenn sie in der Lage ist, das Fortschreiten der Erkrankung und damit ihrer Begleitsymptome relevant hinauszuzögern, ein Tumoransprechen zu erreichen, und darüber hinaus auch bei Einsatz über einen langen Zeitraum hinweg gut verträglich zu sein.

Die Standardtherapie beim MTC ist die Operation in kurativer Intention. Für das fortgeschrittene bzw. metastasierte MTC existiert zu diesem Zeitpunkt in Deutschland neben Vandetanib keine weitere zugelassene bzw. allgemein empfohlene Therapie.

Die aktuellen Behandlungsempfehlungen und Guidelines der *European Society for Medical Oncology* (ESMO; [35]), der European Thyroid Association [36] und der *American Thyroid Association* (ATA; [37]) raten vom Einsatz zytotoxischer Chemotherapie zur Behandlung von Patienten mit fortgeschrittenem MTC aufgrund der unbefriedigenden Resultate ab. Als Erstlinientherapie empfehlen die Fachgesellschaften stattdessen die Behandlung der Patienten mit TKI im Rahmen klinischer Studien und bezeichnen zum Teil ebenso wie die aktuelle Leitlinie (Version 1.2013) des *National Comprehensive Cancer Network* (NCCN) zum Schilddrüsenkarzinom [38] explizit Vandetanib als in Betracht zu ziehende zugelassene Behandlungs-Option.

Die Therapie mit Vandetanib erfolgt durch die Einnahme einer Tablette täglich und ermöglicht dem Patienten somit ein weitgehend normales Leben, in dem wesentlich geringere Einschränkungen durch regelmäßige vielstündige Arzt- oder Klinikbesuche bestehen, als dies z. B. bei einer Chemotherapie der Fall ist.

Vandetanib stellt eine medizinische Innovation dar. Es ist die erste wirksame medikamentöse Behandlungsmöglichkeit für Patienten mit aggressivem und symptomatischem, nicht resektablem, fortgeschrittenem oder metastasiertem MTC. In der palliativen Situation weist die Therapie mit Vandetanib einen bedeutsamen Zusatznutzen für die betroffenen Patienten auf. Ihre Morbidität wird durch das um mehr als 11 Monate (in der genäherten Zielpopulation mit symptomatischer und progressiver Erkrankung um etwa ein Jahr) verlängerte progressionsfreie Überleben, eine Verringerung der Tumorlast und eine Verbesserung der Schmerzsituation im Vergleich zu BSC relevant reduziert. Dem existenziellen Bedürfnis der betroffenen Patienten nach Verminderung und Verzögerung von Leid und Verlust an Selbstbestimmung kann somit künftig erheblich besser Rechnung getragen werden als bisher. Auch nach Einschätzung der DGHO ist Vandetanib aus klinischer Sicht ein „absoluter Durchbruch“, da bis vor kurzem keine Therapieoption zur Verfügung stand und mit

Vandetanib jetzt eine hoch effektive Behandlungsmöglichkeit existiert. Es gebe Patienten, die seit mehr als drei Jahren stabil seien, denen es extrem gut gehe und die ohne Therapie mittlerweile wahrscheinlich verstorben seien – das sei ein „unglaublicher Benefit“ [21].

Dem gegenüber steht ein Beleg für einen größeren, nicht quantifizierbaren Schaden in Bezug auf SUE und Abbrüche aufgrund unerwünschter Ereignisse. Die im Rahmen der RCT unter Vandetanib beobachteten unerwünschten Ereignisse erwiesen sich in der Regel jedoch als transient und durch eine geeignete medikamentöse Behandlung gut kontrollierbar. Das Risiko von QTc-Verlängerungen und Torsade de Pointes wurde durch entsprechende Warnhinweise in der Fachinformation und Maßnahmen im Risk-Management-Plan adressiert.

Die in der G-BA VerfO vorgegebene Einstufung zur Klassifizierung des Zusatznutzen-Ausmaßes lässt sich nach Auffassung von AstraZeneca nur bedingt auf neue Arzneimittel übertragen, welche zur gezielten Behandlung einer nicht mehr heilbaren, fortschreitenden Erkrankung zugelassen werden. Hier besteht möglicherweise die Gefahr einer systematischen Benachteiligung solcher Arzneimittel:

- Erstens ist bei Patienten im Palliativstadium einer Krebserkrankung keine Heilung mehr zu erwarten; stattdessen besteht das Ziel darin, die Lebensqualität und die Selbstbestimmung dieser schwerkranken Menschen zu erhalten, zu fördern und zu verbessern. Im besten Fall wird sich – wie im Falle der Behandlung mit Vandetanib – der fortschreitende Verlauf der Erkrankung für eine begrenzte zusätzliche Zeitspanne aufhalten lassen.

- Zweitens dürfte für ein solches Arzneimittel auch kaum jemals eine nachhaltige große Verbesserung nachweisbar sein, wenn der G-BA – wie aus der zusammenfassenden Dokumentation zum Ticagrelor-Beschluss vom 15.12.2011 hervorzugehen scheint – mit dem Kriterium der Nachhaltigkeit grundsätzlich den Anspruch auf langfristige symptomfreie Zeiträume von mehreren Jahren verbindet. Die Patienten aus der Zielpopulation von Vandetanib haben nach derzeitigem Kenntnisstand eine mediane Restlebenserwartung von insgesamt nur noch 2 bis 3 Jahren; vor diesem Hintergrund müssen eine PFS-Verlängerung um ca. ein Jahr und eine Verlängerung der Zeitdauer bis zur Schmerzprogression um ca. 5 Monate, in der genäherten Zielpopulation um fast 8 Monate, aus Sicht der Patienten und der sie behandelnden Ärzte als bereits sehr nachhaltig erscheinen.

Unter Berücksichtigung der obigen Ausführungen und angesichts der Schwere der Erkrankung und des erheblichen *medical need* in der Zielpopulation, für die derzeit keine Behandlungsalternative existiert, ist der Zusatznutzen von Vandetanib auch nach Abwägung der Vorteile (erheblicher Zusatznutzen in fast jedem der betrachteten Endpunkte) und des nicht quantifizierbaren Schadens in der Gesamtschau zum gegenwärtigen Zeitpunkt als **erheblich** einzustufen.

Dem Zusatznutzen von Vandetanib steht zwar eine höhere als unter BSC gemessene Rate unerwünschter Ereignisse gegenüber, gleichwohl erscheint es aber nicht angemessen, den

aufgezeigten positiven Nutzen deshalb abzuwerten. Dafür sprechen maßgeblich drei Argumente:

1. Aus klinisch-therapeutischer Sicht dürfte unter allen direkt gemessenen Endpunkten die mindestens 11-monatige, in der genährten Zielpopulation ca. 1-jährige, Verlängerung des progressionsfreien Überlebens für diese unheilbaren Patienten von vorrangigem Wert sein.

2. Die Zeit bis zur Schmerzprogression (TWP) konnte in beiden Populationen maßgeblich verlängert werden, in der genäherten Zielpopulation in der Vandetanib-Gruppe sogar um fast 8 Monate im Vergleich zur Placebogruppe. Die Vermeidung/Hinauszögerung von Tumorschmerzen muss als herausragendes Therapieziel bei der Behandlung von Krebspatienten in der Palliativsituation gelten.

3. Die unerwünschten Ereignisse waren unter Vandetanib in der Regel hinreichend gut beherrschbar und nur 12% der Patienten sowohl der Studien- als auch der genäherten Zielpopulation brachen die Behandlung wegen unerwünschter Ereignisse ab. Zudem waren auch die im Kontrollarm eingesetzten Maßnahmen der BSC von unerwünschten Nebenwirkungen begleitet.

EMA und FDA haben die Effektivität von Vandetanib im Rahmen der europäischen und der US-amerikanischen Zulassung auf der Basis der gesamten Studienpopulation bewertet. Die für die EMA für die Teilpopulationen „progressiv“, „symptomatisch“, „progressiv oder symptomatisch“ und „progressiv und symptomatisch“ zum primären Endpunkt PFS sowie dem sekundären Endpunkt ORR durchgeführten Subgruppenanalysen bestätigten die Konsistenz der Daten über die betrachteten Teilpopulationen. Die Beschränkung der Indikation auf die Patienten mit aggressiver und symptomatischer Erkrankung durch die EMA erfolgte aufgrund einer Nutzen-Risiko-Abwägung, nicht aufgrund unterschiedlicher Wirksamkeit [2].

Im Rahmen der Nutzenbewertung von Vandetanib durch das IQWiG anhand der ursprünglichen Version des vorliegenden Dossiers kam das Institut zu der Einschätzung, dass der Einschränkung der Indikation auf Patienten mit aggressiver **und** symptomatischer Erkrankung – entgegen der Interpretation von AstraZeneca – durch ausschließliche Betrachtung der gesamten Patientenpopulation der Zulassungsstudie nicht ausreichend Rechnung getragen worden sei, dass das Dossier somit unvollständig und ein Zusatznutzen nicht belegt sei [3]. Um dieser Einschätzung entgegenzuwirken, wurde im vorliegenden neuen Dossier das ursprüngliche Modul 4 A des Dossiers um sämtliche verfügbaren Daten und durchführbaren Analysen zu der durch das IQWiG als Näherung für die Zielpopulation akzeptierten Teilpopulation der Studie mit symptomatischer **und** progressiver Erkrankung ergänzt und diese wurden den ursprünglichen Ergebnissen zur Gesamtpopulation vergleichend gegenüber gestellt. Wie aus den in Abschnitt 4.3.1.3 dargestellten und in den vorangehenden Ausführungen zusammengefassten Resultaten deutlich wird, ändert sich an

der ursprünglich anhand der Gesamtpopulation vorgenommenen Quantifizierung des Zusatznutzens von Vandetanib gegenüber BSC durch die gesonderte Betrachtung der beschriebenen genäherten Zielpopulation nichts Wesentliches: Alle Effekte sind im Vergleich der beiden Populationen gleichgerichtet und im Falle der Endpunkte zur Wirksamkeit signifikant zugunsten von Vandetanib. Die Signifikanz bleibt trotz des erheblichen Verlustes an statistischer *Power*, bedingt durch die im Vergleich zur Gesamtpopulation um 44% reduzierte Patientenzahl in der genäherten Zielpopulation, für alle Ergebnisse erhalten, obwohl die jeweiligen Konfidenzintervalle zwangsläufig breiter werden. Die generelle Übertragbarkeit der Ergebnisse der Gesamtpopulation auf die genäherte Zielpopulation wird auch durch die in Abschnitt 4.3.1.3.2 dargestellten Interaktionstests bestätigt.

Das IQWiG hat im Rahmen der Nutzenbewertung zum Wirkstoff Ticagrelor einen Vorschlag zur Operationalisierung des Zusatznutzens gemäß 5. Kapitel § 5 Absatz 7 VerfO des Gemeinsamen Bundesausschusses [10] anhand von Schwellenwerten, die sich an der Obergrenze der jeweiligen Konfidenzintervalle für den betrachteten Endpunkt orientieren, vorgelegt [105]. Dieser Vorschlag hat sich im Laufe der weiteren seitdem durchgeführten Bewertungsverfahren jedoch nicht als Standard durchgesetzt. In die Methodik des Instituts für die vorliegende Quantifizierung des Zusatznutzens fließen nur die statistischen Größen „Effektschätzer“ und „Konfidenzintervall“ ein, der Schweregrad und die Relevanz der Erkrankung für die Zielpopulation werden jedoch nicht berücksichtigt. Dieser Aspekt spielt jedoch gerade bei äußerst seltenen und schweren Erkrankungen wie dem fortgeschrittenen oder metastasierten MTC, bei dem für die Teilpopulation mit aggressiver und symptomatischer Erkrankung ein dringender therapeutischer Handlungsbedarf besteht, eine große Rolle. Insbesondere die Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) fordert verbindlich ein, dass das Ausmaß des Zusatznutzens „unter Berücksichtigung des Schweregrades der Erkrankung gegenüber der zweckmäßigen Vergleichstherapie“ zu quantifizieren ist [20]. Gemessen an dieser Vorgabe sieht AstraZeneca für alle betrachteten Endpunkte angesichts der eindeutigen und statistisch signifikanten Resultate die Anforderungen für einen **erheblichen** Zusatznutzen von Vandetanib gegenüber *Best Supportive Care* wie oben ausführlich dargelegt als erfüllt an. Die Zulassungsstudie zu Vandetanib ist die erste große Phase-III-RCT, die auf dem Gebiet des fortgeschrittenen MTC durchgeführt wurde. Die Studie erfüllt die Anforderungen der EMA an eine einzelne pivotale Studie [99] und bedingt damit für die Einschätzung der Evidenzlage trotz der Vorlage nur einer Studie die Ableitung als **Beleg** (siehe Abschnitt 4.4.1).

Unter Berücksichtigung obiger Erwägungen beansprucht AstraZeneca für dieses Arzneimittel für die gesamte Zulassungspopulation einen erheblichen Zusatznutzen im Vergleich zu BSC. Zusätzlich ist hier anzumerken, dass es sich bei Vandetanib um die erste zugelassene zielgerichtete Behandlung für die betreffende MTC-Teilpopulation mit aggressiver und symptomatischer Erkrankung handelt („Solist“).

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 4-179: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit aggressivem und symptomatischem MTC in nicht resektablem, lokal fortgeschrittenem oder metastasiertem Stadium	erheblich

Wie in Abschnitt 4.4.2 ausführlich dargelegt, rechtfertigen die im vorliegenden Dossier dargestellten Resultate der Zulassungsstudie von Vandetanib, Studie 58, die Zuerkenntnis eines erheblichen Zusatznutzens gegenüber der zweckmäßigen Vergleichstherapie BSC in vollem Umfang. Dies gilt prinzipiell für die gesamte Zulassungspopulation von Patientem mit aggressivem und symptomatischem MTC, da sich in der gesamten Studienpopulation wie auch in der *post-hoc* analysierten genäherten Zielpopulation der Patienten mit progressiver und symptomatischer Erkrankung nahezu ausnahmslos gleichgerichtete Effekte vergleichbaren Ausmaßes zeigten und sich die Patienten der Zulassungspopulation in den beiden ausgewerteten Populationen wiederfinden.

Laut Text der Zulassung vom 17.02.2012 sollte bei Patienten, deren RET-Mutationsstatus nicht bekannt oder negativ ist, bei der Entscheidung über eine individuelle Behandlung mit Vandetanib ein möglicherweise geringerer Nutzen berücksichtigt werden. Dennoch lässt sich nach Auswertung der Ergebnisse zum PFS aus Studie 58 mittels einer nach RET-Mutationsstatus aufgeschlüsselten Subgruppenanalyse die Überlegenheit von Vandetanib gegenüber der Placebo-Gruppe (BSC) auch in der Patientengruppe mit unbekanntem RET-Mutationsstatus statistisch signifikant belegen, und zwar in einem Ausmaß, das tatsächlich leicht unterhalb der Gruppe mit positivem RET-Mutationsstatus liegt, jedoch immer noch hoch genug ist, um die Anforderungen für eine Klassifizierung als erheblicher Zusatznutzen zu erfüllen (siehe Abschnitt 4.3.1.3.2.1). Der Zusatznutzen speziell in der Patientengruppe mit RET-negativem Status ließ sich mittels der in Studie 58 gewonnenen Daten aufgrund der geringen Zahl dieser Patienten zunächst nicht statistisch bewerten. Aufgrund des Wirkprofils von Vandetanib gegen weitere in die Entstehung solider Tumore sehr häufig ursächlich involvierte Tyrosinkinasen (VEGFR-1, -2 und -3 sowie EGFR; vgl. die Ausführungen in Modul 2) besteht jedoch berechtigter Grund zur Annahme, dass auch diese Patienten von Vandetanib profitieren. In einer *post-hoc* durchgeführten Zusatzanalyse wurden die

Ergebnisse zu den Wirksamkeits-Endpunkten der Studie 58 noch einmal nachträglich für 79 Patienten der Gesamtpopulation (51 der genäherten Zielpopulation) zusammengestellt, für die in Nachbewertungen der RET-Mutationsstatus hinsichtlich der mit Abstand häufigsten RET-Mutation, M918T, sowie in den sechs weiteren am häufigsten betroffenen Exons als negativ beurteilt wurde. 71 dieser Patienten waren in der ursprünglich innerhalb der Studie durchgeführten Analyse als „RET unbekannt“ eingestuft worden. Beim Vergleich der Resultate der nach den „neuen“ Kriterien als RET-negativ klassifizierten Patienten mit denen der ursprünglichen Auswertung für die nachweislich RET-positiven Patienten stellte sich heraus, dass diese, obgleich aufgrund der geringen Patientenzahl statistisch nicht signifikant, in Richtung und Ausmaß vergleichbar waren [2] (siehe auch Subgruppenanalysen in Abschnitt 4.3.1.3.2.3). Die oben getroffenen Annahmen hinsichtlich der Wirksamkeit von Vandetanib auch bei Patienten mit negativem RET-Mutationsstatus werden dadurch unterstützt, und es besteht weiterhin kein Grund, anhand des RET-Status Patientengruppen mit einem von der gesamten Zielpopulation abweichenden Zusatznutzen zu definieren.

Um für diese Fragestellung weitere Informationen zu sammeln, initiiert AstraZeneca eine offene Studie, die auf einem mit der EMA abgestimmten und genehmigten Protokoll basiert. In dieser Studie sollen RET-negative und RET-positive Patienten verglichen werden, die ein sporadisches medulläres Schilddrüsenkarzinom haben und mit Vandetanib behandelt werden. Diese Studie ist essentieller Bestandteil der Auflagen, die bei Erteilung der Zulassung des Arzneimittels unter „Besonderen Bedingungen“ gemäß Artikel 14(7) der Verordnung (EG) Nr. 726/2004 [106] von der EMA für Vandetanib festgelegt wurden [2] und wird ungefähr 60% der Patienten umfassen, die innerhalb der EU mit Vandetanib behandelt werden (siehe Modul 3 A, Abschnitt 3.4.4).

4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens

Für Arzneimittel, die zur Behandlung eines seltenen Leidens zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt.

Für diese Arzneimittel sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, im vorliegenden Abschnitt einzureichen. Dabei soll zunächst das Ausmaß des Zusatznutzens (in der Definition gemäß AM-NutzenV) begründet beschrieben werden. Unter Berücksichtigung des Ausmaßes des Zusatznutzens sollen Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, benannt werden. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Beschreiben Sie begründet das Ausmaß des Zusatznutzens unter Verwendung der Kategorisierung gemäß AM-NutzenV. Berücksichtigen Sie bei den Aussagen zum

Zusatznutzen nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen. Verweisen Sie, wo notwendig, auf weitere Abschnitte von Modul 4.

nicht anwendbar

Geben Sie auf Basis des Ausmaßes des Zusatznutzens 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 4-180: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

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.

Es wurden keine indirekten Vergleiche durchgeführt.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

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

Über die Informationen aus RCT (Abschnitt 4.3.1.3 und 4.3.2.1.3) hinausgehend werden im Abschnitt 4.3.2.3 unterstützende Informationen aus zwei nicht randomisierten Studien herangezogen. Diese Studien sind gemäß den in Abschnitt 4.2.2 angegeben Ein- bzw. Ausschlusskriterien nicht geeignet, die Fragestellung gemäß 4.2.1 zu beantworten. Sie wurden

dennoch in Abschnitt 4.3.2.3 sowie ebenfalls in Anhang 4-F dargestellt, um die Evidenz für Vandetanib im MTC zu vervollständigen. Auch die in diesen Studien erhobenen Endpunkte wurden der Vollständigkeit halber in Anhang 4-G hinsichtlich ihres Verzerrungspotentials bewertet. Ihre Resultate wurden allerdings **nicht** für die Bewertung des medizinischen Nutzens und des Zusatznutzens von Vandetanib berücksichtigt. Aus diesem Grund werden in den Bewertungsbögen zu Studie D4200C00008 (Tabelle 4-191) und Studie D4200C00068 (Tabelle 4-192) patientenrelevante Endpunkte im Sinne von Tabelle 4-1 (siehe oben) nur aufgeführt, insoweit sie in diesen Studien auch erhoben wurden.

Die Studien wurden in der Registersuche zu Vandetanib als weitere Studien im Indikationsgebiet identifiziert. Beide Studien sind Studien des pharmazeutischen Unternehmers und waren Teil des Zulassungsdossiers fürn die europäische Zulassungsbehörde.

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.

Alle Daten zu den in diesem Dossier für die Quantifizierung des Zusatznutzens von Vandetanib gegenüber BSC herangezogenen Endpunkten lagen vor.

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 2005[13], Molenberghs 2010[14]). 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)

[13] *Burzykowski T (Ed.: The evaluation of surrogate endpoints. New York: Springer; 2005.*
[14] *Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.*

(Burzykowski 2006[15]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[16]) 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.

Nach Überzeugung von AstraZeneca sind die betrachteten Endpunkte PFS, TWP, ORR, DCR, DOR, HRQoL und alle aufgeführten explorativen Morbititätsendpunkte in dieser Indikation patientenrelevant und werden entsprechend in den Abschnitten 4.1 bis 4.4 beschrieben, obgleich AstraZeneca die Einschätzungen des IQWiG [89] aus vorausgegangenen Bewertungen bekannt sind. Der Endpunkt „Biochemisches Ansprechen“ als prognostischer Marker für die Progression und als Indikator für ein Anpsrechen auf die Therapie wird als Surrogat innerhalb der patientenrelevanten Endpunktkategorie krankheitsbedingte Symptome (Morbidität) zur Unterstützung der Ergebnisse zu PFS und ORR herangezogen. Das biochemische Ansprechen, gemessen an den Serumspiegeln von CTN und CEA, stellt gerade im Falle des MTC gut geeignete Marker für den klinischen Endpunkt Progression sowie einen Hinweis auf ein Therapieansprechen zur Verfügung. Die MTC-Zellen sezernieren Calcitonin (CTN) und Carcinoembryonales Antigen (CEA). Hohe Blutkonzentrationen von CTN gehen mit einem Fortschreiten der Erkrankung sowie mit Durchfall und *Flush*-Syndrom einher. Die Durchfälle können bei Patienten mit CTN-

[15] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*
[16] *Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.*

Konzentrationen über 30.000 pg/ml besonders schwerwiegende Ausmaße annehmen und regelmäßige ärztliche Intervention erfordern, was die Lebensqualität für die Patienten drastisch reduziert. In der klinischen Praxis wird die Messung des CTN-Spiegels routinemäßig zur Verfolgung der Krankheitsprogression angewandt, da dieser Parameter direkt mit der Belastung des Organismus durch den Tumor verknüpft ist [107]. Auch CEA ist für diesen Tumor ein geeigneter Marker; erhöhte Plasmakonzentrationen dieses Parameters deuten auf ein Fortschreiten der Erkrankung und die Bildung von Metastasen hin [108].

Eine formelle Validierung dieses Surrogatparameters wurde bislang im Bereich des MTC – und unseres Wissens auch in anderen onkologischen Indikationen – noch nicht durchgeführt. Dennoch werden die im Rahmen der Zulassungsstudie von Vandetanib für das biochemische Ansprechen erhaltenen Resultate zur Unterstützung der Ergebnisse zu PFS und ORR im vorliegenden Dossier mit aufgeführt und bewertet, um alle verfügbaren Informationen zur Wirksamkeit von Vandetanib, ob aus validierten Endpunkten oder nicht, zumindest dargestellt zu haben.

