

Dokumentvorlage, Version vom 20.01.2011

**Dossier zur Nutzenbewertung  
gemäß § 35a SGB V**

*Vandetanib (Caprelsa®)*

AstraZeneca GmbH Deutschland

**Modul 4 A**

*Medulläres Schilddrüsenkarzinom*

Medizinischer Nutzen und  
medizinischer Zusatznutzen,  
Patientengruppen mit therapeutisch  
bedeutsamem Zusatznutzen

Stand: 14.03.2012

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>4</b>
<b>Abbildungsverzeichnis</b> .....	<b>9</b>
<b>Abkürzungsverzeichnis</b> .....	<b>10</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	15
4.2 Methodik.....	25
4.2.1 Fragestellung.....	25
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	29
4.2.3 Informationsbeschaffung.....	32
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	32
4.2.3.2 Bibliografische Literaturrecherche.....	32
4.2.3.3 Suche in Studienregistern.....	33
4.2.3.4 Selektion relevanter Studien.....	34
4.2.4 Bewertung der Aussagekraft der Nachweise.....	34
4.2.5 Informationssynthese und -analyse.....	37
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	37
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	37
4.2.5.3 Meta-Analysen.....	42
4.2.5.4 Sensitivitätsanalysen.....	44
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	45
4.2.5.6 Indirekte Vergleiche.....	46
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	48
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	48
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	48
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	48
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	49
4.3.1.1.3 Studien aus der Suche in Studienregistern.....	52
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	52
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	53
4.3.1.2.1 Studiendesign und Studienpopulationen.....	53
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	59
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	61
4.3.1.3.1 Gesamtüberleben (OS) – RCT.....	61
4.3.1.3.2 Progressionsfreies Überleben (PFS) – RCT.....	66
4.3.1.3.3 Objektive Ansprechrate (ORR) – RCT.....	72
4.3.1.3.4 Biochemisches Ansprechen (CTN; CEA) – RCT.....	77
4.3.1.3.5 Zeit bis zur Schmersverschlimmerung (TWP) – RCT.....	82
4.3.1.3.6 Gesundheitsbezogene Lebensqualität (QoL) gemessen anhand des FACT-G – RCT.....	87

4.3.1.3.7	Unerwünschte Ereignisse (UE) – RCT .....	93
4.3.1.3.8	Subgruppenanalysen – RCT.....	98
4.3.1.3.8.1	Progressionsfreies Überleben (PFS).....	100
4.3.1.3.8.2	Objektive Ansprechrte (ORR) .....	105
4.3.1.3.8.3	Biochemisches Ansprechen (CTN; CEA .....	105
4.3.1.3.8.4	Zeit bis zur Schmerzverschlimmerung (TWP).....	105
4.3.1.4	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien .....	106
4.3.2	Weitere Unterlagen.....	108
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	108
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	108
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	108
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	109
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	109
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	111
4.3.2.2	Nicht randomisierte vergleichende Studien.....	112
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	112
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	112
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	113
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	113
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	114
4.3.2.3	Weitere Untersuchungen.....	114
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	114
4.3.2.3.1.1	Studien des pharmazeutischen Unternehmers .....	115
4.3.2.3.1.2	Studien aus der bibliografischen Literaturrecherche .....	116
4.3.2.3.1.3	Studien aus der Suche in Studienregistern .....	118
4.3.2.3.1.4	Resultierender Studienpool: weitere Untersuchungen .....	118
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	118
4.3.2.3.2.1	Studiendesign und Studienpopulationen .....	118
4.3.2.3.2.2	Verzerrungspotenzial auf Studienebene .....	121
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen.....	122
4.3.2.3.3.1	Objektive Ansprechrte (ORR) – weitere Untersuchungen .....	122
4.3.2.3.3.2	Progressionsfreies Überleben (PFS) – weitere Untersuchungen...	123
4.3.2.3.3.3	WHO (PS) – weitere Untersuchungen.....	124
4.3.2.3.3.4	Lebensqualität (QoL) – weitere Untersuchungen.....	125
4.3.2.3.3.5	Unerwünschte Ereignisse (UE) – weitere Untersuchungen.....	126
4.3.2.3.3.6	Subgruppenanalysen – weitere Untersuchungen .....	127
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	127
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	128
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	128
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	130
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	138