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 Datenquellen auf.

Alle identifizierten Studien (Abschnitt 4.3.1.1) mit Vandetanib wurden in die Quantifizierung des Zusatznutzens gegenüber der zweckmäßigen Vergleichstherapie BSC einbezogen.

Die Charakteristika und die Ergebnisse der RCT D4200C00058 (Studie 58) wurden in den Abschnitten 4.3.1.2 und 4.3.1.3 beschrieben.

Als weitere Untersuchungen wurden die Ergebnisse der Studien D4200C00008 und D4200C00068 einbezogen, um die Evidenz für Vandetanib zu vervollständigen. Die Charakteristika und die Ergebnisse dieser nicht-randomisierten, offenen Studien wurden im Abschnitt 4.3.2.3 dargestellt.

Tabelle 4-181: Für die Nutzenbewertung berücksichtigte Studien und Untersuchungen

Studie	Titel	Datenquelle
D4200C00058	An International, Phase III*, Randomized, Double-Blinded, Placebo-Controlled, Multi-Center Study to Assess the Efficacy of ZD6474 (ZACTIMA™) Versus Placebo in Subjects With Unresectable, Locally Advanced or Metastatic Medullary Thyroid Cancer	[11, 22, 91, 92]
D4200C00008	A Phase II, Open-Label Study To Assess The Efficacy and Tolerability of ZD6474 (ZACTIMA™) 100 mg Monotherapy In Subjects With Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer	[97, 109]
D4200C00068	An Open Label, Two Stage, Phase II Study to Evaluate the Efficacy and Tolerability of ZD6474 in Patients With Locally Advanced or Metastatic Hereditary Medullary Thyroid Carcinoma	[98, 110]

*in clinicaltrial.gov Fehler in der Auszeichnung, im Studienprotokoll steht korrekt „Phase III“

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

[1] Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV. Beratungsanforderung 2011-B-025. Berlin: Gemeinsamer Bundesausschuss, 2011.

[2] Assessment Report Caprelsa [Online-Datenquelle]. European Medicines Agency. 2011. Verfügbar unter: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002315/WC500123603.pdf.

[3] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Vandetanib: Nutzenbewertung gemäß § 35a SGB V - Auftrag A12-09. Köln: IQWiG-Berichte. 2012.

[4] 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 - Vandetanib, 2012.

[5] Deutscher Bundestag. Beschlussempfehlung und Bericht des Ausschusses für Gesundheit (14. Ausschuss). In: Ausschuss für Gesundheit (14. Ausschuss), Hrsg. Berlin: Deutscher Bundestag; 2012. S. 83.

[6] Prävalenz seltener Krankheiten: Bibliographische Angaben. Alphabetische Liste der Krankheiten oder Krankheitsgruppen [Online-Datenquelle]. 2012 [Zugriffsdatum: 24.10.2012]. Verfügbar unter: http://www.orpha.net/orphacom/cahiers/docs/DE/Pravalenzen_seltener_Krankheiten_Alphabetische_Liste.pdf.

[7] VERORDNUNG (EG) Nr. 141/2000 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 16. Dezember 1999 über Arzneimittel für seltene Leiden (ABl. L 18 vom 22.1.2000, S. 1). Geändert durch Verordnung (EG) Nr. 596/2009 des Europäischen Parlaments und des Rates vom 18. Juni 2009, ABl. L 18 2009.

[8] European Medicines Agency. EUROPEAN MEDICINES AGENCY DECISION of 3 November 2008 on the application for agreement of a Paediatric Investigation Plan for vandetanib, EMEA-000052-PIP01-07 in accordance with Regulation (EC) No 1901/2006 of the European Parliament and of the Council as amended. European Medicines Agency, 2008 Contract No.: EMEA/562638/2008 P/94/2008.

[9] Committee for Orphan Medicinal Products. Public summary of opinion on orphan designation - vandetanib for the treatment of medullary thyroid carcinoma. European Medicines Agency (EMA), 2010 Contract No.: EMA/COMP/20891/2006 Rev.1.

[10] Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses. 5 Kapitel: Bewertung des Nutzens von Arzneimitteln nach § 35a SGB V. Berlin: Gemeinsamer Bundesausschuss; 2013.

[11] AstraZeneca Pharmaceuticals LP. An International, Phase III, Randomized, Double-Blinded, Placebo-Controlled, Multi-Center Study to Assess the Efficacy of ZD6474 versus Placebo in Subjects with Unresectable Locally Advanced or Metastatic Medullary Thyroid Cancer - Consolidated post-hoc Analyses for benefit assessment pursuant to Section 35a of the German Social Code Book V (SGB V). Post-hoc Analyses D4200C00058. 2013.

[12] Carlomagno F, Vitagliano D, Guida T, Ciardiello F, Tortora G, Vecchio G, et al. ZD6474, an orally available inhibitor of KDR tyrosine kinase activity, efficiently blocks oncogenic RET kinases. Cancer Res. 2002;62(24):7284-90. Epub 2002/12/25.

[13] Knowles PP, Murray-Rust J, Kjaer S, Scott RP, Hanrahan S, Santoro M, et al. Structure and chemical inhibition of the RET tyrosine kinase domain. J Biol Chem. 2006;281(44):33577-87. Epub 2006/08/25.

[14] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Brustkrebs mit Metastasen: Wie Frauen auch mit einer schweren Erkrankung leben können. Köln: IQWiG Pressemitteilungen 2012. 2012.

[15] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) und Deutsche Gesellschaft für Hämatologie und Onkologie e.V. (DGHO). Nutzenbewertung von Medikamenten in der Onkologie - Diskussionspapier AWMF und DGHO zu den Fragen des Gemeinsamen Bundesausschusses. 2010.

[16] Ehninger G, Folprecht G, Arnold A. Frühe Nutzenbewertung onkologischer Arzneimittel - Diskussionspapier der DGHO zu den Fragen des Gemeinsamen Bundesausschusses anlässlich der Diskussionsrunde am 29. November 2010. In: Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO), Hrsg. Berlin: Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO); 2010.

[17] Haerting J, Victor A, Hauschke D. Kommentar zum IQWiG-Bericht A10-05 "Aussagekraft von Surrogatendpunkten in der Onkologie" (Version 1.0, Stand: 31.01.2011). In: Deutsche Gesellschaft für Medizinische Informatik Biometrie und Epidemiologie (gmds) e. V., Hrsg. Köln: gmds; 2011.

[18] Committee for Medicinal Products for Human Use (CHMP). Guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/205/95. London: European Medicines Agency; 2012.

[19] Stuver SO, Isaac T, Weeks JC, Block S, Berry DL, Davis RB, et al. Factors associated with pain among ambulatory patients with cancer with advanced disease at a

comprehensive cancer center. Journal of oncology practice / American Society of Clinical Oncology. 2012;8(4):e17-23. Epub 2012/11/28.

[20] Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV), 2010.

[21] Gemeinsamer Bundesausschuss. Mündliche Anhörung gemäß 5. Kapitel, § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesauschusses - hier: Wirkstoff Vandetanib. Berlin: Gemeinsamer Bundesausschuss, 2012.

[22] AstraZeneca Pharmaceuticals LP. An International, Phase III, Randomized, Double-Blinded, Placebo-Controlled, Multi-Center Study to Assess the Efficacy of ZD6474 versus Placebo in Subjects with Unresectable Locally Advanced or Metastatic Medullary Thyroid Cancer. D4200C00058 Clinical Study Report. 2010.

[23] Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. Journal of the National Cancer Institute. 2000;92(3):205-16. Epub 2000/02/03.

[24] Center for Drug Evaluation and Research (CBER). Guidance for Industry - Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. Food and Drug Administration (FDA). 2007:1-19.

[25] McCain jr. JA. The Ongoing Evolution of Endpoints in Oncology. Managed Care [Internet]. 2010 [Zugriffsdatum: 11.10.2011]; May, 2010(Supplement). Verfügbar unter: http://www.nxtbook.com/nxtbooks/medimedia/oncologyendpoints/index.php?startid=Cover1.

[26] Pazdur R. Endpoints for assessing drug activity in clinical trials. Oncologist. 2008;13 Suppl 2:19-21. Epub 2008/06/17.

[27] Gemeinsamer Bundesausschuss (G-BA). Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Ticagrelor. Berlin: Gemeinsamer Bundesausschuss (G-BA), 2011.

[28] Dralle H. Maligne Schilddrüsentumoren. In: Deutsche Krebsgesellschaft, Hrsg. Kurzgefasste interdisziplinäre Leitlinien: W. Zuckschwerdt Verlag; 2008.

[29] Cerrato A, De Falco V, Santoro M. Molecular genetics of medullary thyroid carcinoma: the quest for novel therapeutic targets. Journal of molecular endocrinology. 2009;43(4):143-55. Epub 2009/04/23.

[30] Sippel RS, Kunnimalaiyaan M, Chen H. Current management of medullary thyroid cancer. Oncologist. 2008;13(5):539-47. Epub 2008/06/03.

[31] Modigliani E, Cohen R, Campos JM, Conte-Devolx B, Maes B, Boneu A, et al. Prognostic factors for survival and for biochemical cure in medullary thyroid carcinoma: results in 899 patients. The GETC Study Group. Groupe d'etude des tumeurs a calcitonine. Clin Endocrinol (Oxf). 1998;48(3):265-73. Epub 1998/05/14.

[32] Roman S, Lin R, Sosa JA. Prognosis of medullary thyroid carcinoma: demographic, clinical, and pathologic predictors of survival in 1252 cases. Cancer. 2006;107(9):2134-42. Epub 2006/10/05.

[33] Cohen R, Buchsenschutz B, Estrade P, Gardet P, Modigliani E. [Causes of death in patients suffering from medullary thyroid carcinoma: report of 119 cases] - English Abstract. Presse Med. 1996;25(37):1819-22. Epub 1996/11/30. Causes de mortalite

chez les patients atteints de cancer Medullaire de la thyroide. GETC. Groupe d'Etude des Tumeurs a Calcitonine.

[34] Eng C, Mulligan LM, Healey CS, Houghton C, Frilling A, Raue F, et al. Heterogeneous mutation of the RET proto-oncogene in subpopulations of medullary thyroid carcinoma. Cancer Res. 1996;56(9):2167-70. Epub 1996/05/01.

[35] Pacini F, Castagna MG, Brilli L, Pentheroudakis G. Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012;23(Suppl 7):vii110-vii9.

[36] Schlumberger M, Bastholt L, Dralle H, Jarzab B, Pacini F, Smit JWA. 2012 European Thyroid Association Guidelines for Metastatic Medullary Thyroid Cancer. European Thyroid Journal. 2012;1(1):5-14.

[37] Kloos RT, Eng C, Evans DB, Francis GL, Gagel RF, Gharib H, et al. Medullary thyroid cancer: management guidelines of the American Thyroid Association. Thyroid. 2009;19(6):565-612. Epub 2009/05/28.

[38] Tuttle RM, Ball DW, Byrd D, Dickson P, Duh Q-Y, Ehya H, et al. Thyroid Carcinoma. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines). 2013;Version 1.2013:1-98.

[39] Salzwimmer M. Best Supportive Care bei Patienten mit Kopf-Hals-Tumoren. Wien Med Wochenschr. 2008;158(9-10):278-82. Epub 2008/06/19.

[40] Gawlik T, d'Amico A, Szpak-Ulczok S, Skoczylas A, Gubala E, Chorazy A, et al. The prognostic value of tumor markers doubling times in medullary thyroid carcinoma - preliminary report. Thyroid research. 2010;3(1):10. Epub 2010/11/05.

[41] Ziegler A, König IR. Leitlinien für Forschungsberichte: Deutschsprachige Übersetzungen von CONSORT 2010, PRISMA und STARD. Dtsch Med Wochenschr. 2011;136(8):357-8. Epub 2011/02/19.

[42] Fach- und Gebrauchsinformation Caprelsa® 100 mg/300 mg Filmtabletten [Online-Datenquelle]. 2013. Verfügbar unter: http://www.ema.europa.eu/docs/de_DE/document_library/EPAR_-_Product_Information/human/002315/WC500123555.pdf.

[43] Center for Drug Evaluation and Research (CBER). Approval Package for Application Number 022405Orig1s000. In: Department of Health and Human Services, Hrsg. Silver Spring, MD: Food and Drug Administration; 2011.

[44] Häuser W. Pharmakotherapie in der Psychoonkologie. Onkologische Welt. 2010:32-4.

[45] Verband Forschender Arzneimittelhersteller e.V. (vfa). vfa-Beitrag zur G-BA Diskussionsrunde "Frühe Nutzenbewertung onkologischer Arzneimittel". Berlin: Verband Forschender Arzneimittelhersteller e.V.; 2010.

[46] Verband Forschender Arzneimittelhersteller e.V. (vfa). vfa-Position zu Endpunkten bei onkologischen Therapien - Positionspapier. Berlin: Verband Forschender Arzneimittelhersteller e.V.; 2012.

[47] Walker MS, Hasan M, Yim YM, Yu E, Stepanski EJ, Schwartzberg LS. Retrospective study of the effect of disease progression on patient reported outcomes in HER-2 negative metastatic breast cancer patients. Health and quality of life outcomes. 2011;9:46. Epub 2011/06/22.

[48] Bundesverfassungsgericht (BVerfG): Leitsatz zum Beschluss des Ersten Senats vom 6. Dezember 2005 - 1 BvR 347/98 -, 1 BvR 347/98 2005.

[49] LSG Baden-Württemberg: Möglichkeit der Inanspruchnahme einer Substanz ohne oder außerhalb ihrer arzneimittelrechtlichen Zulassung (off-label-use) zu Lasten der gesetzlichen Krankenversicherung - Bestehen eines Anspruchs auf Versorgung mit

dem Medikament Herceptin - Behandlung von Patienten mit metastasierendem Mammakarzinom, L 5 KR 4180/05 ER-B 2005.

[50] Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. Journal of the National Cancer Institute. 2009;101(23):1642-9. Epub 2009/11/12.

[51] Aidelsburger P, Wasem J. Kosten-Nutzen-Bewertungen von onkologischen Therapien. Deutsche Krebsgesellschaft, 2008.

[52] Amit O, Bushnell W, Dodd L, Roach N, Sargent D. Blinded independent central review of the progression-free survival endpoint. Oncologist. 2010;15(5):492-5. Epub 2010/05/22.

[53] Bureau of Metabolism Oncology and Reproductive Science (BMORS), Oncology Division. The use of progression-free survival as the efficacy endpoint for approval of targeted and chemotherapeutic agents for advanced cancer. Health Canada, Issues Analysis Summary. 2007.

[54] Mayfield E. Progression-Free Survival: Patient Benefit or Lower Standard? NCI Cancer Bulletin. 2008.

[55] Bast RC, Thigpen JT, Arbuck SG, Basen-Engquist K, Burke LB, Freedman R, et al. Clinical trial endpoints in ovarian cancer: report of an FDA/ASCO/AACR Public Workshop. Gynecol Oncol. 2007;107(2):173-6. Epub 2007/10/24.

[56] Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol. 2007;25(33):5218-24. Epub 2007/11/21.

[57] Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomised, double-blind, phase 3 study. Lancet. 2009;374(9699):1432-40. Epub 2009/09/22.

[58] Committee for Medicinal Products for Human Use (CHMP). METHODOLOGICAL CONSIDERATIONS FOR USING PROGRESSION-FREE SURVIVAL (PFS) AS PRIMARY ENDPOINT IN CONFIRMATORY TRIALS FOR REGISTRATION - APPENDIX 1 TO THE GUIDELINE ON THE EVALUATION OF ANTICANCER MEDICINAL PRODUCTS IN MAN. London: European Medicines Agency, 2008.

[59] Di Leo A, Bleiberg H, Buyse M. Overall survival is not a realistic end point for clinical trials of new drugs in advanced solid tumors: a critical assessment based on recently reported phase III trials in colorectal and breast cancer. J Clin Oncol. 2003;21(10):2045-7. Epub 2003/05/14.

[60] Gastineau T. FDA Guidance Document Focuses on Cancer Trial Endpoints. Applied Clinical Trials Online [Internet]. 2006 [Zugriffsdatum: 01.12.2011]. Verfügbar unter: http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/article/articleDetail.jsp?id=324334.

[61] Radbruch L, Loick G, Kiencke P, Lindena G, Sabatowski R, Grond S, et al. Validation of the German version of the Brief Pain Inventory. J Pain Symptom Manage. 1999;18(3):180-7. Epub 1999/10/12.

[62] Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. Pain. 1983;17(2):197-210. Epub 1983/10/01.

[63] Saad ED, Katz A. Progression-free survival and time to progression as primary end points in advanced breast cancer: often used, sometimes loosely defined. Ann Oncol. 2009;20(3):460-4. Epub 2008/12/20.

[64] Dy GK, Krook JE, Green EM, Sargent DJ, Delaunoit T, Morton RF, et al. Impact of complete response to chemotherapy on overall survival in advanced colorectal cancer: results from Intergroup N9741. J Clin Oncol. 2007;25(23):3469-74. Epub 2007/08/10.
[65] Li X, Liu S, Gu H, Wang D. Surrogate end points for survival in the target treatment of advanced non-small-cell lung cancer with gefitinib or erlotinib. J Cancer Res Clin Oncol. 2012;138(11):1963-9. Epub 2012/07/06.
[66] Asahina H, Oizumi S, Inoue A, Kinoshita I, Ishida T, Fujita Y, et al. Phase II study of gefitinib readministration in patients with advanced non-small cell lung cancer and previous response to gefitinib. Oncology. 2010;79(5-6):423-9. Epub 2011/04/09.
[67] Ciliberto D, Botta C, Correale P, Rossi M, Caraglia M, Tassone P, et al. Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: A meta-analysis of randomised trials. Eur J Cancer. 2013;49:593-603. Epub 2012/09/20.
[68] Tsao AS, Liu S, Lee JJ, Alden C, Blumenschein G, Herbst R, et al. Clinical Outcomes and Biomarker Profiles of Elderly Pretreated NSCLC Patients from the BATTLE Trial. J Thorac Oncol. 2012;7(11):1645-52. Epub 2012/10/13.
[69] Barnes J, Jacobsen E, Feng Y, Freedman A, Hochberg E, Lacasce A, et al. Everolimus in combination with rituximab induces complete responses in heavily pretreated diffuse large B-cell lymphoma. Haematologica. 2012. Epub 2012/11/13.
[70] Jimenez B, Trigo JM, Pajares BI, Saez MI, Quero C, Navarro V, et al. Efficacy and safety of weekly paclitaxel combined with cetuximab in the treatment of pretreated recurrent/metastatic head and neck cancer patients. Oral oncology. 2013;49:182-5. Epub 2012/10/03.
[71] Savvides P, Nagaiah G, Lavertu PN, Fu P, Wright JJ, Chapman R, et al. Phase II Trial of Sorafenib in Patients with Advanced Anaplastic Carcinoma of the Thyroid. Thyroid. 2012. Epub 2012/11/02.
[72] Siegel DS, Martin T, Wang M, Vij R, Jakubowiak AJ, Lonial S, et al. A phase 2 study of single-agent carfilzomib (PX-171-003-A1) in patients with relapsed and refractory multiple myeloma. Blood. 2012;120(14):2817-25. Epub 2012/07/27.
[73] Sheikh N, Chambers CR. Efficacy vs. effectiveness: Erlotinib in previously treated non-small-cell lung cancer. J Oncol Pharm Pract. 2012. Epub 2012/11/24.
[74] Van Damme V, Govaerts E, Nackaerts K, Dooms C, Wauters I, Vansteenkiste J. Clinical factors predictive of long-term survival in advanced non-small cell lung cancer. Lung cancer. 2013;79:73-6. Epub 2012/10/23.
[75] Kumar NB, Kazi A, Smith T, Crocker T, Yu D, Reich RR, et al. Cancer cachexia: traditional therapies and novel molecular mechanism-based approaches to treatment. Curr Treat Options Oncol. 2010;11(3-4):107-17. Epub 2010/12/04.
[76] Palesty JA, Dudrick SJ. What we have learned about cachexia in gastrointestinal cancer. Digestive diseases. 2003;21(3):198-213. Epub 2003/10/23.
[77] Mann B, Karsten C, Hotz H, Buhr HJ. Medulläres Schilddrüsenkarzinom. Der Onkologe. 2000;6:651-9.
[78] Karnofsky DA, Burchenal JH. The Clinical Evaluation of Chemotherapeutic Agents in Cancer. In: MacLeod CM, Hrsg. Evaluation of Chemotherapeutic Agents: Columbia Univ Press; 1949. S. 196.
[79] Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky Performance Status Scale. An examination of its reliability and validity in a research setting. Cancer. 1984;53(9):2002-7. Epub 1984/05/01.

[80] Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649-55. Epub 1982/12/01.

[81] Buccheri G, Ferrigno D, Tamburini M. Karnofsky and ECOG performance status scoring in lung cancer: a prospective, longitudinal study of 536 patients from a single institution. Eur J Cancer. 1996;32A(7):1135-41. Epub 1996/06/01.

[82] Plante GE, VanItallie TB. Opioids for cancer pain: the challenge of optimizing treatment. Metabolism: clinical and experimental. 2010;59 Suppl 1:S47-52. Epub 2010/09/21.

[83] Thapa D, Rastogi V, Ahuja V. Cancer pain management-current status. Journal of anaesthesiology, clinical pharmacology. 2011;27(2):162-8. Epub 2011/07/21.

[84] Saad ED, Katz A, Hoff PM, Buyse M. Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. Ann Oncol. 2010;21(1):7-12. Epub 2009/11/11.

[85] Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993;11(3):570-9. Epub 1993/03/01.

[86] Schlesinger-Raab A, Schubert-Fritschle G, Hein R, Stolz W, Volkenandt M, Holzel D, et al. Quality of life in localised malignant melanoma. Ann Oncol. 2010;21(12):2428-35. Epub 2010/05/25.

[87] Committee for Proprietary Medicinal Products (CPMP). ICH Topic E 9: Statistical Principles for Clinical Trials - Step 5: Note for Guidance on Statistical Principles for Clinical Trials (CPMP/ICH/363/96). London: European Medicines Agency; 1998.

[88] Higgins JPT, Green Se. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, West Sussex , England: The Cochrane Collaboration and John Wiley & Sons Ltd.; 2008.

[89] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden - Version 4.0 vom 23.09.2011. In: IQWiG, Hrsg. Köln: IQWiG; 2011.

[90] An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer [Online-Datenquelle]. 2012 [Zugriffsdatum: 27.11.2012]. Verfügbar unter: http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00410761.

[91] An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer [Online-Datenquelle]. 2012 [Zugriffsdatum: 11.02.2013]. Verfügbar unter: http://www.clinicaltrials.gov/ct2/show/NCT00410761.

[92] Wells SA, Jr., Robinson BG, Gagel RF, Dralle H, Fagin JA, Santoro M, et al. Vandetanib in Patients With Locally Advanced or Metastatic Medullary Thyroid Cancer: A Randomized, Double-Blind Phase III Trial. J Clin Oncol. 2011. Epub 2011/10/26.

[93] Committee for Medicinal Products for Human Use (CHMP). Rapporteurs' Day 150 Joint Response Assessment Report Caprelsa. European Medicines Agency, 2011.

[94] FACIT.org. FACT-G Quality of Life Questionnaire (Version 4) - German. Elmhurst, IL, USA: FACIT.org; 2006.

[95] FACIT.org. FACT-G Scoring Guidelines (Version 4). Elmhurst, IL, USA: FACIT.org; 2007.

[96] FACIT.org. FACIT Administration and Scoring Guidelines - Scoring the FACT-G. Elmhurst, IL, USA: FACIT.org; 2010.

[97] Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer [Online-Datenquelle]. 2012 [Zugriffsdatum: 11.02.2013]. Verfügbar unter: http://www.clinicaltrials.gov/ct2/show/NCT00098345.

[98] A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer [Online-Datenquelle]. 2012 [Zugriffsdatum: 11.02.2013]. Verfügbar unter: http://www.clinicaltrials.gov/ct2/show/NCT00358956.

[99] Committee for Proprietary Medicinal Products (CPMP). Points to Consider on Application with 1. Meta-Analyses; 2. One Pivotal Study. London: European Medicines Agency, 2001.

[100] Clarke JT. Is the current approach to reviewing new drugs condemning the victims of rare diseases to death? A call for a national orphan drug review policy. CMAJ. 2006;174(2):189-90. Epub 2006/01/18.

[101] Wörmann B, Schmoll HJ, Fassnacht M, Müller S, Spitzweg C. Stellungnahme zur Nutzenbewertung des IQWiG gemäß § 35a SGB V - Vandetanib. In: Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO), Hrsg. Berlin: Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO); 2012.

[102] Robert-Koch-Institut. Krebs in Deutschland 2007/2008. Berlin: Robert-Koch-Institut, Gesundheitsberichterstattung des Bundes. 2012.

[103] Raue F, Frank-Raue K. Multiple endocrine neoplasia type 2: 2007 update. Hormone research. 2007;68 Suppl 5:101-4. Epub 2008/02/07.

[104] Reiners C, Dietlein M, Luster M. Struma maligna - Schilddrüsenkarzinome. Deutsche Medizinische Wochenschrift. 2008;133(43):2215-28.

[105] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Ticagrelor: Nutzenbewertung gemäß § 35a SGB V - Anhang A: Erläuterungen zur Operationalisierung des Ausmaßes des Zusatznutzens gemäß AM-NutzenV auf Endpunktebene. Köln: 2011.

[106] VERORDNUNG (EG) Nr. 726/2004 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 31. März 2004 zur Festlegung von Gemeinschaftsverfahren für die Genehmigung und Überwachung von Human- und Tierarzneimitteln und zur Errichtung einer Europäischen Arzneimittel-Agentur, 2004.

[107] Giraudet AL, Al Ghulzan A, Auperin A, Leboulleux S, Chehboun A, Troalen F, et al. Progression of medullary thyroid carcinoma: assessment with calcitonin and carcinoembryonic antigen doubling times. Eur J Endocrinol. 2008;158(2):239-46. Epub 2008/01/31.

[108] Barbet J, Campion L, Kraeber-Bodere F, Chatal JF. Prognostic impact of serum calcitonin and carcinoembryonic antigen doubling-times in patients with medullary thyroid carcinoma. J Clin Endocrinol Metab. 2005;90(11):6077-84. Epub 2005/08/11.

[109] AstraZeneca Pharmaceuticals LP. An Open-Label, Two-Stage, Phase II Study to Evaluate the Efficacy and Tolerability of ZD6474 in Patients With Locally Advanced or Metastatic Hereditary Medullary Thyroid Carcinoma. D4200C00008 Clinical Study Report. 2009.

[110] AstraZeneca Pharmaceuticals LP. A Phase II, Open-Label Study To Assess The Efficacy and Tolerability of ZD6474 (ZACTIMA™) 100 mg Monotherapy In Subjects with Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer. D4200C00068 Clinical Study Report. 2008.

[111] Wong SS, 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):451-5. Epub 2006/11/04.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. 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. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Datenbankname	EMBASE
Suchoberfläche	Ovid
Datum der Suche	08.12.2010
Zeitsegment	1980 to 2010 week 50
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[17]] – Strategy minimizing difference between sensitivity and specificity

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

Die für die Neueinreichung des vorliegenden Dossiers erforderlich gewordene Literaturrecherche wurde im Internetauftritt des DIMDI mit den nachfolgend einzeln für die jeweilig durchsuchten Datenbanken spezifizierten Suchstrategien durchgeführt:

[17] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Datenbankname	Cochrane-Datenbanken (CDSR, CDAR, CCTR, INAHTA, NHSEED)
Suchoberfläche	DIMDI Smart Search
Datum der Suche	11.01.2013
Zeitsegment	Cochrane Central Register of Controlled Trials (CCTR): 1948 bis aktuell Cochrane Database of Systematic Reviews (CDSR): 1995 bis aktuell Database of Abstracts of Reviews of Effects (CDAR): 2002 bis aktuell Health Technology Assessment Database (INAHTA): 2003 bis aktuell NHS Economic Evaluation Database (NHSEED): 1995 bis aktuell
Suchfilter	Keiner verwendet

#	Suchbegriffe	Ergebnis
1	CCTR93 CDAR94 CDSR93 INAHTA NHSEED	727444
2	(((CT D "thyroid ca"? OR UT="thyroid ca"? OR IT="thyroid ca"? OR SH="thyroid ca"?) OR (CTG D "thyroid ca"? OR UTG="thyroid ca"? OR ITG="thyroid ca"? OR SHG="thyroid ca"?)) OR FT=thyroid ca?) OR TI=thyroid ca?	447
3	(((CT D "thyroid neopla"? OR UT="thyroid neopla"? OR IT="thyroid neopla"? OR SH="thyroid neopla"?) OR (CTG D "thyroid neopla"? OR UTG="thyroid neopla"? OR ITG="thyroid neopla"? OR SHG="thyroid neopla"?)) OR FT=thyroid neopla?) OR TI=thyroid neopla?	373
4	(((CT D "struma maligna" OR UT="struma maligna" OR IT="struma maligna" OR SH="struma maligna") OR (CTG D "struma maligna" OR UTG="struma maligna" OR ITG="struma maligna" OR SHG="struma maligna")) OR FT=struma maligna) OR TI=struma maligna	0
5	(((CT D "schilddr"##"sen"#"ar"#"inom"? OR UT="schilddr"##"sen"#"ar"#"inom"? OR IT="schilddr"##"sen"#"ar"#"inom"? OR SH="schilddr"##"sen"#"ar"#"inom"?) OR (CTG D "schilddr"##"sen"#"ar"#"inom"? OR UTG="schilddr"##"sen"#"ar"#"inom"? OR ITG="schilddr"##"sen"#"ar"#"inom"? OR SHG="schilddr"##"sen"#"ar"#"inom"?)) OR FT=schilddr##sen#ar#inom?) OR TI=schilddr##sen#ar#inom?	4
6	(((CT D "schilddr"##"senkrebs"? OR UT="schilddr"##"senkrebs"? OR IT="schilddr"##"senkrebs"? OR SH="schilddr"##"senkrebs"?) OR (CTG D "schilddr"##"senkrebs"? OR UTG="schilddr"##"senkrebs"? OR ITG="schilddr"##"senkrebs"? OR SHG="schilddr"##"senkrebs"?)) OR FT=schilddr##senkrebs?) OR TI=schilddr##senkrebs?	0

7	2 OR 3 OR 4 OR 5 OR 6	452
8	(((CT D "medull"##"r"? OR UT="medull"##"r"? OR IT="medull"##"r"? OR SH="medull"##"r"?) OR (CTG D "medull"##"r"? OR UTG="medull"##"r"? OR ITG="medull"##"r"? OR SHG="medull"##"r"?)) OR FT=medull##r?) OR TI=medull##r?	660
9	7 AND 8	36
10	(((CT D "mtc" OR UT="mtc" OR IT="mtc" OR SH="mtc") OR (CTG D "mtc" OR UTG="mtc" OR ITG="mtc" OR SHG="mtc")) OR FT=mtc) OR TI=mtc	122
11	9 OR 10	144
12	(((CT D "caprelsa" OR UT="caprelsa" OR IT="caprelsa" OR SH="caprelsa") OR (CTG D "caprelsa" OR UTG="caprelsa" OR ITG="caprelsa" OR SHG="caprelsa")) OR FT=caprelsa) OR TI=caprelsa	0
13	(((CT D "zactima" OR UT="zactima" OR IT="zactima" OR SH="zactima") OR (CTG D "zactima" OR UTG="zactima" OR ITG="zactima" OR SHG="zactima")) OR FT=zactima) OR TI=zactima	3
14	(((CT D "vandetanib" OR UT="vandetanib" OR IT="vandetanib" OR SH="vandetanib") OR (CTG D "vandetanib" OR UTG="vandetanib" OR ITG="vandetanib" OR SHG="vandetanib")) OR FT=vandetanib) OR TI=vandetanib	31
15	(((CT D "ZD6474" OR UT="ZD6474" OR IT="ZD6474" OR SH="ZD6474") OR (CTG D "ZD6474" OR UTG="ZD6474" OR ITG="ZD6474" OR SHG="ZD6474")) OR FT=ZD6474) OR TI=ZD6474	7
16	(((CT D "ZD 6474" OR UT="ZD 6474" OR IT="ZD 6474" OR SH="ZD 6474") OR (CTG D "ZD 6474" OR UTG="ZD 6474" OR ITG="ZD 6474" OR SHG="ZD 6474")) OR FT=ZD 6474) OR TI=ZD 6474	0
17	(((CT D "ZD-6474" OR UT="ZD-6474" OR IT="ZD-6474" OR SH="ZD-6474") OR (CTG D "ZD-6474" OR UTG="ZD-6474" OR ITG="ZD-6474" OR SHG="ZD-6474")) OR FT=ZD-6474) OR TI=ZD-6474	0
18	12 OR 13 OR 14 OR 15 OR 16 OR 17	36
19	11 AND 18	2

Datenbankname	MEDLINE
Suchoberfläche	DIMDI Smart Search
Datum der Suche	11.01.2013
Zeitsegment	1960 bis aktuell
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [111] – Strategy minimizing difference between sensitivity and specificity – übersetzt in die DIMDI-Syntax

#	Suchbegriffe	Ergebnis
1	ME60	22410514
2	(((CT D "thyroid ca"? OR UT="thyroid ca"? OR IT="thyroid ca"? OR SH="thyroid ca"?) OR (CTG D "thyroid ca"? OR UTG="thyroid ca"? OR ITG="thyroid ca"? OR SHG="thyroid ca"?)) OR FT=thyroid ca?) OR TI=thyroid ca?	42574
3	(((CT D "thyroid neopla"? OR UT="thyroid neopla"? OR IT="thyroid neopla"? OR SH="thyroid neopla"?) OR (CTG D "thyroid neopla"? OR UTG="thyroid neopla"? OR ITG="thyroid neopla"? OR SHG="thyroid neopla"?)) OR FT=thyroid neopla?) OR TI=thyroid neopla?	36240
4	(((CT D "struma maligna" OR UT="struma maligna" OR IT="struma maligna" OR SH="struma maligna") OR (CTG D "struma maligna" OR UTG="struma maligna" OR ITG="struma maligna" OR SHG="struma maligna")) OR FT=struma maligna) OR TI=struma maligna	115
5	(((CT D "schilddr"##"sen"#"ar"#"inom"? OR UT="schilddr"##"sen"#"ar"#"inom"? OR IT="schilddr"##"sen"#"ar"#"inom"? OR SH="schilddr"##"sen"#"ar"#"inom"?) OR (CTG D "schilddr"##"sen"#"ar"#"inom"? OR UTG="schilddr"##"sen"#"ar"#"inom"? OR ITG="schilddr"##"sen"#"ar"#"inom"? OR SHG="schilddr"##"sen"#"ar"#"inom"?)) OR FT=schilddr##sen#ar#inom?) OR TI=schilddr##sen#ar#inom?	36071
6	(((CT D "schilddr"##"senkrebs"? OR UT="schilddr"##"senkrebs"? OR IT="schilddr"##"senkrebs"? OR SH="schilddr"##"senkrebs"?) OR (CTG D "schilddr"##"senkrebs"? OR UTG="schilddr"##"senkrebs"? OR ITG="schilddr"##"senkrebs"? OR SHG="schilddr"##"senkrebs"?)) OR FT=schilddr##senkrebs?) OR TI=schilddr##senkrebs?	36058
7	2 OR 3 OR 4 OR 5 OR 6	42727

8	(((CT D "medull"##"r"? OR UT="medull"##"r"? OR IT="medull"##"r"? OR SH="medull"##"r"?) OR (CTG D "medull"##"r"? OR UTG="medull"##"r"? OR ITG="medull"##"r"? OR SHG="medull"##"r"?)) OR FT=medull##r?) OR TI=medull##r?	70659
9	7 AND 8	6176
10	(((CT D "mtc" OR UT="mtc" OR IT="mtc" OR SH="mtc") OR (CTG D "mtc" OR UTG="mtc" OR ITG="mtc" OR SHG="mtc")) OR FT=mtc) OR TI=mtc	2866
11	9 OR 10	7396
12	(((CT D "caprelsa" OR UT="caprelsa" OR IT="caprelsa" OR SH="caprelsa") OR (CTG D "caprelsa" OR UTG="caprelsa" OR ITG="caprelsa" OR SHG="caprelsa")) OR FT=caprelsa) OR TI=caprelsa	6
13	(((CT D "zactima" OR UT="zactima" OR IT="zactima" OR SH="zactima") OR (CTG D "zactima" OR UTG="zactima" OR ITG="zactima" OR SHG="zactima")) OR FT=zactima) OR TI=zactima	37
14	(((CT D "vandetanib" OR UT="vandetanib" OR IT="vandetanib" OR SH="vandetanib") OR (CTG D "vandetanib" OR UTG="vandetanib" OR ITG="vandetanib" OR SHG="vandetanib")) OR FT=vandetanib) OR TI=vandetanib	263
15	(((CT D "ZD6474" OR UT="ZD6474" OR IT="ZD6474" OR SH="ZD6474") OR (CTG D "ZD6474" OR UTG="ZD6474" OR ITG="ZD6474" OR SHG="ZD6474")) OR FT=ZD6474) OR TI=ZD6474	167
16	(((CT D "ZD 6474" OR UT="ZD 6474" OR IT="ZD 6474" OR SH="ZD 6474") OR (CTG D "ZD 6474" OR UTG="ZD 6474" OR ITG="ZD 6474" OR SHG="ZD 6474")) OR FT=ZD 6474) OR TI=ZD 6474	7
17	(((CT D "ZD-6474" OR UT="ZD-6474" OR IT="ZD-6474" OR SH="ZD-6474") OR (CTG D "ZD-6474" OR UTG="ZD-6474" OR ITG="ZD-6474" OR SHG="ZD-6474")) OR FT=ZD-6474) OR TI=ZD-6474	7
18	12 OR 13 OR 14 OR 15 OR 16 OR 17	388
19	11 AND 18	65
20	DT="RANDOMIZED CONTROLLED TRIAL"	336919
21	Randomized/(Ti;Ab;Ct;TE)	331013
22	CT=Randomized*	151899
23	Placebo/(Ti;Ab;Ct;TE)	146331
24	CT=placebo*	33481
25	20 OR 21 OR 22 OR 23 OR 24	629887
26	19 AND 25	19
27	26 AND LA=(ENGLISH; GERMAN) AND (pps=Mensch OR pps=Krebserkrankungen)	18

Datenbankname	EMBASE und EMBASE Alert
Suchoberfläche	DIMDI Smart Search
Datum der Suche	11.01.2013
Zeitsegment	1947 bzw. 2008 bis aktuell
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [111] - Strategy minimizing difference between sensitivity and specificity – übersetzt in die DIMDI-Syntax

#	Suchbegriffe	Ergebnis
1	EM47 EA08	26393397
2	((CT D "thyroid ca"? OR UT="thyroid ca"? OR IT="thyroid ca"? OR SH="thyroid ca"?) OR FT=thyroid ca?) OR TI=thyroid ca?	43996
3	((CT D "thyroid neopla"? OR UT="thyroid neopla"? OR IT="thyroid neopla"? OR SH="thyroid neopla"?) OR FT=thyroid neopla?) OR TI=thyroid neopla?	51995
4	((CT D "struma maligna" OR UT="struma maligna" OR IT="struma maligna" OR SH="struma maligna") OR FT=struma maligna) OR TI=struma maligna	174
5	((CT D "schilddr"##"sen"#"ar"#"inom"? OR UT="schilddr"##"sen"#"ar"#"inom"? OR IT="schilddr"##"sen"#"ar"#"inom"? OR SH="schilddr"##"sen"#"ar"#"inom"?) OR FT=schilddr##sen#ar#inom?) OR TI=schilddr##sen#ar#inom?	693
6	((CT D "schilddr"##"senkrebs"? OR UT="schilddr"##"senkrebs"? OR IT="schilddr"##"senkrebs"? OR SH="schilddr"##"senkrebs"?) OR FT=schilddr##senkrebs?) OR TI=schilddr##senkrebs?	22
7	2 OR 3 OR 4 OR 5 OR 6	58649
8	((CT D "medull"##"r"? OR UT="medull"##"r"? OR IT="medull"##"r"? OR SH="medull"##"r"?) OR FT=medull##r?) OR TI=medull##r?	118926
9	7 AND 8	8843
10	((CT D "mtc" OR UT="mtc" OR IT="mtc" OR SH="mtc") OR FT=mtc) OR TI=mtc	3848
11	9 OR 10	10893
12	((CT D "caprelsa" OR UT="caprelsa" OR IT="caprelsa" OR SH="caprelsa") OR FT=caprelsa) OR TI=caprelsa	2264
13	((CT D "zactima" OR UT="zactima" OR IT="zactima" OR SH="zactima") OR FT=zactima) OR TI=zactima	2267
14	((CT D "vandetanib" OR UT="vandetanib" OR IT="vandetanib" OR SH="vandetanib") OR FT=vandetanib) OR TI=vandetanib	2313

15	((CT D "ZD6474" OR UT="ZD6474" OR IT="ZD6474" OR SH="ZD6474") OR FT=ZD6474) OR TI=ZD6474	2279
16	((CT D "ZD 6474" OR UT="ZD 6474" OR IT="ZD 6474" OR SH="ZD 6474") OR FT=ZD 6474) OR TI=ZD 6474	2264
17	((CT D "ZD-6474" OR UT="ZD-6474" OR IT="ZD-6474" OR SH="ZD-6474") OR FT=ZD-6474) OR TI=ZD-6474	1193
18	12 OR 13 OR 14 OR 15 OR 16 OR 17	2323
19	11 AND 18	247
20	random?/(ti;ab)	691469
21	placebo?/(ti;ab;te;ct)	283601
22	double-blind?/(ti;ab)	126501
23	20 OR 21 OR 22	888969
24	19 AND 23	80
25	24 AND LA=(ENGLISH; GERMAN) AND (pps=Mensch OR pps=Krebserkrankungen)	76

Verlauf und Ergebnisse der weiteren Selektion werden in Abschnitt 4.3.1.1.2 ausführlich dargestellt.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

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

Um sicherzustellen, dass alle verfügbaren Studiendaten für die Bewertung berücksichtigt worden sind, wurde am 05. Januar 2012 eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien in den gängigen medizinischen Studienregistern durchgeführt. Es wurde gemäß den Vorgaben der Verfahrensordnung eine Studienregistersuche in den Datenbanken clinicaltrials.gov, clinicalstudyresults.org sowie über das *International Clinical Trials Registry Platform Search Portal* (ICTRP *Search Portal*, Suchportal der WHO) durchgeführt.

Alle in den einzelnen Studienregistern identifizierten Studien wurden gemäß den in Abschnitt 4.2.1 prädefinierten Kriterien selektiert, um für die Zusatznutzenbewertung von Vandetanib relevante Studien zu selektieren. Relevante Studien sind in Abschnitt 4.3.1.1.3 dokumentiert und wurden für den der Bewertung von Vandetanib zugrundeliegenden Studienpool aufgenommen.

Studienregister	*International Clinical Trials Registry Platform Search Portal*
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07. Januar 2013
Suchstrategie	metastatic medullary thyroid cancer OR medullary thyroid cancer OR thyroid cancer OR thyroid carcinoma OR thyroid neoplasm OR thyroid nodule OR thyroid disease [condition] AND vandetanib OR ZD 6474 OR ZD6474 OR ZD-6474 OR zactima OR caprelsa [intervention]
Treffer	22

In der Datenbank *International Clinical Trials Registry Platform Search Portal* konnten 22 Studien identifiziert werden (Tabelle 4-182).