4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	139
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	140
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	140
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	140
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	141
4.5.4	Verwendung von Surrogatendpunkten.....	141
4.6	Liste der eingeschlossenen Studien.....	143
4.7	Referenzliste.....	143
	<b>Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche .....</b>	<b>149</b>
	<b>Anhang 4-B : Suchstrategien – Suche in Studienregistern.....</b>	<b>150</b>
	<b>Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund .....</b>	<b>155</b>
	<b>Anhang 4-D : Liste der abgebrochenen Studien .....</b>	<b>156</b>
	<b>Anhang 4-E : Liste der laufenden Studien.....</b>	<b>157</b>
	<b>Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....</b>	<b>158</b>
	<b>Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>	<b>202</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Vandetanib .....	17
Tabelle 4-2: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Vandetanib .....	30
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	49
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	49
Tabelle 4-5: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel .....	52
Tabelle 4-6: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-7: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	54
Tabelle 4-8: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	55
Tabelle 4-9: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 1).....	55
Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 2: Altersgruppen) (Alle Angaben in %) .....	55
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 ) .....	55
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 3b: Systemische Vortherapie gegen MTC) (Alle Angaben in %).....	56
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 4: Primärtumor) (Alle Angaben in %).....	56
Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 5: Regionäre Lymphknoten) (Alle Angaben in %) .....	56
Tabelle 4-15: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 6: Fernmetastasen/ Stadienklassifikation) (Alle Angaben in %) .....	56
Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 7: Familiengeschichte mit MTC/ Begleitende Syndrome) (Alle Angaben in %).....	57
Tabelle 4-17: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 8: RET-Mutationsstatus sporadisch) (Alle Angaben in %).....	57
Tabelle 4-18: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 9: RET-Mutationsstatus hereditär) (Alle Angaben in %).....	57

Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %) .....	58
Tabelle 4-20: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 11: WHO-PS) (Alle Angaben in %).....	58
Tabelle 4-21: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	60
Tabelle 4-22: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	61
Tabelle 4-23: Operationalisierung von Gesamtüberleben (OS).....	62
Tabelle 4-24: Bewertung des Verzerrungspotenzials für Gesamtüberleben (OS) in RCT mit dem zu bewertenden Arzneimittel .....	63
Tabelle 4-25: Ergebnisse für Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel.....	64
Tabelle 4-26: Operationalisierung von Progressionsfreies Überleben (PFS) .....	66
Tabelle 4-27: Bewertung des Verzerrungspotenzials für Progressionsfreies Überleben (PFS) in RCT mit dem zu bewertenden Arzneimittel .....	67
Tabelle 4-28: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel .....	68
Tabelle 4-29: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel .....	69
Tabelle 4-30: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode).....	69
Tabelle 4-31: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel .....	70
Tabelle 4-32: Operationalisierung der objektiven Ansprechrate (ORR) .....	73
Tabelle 4-33: Bewertung des Verzerrungspotenzials für die objektive Ansprechrate (ORR) in RCT mit dem zu bewertenden Arzneimittel .....	73
Tabelle 4-34: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel .....	75
Tabelle 4-35: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem zu bewertenden Arzneimittel .....	75
Tabelle 4-36: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel .....	76
Tabelle 4-37: Operationalisierung des biochemischen Ansprechens (CTN; CEA).....	77
Tabelle 4-38: Bewertung des Verzerrungspotenzials für biochemisches Ansprechen (CTN; CEA) in RCT mit dem zu bewertenden Arzneimittel.....	78
Tabelle 4-39: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel .....	79
Tabelle 4-40: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel .....	80