Tabelle 4-182 (Anhang): Im *International Clinical Trials Registry Platform Search Portal* identifizierte Studien

Recruitment status	***Main ID***	***Public Title***	***Date of Registra-tion***
Not recruiting	NCT01661179	Evaluate the Safety and Tolerability of Vandetanib in Japanese Patients With Medullary Thyroid Carcinoma	17/07/2012
Recruiting	NCT01496313	To Compare The Effects Of Two Doses Of Vandetanib In Patients With Advanced Medullary Thyroid Cancer	27.12.2011
Authorised	EUCTR2011-000661-12-GB	A study of vandetanib in combination with whole brain radiotherapy in the treatment of metastatic melanoma which has spread to the brain.	14/07/2011
Authorised	EUCTR2010-021951-26-GB	Clinical trial comparing combination chemotherapy(gemcitabine and vandetanib) therapy with gemcitabine therapy alone in advanced pancreas cancer	22/02/2011
Authorised	EUCTR2010-023428-26-DE	Study to determine if contacting patients with MTC more frequently results in earlier detection and treatment of signs and symptoms of AEs and thus a decrease in the percentage of time patients experience AEs during first 12 months on vandetanib treatment	22/11/2010
Not recruiting	ISRCTN68146831	Vandetanib in non-cisplatin fit patients with urothelial cancers	20.07.2009
Not recruiting	NCT00923247	A Targeted Phase I/II Trial of ZD6474 (Vandetanib; ZACTIMA) Plus the Proteasome Inhibitor, Bortezomib (Velcade(Registered Trademark)), in Adults With Solid Tumors With a Focus on Hereditary or Sporadic, Locally Advanced or Metastatic Medullary Thyroi...	17.06.2009
Not recruiting	NCT00811369	Trial to Evaluate the Therapeutic Benefit of Fulvestrant in Combination With ZACTIMA in Postmenopausal Women With Bone Predominant, Hormone Receptor Positive Metastatic Breast Cancer	15.12.2008
Not recruiting	NCT00537095	Efficacy and Safety of Zactima™ in Patients With Metastatic Papillary or Follicular Thyroid Cancer	27.09.2007
Not recruiting	NCT00507091	Phase I Irinotecan, 5-Fluorouracil and Leucovorin Combination	10.07.2007
Not recruiting	NCT00498797	E3-Hormone Refractory Prostrate Cancer Taxotere Combination	09.07.2007

Not recruiting	NCT00496665	ZD6474 (Zactima) and Metronomic Chemotherapy in Advanced Breast Cancer	03.07.2007
Authorised	EUCTR2007-001890-27-FR	A Randomized, Double Blind, placebo-controlled Phase II, Multi-Centre Study to Assess the Efficacy and Safety of Zactima™ in Patients with advanced or metastatic papillary or follicular Thyroid Carcinoma failing or unsuitable for Radioiodine therapy	10/05/2007
Authorised	EUCTR2007-001891-35-FR	A randomized, double-blind phase II trial to assess the efficacy and safety of bicalutamide (Casodex®) associated to ZD6474 (Zactima™) or to placebo in patients with castration-refractory metastatic prostate cancer without any clinical symptom related to disease progression.	10/05/2007
Authorised	EUCTR2006-005022-23-FR	A Phase II, Double-Blind, Placebo Controlled, Randomised Study To Assess The Efficacy And Safety Of 2 Doses Of ZACTIMA (ZD6474) In Combination With FOLFOX vs FOLFOX Alone For The Treatment Of Colorectal Cancer In Patients Who Have Failed Therapy With An Irinotecan And Fluoropyrimidine Containing Regimen.	12/01/2007
Not recruiting	NCT00410761	An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer	06.12.2006
Not recruiting	EUCTR2006-003695-35-SE	A Phase III, Randomized, Double-blinded, Parallel Group, Multi-centre Study to Assess the Efficacy and Safety of ZD6474 (ZACTIMA™) in Combination with Pemetrexed (Alimta®) versus Pemetrexed alone in Patients with Locally-Advanced or Metastatic (stage IIIB or IV) Non-Small Cell Lung Cancer (NSCLC) after Failure of 1st Line Anti-cancer Therapy	13/11/2006
Not recruiting	EUCTR2005-005077-29-DE	A study to learn how hereditary and sporadic thyroid cancer patients, treated with ZD6474 react to the drug, what happens to ZD6474 in the human body, about the side effects of ZD6474 and if ZD6474 can decrease or prevent tumour growth.	18/08/2006
Not recruiting	NCT00358956	A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer	28.07.2006
Authorised	EUCTR2006-001354-28-NL	Not aplicable	20/04/2006
Not recruiting	NCT00098345	Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer	07.12.2004
Not recruiting	NCT00047840	This Study is to Assess the Efficacy and Safety of ZD6474 in Subjects With Non-Small Cell Lung Cancer.	18.10.2002

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	07. Januar 2013
Suchstrategie	metastatic medullary thyroid cancer OR medullary thyroid cancer OR thyroid cancer OR thyroid carcinoma OR thyroid neoplasm OR thyroid nodule OR thyroid disease [condition] AND vandetanib OR ZD 6474 OR ZD6474 OR ZD-6474 OR zactima [intervention]
Treffer	11

In der Datenbank clinicaltrials.gov konnten elf Studien identifiziert werden (Tabelle 4-183).

Tabelle 4-183 (Anhang): In clinicaltrials.gov identifizierte Studien

Status	**Studie**
Active, not recruiting *Has Results*	An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer Condition: Thyroid Cancer Intervention: Drug: ZD6474 (Vandetanib) NCT ID: NCT00410761 Other IDs: D4200C00058
Recruiting *No Results Available*	Evaluate the Safety and Tolerability of Vandetanib in Japanese Patients With Medullary Thyroid Carcinoma Condition: Unresectable Locally Advanced or Metastatic, Medullary Thyroid Carcinoma Intervention: Drug: Vandetanib 300mg NCT ID: NCT01661179 Other IDs: D4200C00098
Active, not recruiting *Has Results*	A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer Condition: Thyroid Cancer Intervention: Drug: ZD6474 (vandetanib) NCT ID: NCT00358956 Other IDs: D4200C00068
Recruiting *No Results Available*	To Compare The Effects Of Two Doses Of Vandetanib In Patients With Advanced Medullary Thyroid Cancer Condition: Thyroid Cancer Intervention: Drug: 300mg vandetanib; Drug: 150mg vandetanib NCT ID: NCT01496313 Other IDs: D4200C00097, 2011-004701-24
Active, not recruiting	Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer

Status	**Studie**
Has Results	Condition: Thyroid Cancer Intervention: Drug: ZD6474 (vandetanib) NCT ID: NCT00098345 Other IDs: D4200C00008
Active, not recruiting *No Results Available*	Study to Determine if Contacting Patients With MTC More Frequently Results in Earlier Detection and Treatment of Signs and Symptoms of AEs and Thus a Decrease in the Percentage of Time Patients Experience AEs During First 12 Months on Vandetanib Treatment Conditions: Locally Advanced or Metastatic Medullary Thyroid Cancer; Medullary Thyroid Cancer Intervention: Behavioral: Patient outreach NCT ID: NCT01298323 Other IDs: D4200C00088, 2010-023428-26
Active, not recruiting *Has Results*	Efficacy and Safety of Vandetanib (ZD6474) in Patients With Metastatic Papillary or Follicular Thyroid Cancer Condition: Thyroid Neoplasms Interventions: Drug: Vandetanib; Other: Placebo NCT ID: NCT00537095 Other IDs: D4200C00079
Active, not recruiting *No Results Available*	A Targeted Phase I/II Trial of ZD6474 (Vandetanib; ZACTIMA) Plus the Proteasome Inhibitor, Bortezomib (Velcade(Registered Trademark)), in Adults With Solid Tumors With a Focus on Hereditary or Sporadic, Locally Advanced or Metastatic Medullary Thyroid Ca (MTC) Conditions: Breast Cancer; Colon Cancer; Prostate Cancer; Lung Cancer; Adrenal Cancer; Renal Cancer; Gastric Cancer; Ovarian Cancer; Cervical Cancer; Thyroid Cancer; Sarcoma Interventions: Drug: VELCADE, PS-341 (Bortezomib, NSC #681239); Drug: ZACTIMA (Vandetanib) (ZD6474) NCT ID: NCT00923247 Other IDs: 090089, 09-C-0089
Recruiting *No Results Available*	Vandetanib to Treat Children and Adolescents With Medullary Thyroid Cancer Conditions: Multiple Endocrine Neoplasia Type 2A; Multiple Endocrine Neoplasia Type 2B Intervention: Drug: ZACTIMA (Vandetanib) NCT ID: NCT00514046 Other IDs: 070189, 07-C-0189
Recruiting *No Results Available*	Vandetanib Risk Minimisation Effectiveness Conditions: Medullary Thyroid Cancer Intervention: NCT ID: NCT01757470
	Other IDs: D4200L00059
Withdrawn	Vandetanib and Docetaxel in Treating Patients With Advanced Solid Tumors

Status	**Studie**	
No Results Available	Conditions:	Breast Cancer; Head and Neck Cancer; Lung Cancer; Prostate Cancer; Unspecified Adult Solid Tumor, Protocol Specific
	Interventions:	Drug: docetaxel; Drug: vandetanib; Genetic: proteomic profiling; Other: laboratory biomarker analysis; Other: pharmacological study
	NCT ID:	NCT00937417
	Other IDs:	CDR0000647578, SWOG-S0716

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund

Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds auf.

Es wurden keine Studien anhand der Sichtung im Volltext ausgeschlossen. Sechs anhand des Abstracts nicht eindeutig auszuschließende Publikationen aus der Literaturrecherche wurden nach Sichtung des Volltextes ausgeschlossen; die nachfolgende Tabelle nennt die Ausschlussgründe:

Tabelle 4-184 (Anhang): im Volltext gesichtete und ausgeschlossene Publikationen mit Ausschlussgrund

Publikation	**Ausschlussgrund**
Bible, K. C. (2012). "Treating advanced radioresistant differentiated thyroid cancer." The Lancet Oncology 13(9): 854-855.	**A3**: Review, keine Primärpublikation
Deshpande, H., V. Marler, et al. (2011a). "Clinical utility of vandetanib in the treatment of patients with advanced medullary thyroid cancer." OncoTargets and Therapy 4: 209-215	**A3**: Review, keine Primärpublikation
Deshpande, H., S. Roman, et al. (2011b). "Vandetanib (ZD6474) in the treatment of medullary thyroid cancer." Clinical Medicine Insights: Oncology 5: 213-221.	**A3**: Review, keine Primärpublikation
Frampton, J. E. (2012). "Vandetanib: In medullary thyroid cancer." Drugs 72(10): 1423-1436.	**A3**: keine Primärpublikation
Hyer, R. (2010). "Vandetanib slows metastatic medullary thyroid cancer." Oncology Report(JULY-AUGUST): 28.	**A3**: Review, keine Primärpublikation
Robinson, B. G., L. Paz-Ares, et al. (2010). "Vandetanib (100 mg) in patients with locally advanced or metastatic hereditary medullary thyroid cancer." The Journal of clinical endocrinology and metabolism 95(6): 2664-2671.	**A3, A4**: keine randomisierte, kontrollierte Studie

Anhang 4-D Liste der abgebrochenen Studien

Die Durchführung einer bibliografischen Literaturrecherche war nicht erforderlich, da die Zulassung für das zu bewertende Arzneimittel Vandetanib weniger als 12 Monate vor dem Zeitpunkt der Dossier-Einreichung erteilt wurde.

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-185 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	**Studienkategorie**			**verfügbare Datenquellen[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)						

a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.

b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

Der in Abschnitt 4.3.1.1 identifizierte resultierende Studienpool enthält nur eine Studie (siehe Tabelle 4-6); diese wurde nicht abgebrochen.

Anhang 4-E: Liste der laufenden Studien

Benennen Sie in der nachfolgenden Tabelle die laufenden Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-186 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	**Studienkategorie**			**verfügbare Datenquellen[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						
D4200C00058	ja	ja	nein	ja [22]	ja [91] (NCT00410761)	ja [92]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						

a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.
b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

Der in Abschnitt 4.3.1.1 identifizierte resultierende Studienpool umfasst lediglich eine für die Nutzenbewertung relevante Studie (siehe Tabelle 4-6); die einzige andere im Zuge der durchgeführten Recherchen identifizierte RCT im Abnwendungsgebiet, Studie D4200C00097, wurde begründet ausgeschlossen (Tabelle 4-3).

Anhang 4-F: 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 4-187 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 4-187 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-187 (Anhang): Studiendesign und -methodik für Studie D4200C00058

Item[a]	Charakteristikum	Studieninformation
-	**Studienziel**	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Bewertung der Wirksamkeit von Vandetanib versus Placebo
-	**Methoden**	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	International, randomisiert, doppelblind, multizentrisch, placebo-kontrolliert; Zuteilungsverhältnis 2:1; Randomisierung wurde nach Standort in 3-er-Blöcken stratifiziert.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Amendment 2 (30 May 2007): • Anpassung der Einschlusskriterien zur Klarstellung, dass für eine Qualifizierung als „messbare Läsion“ Messungen zur Involvierung von Lymphknoten, zur Summe der Durchmesser und zur Größe der Zielläsionen durchgeführt werden. • Anpassung der Ausschlusskriterien: erforderliches Serumkreatinin zwischen ≤50 mL/min und ≤30 mL/min (Änderung der Kreatinin-Clearance auf Grundlage einer AstraZeneca-Studie) Amendment 3 (15 May 2008): • Re-Definition der Studie als Phase III (Zulassungsstudie) • Berechnung der DCR als Prozentsatz der Patienten mit komplettem Ansprechen (CR), partiellem Ansprechen (PR), oder stabilem Krankheitsbild ≥24 Wochen. (Konsistenz mit anderen AstraZeneca-Studien) Amendment 5[b] (18 May 2009): • Änderung der Studienziele bezüglich patientenberichteter Zielgrößen (PRO): TWP wird als einzige PRO als sekundärer Endpunkt der Studie definiert, alle weiteren PRO als explorative Endpunkte • Änderung der Definition des Endpunkts Gewichtszunahme vom sekundären zum explorativen Endpunkt • Streichung der co-primären Analysepopulation der Patienten mit bekannter RET-Mutation aus der Studie (Status konnte bei 41% der Patienten nicht ermittelt werden) • Festlegung einer Kombination aus

Item[a]	Charakteristikum	Studieninformation
		zwei Methoden zur Determinierung des RET-Status: ARMS Assay zur spezifischen Detektion der häufigsten RET-Mutation (M918T) in sporadischem MTC, und direkte DNA-Sequenzierung nach PCR-basierter Amplifizierung des RET-Gens (Exons 10, 11, 13-16) aus Tumor-DNA (zusätzliche Klärung der Verifizierung des RET-Status) Amendment 6 (13 January 2010): • Anpassung des Studienplans um Prüfärzte mit der Möglichkeit zu versehen, Patienten, welche auf verblindeter, randomisierter Therapie verblieben, zu enblinden. Als Konsequenz der Entblindung konnten die Patienten nicht weiter auf verblindeter Therapie verbleiben; sie mussten entweder im Open-Label-Teil der Studie weiterbehandelt werden oder die verblindete Therapie beenden und sich der Nachverfolgung bezüglich des Überlebens unterziehen. Nicht entblindete Patienten mussten so weiterhin in der Studie verbleiben wie im Studienplan vorgesehen. (Amendment auf Basis der Resultate der primären Analyse der Studie: durch die Möglichkeit der Entblindung wurde allen Patienten auf Placebo die Option verschafft, Open Label Vandetanib zu erhalten.)
4	Probanden / Patienten	331 randomisierte Patienten mit nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: • Bereitstellung einer schriftlichen Einverständniserklärung • Weiblich oder männlich im Alter von 18 Jahren und älter • Vorherig bestätigte histologische Diagnose von inoperablem, lokal fortgeschrittenem oder metastasiertem hereditärem oder sporadischem medullärem Schilddrüsenkarzinom. Nachweis musste in der Patientenakte bereitgestellt sein. • Lebenserwartung von 12 Wochen oder länger (*Diese Patienten wurden als am ehesten dazu fähig erachtet, die Studienprozeduren und die Behandlung zu vertragen*)

Item[a]	Charakteristikum	Studieninformation
		• WHO PS 0-2 (*Diese Patienten wurden als am ehesten dazu fähig erachtet, die Studienprozeduren und die Behandlung zu vertragen*) • Fähigkeit, die Studienmedikation zu schlucken • Vorhandensein eines messbaren Tumors (nach vorab spezifizierten Kriterien) definiert (a) durch eine Einzelläsion, die ≥2cm misst oder (b), durch folgende Kriterien bei mehreren Läsionen: Eine Technik mit ≤5 mm- Sektionen: Summe der Durchmesser ≥2 cm (keine Zielläsionen, die <1 cm messen und keine Lymphknoten, die <1,5 cm messen) ODER Eine Technik mit >5 mm Sektionen: Summe der Durchmesser ≥4 cm (keine Zielläsion, die <2 cm misst) • Calcitonin *(CTN, Calcitonin)* ≥500 pg/mL (gebräuchliche Einheit) oder ≥146,3 pmol/l (internationale Standardeinheit) • Alle Patienten (mit Ausnahme von denjenigen mit hereditärem MTC die eine nachgewiesene Keimbahn-RET-Mutation hatten) mussten eine passend gelagerte Tumor-Sammelprobe einreichen. Falls diese Probe nicht früher als 2 Wochen vor der Randomisierung verfügbar war, wurde eine frische Tumorprobe entnommen. Die Tumorprobe musste von der Forschungseinrichtung eingeholt werden und vor der Randomisierung zu ihrem Bestimmungsort gebracht werden • Negativer Schwangerschaftstest für weibliche gebärfähige Patienten Ausschlusskriterien: • Hirnmetastasen oder Rückenmarkskompression, sofern nicht mindestens 4 Wochen vor der ersten Dosis behandelt und seit 10 Tagen stabil ohne Steroidbehandlung (*Dieses Ausschlusskriterium fand aufgrund des potenziell erhöhten Risikos intrakranieller Hämorrhagie bei Patienten mit unbehandelten oder instabilen Hirnmetastasen, die mit VEGF- Inhibitoren behandelt wurden, Anwendung. Außerdem wurden so Patienten ausgeschlossen, die möglicherweise unfähig gewesen*

Item[a]	Charakteristikum	Studieninformation
		wären, in die Studienbewertungen und –prozeduren einzuwilligen) • Jegliche Begleitmedikationen, die das Zeitintervall zwischen dem Start der Q-Welle und dem Ende der T-Welle (korrigiert gegen die Herzrate)**(**QTc) beeinflussen oder die CYP3A4-Funktion auslösen können (mit Ausnahme von Somatostatin oder Somatostatin Analogon) und/oder alle verbotenen Medikationen, die in dem revidierten Studienprotokoll (*CSP, Clinical Study Protocol*), Appendix E (siehe Anhang 12.1.1., Studienbericht), referenziert werden • Bedeutende Operation innerhalb 4 Wochen vor Randomisierung (*Dieses Ausschlusskriterium fand Anwendung, um das potenzielle Risiko der Wundheilungstoxizität für Patienten zu reduzieren*) • Die letzte Dosis der vorhergehenden Chemotherapie wurde weniger als 4 Wochen vor der Randomisierung erhalten (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziellen Risikos für überschneidende Toxizität und potenzielle Schwierigkeit der Bewertung der Wirksamkeit von Vandetanib. Um den Effekt von früherer palliativer Strahlentherapie auf Bewertungskriterien zum Ansprechen bei soliden Tumoren (RECIST, Response Evaluation Criteria in Solid Tumors) zu minimieren, wurden Läsionen, die weniger als 12 Wochen vor den Basis-Scans bestrahlt wurden, davon ausgeschlossen, als Zielläsionen gewählt zu werden. Diese Läsionen konnten nur als Nicht-Zielläsionen gewählt werden. Des Weiteren wurden Läsionen, die mindestens 12 Wochen vor den Basis-Scans bestrahlt wurden, nicht als Zielläsionen gewählt, es sei denn keine anderen passenden Läsionen waren verfügbar.*) • Strahlentherapie innerhalb der letzten 4 Wochen vor der Randomisierung (mit der Ausnahme von palliativer Strahlentherapie) (*Dieses Ausschlusskriterium fand Anwendung, um ein potenzielles Risiko von überschneidender Toxizität durch Strahlentherapie zu*

Item[a]	Charakteristikum	Studieninformation
		reduzieren) • Serum Bilirubin >1,5 x Obergrenze des Referenzbereichs (ULRR, *Upper Limit of Reference Range)* (*Dies dient als generelles Sicherheits-Ausschlusskriterium*) • Kreatinin-*Clearance* <30 mL/min (bestimmt durch die Cockcroft-Gault Formel) (*Dieses Aussschlusskriterium fand Anwendung aufgrund des Möglichkeit für eine erhöhte Belastung durch Vandetanib bei Patienten mit schwerem Nierenversagen*) • Kalium <4,0 mmol/L trotz Zusatzzufuhr oder über der Obergrenze der CTCAE Grad 1. Magnesium unter dem Normbereich trotz Zusatzzufuhr oder über der Obergrenze der CTCAE Grad 1. Serum Calcium über der Obergrenze der CTCAE Grad 1. In Fällen, in denen das Serum Calcium unter dem Normbereich lag, musste das für Albumin korrigierte Calcium eingeholt werden und an die Stelle des gemessenen Serumwerts gesetzt werden. Der Ausschluss musste dann darauf gründen, ob das Albumin-korrigierte Calcium unter die Normgrenze fällt. Korrigiertes Calcium= Ca + 0,8 X (4-serum Albumin) (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT Verlängerung entstehen*) • Alanin-Aminotransferase, -Aspartat-Aminotransferase oder Alkalische Phosphatase >2,5 x ULRR, oder >5,0 x ULRR, wenn der Prüfarzt dieses mit Lebermetastasen zusammenhängend beurteilt. (*Dies dient als generelles Sicherheits-Ausschlusskriterium*) • Signifikantes kardiales Ereignis (z. B. Myokardinfarkt), Vena-cava-superior-Syndrom, NYHA Klassifikation der Herzkrankheit ≥2, innerhalb von 12 Wochen vor der Randomisierung oder Vorhandensein einer Herzerkrankung, die, laut Meinung des Prüfarztes, das Risiko ventrikulärer Arrhythmien erhöht (*Dieses Ausschlusskriterium fand*

Item[a]	Charakteristikum	Studieninformation
		Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien, die durch Vandetanib-assoziierte QT-Verlängerung entstehen) • Vorgeschichte von Arrhythmien (multifokale ventrikuläre Extrasystolen, Bigeminus, Trigeminus, ventrikuläre Tachykardie), die symptomatisch waren oder Behandlung erforderten (CTCAE Grad 3), symptomatisches oder unkontrolliertes Vorhofflimmern trotz Behandlung oder asymptomatisch fortwährende ventrikuläre Tachykardie. Patienten mit Vorhofflimmern, welches durch Medikation kontrolliert wurde, wurden zugelassen. (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien, die durch Vandetanib-assoziierte QTVerlängerung entstehen*) • Angeborenes Long-QT-Syndrom oder Verwandter ersten Grades mit unerklärtem plötzlichen Tod mit unter 40 Jahren (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen*) • QT-Verlängerung mit anderen Medikationen die eine Einstellung der Medikation erforderten (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen*) • Vorhandensein eines Linksschenkelblocks (*Dieses Ausschlusskriterium fand Anwendung aufgrund von Schwierigkeiten, das QT-Intervall bei Patienten mit Linksschenkelblock akkurat zu messen.*) • QTc mit Bazett's Formel unmessbar oder ≥480 ms beim *Screening*-EKG. Zu beachten: Wenn ein Patient ein QTc-Intervall ≥480 ms beim *Screening*-EKG hatte, konnte das *Screening*-EKG 2 Mal wiederholt werden (mindestens 24 Stunden

Item[a]	Charakteristikum	Studieninformation
		verschoben) für insgesamt 3 EKGs. (Die durchschnittliche QTc von den 3 *Screening*-EKGs musste <480 ms liegen, damit der Patient für die Studie qualifiziert ist). Wenn ein Patient vor dem Studieneintritt eine Medikation mit möglicher Assoziation zu Torsades de Pointes (siehe Appendix E, Tabelle 2 des revidierten CSP [Appendix 12.1.1.]) erhielt und die Medikation konnte vor der Studienbehandlung nicht eingestellt werden, musste das *Screening*-QTc < 460 ms sein. (*Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen*) • Hypertonie, die nicht durch medikamentöse Therapie kontrolliert wird (systolischer Blutdruck >160 Millimeter Quecksilbersäule (mmHg) oder diastolischer Blutdruck >100 mmHg) (*Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeitvon Vandetanib, den Blutdruck zu erhöhen)* • Vorherige oder aktuelle Malignität anderer Histologien innerhalb der letzten 5 Jahre, mit Ausnahme von Tumoren assoziiert mit MEN2a und MEN2b, Carcinoma in situ der Zervix oder adäquat behandeltes Basalzellen- oder Plattenepithelkarzinom der Haut (*Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit einer vorherigen oder aktuellen Malignität, die Studienendpunktezu beeinflussen*) • Jegliche ungeklärte chronische Toxizität durch eine vorherige Anti-Krebs-Therapie, größer als CTCAE Grad 2 (*Dies dient als generelles Sicherheitsausschlusskriterium*) • Teilnahme an einer klinischen Studie und /oder Erhalt eines investigativen Medikaments während der letzten 30 Tage (Teilnahme an der Überleben-*Follow-up*- Periode einer Studie war kein Ausschlusskriterium) (Dieses *Sicherheitsausschlusskriterium fand Anwendung, um potenzielle Interaktionen zwischen investigativen Wirkstoffen zu*