Tabelle 4-41: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel .....	80
Tabelle 4-42: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel .....	81
Tabelle 4-43: Operationalisierung von TWP .....	82
Tabelle 4-44: Bewertung des Verzerrungspotenzials für Verzögerung der TWP in RCT mit dem zu bewertenden Arzneimittel .....	83
Tabelle 4-45: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel .....	85
Tabelle 4-46: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel.....	85
Tabelle 4-47: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel .....	86
Tabelle 4-48: Operationalisierung des Endpunktes Gesundheitsbezogene Lebensqualität .....	88
Tabelle 4-49: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel .....	88
Tabelle 4-50: Ergebnisse für gesundheitsbezogene Lebensqualität(HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel .....	89
Tabelle 4-51: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) zum Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel.....	91
Tabelle 4-52: Operationalisierung von Unerwünschte Ereignisse (UE).....	94
Tabelle 4-53: Bewertung des Verzerrungspotenzials für Unerwünschte Ereignisse in RCT mit dem zu bewertenden Arzneimittel .....	94
Tabelle 4-54: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	95
Tabelle 4-55: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	95
Tabelle 4-56: Ergebnisse für das Auftreten von QTc-bedingten Ereignissen und UEs der bevorzugten Bezeichnung „ <i>electrocardiogram QT prolonged</i> “ (nach MedDRA) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	96
Tabelle 4-57: Individuelle Interaktionstests (post-hoc) - PFS .....	100
Tabelle 4-58: PFS Subgruppenanalysen nach <i>Baseline</i> -Charakteristika .....	100
Tabelle 4-59: PFS Subgruppenanalysen nach Biomarkern.....	101
Tabelle 4-60: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung).....	102
Tabelle 4-61: PFS Subgruppenanalysen nach Plasma Biomarkern .....	103
Tabelle 4-62: PFS Subgruppenanalyse nach Altersgruppe .....	103
Tabelle 4-63: TWP Subgruppenanalyse nach Schmerzmittelgebrauch .....	105
Tabelle 4-64: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	109

Tabelle 4-65: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	109
Tabelle 4-66: Operationalisierung von <Endpunkt xxx>.....	110
Tabelle 4-67: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	110
Tabelle 4-68: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	111
Tabelle 4-69: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	113
Tabelle 4-70: Operationalisierung von <Endpunkt xxx>.....	113
Tabelle 4-71: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	114
Tabelle 4-72: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen.....	115
Tabelle 4-73: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – weitere Untersuchungen .....	116
Tabelle 4-74: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen mit dem zu bewertenden Arzneimittel .....	118
Tabelle 4-75: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen..	119
Tabelle 4-76: Charakterisierung der Intervention – weitere Untersuchungen .....	120
Tabelle 4-77: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 1).....	120
Tabelle 4-78: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 2).....	120
Tabelle 4-79: Verzerrungspotenzial auf Studienebene – Weitere Untersuchungen mit dem zu bewertenden Arzneimittel .....	121
Tabelle 4-80: Operationalisierung der objektiven Ansprechrate .....	122
Tabelle 4-81: Ergebnisse für die objektive Ansprechrate (ORR) – Weitere Untersuchungen.....	123
Tabelle 4-82: Operationalisierung des progressionsfreien Überleben .....	123
Tabelle 4-83: Ergebnisse für das Progressionsfreies Überleben (PFS) – Weitere Untersuchungen.....	124
Tabelle 4-84: Operationalisierung des WHO <i>Performance Status</i> .....	124
Tabelle 4-85: Ergebnisse des WHO <i>Performance Status</i> (WHO PS) – Weitere Untersuchungen.....	124
Tabelle 4-86: Operationalisierung der gesundheitsbezogenen Lebensqualität .....	125
Tabelle 4-87: Ergebnisse zur Gesundheitsbezogenen Lebensqualität – Weitere Untersuchungen.....	125
Tabelle 4-88: Operationalisierung der unerwünschten Ereignisse.....	126
Tabelle 4-89: Unerwünschte Ereignisse – Weitere Untersuchungen.....	126

Tabelle 4-90: Ausmaß des Zusatznutzens auf Endpunktebene: Vandetanib vs. Placebo .....	131
Tabelle 4-91: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	138
Tabelle 4-92: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens .....	140
Tabelle 4-93: Für die Nutzenbewertung berücksichtigte Studien und Untersuchungen.....	143
Tabelle 4-94 (Anhang): Im <i>International Clinical Trials Registry Platform Search Portal</i> identifizierte Studien .....	151
Tabelle 4-95 (Anhang): In <i>clinicaltrials.gov</i> identifizierte Studien .....	152
Tabelle 4-96 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	156
Tabelle 4-97 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	157
Tabelle 4-98 (Anhang): Studiendesign und -methodik für Studie D4200C00058.....	159
Tabelle 4-99 (Anhang): Studiendesign und -methodik für Studie D4200C00008.....	187
Tabelle 4-100 (Anhang): Studiendesign und -methodik für Studie D4200C00068.....	195
Tabelle 4-101 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00058 .....	202
Tabelle 4-102 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00008 .....	231
Tabelle 4-103 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie D4200C00068 .....	253

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	51
Abbildung 2: Kaplan-Meier-Kurve für Gesamtüberleben (ITT-Population).....	65
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	65
Abbildung 4: Kaplan-Meier-Kurve für Progressionsfreies Überleben (ITT-Population).....	70
Abbildung 5: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	72
Abbildung 6: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	77
Abbildung 7: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	81
Abbildung 8: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population)...	86
Abbildung 9: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	87
Abbildung 10: FACT-G über die Zeit – Veränderung der gesundheitsbezogenen Lebensqualität zum Ausgangswert.....	92
Abbildung 11: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	93
Abbildung 12: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	98
Abbildung 13: <i>Forest-Plot</i> der Subgruppenanalyse für PFS – <i>Baseline</i> - Charakteristiken und Krankheitsstatus .....	104
Abbildung 14: <i>Forest-Plot</i> der Subgruppenanalyse für PFS – Biomarker .....	104
Abbildung 15: <i>Forest-Plot</i> der Subgruppenanalyse für TWP.....	106
Abbildung 16: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen.....	117
Abbildung 17: Flussdiagramm Patientenverteilung Studie D4200C00058 .....	185
Abbildung 18: Flussdiagramm Studiendesign Studie D4200C00058.....	186
Abbildung 19: Flussdiagramm Studiendesign Studie D4200C00008.....	194
Abbildung 20: Flussdiagramm Studiendesign Studie D4200C00068.....	201

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
ALP	Alkaline Phosphatase
ALT	Alaninaminotransferase
AM-NutzenV	Arzneimittelnutzenverordnung
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)
DIMDI	Deutsches Institut für Medizinische Dokumentation
DOR	Dauer des Ansprechens (Duration of Response)
eCSF	electronic Case Report Form
EG	Europäische Gemeinschaft
EGF	Epidermaler Wachstumsfaktor (Epidermal Growth Factor)
EGFR	Rezeptor für den epidermalen Wachstumsfaktor (Epidermal Growth Factor Receptor)
ECHO	Echokardiogramm

<b>Abkürzung</b>	<b>Bedeutung</b>
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 Gesundheitswesen
ITT	Intention to Treat
IVRS	Interactive Voice Response Services
KI	Konfidenzintervall
LVEF	linksventrikuläre Ejektionsfraktion
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
NCCN	National Comprehensive Cancer Network
NYHA	New York Heart Association
OR	Odds Ratio (Quotenverhältnis)
ORR	Objektive Ansprechrate (Objective Response Rate)
OS	Gesamtüberleben (Overall Survival)

<b>Abkürzung</b>	<b>Bedeutung</b>
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)
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

<b>Abkürzung</b>	<b>Bedeutung</b>
	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)

#### 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 metastatischer 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 des MTC in der durch die Zulassung festgelegten Zielpopulation 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 derzeit verfügbare Behandlungsmöglichkeit verstanden. Es ist dabei darauf hinzuweisen, dass die Behandlung mit dem zu bewertenden Arzneimittel Vandetanib den Einsatz von BSC nicht ersetzen kann, sondern zusätzlich dazu erfolgt.