Item[a]	Charakteristikum	Studieninformation
		verhindern und um Komplikationen bei der Interpretation der Ergebnisse zu vermeiden) • Vorherige Exposition gegenüber Vandetanib (*Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit, dass Vandetanib bei Patienten, die eine vorherige Behandlung mit dem Medikament erhalten haben, weniger wirksam sein könnte*) • Aktuell schwanger oder stillend (*Diese Ausschlusskriterium fand Anwendung aufgrund unbekannter Risiken, die während einer Schwangerschaft oder Stillzeit mit einer Vandetanib-Behandlung assoziiert sein könnten*) • Mitwirkung an der Planung und Durchführung der Studie (trifft auf AstraZeneca Mitarbeiter und den Mitarbeiterstab der Forschungseinrichtung zu) (*Generelles prozedurales Ausschlusskriterium fand Anwendung, um Interessenskonflikte zu vermeiden*) • Vorherige Randomisierung oder Behandlung in der gegenwärtigen Studie Restriktionen: • Aufgrund des experimentellen Charakters von Vandetanib mussten Frauen 1 Jahr postmenopausal, chirurgisch unfruchtbar sein oder eine akzeptable Methode der Kontrazeption (definiert als mechanische Verhütungsmethode in Verbindung mit einem Spermizid) benutzen, um für die Dauer der Studie (von dem Moment an, in dem sie die Einverständniserklärung unterschrieben haben und bis 2 Monate nach der letzten Dosis) unter Vandetanib oder Placebo eine Schwangerschaft zu vermeiden. Des Weiteren waren orale Kontrazeptiva, zugelassene kontrazeptive Implantate, injizierbarer Langzeitempfängnisschutz, die Spirale oder Tubenligatur erlaubt. Orale Kontrazeption alleine war nicht akzeptabel; zusätzliche mechanische Verhütungsmethoden in Verbindung mit Spermizid

Item[a]	Charakteristikum	Studieninformation
		mussten benutzt werden. • Männliche Patienten mussten chirurgisch unfruchtbar sein oder eine akzeptable Methode der Kontrazeption (definiert als mechanische Verhütungsmethode in Verbindung mit einem Spermizid) verwenden, um für die Dauer der Studie (von dem Moment an, in dem sie die Einverständniserklärung unterschrieben haben und bis 2 Monate nach der letzten Dosis) unter Vandetanib oder Placebo, eine Schwangerschaft des Partners zu vermeiden. • Patienten, die Blutspender waren, durften während der Studie und in den folgenden 12 Wochen nach der letzten Dosis des Studienwirkstoffs kein Blut spenden. • Jegliche Begleitmedikationen, die die QTc hätten beeinflussen oder die CYP3A4-Funktion hätten auslösen können (siehe revidierter CSP, [Appendix 12.1.1]) • Der Gebrauch von Somatostatin (oder Somatostatin Analogon) vor Behandlungsbeginn war zur Kontrolle von Diarrhö erlaubt; allerdings durfte die Dosis nicht erhöht werden sobald der Patient mit dem Studienmedikament begonnen hat. Dosisreduktion war erlaubt.
4b	Studienorganisation und Ort der Studiendurchführung	63 Studienzentren in 24 Ländern: Australien, Belgien, Brasilien, Dänemark, Deutschland, Frankreich, Indien, Italien, Kanada, Korea, Niederlande, Mexico, Österreich, Polen, Portugal, Rumänien, Russland, Serbien, Spanien, Schweden, Schweiz, Tschechien, Ungarn und in den USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Einmal täglich, verblindet, orale Einnahme einer 300 mg Tablette Vandetanib oder einer 300 mg Placebo-Tablette bis eine Progression der Erkrankung eintritt; ab Progression der Erkrankung werden Patienten aus der verblindeten Studie entlassen und haben die Option, eine offene Behandlung mit Vandetanib zu beginnen; bei Auftreten von Grad 3 oder höheren unerwünschten Ereignissen: Möglichkeit der Dosisreduzierung auf 200 mg Vandetanib pro Tag und danach auf 100 mg pro Tag.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden	Im Folgenden steht nach dem Zielkriterium fett gedruckt in Klammern, um welche Art von Zielkriterium es sich handelt.

Item[a]	Charakteristikum	Studieninformation
	(z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primäres Zielkriterium: Nachweis einer Verbesserung des PFS mit Vandetanib verglichen mit Placebo bei Patienten mit nichtresektablem, lokal fortgeschrittenem oder metastasiertem Schilddrüsenkarzinom (**Wirksamkeit**) Sekundäre Zielkriterien: 1. Nachweis einer Verbesserung der objektiven *Ansprechrate* **(**ORR, *Objective Response Rate)*, Krankheitskontrollrate (DCR, *Disease Control Rate*) und Dauer des Ansprechens **(**DOR, *Duration of Response)* mit Vandetanib verglichen mit Placebo (**Wirksamkeit**) 2. Nachweis einer Verbesserung des OS bei Patienten mit MTC die mit Vandetanib behandelt wurden verglichen mit Placebo (**Wirksamkeit**) 3. Nachweis einer Verbesserung des biochemischen Ansprechens mit Vandetanib verglichen mit Placebo (gemessen durch CTN und Carcinoembryonales Antigen (CEA)) (**Wirksamkeit**) 4. Nachweis einer Verzögerung der TWP unter Patienten mit MTC nach Behandlung mit Vandetanib verglichen mit Placebo (**Patienten berichteter Endpunkt** *(***PRO**, *Patient Reported Outcome))* 5. Bestimmung der Pharmakokinetik (PK) von Vandetanib in dieser Patientenpopulation und Untersuchung jeglichen Einflusses von Demografie und Pathophysiologie des Patienten auf die PK (**PK**) 6. Bewertung der Beziehung zwischen PK und QTc, der Sicherheit, der Wirksamkeit und Biomarkern (**PK/ Pharmakodynamik**) 7. Bestimmung der Sicherheit und Verträglichkeit der Vandetanib-Behandlung bei MTC Patienten (**Sicherheit**)

Item[a]	Charakteristikum	Studieninformation
		8. Bestimmung des mutationellen Status des RET-Protoonkogens in der DNA, entnommen aus Tumorproben (**Biomarker**) Explorative Zielkriterien: 1. Untersuchung des Behandlungseffekts von Vandetanib verglichen mit Placebo auf Diarrhö bei Patienten mit MTC. (**PRO**) 2. Untersuchung von Änderungen im Plasma-VEGF, VEGFR-2 und Basis-Fibroblastenwachstumsfaktor (bFGF) -Level bei Patienten, die mit Vandetanib behandelt werden verglichen mit Placebo und ihre Beziehung zur Wirksamkeit (**Biomarker**) 3. Untersuchung von Änderungen in Serum-Proteinprofilen bei Patienten, die mit Vandetanib behandelt werden verglichen mit Placebo und deren Beziehung zur Wirksamkeit und Progression der Erkrankung (**Biomarker**) 4. Messung des EGFR-Expressionslevels im Tumorgewebe bei Patienten, die mit Vandetanib behandelt werden verglichen mit Placebo und Bestimmung der Beziehung zwischen Expressionslevel und Wirksamkeit (**Biomarker**) 5. Untersuchung von Änderungen bei Tumor-Biomarkern der Inhibition von RET, VEGFR und EGFR Signalbahnen (**Biomarker**) 6. Nachweis einer Verzögerung des TWP bei Patienten mit MTC, die an der *Baseline* keine Schmerzen haben, definiert durch das Benötigen von <10 mg/Tag Morphinsulfat oder Äquivalentem, nach der Behandlung mit Vandetanib verglichen mit Placebo (**PRO**) 7. Nachweis einer Verzögerung des TWP bei Patienten mit MTC, die an der *Baseline* Schmerzen haben, definiert durch das Benötigen von ≥10 mg/Tag Morphinsulfat oder Äquivalentem, nach der

Item[a]	Charakteristikum	Studieninformation
		Behandlung mit Vandetanib verglichen mit Placebo (**PRO**) 8. Nachweis einer Reduktion des Gebrauchs von opioidanalgetischer Medikation bei Patienten mit MTC, die an der *Baseline* Schmerzen haben, definiert durch das Benötigen von ≥10 mg/Tag Morphinsulfat oder Äquivalentem, nach der Behandlung mit Vandetanib verglichen mit Placebo (**PRO**) 9. Nachweis einer Gewichtssteigerung bei Patienten mit MTC die mit Vandetanib behandelt werden verglichen mit Placebo (**Wirksamkeit**) 10. Nachweis einer Verzögerung der Zeit bis zur Verschlechterung des WHO PS **(**TDPS, *Time to Decline in Patient WHO Performance Status)* bei Patienten die mit Vandetanib behandelt werden verglichen mit Placebo (**Wirksamkeit**) 11. Untersuchung des Effekts von Vandetanib, -verglichen mit Placebo, auf die Lebensqualität **(**QoL, *Quality of life)* des Patienten, gemessen durch die *Functional Assessment of Cancer Therapy General Scale* (FACT-G) (**PRO**) • Radiologische Auswertung alle 12 Wochen mittels modifiziertem RECIST. Patienten wurden bewertet bis objektive Progression eintrat. • Alle medizinischen Bilder wurden durch die Einrichtung und ein zentral ernanntes Auftragsforschungsinstitut (BioClinica [ehemals Bio-Imaging Technologies, Inc.], Newtown, PA USA) bewertet. Für weitere Informationen, siehe Appendix F des revidierten CSP (Appendix 12.1.1) • Alle Patienten mit sporadischem MTC mussten eine passend gelagerte Tumorprobe vor der Randomisierung abgeben. Die Resultate waren unbekannt bis die Patienten randomisiert wurden. Wenn bis 2 Wochen vor der Randomisierung keine Tumorprobe verfügbar war, wurde stattdessen eine aktuelle Tumorprobe vor der Randomisierung erhoben. • Bei Patienten mit Tumoren, die

Item[a]	Charakteristikum	Studieninformation
		aktivierende Mutationen im RET-Gen beherbergen, wurden Tumorproben gesammelt, um den Mutationsstatus des RET-Gens zu ermitteln und um dies zu den klinischen Ergebnissen in Beziehung zu setzen. • Bei Patienten mit hereditärem MTC, die eine nachgewiesene Keimbahn-RET-Mutation haben, war es nicht erforderlich, eine vorgeschriebene Tumor- Sammelprobe einzureichen. Sie waren aber zulässig für optionale Biopsien. Von Patienten, die in diesen explorativen Teil der Studie eingewilligt haben, wurden optional frische Tumorbiopsien gesammelt, um den Effekt der Vandetanib-Dosierung auf RET, EGFR und VEGFR Signalbahnen in Tumoren zu beschreiben. Tumor-Biopsieproben wurden per Standard-*Core*-Biopsietechniken oder per Feinnadelbiopsieproben in Paraffinblöcken entnommen. 1. Biopsie: Vor der ersten Dosis des Studienwirkstoffs 2. Biopsie: In Woche 12 nach der Randomisierung (solange der Patient sein Einverständnis nicht zurückgezogen hat) Alle Tumorproben wurden bei Raumtemperatur zu dem Zentrallabor geschickt, welches vertraglich von AstraZeneca bestimmt wurde. <u>Erhebungszeitpunkte:</u> *Screening* (Tag -21 bis 0 vor Randomisierung), die ersten drei Visiten wöchentlich, vierte Visite 2 Wochen später, 5. und 6. Visite alle vier Wochen, danach alle 12 Wochen, Nach Austritt aus verblindeter Studie: 60-Tage *Follow-up*, Überleben-*Follow-up* <u>Erhobene Parameter:</u> Einmalig erhoben: Einverständniserklärung, Demografien, medizinische / chirurgische Geschichte, Einschluss- / Ausschlusskriterien, Krankheitsausmaß, Schwangerschaftstest für gebärfähige Frauen, vorgeschriebene Tumor-Sammelprobe, Randomisierung zu Vandetanib bzw. Placebo, Verabreichung der 1. Dosis von Vandetanib bzw. Placebo Im Studienverlauf mehrmals erhoben (genaue Erhebungszeitpunkte siehe Studienbericht):

Item[a]	Charakteristikum	Studieninformation
		Ärztliche Untersuchung*, RECIST Bewertung*, WHO PS, 12-Kanal-EKG*, Vitalzeichen (Blutdruck, Puls, Temperatur)*, Begleitmedikation*, Nebenwirkungsüberprüfung*, Gewicht*, Blutproben für klinische Labortests*, 24-Stunden-Urinanalyse, Urinanalyse*, Blutproben für CTN/CEA*, Blutproben für Biomarker*, PK Blutproben für Vandetanib*, optionale frische Tumorbiopsie*, Vandetanib / Placebo Ausgabe / Rückgabe (alle 12 Wochen wurden Vandetanib- bzw. Placebo-Rationen für die nächsten 12 Wochen ausgegeben, unbenutzte Studienmedikation musste bei jeder Visite zurückgegeben werden), FACT-G QoL Fragebogen*, patientenberichtete Schmerzerfassung, Häufigkeit des Stuhlgangs (Patienten-erhoben), Schmerzmittelgebrauch (Patienten-erhoben), augenärztliche Untersuchung* *= Parameter, die auch in der offenen Phase nach Progression erhoben wurden (zusätzlich: Rückgabe der verblindeten Medikamente, offene Vandetanib / Placebo Ausgabe / Rückgabe)
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	18. Mai 2009 • PRO-Variablen wurden korrigiert, was sekundäre und explorative Zielkriterien der Studie beeinflusst hat. PRO-Variablen und statistische Methoden für PRO-Variablen wurden auch für die Konsistenz bei Änderungen der Studien-Zielkriterien überarbeitet. Begründung: Um TWP als einziges sekundäres PRO-Zielkriterium und alle anderen PRO-Zielkriterien als explorative Zielkriterien zu identifizieren. • Patientengewicht wurde von sekundärem zu explorativem Zielkriterium geändert. Begründung: Konsistent mit oben erwähnter Änderung sekundäre Zielkriterien bis auf ORR, DCR, DOR, biochemischem Ansprechen und TWP einzuschränken
7	Fallzahl	Anfänglich angestrebte Anzahl zu rekrutierender Patienten: 232, tatsächlich angemeldete Patienten: 437, Randomisierung erhalten: 331 Patienten, davon 231 Patienten zu Vandetanib randomisiert, 100 Patienten zu Placebo randomisiert
7a	Wie wurden die Fallzahlen bestimmt?	Annahmen: Eine 2:1 Randomisierung annehmend, werden, um eine Verdoppelung

Item[a]	Charakteristikum	Studieninformation
		des PFS bei 2-seitigem α = 0.05 Level und 80% Power zu ermitteln, mindestens 90 Ereignisse benötigt. Angenommen, ein medianes PFS von 12 Monaten in der Kontrollgruppe, eine nichtlineare Rekrutierungsperiode von 22 Monaten und eine minimale *Follow-up* Zeit von 6, 7 Monaten, müssten 232 Patienten für die Studie rekrutiert werden. D. h. die totale Länge der Studie wurde auf 28,7 Monate geschätzt, um 90 Progressions-Ereignisse zu beobachten.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurde keine Interim-Analyse geplant und durchgeführt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Das Randomisierungsschema wird von einer Computersoftware produziert, welches eine Standardprozedur zur Zufallszahlengenerierung einbindet. Dieses Randomisierungsschema wird nach Zentrum stratifiziert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung wurde nach Standort in 3er Blöcken stratifiziert.
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	Sobald Patienten für die Studie überprüft werden, müssen sie einen Einschreibungscode (E-code, *Enrollment Code*) zugewiesen bekommen. Dieser E-code setzt sich zusammen aus der Zentrumsnummer und der Patientennummer innerhalb des jeweiligen Zentrums (z. B.: erster Patient, der in Zentrum Nummer 0001 überprüft wird, bekommt E-code E0001001 zugewiesen, der zweite überprüfte Patient E0001002, usw.). Diese Nummer ist die persönliche, unverwechselbare Kennung und wird zur Identifizierung des Patienten auf den elektronischen Prüfbögen (eCRF, *electronic Case Report Form*) benutzt. Bei Eintritt eines Patienten in die Überprüfung sollte der Prüfarzt die zentrale Registrierung / das Randomisierungszentrum per Telefon anrufen, um den Patienten zu registrieren. Ablauf des gesamten Prozesses: • Patient wird identifiziert zur zentralen Registrierung / Randomisationszentrum mittels Initialen, E-code und Geburtstag. • IVRS (*Interactive Voice Response Services*) Die zentraleRegistrierung / Randomisierungszentrum informiert den Prüfarzt, die MedID-Nummer dem Patienten beim Randomisierungsbesuch zuzuweisen. • Der Prüfarzt wird für jede

Item[a]	Charakteristikum	Studieninformation
		nachfolgende Ausgabe-Visite IVRS hinzuziehen, um eine neue MedID-Nummer zuzuweisen Bei den Visiten zugewiesene MedID-Nummer entspricht der Behandlung zu der der Patient ursprünglich randomisiert wurde oder er wird eine Dosisreduzierung erhalten, wenn nötig.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	• Bei Eintritt eines Patienten in die Studie sollte der Prüfarzt die zentrale Registrierung / das Randomisierungszentrum per Telefon anrufen, um den Patienten zu registrieren. • Zuteilung zur Behandlung durch Randomisierungsschema, welches von einer Computersoftware produziert wird.
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?	Studienmedikation wird durch MedID-Nummer gekennzeichnet, die mit dem Randomisierungsschema verknüpft ist. Patienten und Personal des Prüfzentrums waren verblindet, sonst keine weitere Informationen
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Vandetanib- und Placebo-Tabletten sehen identisch aus. Auch die Tablettenverpackungen sind identisch.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Es wurden verschiedene Analysepopulationen festgelegt: 1. ITT-Population – alle randomisierten Patienten 2. PP-Population (per Protokoll Population) – alle randomisierten Patienten, ausgenommen die, die mindestens eine signifikante Protokollabweichung hatten von denen der Sponsor glaubt, dass sie einen potenziellen Einfluss auf die Wirksamkeits-Zielkriterien der Studie haben 3. Sicherheitspopulation – alle randomisierten Patienten, die mindestens eine Dosis der randomisierten Behandlung erhalten haben (d. h. mindestens eine Dosis von Vandetanib/Placebo) 4. PK-Population – alle randomisierten Patienten mit validen Plasma-Konzentrationen von Vandetanib, die zur Vandetanib-Gruppe randomisiert wurden

Item[a]	Charakteristikum	Studieninformation
		5. Offene Population– alle randomisierten Patienten, die mindestens eine Dosis der offenen Behandlung bekommen haben Variablen zum Zielkriterium Wirksamkeit: a. PFS <u>Primäranalyse</u> PFS wurde durch alle verfügbaren, durch eine zentrale Beurteilungsstelle ausgewerteten, RECIST-Bewertungen abgeleitet. Die RECIST-Kriterien wurden leicht modifiziert auf Basis von bestimmten radiografischen Charakteristika, hypodensen Läsionen und verkalkten Läsionen benutzt. Folgende Hypothesen wurden getestet: H0: Kein Unterschied zwischen 300 mg und Placebo H1: Unterschied zwischen 300 mg und Placebo • Basierend auf ITT-Population wurde ein Log-Rank-Test durchgeführt • Angabe der Punktschätzer für das mediane PFS für jede Behandlungsgruppe (grafische Darstellung mit Kaplan-Meier) • Zusammenfassung des Progressionsstatus von Patienten zur Zeit der Analyse • Zusammenfassung der Anzahl an Patienten, die eine Krebs-Folgetherapie nach der Beendigung der randomisierten Behandlung erhielten (Zusätzlich Zusammenfassung über diese ersten Folgetherapien) • Zusammenfassung der medianen Dauer des *Follow-Up* <u>Sensitivitätsanalysen</u> Für PFS: • PP-Analyse – Log-Rank-Test basierend auf PP-Population • PFS von zentraler Beurteilungsstelle, Daten in offener Phase

Item[a]	Charakteristikum	Studieninformation
		ausgeschlossen – Log-Rank-Test • PFS von einrichtungseigener Beurteilungsstelle – Log-Rank-Test • Cox-Modell mit proportionalem Ausfallrisiko (Cox–PH-Modell) basierend auf ITT-Population mit Behandlungseffekt, RET-Mutationsstatus, CTN-Verdopplungszeit, CEA-Verdopplungszeit, Anzahl vorheriger Therapien, Ansprechen auf jüngste vorherige Therapie und MTC-Status • PFS von zentraler Beurteilungsstelle, keine Adjustierung für verkalkte Läsionen – Log-Rank-Test • PFS von zentraler Beurteilungsstelle, keine Adjustierung für hypodense oder hypointensive Läsionen – Log-Rank-Test • Whitehead's Methode (Methode zum gruppierten Überleben) um den Einfluss einer unterschiedlichen Häufigkeit von Bewertungen zwischen Behandlungsgruppen zu schätzen • Globaler Interaktionstest mit 1%-Signifikanzlevel für das Cox-Modell <u>Subgruppenanalysen:</u> Mittels Log-Rank-Test, grafische Darstellung mittels *Forest Plot* Folgende Subgruppen wurden vor der Entblindung vorgegeben: • RET-Mutationsstatus • CTN-Verdopplungszeit • CEA-Verdopplungszeit • Anzahl an vorherigen Therapien • Ansprechen auf die jüngste vorherige Therapie • MTC-Status • Geschlecht

Item[a]	Charakteristikum	Studieninformation
		• Phase/Ausmaß der Erkrankung bei Eintritt • *Baseline* WHO PS • ethnische Zugehörigkeit • *Baseline* Werte für p-VEGF, p-VEGFR2 und p-bFGF b. ORR <u>Primäranalyse</u> Logistische Regression mit Kovariable „Behandlung“ (als einzige Kovariable, die auf der ITT-Population basiert) Resultate in Form einer Schätzung des *Odds Ratio* <u>Sensitivitätsanalyse</u> Logistische Regression mit Behandlungseffekt, RET-Mutationsstatus, CTN-Verdopplungszeit, CEA-Verdopplungszeit, Anzahl vorheriger Therapien, Ansprechen auf jüngste vorherige Therapie und MTC-Status c. DCR <u>Primäranalyse</u> Logistische Regression mit Kovariable „Behandlung“ (als einzige Kovariable, die auf der ITT-Population basiert) Resultate in Form einer Schätzung des *Odds Ratio* <u>Sensitivitätsanalyse</u> Logistische Regression mit Behandlungseffekt, RET-Mutationsstatus, CTN-Verdopplungszeit, CEA-Verdopplungszeit, Anzahl vorheriger Therapien, Ansprechen auf jüngste vorherige Therapie und MTC-Status Resultate in Form einer Schätzung des *Odds Ratio* d. DOR 2 Arten der Definition: • Zeit von der Randomisierung bis zum Tag der nachgewiesenen Krankheitsprogression laut zentraler Beurteilungsstelle oder Tod in Abwesenheit einer Krankheitsprogression

Item[a]	Charakteristikum	Studieninformation
		• Zeit vom Tag des ersten nachgewiesenen Ansprechens bis zum Tag der nachgewiesenen Progression laut zentraler Beurteilungsstelle oder Tod in Abwesenheit einer Krankheitsprogression DOR wurde zusammengefasst als Median des DORs nach beiden Definitionen Kaplan-Meier-Kurven für DOR wurden erstellt. e. OS Ist zum Zeitpunkt analysiert worden, an dem auch alle anderen Endpunkte analysiert worden und noch ein zweites Mal, wenn mindestens 50% der Patienten gestorben sind. Zum Zeitpunkt der ersten Analyse sind 48 Patienten gestorben und es wird erwartet, dass zum Zeitpunkt der zweiten Analyse 166 Patienten gestorben sein würden. Das Signifikanzniveau war für die erste Analyse 0,02% mit 99,98% Konfidenzintervallen. Bei der geplanten Überlebensfortschreibung war das Signifikanzniveau 4,98% mit 95,02% Konfidenzintervallen <u>Primäranalyse</u> Log-Rank-Test basierend auf der ITT-Population, grafische Darstellung durch Kaplan-Meier-Kurven. <u>Sensitivitätsanalysen</u> • Log-Rank-Test basierend auf der PP-Population • Cox-PH-Modell basierend auf ITT-Population (mit den gleichen Termen wie Cox- PH-Modell für PFS) Globaler Interaktionstest wurde basierend auf dem Cox-Modell durchgeführt f. Biochemisches Ansprechen Bestes CEA- und CTN- Ansprechen wird an der *Baseline* und während der Behandlung bewertet. Therapieansprecher sind die Patienten, die das beste

Item[a]	Charakteristikum	Studieninformation
		biochemische Ansprechen bei CR und PR haben Definitionen: • Komplettes Ansprechen (CR, *Complete Response*): Komplette Normalisierung der CEA/CTN-Level nach der Behandlung, bestätigt durch eine wiederholte Prüfung des CEA/CTN-Level • Partielles Ansprechen (PR, *Partial Response*) : Mindestens 50% Reduzierung des CEA/CTN-Levels (Referenz = *Baseline* - Level) • Progressive Erkrankung (PD, *Progressive Disease*): Mindestens 50% Erhöhung des CEA/CTN-Levels (Referenz = *Baseline* - Level) • Stabile Erkrankung (*Stable Disease*): Weder ausreichende Normalisierung, um als PR zu gelten, noch ausreichende Normalisierungserhöhung, um als PD zu gelten (Referenz = *Baseline* - Level) g. Klinische Gewichtszunahme Definiert als Gewichtszunahme um >5% gegenüber des *Baseline*-Werts und nicht gefolgt von einer Gewichtsüberprüfung, die eine Gewichtsabnahme gegenüber dem *Baseline*-Wert anzeigt. Dieser Zustand muss mindestens 4 Wochen anhalten ohne dass ein Zusammenhang mit Wassereinlagerungen oder Wassersucht besteht. Klinische Gewichtszunahmerate entspricht dem Anteil an randomisierten Patienten, die die oben genannten Kriterien erfüllen Analyse mittels logistischer Regression mit „Behandlung“ als einziger Kovariable, die auf der ITT-Population basiert.