Die Quantifizierung des Zusatznutzens wird anhand der folgenden patientenrelevanten Endpunkte vorgenommen:

- Mortalität
  - Gesamtüberleben (*Overall Survival, OS*)
- Morbidität
  - Progressionsfreies Überleben (*Progression-Free Survival, PFS*)
  - Objektive Ansprechrates (*Objective Response Rate, ORR*)
  - Zeit bis zur Schmerzprogression (*Time to Worsening of Pain, TWP*)
- 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 für den klinischen Endpunkt Progression die Parameter CTN und CEA für das biochemische Ansprechen untersucht. Diese Endpunkte werden im Abschnitt 4.5.4 beschrieben und ihre Verwendung begründet.

### Datenquellen

MTC ist eine Erkrankung, die die Kriterien der Verordnung (EG) Nr.141/2000 des EU-Parlaments für seltene Leiden (2) in vollem Umfang erfüllt. 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 (3). 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 auf Antrag von AstraZeneca von der EMA zurückgezogen (4). Aus diesen Gründen besitzt Vandetanib gemäß § 12 Verfo des G-BA (5) nicht den Status eines Arzneimittels für seltene Leiden (*Orphan Drug*).

Die Bewertung soll, soweit bei dieser äußerst seltenen Indikation möglich, auf der Basis randomisierter kontrollierter Studien (RCT) erfolgen, die, 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 der Phase-III-RCT D4200C00058 („Studie 58“), die von AstraZeneca im Rahmen der Vorbereitung der amerikanischen und europäischen Zulassung an 331 Patienten durchgeführt worden ist. Daneben werden aufgrund der Seltenheit der Erkrankung nichtkontrollierte 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 mit Vandetanib beim MTC für die Bewertung berücksichtigt werden, ist darüber hinaus eine Studienregistersuche nach abgeschlossenen, abgebrochenen und laufenden Studien mit dem zu bewertenden Arzneimittel Vandetanib in den gängigen medizinischen Studienregistern durchgeführt worden (Abschnitt 4.2.3.3).

### Selektionskriterien für Studien

Die Selektionskriterien für den Ein- bzw. Ausschluss von für die Nutzenbewertung relevanten Studien sind in Tabelle 4-1 dargestellt.

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)	Erwachsene Patienten beiderlei Geschlechts mit aggressivem und symptomatischem medullärem Schilddrüsenkarzinom (MTC) in nicht resektablem lokal fortgeschrittenem oder metastatischem Stadium	Patienten unter 18 Jahren; Indikationen außerhalb des MTC
Intervention (I)	Behandlung mit Vandetanib 300 mg einmal täglich oral	Behandlung mit abweichenden Dosierungen von Vandetanib
Kontrollgruppe (C)	Behandlung mit Placebo/ <i>Best Supportive Care</i>	Vom Einschlusskriterium abweichende/keine Kontrollgruppe
Endpunkte (O)	Mindestens <b>einer</b> der patientenrelevanten Endpunkte wird berichtet: <ul style="list-style-type: none"> <li>• Gesamtüberleben (OS)</li> <li>• Progressionsfreies Überleben (PFS)</li> <li>• Objektive Ansprechrates (ORR)</li> <li>• Zeit bis zur Schmerzprogression (TWP)</li> <li>• Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G</li> <li>• Unerwünschte Ereignisse (UE)</li> </ul>	
Studientyp	Randomisierte kontrollierte Studie (RCT)	Vom Einschlusskriterium abweichendes Studiendesign
Studiendauer	Keine Einschränkungen; ein medianes <i>Follow-Up</i> für primäre Endpunktanalyse von möglichst etwa 24 Monaten (Hinweis: <i>Follow-Up</i> 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

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

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 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

#### *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“ Verzerrungspotenzial 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.

Die für die Nutzenbewertung von Vandetanib herangezogene randomisierte kontrollierte Studie wurde 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 RCT hinausgehend wurden unterstützende Informationen aus zwei nicht-randomisierten Studien herangezogen. Diese Studien sind gemäß der in Abschnitt 4.2.2 angegebenen 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 4-102) und Studie D4200C00068 (Tabelle 4-103) patientenrelevante Endpunkte im Sinne von Tabelle 4-1 (siehe oben) nur erfasst, insoweit sie in diesen Studien auch erhoben wurden.

### **