Item[a]	Charakteristikum	Studieninformation
		h. TDPS Zu gegebenem Zeitpunkt war eine Verschlechterung des WHO PS als ≥1 Punkt-Steigerung des *Baseline*-Werts definiert. TDPS = Intervall vom Tag der Randomisierung bis zur ersten Feststellung einer Verschlechterung <u>Primäranalyse</u> Log-Rank-Test basierend auf ITT-Population Resultate als *Hazard Ratios* Zusätzlich: Punktschätzer des medianen TDPS. Grafische Darstellung: Kaplan-Meier-Kurven Variablen zu PRO: a. Patientenberichtete Schmerzerfassung (BPI, *Brief Pain Inventory*) Bietet Informationen über Intensität der Schmerzen (sensorische Dimension) und über den Grad, in dem der Schmerz die Funktion beeinträchtigt (Reaktive Dimension). Beurteilungsskala von 0 bis 10 des Grad des Schmerzes und der Beeinträchtigung genereller Aktivität, Stimmung, Laufen und anderer physikalischer Aktivität, Arbeit, sozialer Aktivität , Beziehungen zu anderen und Schlaf. BPI ist selbst-berichtend, für 4 Tage täglich, um einen *Baseline*-Wert zu erstellen, dann einmal wöchentlich bis Entblindung. Erhobene Bewertungen: • Einzelpostenbewertung des Postens „Schlimmster Schmerz“ • Arithmetisches Mittel der Bewertungen des 4-Posten-Schmerzstärke-Index • Arithmetisches Mittel der Bewertungen des 7-Posten Schmerz-Beeinträchtigungs-Index b. Opioidanalgetika-Gebrauch • Selbst-berichtend • Konvertiert zum Äquivalent

Item[a]	Charakteristikum	Studieninformation
		zu Morphinsulfat in mg/Tag • Für jeden Patienten individuell aufgelistet • Zusammenfassung des Gebrauchs und der Änderungen im Gebrauch ausgehend vom *Baseline*-Wert Liniendiagramm der mittleren Änderung vom *Baseline*-Wert c. TWP • Abgeleitet vom Posten „Schlimmster Schmerz“ des BPI-SF oder vom Opioidanalgetika-Gebrauch. • Grad der Schmerzprogression entsprach entweder einer Erhöhung von ≥2 Punkten vom *Baseline*-Wert des Postens „Schlimmster Schmerz“ oder einer Erhöhung des Opioidanalgetika Gebrauchs um ≥10 mg/Tag vom *Baseline*-Wert. • Grad der Verbesserung des Schmerzes entsprach (a) entweder einer Abnahme von ≥2 Punkten vom *Baseline*-Wert des Postens „Schlimmster Schmerz“, während der Gebrauch von Opioidanalgetika um nicht mehr als 10 mg/Tag gegenüber dem *Baseline*-Wert zugenommen hat oder (b) einer Abnahme im Opioidanalgetika-Gebrauch gegenüber der vorherigen Visite um >50% ohne eine Zunahme des Postens „Schlimmster Schmerz“ um ≥2 Punkte vom *Baseline*-Wert • Wenn der Patient keine bestätigte Verschlechterung zum Zeitpunkt der Analyse hatte, wurde er am Tag des letzten auswertbaren OS (Kombination des Ansprechens von Opioidanalgetika-Gebrauch und „Schlimmstem Schmerz“) zensiert.

Item[a]	Charakteristikum	Studieninformation
		Primäranalyse • Log-Rank-Test für TWP • Resultate als *Hazard Ratios* • Punktschätzer des medianen TWP für jede Behandlungsgruppe und grafische Darstellung des TWP in Kaplan-Meier - Kurven Sensitivitätsanalysen • Log-Rank-Test (basierend nur auf „Schlimmster Schmerz“ und nicht auf Opioidanalgetika-Gebrauch) • Cox-PH-Modell basierend auf ITT-Population (mit den gleichen Termen wie Cox-PH-Modell für PFS) • Log-Rank-Test basierend nur auf „Schlimmster Schmerz“ und nicht auf Opioidanalgetika-Gebrauch, wobei die „Schlimmster Schmerz“ Bewertungen für Patienten, die für TWP zensiert wurden, extrapoliert wurden, um zu bestimmen, ob sie sich, basierend auf den gesammelten Daten, verschlechtert haben könnten. Falls ja, wurde eine kalkulatorische TWP in der Analyse benutzt. Die Subgruppenanalyse wurde in der ITT-Population mittels Log-Rank-Test für Opioidanalgetika-Gebrauch durchgeführt d. Reduktion von Opioidanalgetika-Gebrauch Analyse mittels Logistischer Regression mit Kovariable Behandlung als einzige Kovariable, die auf der ITT Population basiert Resultate in Form von *Odds Ratios* e. Dauer des Opioidanalgetika-Gebrauchs Als Liste und Zusammenfassung f. FACT-G Mittels FACT-G Fragebogen, der

Item[a]	Charakteristikum	Studieninformation
		alle 12 Wochen bis Beendigung der randomisierten Studie vom Patienten ausgefüllt wurde. Berechnet wurden: Totaler FACT-G Wert und die 4 Subskalen Physisches Wohlbefinden (PWB), soziales Wohlbefinden, emotionales Wohlbefinden (EWB) und funktionales Wohlbefinden (FWB)
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Siehe 12a
-	**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	a) Vandetanib: 231, Placebo: 100 b) Vandetanib: 231, Placebo: 99 c) Patienten in ITT-Population: Vandetanib: 231, Placebo: 100 Patienten in PP-Population: Vandetanib: 215, Placebo: 91
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	1 Patient (1%, Placebo-Gruppe) erhielt keine randomisierte Behandlung aufgrund Tod Protokoll Abweichungen (Vandetanib (V):16, Placebo (P):9): • Abweichungen von den Basis RECIST-Kriterien (V:2 (0,9%), P:2 (2%)): o Basis RECIST Überprüfung mehr als 28 Tage vor der ersten Dosis der randomisierten Behandlung: V:2 (0,9%), P:2 (2%) • Abweichung von den Einschlusskriterien (V:10 (4,3%), P:5 (5%)): o Kein messbarer Tumor am *Baseline*: V:9 (3,9%), P:5 (5%) o Keine vorherig bestätigte histologische Diagnose von inoperablem, lokal fortgeschrittenem oder metastasiertem, hereditärem oder sporadischem medullärem Schilddrüsenkarzinom: V:1 (0,4%), P:0 (0%) • Abweichungen von den

Item[a]	Charakteristikum	Studieninformation
		Ausschlusskriterien (V:1 (0,4%), P:0(0%)): o Vorherige oder aktuelle Malignitäten anderer Histologien innerhalb der letzten 5 Jahre (mit Ausnahme der im Protokoll spezifizierten): V:1 (0,4%), P:0 (0%) • Abweichungen von der Begleitmedikation (V:0 (0%), P:1(1%)): o Einnahme von CYP3A4 Induzierern für mindestens 14 Tage während der randomisierten Behandlung; V:0 (0%), P:1(1%) • Abweichungen von der Randomisierung (V:4 (1,7%), P:1 (1%)): o Mindestens eine Dosis der falschen randomisierten Behandlung erhalten: V:4 (1,7%), P:0 (0%) o Wurde randomisiert, hat aber keine randomisierte Behandlung erhalten: V:0 (0%), P:1 (1%)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Erster Patient eingeschrieben: 23. November 2006 Letzter Patient eingeschrieben: 19. Oktober 2007 Daten C*ut-off*: 31. Juli 2009 Mediane Dauer des *Follow-up* nach Daten *Cut-off*: 103 Wochen
14b	Informationen, warum die Studie endete oder beendet wurde	Studie endet: voraussichtlich Dezember 2016
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

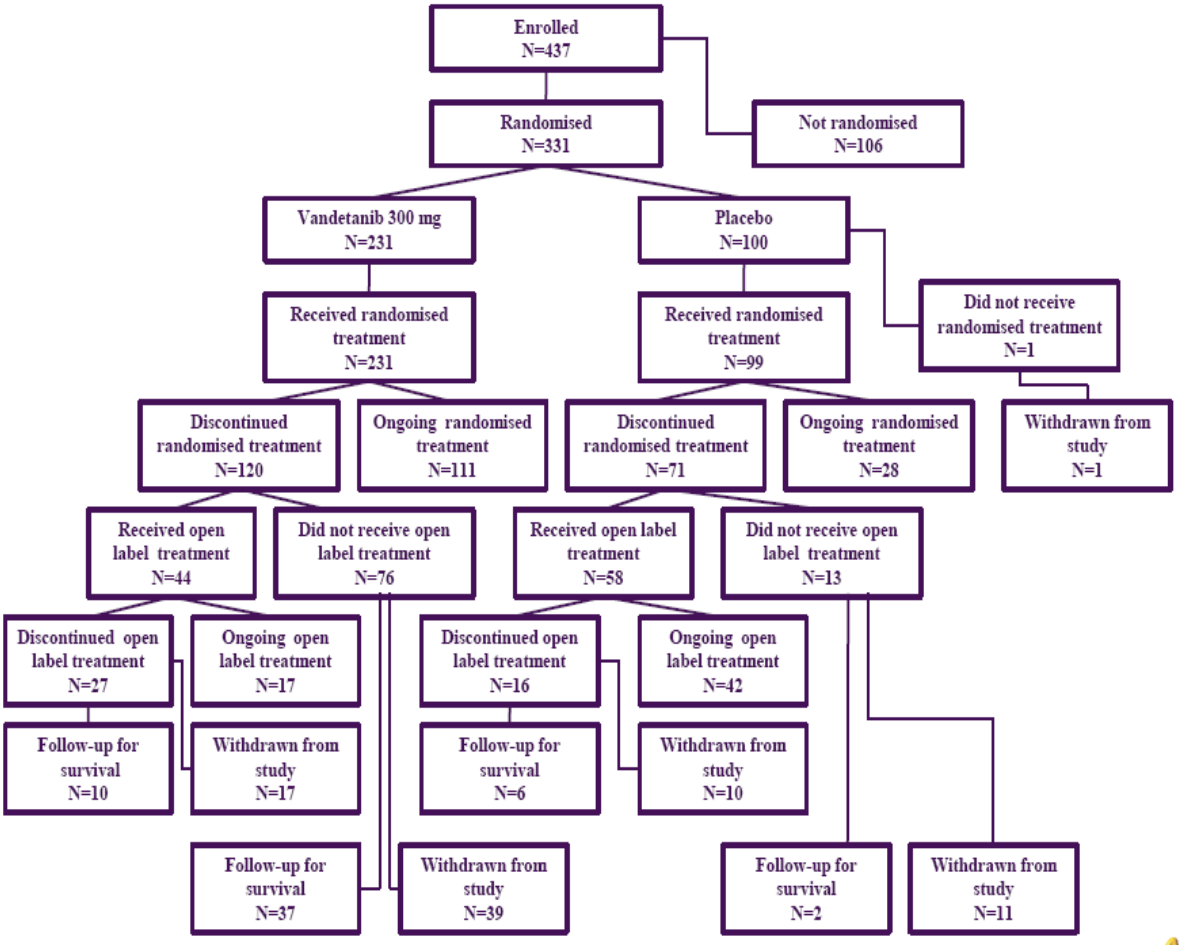


Abbildung 48: Flussdiagramm Patientenverteilung Studie D4200C00058

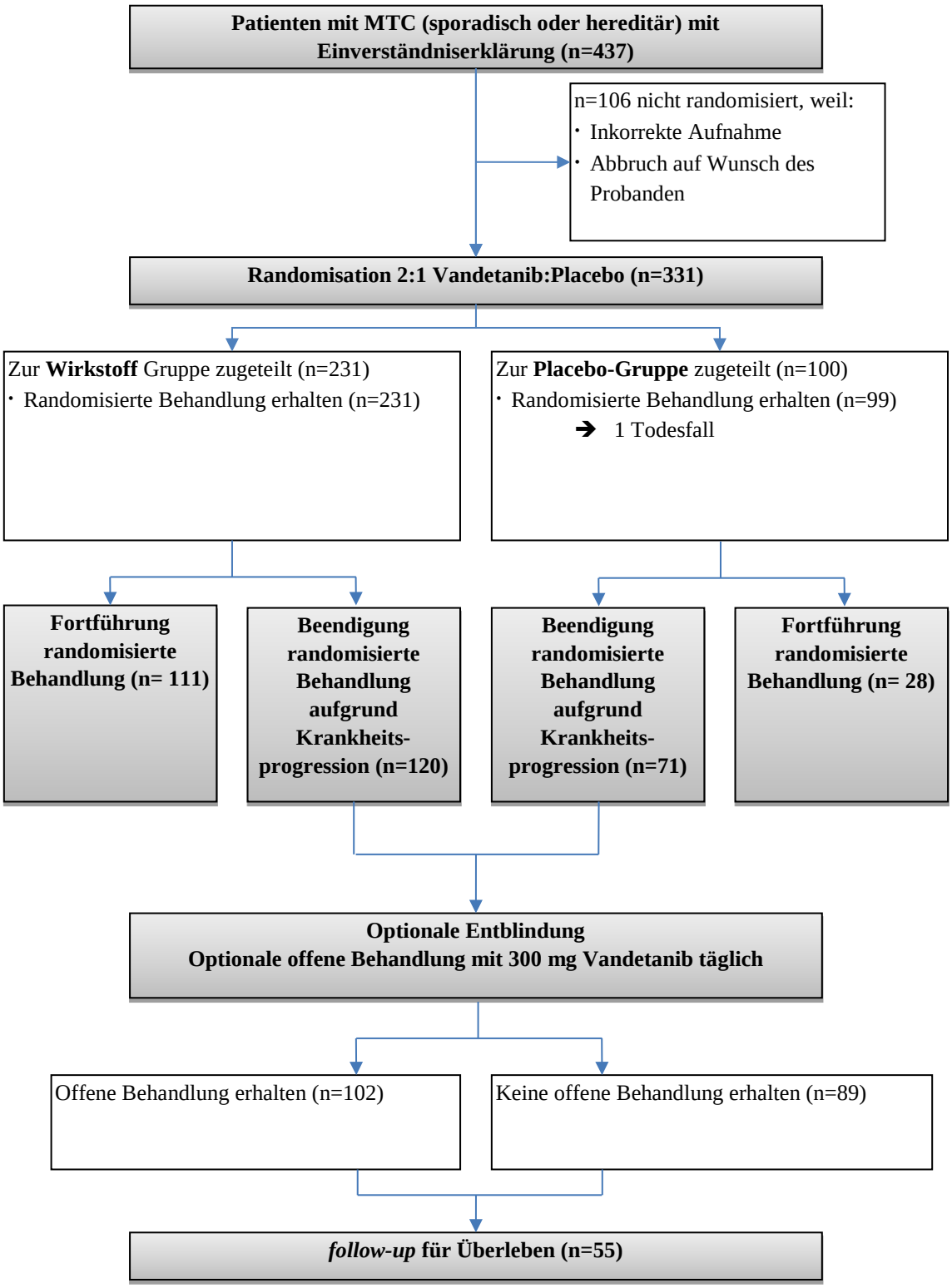


Abbildung 49: Flussdiagramm Studiendesign Studie D4200C00058

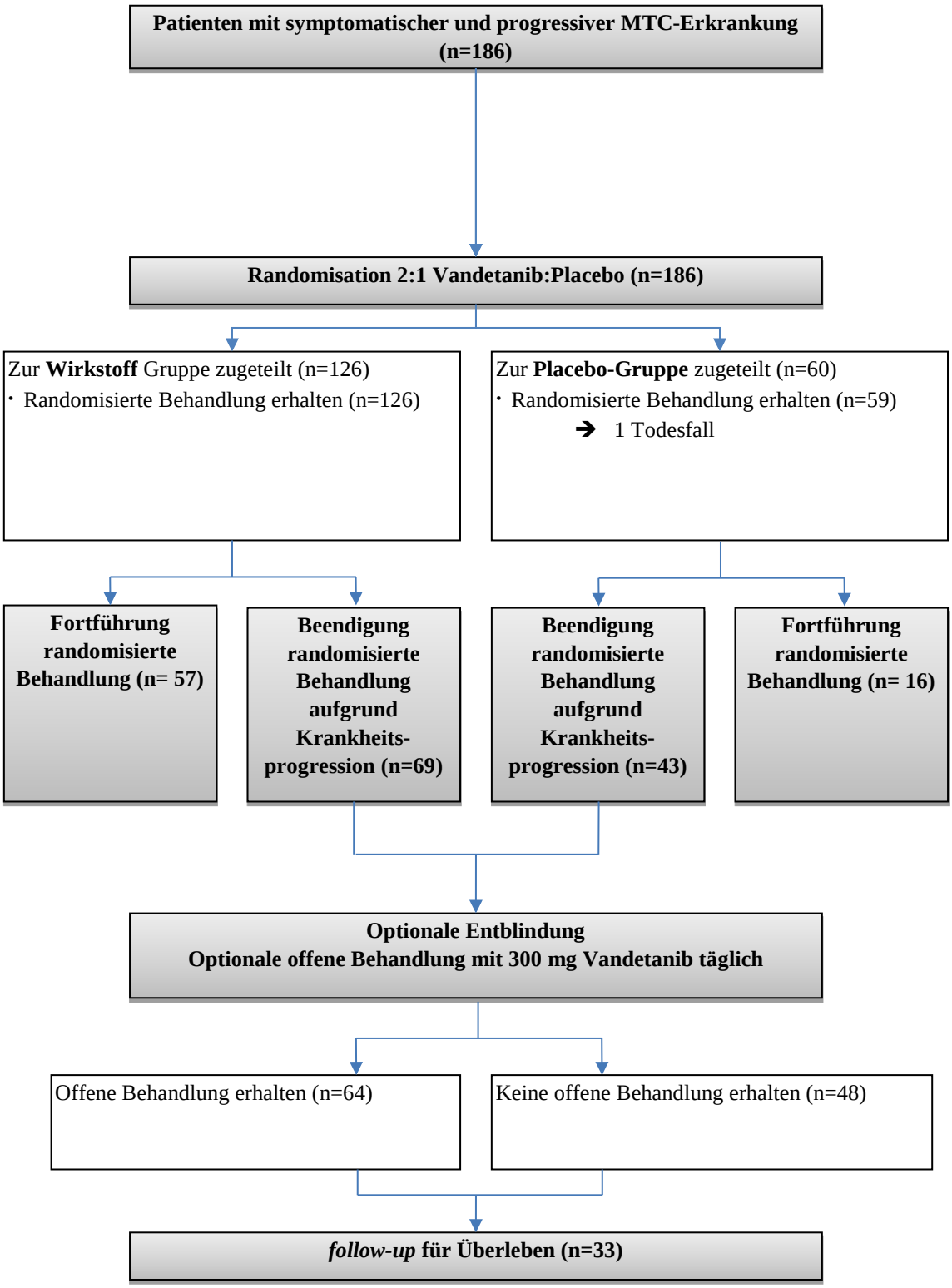


Abbildung 50: Flussdiagramm Teilpopulation mit symptomatischer und progressiver Erkrankung in Studie D4200C00058

Tabelle 4-188 (Anhang): Studiendesign und -methodik für Studie D4200C00008

Item[a]	Charakteristikum	Studieninformation
-	**Studienziel**	
3	Zielsetzungen	Primäres Ziel: Bewertung der objektiven Ansprechrate von Patienten mit lokal fortgeschrittenem oder metastasierendem hereditärem MTC, die mit 300 mg Vandetanib als Monotherapie behandelt wurden. Eine beobachtete Ansprechrate größer 0% wurde als Studienziel definiert. Sekundäre Ziele: 1. Untersuchung des biochemischen Ansprechens von Calcitonin (CTN) bei Behandlung mit Vandetanib 2. Untersuchung, ob Vandetanib zu einem Rückgang der Stuhlgangshäufigkeit und Verbesserung der Stuhlgangkonsistenz bei Patienten mit symptomatischem Durchfall in Verbindung mit vererbbaren hereditärem MTC vorliegt 3. Untersuchung des progressionsfreien Überlebens (*Progression-Free Survival*, PFS) 4. Untersuchung der Krankheitskontrollrate (*Disease Control Rate*, DCR), der Dauer des objektiven Ansprechens (*Duration of Objektive Response*, DOR) und der Dauer der Krankheitskontrolle (*Duration of Disease Control*) 5. Bewertung des WHO *Performance Status* (WHO PS) 6. Beschreibung der Pharmakokinetik in dieser Population 7. Beschreibung des pharmakokinetischen - pharmakodynamischen Zusammenhangs zwischen der Vandetanib Exposition und Veränderungen der QTc Verlängerung, der seltenen Ereignisse, des Ansprechens, der TTP und Veränderungen der CTN- und CEA-Konzentration 8. Untersuchung der Sicherheit und Verträglichkeit von Vandetanib in dieser Patientenpopulation Explorative Ziele: 1. Untersuchung des basalen CTN Levels nach einem *calcium-pentagastrin stimulation test* in ausgewählten Patienten mit einem kompletten biochemischen Ansprechen. 2. Untersuchung, ob eine Veränderung des Baseline Plasma Level des *Vascular Endothelial Growth Factor* (VEGF), *Basic Fibroblast Growth Factor* (bFGF) und CEA prognostischen Wert bei Patienten mit MTC das Ansprechen von Vandetanib verändert. 3. Effekt von Vandetanib auf DNA Mikroarray Muster 4. Effekt von Vandetanib auf Lebensqualität (QOL) und auf vom Tumor hervorgerufener Diarrhö unter Verwendung des FACIT-D (Functional Assessment of Chronic Illness Therapy – Diarrhea)
-	**Methoden**	
4	Studiendesign	Internationale, einarmige, multizentrische, offene, zwei-

Item[a]	Charakteristikum	Studieninformation
		stufige Phase-II-Studie. 30 Patienten waren insgesamt geplant. Falls nach 15 Patienten kein einziges Ansprechen vorlag, wurden keine weiteren Patienten rekrutiert.
5	Rahmen (Setting und Ort der Studie, relevante zeitliche Angaben, einschließlich der Zeiträume der Rekrutierung, der Exposition, der Nachbeobachtung und der Datensammlung)	Prüfzentren in der USA und Frankreich Erster Patient aufgenommen: November 2004 Letzter Patient abgeschlossen: Februar 2008 Exposition: Täglich 300 mg Vandetanib, oral, bis zur Progression
6	Studienteilnehmer	
6a	*Kohortenstudie* (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl an; Methoden der Nachbeobachtung) *Fallkontrollstudie* (Einschlusskriterien, Herkunft der Fälle und Kontrollen, Methoden, mit denen die Fälle erhoben und die Kontrollen ausgewählt wurden; Begründung (Rationale) für die Auswahl der Fälle und Kontrollen) *Querschnittsstudie* (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl)	Einschlusskriterien: 1. Einverständniserklärung unterschrieben 2. Männer oder Frauen im Alter ≥ 18 3. Histologisch bestätigte Diagnose: MEN 2a, MEN 2b oder FMTC mit einer charakteristischen Keimzellmutation im RET Protoonkogen (Tumorblock, Polymerase Chain Reaction [PCR] Analyse, und / oder Familienhistorie wird angefordert, um die Diagnose zu bestätigen). Die Unterlagen müssen in der Patientenakte vorliegen. 4. WHO *Performance Status* 0-2 5. Fähig, die Studienmedikation zu schlucken 6. Mindestens eine messbare Läsion von mindestens 1 cm Durchmesser im CT-Aufnahme oder 2 cm mit konventionellen Techniken 7. Negativer Serum-Schwangerschafts-test für Frauen im gebärfähigen Alter. Frauen: mind. 1 Jahr postmenopausal, Sterilisation oder Verwendung einer geeigneten Verhütungsmethode (oral verabreichtes Verhütungsmittel, Barrieremethoden, zugelassenes Implantat, Langzeitinjektion, intra-uterine Methode oder Ligatur der Eileiter) Männer: Vasektomie oder Verwendung einer geeigneten Verhütungsmethode während der Studiendauer 8. Schlucken der Medikation als ganze Tablette ohne Kauen, Zerkleinern oder Teilen Ausschlusskriterien: 1. Gehirnmetastasen oder Rückenmarkskompression, sofern nicht mind. 4 Wochen vor der ersten Dosis bestrahlt und stabil ohne Behandlung mit Steroiden für mind. eine Woche 2. Nachweis eines Phäochromozytoms, basierend auf einem Vorkommen von Katecholaminen im Urin in der 24h Untersuchung 3. Weniger als 1.5x 10^9/L Neutrophile oder weniger als 100,000/mm^3 Platelets 4. Serum Bilirubin größer als 1.5 x oberes Limit des Referenzwertes (ULRR)

Item[a]	Charakteristikum	Studieninformation
		5. Serum Kreatinin größer als 1.5 x ULRR 6. Kaliumkonzentration kleiner 3.5 mEq/l, Kalzium (ionisiertes Kalzium oder für Albumin adjustiertes), oder Magnesium Konzentrationen außerhalb der Normalwerte, sofern nicht von AstraZeneca genehmigt. Zugabe von Elektrolyten ist möglich. 7. ALT oder Aspartate Aminotransferase (AST) größer als 2.5 x ULRR, bzw. ALT oder AST größer als 5.0 x ULRR, falls der Prüfarzt einen Zusammenhang mit Lebermetastasen diagnostiziert oder alkalische Phosphatase (ALP) größer als 2.5 x ULRR 8. Vorliegen einer schweren oder unkontrollierten systemischen Krankheit, einschließlich Hepatitis B oder HIV. Ein Screening für chronische Krankheiten ist nicht erforderlich, obwohl Patienten mit bekannten chronischen Krankheiten nicht eingeschlossen werden sollten. 9. Signifikanter kardiologischer Vorfall (einschließlich symptomatischer Herzfehler, Herzinfarkt, oder Angina) innerhalb der letzten 3 Monate der ersten Studiendosis oder Vorhandensein einer kardiologischen Krankheit, die nach Einschätzung des Prüfarztes das Risiko einer ventrikulären Arrhythmie erhöht 10. Vorgeschichte von Arrhythmien (multifokale ventrikuläre Extrasystolen, Bigeminus, Trigeminus, ventrikuläre Tachykardie, Bradykardie), die symptomatisch waren oder Behandlung erforderten (CTCAE Grad 3) oder Vorgeschichte einer symptomatischen oder fortwährenden ventrikulären Tachykardie. Klinische Signifikanz wird durch den Prüfarzt bestimmt. 11. Chronisches Vorhofflimmern 12. Vorgeschichte von QTc Intervall Verlängerungen bei anderen Medikamenten 13. Angeborenes langes QT Syndrom 14. QTc nach Bazett-Korrektur nicht messbar oder ≥ 480 ms bei der ECG Screening-Untersuchung. Zu beachten: Wenn ein Patient ein QTc Intervall ≥460 ms beim Screening EKG hatte, konnte der Screening EKG nach mindestens 24h wiederholt werden. Die durchschnittliche QTc beider EKGs musste dann unter 460 ms liegen. Falls ein Patient nach diesem Kriterium einschließbar ist, wird als Baseline EKG Wert der Durchschnitt aus diesen beiden EKGs und dem EKG direkt vor erster Studienmedikation verwendet. 15. Verbotene Begleitmedikation in dieser Studie einschließlich 5HT3 Antagonisten: starke und mittlere CYP3A4 Hemmer, starke und mittlere CYP3A4Induktoren, Medikationen, die das QTc Intervall verlängern oder *Torsades de Pointes* hervorrufen, sofern nicht von AstraZeneca zugelassen. Siehe auch Kapitel 3.4.9.1 „*Cancer and other concomitant treatment*“ im Studienprotokoll. 16. Vorgeschichte von linksventrikulärer Ejektionsfraktion (LVEF) weniger als 45%,

Item[a]	Charakteristikum	Studieninformation
		gemessen durch *multi-gated acquisition scan* (MUGA) oder Echokardiogramm (ECHO) für Patienten mit früherer Anthrazyklin Therapy (Gesamtdosis > 450 mg/m^2) oder signifikante kardiovaskuläre Krankheit oder Brustbestrahlung, wie vom Prüfarzt bestimmt. 17. Unkontrollierte Hypertension (systolischer Blutdruck > 160 mm Hg oder diastolischer Blutdruck > 100 mm Hg) 18. Jegliche andere Malignität oder Vorgeschichte einer Malignität in den letzten 5 Jahren, sofern nicht von AstraZeneca zugelassen (mit Ausnahme von Nichtmelanom-Hautkrebs oder *Carcinoma in situ* des Zervix) 19. Chemotherapie und / oder Bestrahlungstherapie, falls diese nicht in den letzten 4 Wochen vor erster Studienmedikation endeten 20. Jegliche ungelöste chronische Toxizität mit CTCAE Grad größer 2 von vorherigen Antitumor Therapie, sofern nicht von AstraZeneca zugelassen. 21. Teilnahme an einer klinischen Studie in den letzten 30 Tagen (abgesehen von Studien zur Lebensqualität) 22. Aktuell schwanger oder stillend Herkunft: Patienten aus der USA und Frankreich
6b	*Kohortenstudie* (Matching (Paarbildung) und Matchingkriterien, Anzahl der exponierten und der nicht exponierten Teilnehmer) *Fallkontrollstudie* (Matching (Paarbildung) und Matchingkriterien, Anzahl der Kontrollen pro Fall)	Nicht zutreffend
7	Variablen	Primäres Zielkriterium: Objektive Ansprechrate (ORR) nach dem modifizierten RECIST-Kriterium. ORR ist definiert als der Prozentsatz an Patienten, die eine komplette oder partielle Remission (bestätigt nach mindestens 4 Wochen) hatten. Sekundäre Zielkriterien: 1. Biochemisches Ansprechen: Änderung gegenüber des Baselinewertes bei Calcitonin (CTN) nach Behandlung mit Vandetanib 2. Symptomatisches Ansprechen: Änderung gegenüber des Baselinewertes der Stuhlgangshäufigkeit und -konsistenz bei Patienten mit symptomatischem Durchfall in Verbindung mit vererbbaren hereditärem MTC nach Behandlung mit Vandetanib 3. Progressionsfreies Überleben (*Progression Free Survival*, PFS) gemessen von der ersten Dosisgabe bis Progression oder Tod 4. *Disease Control Rate* (DCR; Rate der Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 24 Wochen)

Item[a]	Charakteristikum	Studieninformation
		Dauer des objektiven Ansprechens (*Duration of Objective Response*, DOR) gemessen vom Beginn des Ansprechens bis zur Progression. Weitere Berechnungsweise, die ebenfalls ausgewertet wird, ist die Dauer von der ersten Dosis bis zur Progression für Patienten mit einem objektiven Ansprechen. Dauer der Krankheitskontrolle (*Duration of Disease Control*) gemessen von der ersten Dosisgabe bis zur Progression. 5. Bewertung des WHO *Performance Status* (WHO-PS) Explorative Zielkriterien: 1. FDG-PET (*2-[F-18] fluor-2-deoxy-D-glucose positron emission tomography*) 2. *Calcium-pentagastrin stimulation test* 3. CEA levels (*Carcinoembryonic antigen*) 4. DNA (*Deoxyribonucleic acid*) Mikroarray Muster 5. FACIT-D (*Functional Assessment of Chronic Illness Therapy – Diarrhea*) 6. RET (*Rearranged during Transfection (proto-oncogene)*), phospho-RET, EGFR (*Epidermal growth factor receptor*), phospho-EGFR, VEGFR-2 (*Vascular endothelial growth factor receptor*), phospho-VEGFR-2, MAPK (*Mitogen-activated protein kinase*), phospho-MAPK, SHC protein (*Src homology 2 domain containing*), phospho-SHC protein
8	Datenquellen/Messmethoden	Für das primäre Zielkriterium wurden radiologische Untersuchungen durchgeführt und entsprechend den RECIST Kriterien bewertet.
9	Bias	Ein zentraler Review der CT/MRI *scans* für den primären Endpunkt wurde durchgeführt um einen Zentrumsbias zu vermeiden. Beim FACIT-D wurden entsprechende Maßnahmen getroffen, falls sich Hinweise auf eine Symptomatik bei den fehlenden Werten ergaben, um potentiellen Bias zu vermeiden.
10	Studiengröße	Die Studie hatte ein zweistufiges Design mit jeweils 15 geplanten Patienten pro Stufe. Falls nach den ersten 15 Patienten kein einziges Ansprechen vorlag, wurden keine weiteren Patienten rekrutiert. 30 Patienten wurden eingeschlossen.
11	Quantitative Variablen	Siehe Item 7
12	Statistische Methoden	
12a	Statistische Methoden, einschließlich der Methoden für die Kontrolle von *Confounding*	Grundlage: *Intention-To-Treat*-Prinzip Aufgrund des einarmigen Studiendesigns wurden keine statistischen Vergleiche durchgeführt. Berechnung der Zielkriterien Berechnung von exakten 95%-Konfidenzintervallen zu den Endpunkten ORR, CTN-Response, DCR, Darstellung der Kaplan-Meier-Kurve zum Endpunkt „PFS“ „*Duration of Disease Control*“ und „DOR“, einschließlich

Item[a]	Charakteristikum	Studieninformation
		dem medianen Überleben und entsprechendem 95% Konfidenzintervall. Berechnung des Mittelwerts, Medians, der Standardabweichung sowie Maximum und Minimum für Endpunkte mit stetigen Zielparametern (z. B. FACIT-D Score und die Laborwerte). Zusammenhänge verschiedener Parameter (z. B. Vandetanib-Plasmakonzentration und unerwünschte Ereignisse) wurden anhand von Korrelationen dargestellt.
12b	Subgruppen und Interaktionen	Nicht zutreffend
12c	Umgang mit fehlenden Daten	Fehlende Werte beim FACIT-D wurden entsprechend der üblichen Vorgaben beim FACIT-D behandelt (d.h. 50% der Werte mussten vorhanden sein, damit eine FACIT Subskala ausgewertet wurde).
12d	*Kohortenstudie* (Umgang mit vorzeitigem Ausscheiden („loss to Follow-up“)) *Fallkontrollstudie* (Berücksichtigung des *Matchings* (Paarbildung) von Fällen und Kontrollen bei der Auswertung) *Querschnittsstudie* (Auswertungsmethoden, die die gewählte Strategie zur Stichprobenauswahl (*Sampling strategy*) berücksichtigen)	Patienten mit vorzeitigem Ausscheiden wurden als non-Responder betrachtet.
12e	Sensitivitätsanalysen	Nicht zutreffend
-	**Ergebnisse**	
13	Teilnehmer	
13a	Anzahl der Teilnehmer während jeder Studienphase (Anzahl der Teilnehmer, die potenziell geeignet waren, die auf Eignung untersucht wurden, die als geeignet bestätigt wurden, die tatsächlich an der Studie teilgenommen haben, deren Nachbeobachtung abgeschlossen wurde und deren Daten ausgewertet wurden)	35 Patienten wurden gescreent. 30 Patienten wurden in die Studie eingeschlossen und behandelt. 13 Patienten haben die Studie vorzeitig abgebrochen.
13b	Gründe für die Nicht-Teilnahme in jeder Studienphase	Nicht zutreffend
13c	Flussdiagramm	Siehe Flow Chart unten.
17	Weitere Auswertungen	Nicht zutreffend
a: nach STROBE 2007		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

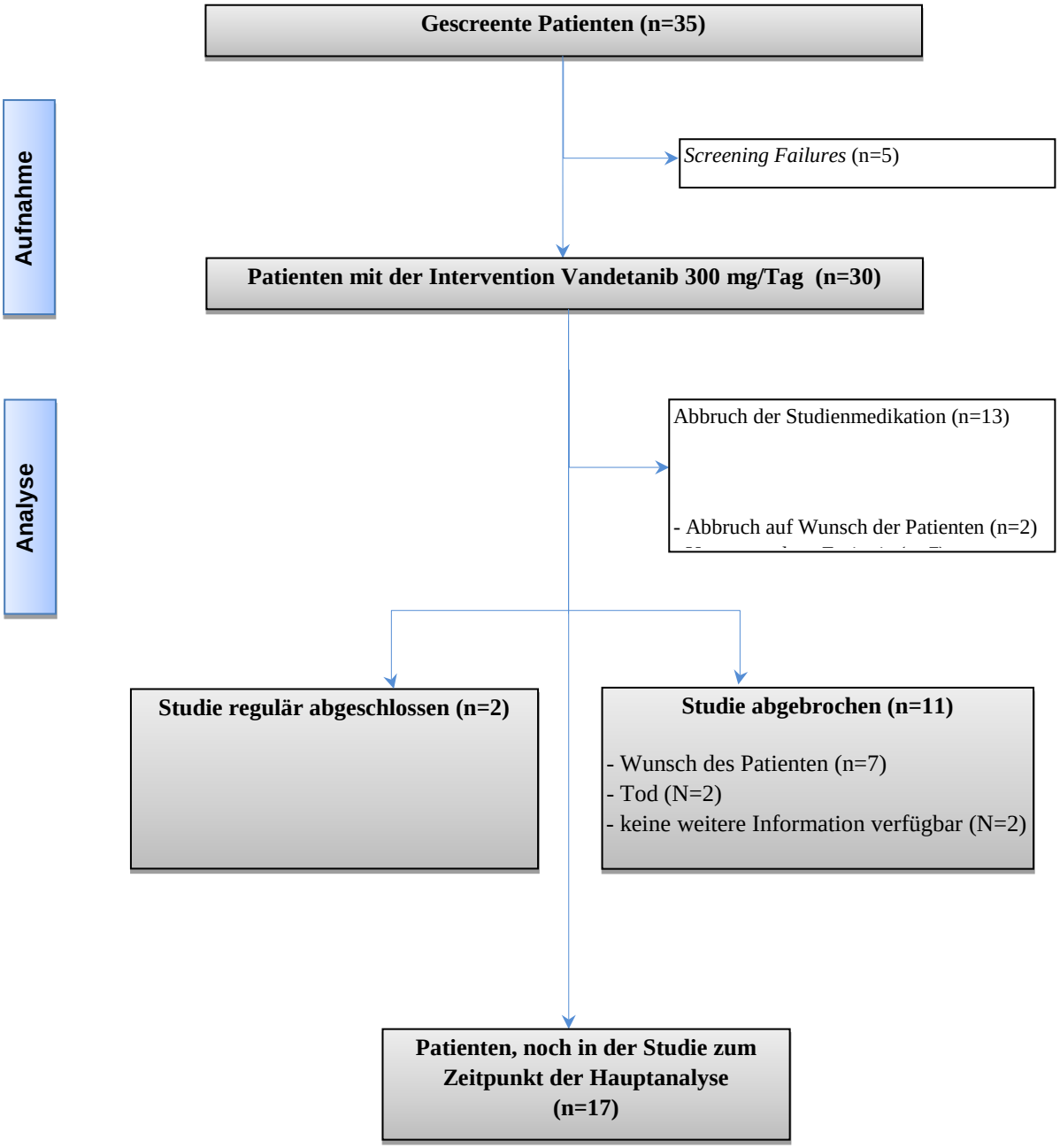


Abbildung 51: Flussdiagramm Studiendesign Studie D4200C00008

Tabelle 4-189 (Anhang): Studiendesign und -methodik für Studie D4200C00068

Item[a]	Charakteristikum	Studieninformation
-	**Studienziel**	
3	Zielsetzungen	Primäres Ziel: Bewertung der objektiven Ansprechrate von Patienten, die mit 100 mg Vandetanib als Monotherapie behandelt wurden. Eine beobachtete Ansprechrate von 20% wurde als Studienziel definiert. Sekundäre Ziele: 1. Untersuchung der Sicherheit und Verträglichkeit der Behandlung in dieser Patientenpopulation 2. Untersuchung des progressionsfreien Überlebens (*Progression-Free Survival*, PFS) 3. Untersuchung der Krankheitskontrollrate (*Disease Control Rate*, DCR), der Dauer bis zum objektiven Ansprechen (*Duration of Objective Response*, DOR) und der Dauer der Krankheitskontrolle (*Duration of Disease Control)* 4. Bewertung der Veränderung des WHO *Performance Status* (WHO (PS)) gegenüber dem Ausgangswert unter Benutzung des World Health Organisation (WHO) PS 5. Untersuchung, ob ein Anstieg der Stuhlgangshäufigkeit bei Patienten mit symptomatischem Durchfall in Verbindung mit hereditärem MTC vorliegt 6. Bewertung des Effekts der Dosis 100 mg Vandetanib auf die Konzentration von Calcitonin (CTN) und des carcinoembryonalen Antigens (CEA) 7. Beschreibung der Pharmakokinetik in dieser Population 8. Beschreibung des pharmakokinetischen - pharmakodynamischen Zusammenhangs zwischen der Vandetanib Exposition und Veränderungen der QTc Zwischenraum, der seltenen Ereignisse, des Ansprechens, des PFS und Veränderungen der CTN- und CEA-Konzentration Explorative Ziele: 1. Untersuchung, ob der Ausgangs-Plasmaspiegel von CTN und/oder CEA eine prognostische Signifikanz für Patienten mit MTC hat und ob die Exposition mit Vandetanib die CTN- und CEA-Konzentrationen erhöht oder erniedrigt 2. Erforschung des Effekts von Vandetanib auf die Lebensqualität unter Benutzung der *Functional Assessment of Cancer Therapy General Scale* (FACIT-G) 3. Einschätzung des Effekts von Vandetanib auf die Aktivität des RET Onkoproteins, des EGFRs und des VEGFR-2 im Tumorgewebe 4. Bestätigung der Keimbahn-RET-Mutation im Tumorgewebe 5. Einschätzung des Effekts von Vandetanib auf zirkulierende Angiogenese-Biomarker, auf die Tumorlast und - auf urinbasierte Biomarker

Item[a]	Charakteristikum	Studieninformation
-	**Methoden**	
4	Studiendesign	Internationale, einarmige, multizentrische, offene Phase-II-Studie
5	Rahmen (Setting und Ort der Studie, relevante zeitliche Angaben, einschließlich der Zeiträume der Rekrutierung, der Exposition, der Nachbeobachtung und der Datensammlung)	9 Prüfzentren in Australien, Italien, Kanada, Niederlande, Rumänien, Schweiz, Spanien, und den USA Erster Patient aufgenommen: 29 August 2006 Letzter Patient abgeschlossen: 31. Januar 2008 Täglich 100 mg Vandetanib, oral, bis zur Progression Post-Progressive Behandlung von 300 mg Vandetanib, falls der Prüfarzt einen Nutzen für den Patienten sieht
6	Studienteilnehmer	
6a	*Kohortenstudie* (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl an; Methoden der Nachbeobachtung) *Fallkontrollstudie* (Einschlusskriterien, Herkunft der Fälle und Kontrollen, Methoden, mit denen die Fälle erhoben und die Kontrollen ausgewählt wurden; Begründung (Rationale) für die Auswahl der Fälle und Kontrollen) *Querschnittsstudie* (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl)	Einschlusskriterien: 1. Bereitstellung einer schriftlichen Einverständniserklärung 2. Weiblich oder männlich im Alter von 18 Jahren und älter 3. Histologisch bestätigte Diagnose: lokal fortgeschrittenes oder metastasierender hereditärer MTC, für den keine Standardtherapie mehr verfügbar ist 4. Lebenserwartung ≥ 12 Wochen 5. WHO Performance Status 0-2 6. Mindestens eine messbare Läsion von mindestens 10 mm Durchmesser an der breitesten Stelle durch Spiral-CT, CT-Scan (5 mm Schichtdicke) oder 20 mm mit konventionellen Techniken (> 5 mm Schichtdicke) gemäß des modifizierten RECIST-Kriteriums 7. Negativer Schwangerschaftstest für Frauen im gebärfähigen Alter. Frauen: mind. 1 Jahr postmenopausal, Sterilisation oder Verwendung einer geeigneten Verhütungsmethode Männer: Vasektomie oder Verwendung einer geeigneten Verhütungsmethode während der Studie 8. Schlucken der Medikation als ganze Tablette ohne Kauen, Zerkleinern oder Teilen Ausschlusskriterien: 1. Gehirnmetastasen oder Rückenmarkskompression, sofern nicht mind. 4 Wochen vor der ersten Dosis bestrahlt und stabil ohne Behandlung mit Steroiden für mind. eine Woche 2. Letzte Dosis der vorherigen Chemotherapie in den letzten 4 Wochen vor Beginn der Studientherapie 3. Bestrahlung in den letzten 4 Wochen vor Beginn der Studientherapie 4. Größere Operationen in den letzten 4 Wochen oder unzureichend geheilter chirurgischer Eingriff vor Beginn der Studientherapie 5. Ungeklärte Toxizitäten mit CTCAE-Grad > 2 (Common Terminology Criteria for Adverse Events, National Institutes of Health, National Cancer Institute) von vorheriger Antitumortherapie

Item[a]	Charakteristikum	Studieninformation
		6. Serum-Bilirubin-Konzentration > 1.5 mal dem oberen Grenzwert des Referenzbereichs (ULRR) 7. Serum-Kreatinin-Konzentration > 1.5 mal ULRR oder Kreatinin-Clearance ≤ 50 ml/min (Berechnet durch die Cockcroft-Gault-Formel) 8. Alaninaminotransferase (ALT), Aspartataminotransferase (AST) oder alkalische Phosphatase (ALP) > 2.5 mal ULRR ohne nachweisbare Lebermetastasen , sonst > 5 mal ULRR falls der Prüfarzt einen Zusammenhang mit Lebermetastasen diagnostiziert 9. Bedeutendes kardiales Ereignis (z. B. Herzinfarkt, Vena-cava-superior-Syndrom, New York Heart Association classification of heart disease > 2) in den letzten 3 Monaten vor Studieneintritt oder Vorhandensein einer Herzerkrankung, die laut Meinung des Prüfarztes, das Risiko ventrikulärer Arrhythmien erhöht 10. Vorgeschichte von Arrhythmien (multifokale vorzeitige ventrikuläre Kontraktion (PVCs), Bigeminus, Trigeminus, ventrikuläre Tachykardie, symptomatisches oder unkontrolliertes Vorhofflimmern), die symptomatisch waren oder Behandlung erforderten (CTCAE Grad 3) oder asymptomatisch fortwährende ventrikuläre Tachykardie. Patienten mit Vorhofflimmern, welches durch Medikation kontrolliert worden ist, werden zugelassen 11. Angeborenes long-QT-Syndrom oder Verwandtschaft ersten Grades mit einem ungeklärten plötzlichen Todesfall unter 40 Jahren 12. QT-Verlängerung mit anderen Medikamenten, die einen Abbruch dieser Therapie erfordern 13. Linksschenkelblock 14. QTc nach Bazett-Korrektur nicht messbar oder ≥T480 ms bei der EKG Screening-Untersuchung. Zu beachten: Wenn ein Patient ein QTc Intervall ≥480 ms beim Screening EKG hatte, konnte das Screening-EKG zweimal wiederholt werden. Die durchschnittliche QTc musste dann unter 480 ms liegen. 15. Kaliumkonzentration < 4.0 mmol/l trotz Ergänzung, Kalzium (ionisiert oder Albumin-korrigiert) oder Magnesium außerhalb des Referenzbereichs trotz Ergänzung 16. Schwangerschaft oder stillend Frauen 17. Begleitende Medikation, welche die QTc-Verlängerung auslösen könnte (mit Ausnahme von Somatostatin oder Somatostatin-Analogon) und / oder verbotene Medikamente laut Studienprotokoll 18. Nichtkontrollierte Hypertonie (systolischer Blutdruck > 160 mm Hg oder diastolischer Blutdruck > 100 mm Hg) 19. Vorangegangene oder aktuelle bösartige Tumore an anderen Stellen in den letzten 5 Jahren mit Ausnahme eines Carcinoma in situ des Zervix oder

Item[a]	Charakteristikum	Studieninformation
		angemessene behandelte Basalzellen- oder Plattenepithelkarzinom der Haut 20. Anzeichen für schwere oder unkontrollierte systemische Erkrankungen; oder der aktuelle Gesundheitszustand ist für eine Studienteilnahme/Einhaltung des Protokolls nach Meinung des Prüfarztes unvorteilhaft. 21. Vorangegangene Teilnahme an der vorliegenden Studie oder vorangegangene Teilnehme an einer Studie mit einer Behandlung durch Vandetanib 22. Mitwirkung an der Planung und Durchführung der Studie (angewandt auf Personal von AstraZeneca oder Personal des Prüfzentrums) 23. Erhalt jeglicher Prüfmedikation innerhalb 30 Tage vor Beginn dieser Studie
6b	*Kohortenstudie* (Matching (Paarbildung) und Matchingkriterien, Anzahl der exponierten und der nicht exponierten Teilnehmer) *Fallkontrollstudie* (Matching (Paarbildung) und Matchingkriterien, Anzahl der Kontrollen pro Fall)	Nicht zutreffend
7	Variablen	Variable für primäres Ziel: Objektive Ansprechrate (ORR) nach dem RECIST-Kriterium. ORR ist definiert als der Anteil (%) an Patienten, die eine komplette oder partielle Remission hatten. Variablen für sekundäre Ziele: - Unerwünschter Ereignissen (UE) - Klinisch signifikante abnormale Laborwerte, EKG-Veränderungen und Vitaldaten - Progressionsfreies Überleben (PFS): gemessen von der ersten Medikamentengabe bis zur Progression - *Disease Control Rate* (DCR; Rate der Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 12 Wochen) - Dauer des objektiven Ansprechen (DOR) - *Duration of Disease Control* (Dauer von erster Dosis bis zur Progression für Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 12 Wochen) - *Performance Status* (PS) basierend auf dem WHO-Kriterium - Häufigkeit des Stuhlgangs bei Patienten mit symptomatischem Durchfall - Änderungen des CTN und CEA Levels - PK und PD von Vandetanib - QTc Intervall Variablen für explorative Ziele: - Lebensqualität (FACT-G) - RET Onkoprotein, RET Genmutation - EGFR und die VEGFR-2 Signalbahnen - Serum Protein - Plasma VEGF, VEGFR2 und bFGF

Item[a]	Charakteristikum	Studieninformation
		- Urinäres Katecholamin, Metanephrin und Normetanephrin
8	Datenquellen/Messmethoden	Für das primäre Zielkriterium wurden radiologische Untersuchungen durchgeführt und entsprechend den RECIST Kriterien bewertet.
9	Bias	Beim FACT-G wurden entsprechende Maßnahmen getroffen, falls sich Hinweise auf eine Symptomatik bei den fehlenden Werten ergaben, um potentiellen Bias zu vermeiden. Falls ungeplanten Tumoruntersuchungen stattfanden, an denen keine Progression festgestellt werden konnte, sollte das die nächste Untersuchung trotzdem zum geplanten Zeitpunkt stattfinden, um keinen Bias, durch unterschiedliche Messpunkte zu erzeugen.
10	Studiengröße	Annahmen: Die Fallzahl von 15 Patienten wurde aus folgendem Grund gewählt: Sofern kein Ansprechen bei 15 Patienten vorliegt, liegt die Wahrscheinlichkeit, dass die ORR ≥ 20% ist, bei weniger als 5%.
11	Quantitative Variablen	Siehe Item 7
12	Statistische Methoden	
12a	Statistische Methoden, einschließlich der Methoden für die Kontrolle von *Confounding*	Grundlage: *Intention-to-Treat*-Prinzip Berechnung der Zielkriterien Berechnung von exakten 95%-Konfidenzintervallen zu den Endpunkten „ORR“, „CTN-Response“, „DCR“, Darstellung der Kaplan-Meier-Kurve zum Endpunkt „PFS“ „*Duration of Disease Control*“ und „DOR“, einschließlich dem medianen Überleben und entsprechendem 95% Konfidenzintervall. Berechnung des Mittelwerts, Medians, der Standardabweichung sowie von Maximum und Minimum für den FACT-G Score und die Laborwerte Berechnung der Korrelation des Serum-Proteinexpressionsprofils und der Tumorbiomarker mit PFS, ORR, DCR und DOR
12b	Subgruppen und Interaktionen	Nicht zutreffend

Item[a]	Charakteristikum	Studieninformation
12c	Umgang mit fehlenden Daten	Fehlende Werte beim FACT-G wurden entsprechend der üblichen Vorgaben beim FACT-G behandelt (d.h. 50% der Werte mussten vorhanden sein, damit eine FACIT Subskala ausgewertet wurde).
12d	*Kohortenstudie* (Umgang mit vorzeitigem Ausscheiden („loss to Follow-up“)) *Fallkontrollstudie* (Berücksichtigung des *Matchings* (Paarbildung) von Fällen und Kontrollen bei der Auswertung) *Querschnittsstudie* (Auswertungsmethoden, die die gewählte Strategie zur Stichprobenauswahl (*Sampling strategy*) berücksichtigen)	Patienten mit vorzeitigem Ausscheiden wurden als non-Responder betrachtet.
12e	Sensitivitätsanalysen	Nicht zutreffend
-	**Ergebnisse**	
13	Teilnehmer	
13a	Anzahl der Teilnehmer während jeder Studienphase (Anzahl der Teilnehmer, die potenziell geeignet waren, die auf Eignung untersucht wurden, die als geeignet bestätigt wurden, die tatsächlich an der Studie teilgenommen haben, deren Nachbeobachtung abgeschlossen wurde und deren Daten ausgewertet wurden)	22 Patienten wurden eingeschlossen. 19 dieser Patienten wurden mit Vandetanib behandelt (siehe auch Flow-Chart unten).
13b	Gründe für die Nicht-Teilnahme in jeder Studienphase	Nicht zutreffend
13c	Flussdiagramm	Siehe Flow Chart unten.
17	Weitere Auswertungen	Nicht zutreffend
a: nach STROBE 2007		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

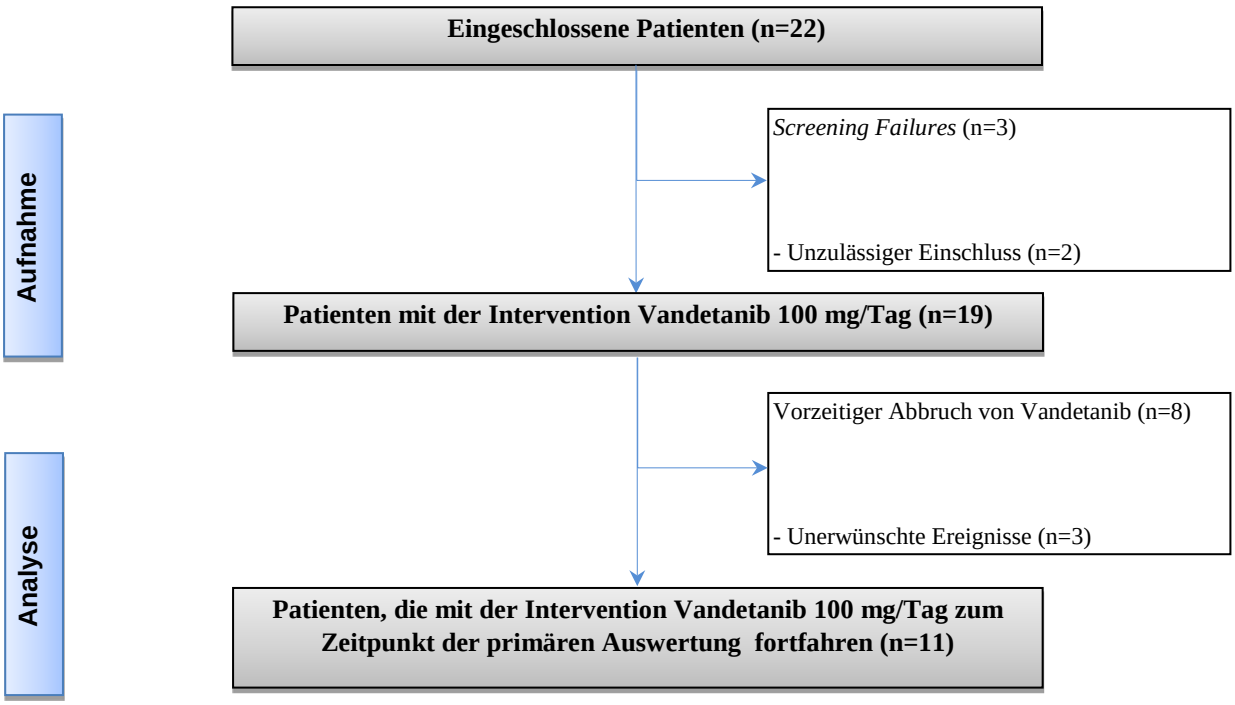


Abbildung 52: Flussdiagramm Studiendesign Studie D4200C00068

Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

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 4-190 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00058

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene**. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials**. 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 werden die Bögen studienbezogen ausgefüllt. 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.

Studie: _ D4200C00058

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Clinical Study Report: An International, Phase III, Randomized, Double-Blinded, Placebo-Controlled, Multi-Center Study to Assess the Efficacy of ZD6474 versus Placebo in Subjects with Unresectable Locally Advanced or Metastatic Medullary Thyroid Cancer	D4200C00058 CSR

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:
Es liegt eine randomisierte kontrollierte Studie vor.

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:

Adäquate Erzeugung der Randomisierungssequenz

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:

Adäquate Beschreibung der Verdeckung der Gruppenzuteilung

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 Anhaltspunkte 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 Behandlern

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:

Da Patienten gemäß *Good Clinical Practice* vor Beginn der Studie über zu erwartende Nebenwirkungen aufgeklärt werden müssen, ist die Gefahr einer Entblindung der Patienten und Behandler bei Auftreten charakteristischer Nebenwirkungen, insbesondere bei einem Vergleich gegen Placebo, gegeben. Daher wird endpunktspezifisch die mögliche Entblindung der Patienten diskutiert. Da für nahezu alle Endpunkte von einer Verblindung ausgegangen wird, wird auf Studienebene der Punkt mit „ja“ bewertet.

Behandler bzw. Weiterbehandler

☒ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Aufgrund des erwartenden Nebenwirkungsprofils ist die Gefahr einer Entblindung der Patienten und Behandler bei Auftreten charakteristischer Nebenwirkungen,

insbesondere bei einem Vergleich gegen Placebo, gegeben. Daher wird endpunktspezifisch die mögliche Entblindung der behandelnden Ärzte diskutiert. Da für nahezu alle Endpunkte von einer Verblindung ausgegangen wird, wird auf Studienebene der Punkt mit „ja“ bewertet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

☒ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

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

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Abweichung der geplanten Fallzahl. Anstelle von 232 Patienten wurden 331 Patienten randomisiert

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.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: _Progressionsfreies Überleben (PFS)_____

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

Die primäre Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt
.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es gibt keine Hinweise auf das Verzerrungspotential beeinflussende Punkte

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Zeit bis zur Schmerzprogression (TWP)_____

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

Eine mögliche partielle Entblindung von Patienten im Vandetanibgruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern und vom Patienten schwer zu differenzieren sind.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise, auf das Verzerrungspotential beeinflussender Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Objektive Ansprechrate (ORR)_____

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:
Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:
Es liegen keine Hinweise, auf das Verzerrungspotential beeinflussender Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Krankheitskontrollrate (DCR)_____

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Ein möglicher Verzerrungsapekt durch die Möglichkeit des *Cross-Over* auf *Open-Label* Vandetanib kann ausgeschlossen werden. Es liegen keine verzerrenden Aspekte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Dauer des objektiven Ansprechens (DOR)_____

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:
Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es gibt keine Hinweise auf, das Verzerrungspotential beeinflussende Punkte.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ explorativer Morbiditätsendpunkt: Gewichtszunahme

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:
Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergbnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beinflussender Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ explorativer Morbiditätsendpunkt: Einfluss der Behandlung auf Diarrhöen_____

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Der Anteil der Patienten, die in der Berechnung deskriptiver Statistiken nicht berücksichtigt wurde, war zu allen Zeitpunkten hoch und war zwischen den Behandlungsgruppen nicht balanciert

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es wurden keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung gefunden.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es wurden keine Hinweise auf, das Verzerrungspotential beinflussende Punkte gefunden

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.

falls hoch, obligate Begründung für die Einstufung:

Aufgrund der mangelhaften Umsetzung des ITT-Prinzips wird das Verzerrungspotential für diesen Endpunkt als „hoch“ bewertet

Endpunkt: _ explorativer Morbiditätsendpunkt: WHO-Performancestatus (WHO PS)__

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

Die Endpunkt-Analyse wurde durch ein unabhängiges, verblindetes, zentrales Reviewverfahren erhoben. Somit ist eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf, eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beeinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: _ explorativer Morbiditätsendpunkt: Zeit bis zur Schmerzprogression (TWP) in Abhängigkeit vom Opioid-Gebrauch zu Studienbeginn (Baseline)

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

Der Endpunkt wurde verbilndet erhoben

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf, eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beeinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: _ explorativer Morbiditätsendpunkt: Reduktion des Opioid-Gebrauchs gegenüber dem Ausgangswert zu Studienbeginn (Baseline)

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

Der Endpunkt wurde verbilndet erhoben

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf, eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beeinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Surrogat: Biochemisches Ansprechen (CTN;CEA)_____

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

Die Analyse der CTN- bzw. CEA-Spiegel wurde durch ein unabhängiges, verblindetes, Zentrallabor durchgeführt, wodurch eine mögliche Entblindung aufgrund des Nebenwirkungsprofils von Vandetanib ausgeschlossen ist

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beeinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: _Gesamtüberleben (OS)_____

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

Der Endpunkt „Tod“ ist stets objektiv zu bewerten und ist durch eine mögliche Entblindung aufgrund des typischen Nebenwirkungsprofils nicht beeinflusst.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Aussagekraft der Ergebnisse durch *Cross-Over* Option auf *Open-Label* Vandetanib in beiden Armen stark beeinflusst

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.

falls hoch, obligate Begründung für die Einstufung:

Aussagekraft der Ergebnisse durch *Cross-Over* Option auf *Open-Label* Vandetanib in beiden Armen stark eingeschränkt

Endpunkt: Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand FACT-G

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

Eine mögliche partielle Entblindung von Patienten im Vandetanibgruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern und vom Patienten schwer zu differenzieren sind.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Der Anteil der Patienten, die in der Berechnung deskriptiver Statistiken nicht berücksichtigt wurde, war zu allen Zeitpunkten hoch (bereits zu Woche 24 größer als 20% in der Vandetanib-Gruppe und fast 30% in der Placebo Gruppe) und war zwischen den Behandlungsgruppen nicht balanciert

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise für eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine Hinweise auf, das Verzerrungspotential beinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Aufgrund der mangelhaften Umsetzung des ITT-Prinzips wird das Verzerrungspotential für diesen Endpunkt als „hoch“ bewertet

Endpunkt: __Unerwünschte Ereignisse_____

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

Eine mögliche partielle Entblindung von Patienten im Vandetanibgruppe durch das Auftreten charakteristischer Nebenwirkungen erscheint unwahrscheinlich, da sich krankheitsbedingte Symptome und therapieassoziierte Nebenwirkungen überlagern und vom Patienten schwer zu differenzieren sind.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Das ITT-Prinzip wurde adäquat umgesetzt.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichtserstattung vor.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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:

Es liegen keine hinweise auf, das Verzerrungspotential beeinflussende Punkte vor.

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.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-191 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00008

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene**. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials**. 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 werden die Bögen studienbezogen ausgefüllt. 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.

Studie: _ D4200C00008**_____**

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Clinical Study Report: An Open-Label, Two-Stage, Phase II Study to Evaluate the Efficacy and Tolerability of ZD6474 in Patients With Locally Advanced or Metastatic Hereditary Medullary Thyroid Carcinoma	D4200C00008 CSR

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Keine Randomisierung fand aufgrund des einarmigen Studiendesigns statt.

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:

Die Studie wurde einarmig durchgeführt.

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 Anhaltspunkte 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:

Die Studie wurde einarmig durchgeführt.

3. Verblindung von Patienten und Behandlern

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:

Die Studie wurde einarmig und dadurch offen durchgeführt.

Behandler bzw. Weiterbehandler

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde einarmig und dadurch offen durchgeführt.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

☒ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

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

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

Alle patientenrelevanten Endpunkte wurden angemessen und wie geplant berichtet.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

_____

_____

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Objektive Ansprechrate (ORR) (primärer 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

   Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

   Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte

Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: _Progressionsfreies Überleben (PFS)_____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ WHO *Perfomance Status* _____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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:

Die Protokollverletzer und Lost-to-follow-up-Patienten sind dargestellt. Fehlende Werte wurden dargestellt.
Das ITT Prinzip kann durch das einarmige Studiendesign nicht verletzt werden.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gesundheitsbezogene Lebensqualität gemessen anhand FACIT-D_____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

---

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Unerwünschte Ereignisse_____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4-192 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00068

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene**. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials**. 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 werden die Bögen studienbezogen ausgefüllt. 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.

Studie: _ D4200C00068_____

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Clinical Study Report: A Phase II, Open-Label Study To Assess The Efficacy and Tolerability of ZD6474 (ZACTIMA™) 100 mg Monotherapy In Subjects with Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer	D4200C00068 CSR

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Keine Randomisierung fand aufgrund des einarmigen Studiendesigns statt.

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:

Die Studie wurde einarmig durchgeführt.

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 Anhaltspunkte 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:

Die Studie wurde einarmig durchgeführt.

3. Verblindung von Patienten und Behandlern

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:

Die Studie wurde einarmig und dadurch offen durchgeführt.

Behandler bzw. Weiterbehandler

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde einarmig und dadurch offen durchgeführt.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

☒ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: _Objektive Ansprechrate (ORR) (primärer 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

   Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

   Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte

Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: _Progressionsfreies Überleben (PFS)_____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ WHO *Perfomance Status* **_____**

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Gesundheitsbezogene Lebensqualität gemessen anhand FACT-G

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: __ Unerwünschte Ereignisse_____

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

Die Studie hatte ein einarmiges Studiendesign, so dass keine Verblindung möglich war.

2. **Adäquate Umsetzung des ITT-Prinzips**

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Patientenanzahlen 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung: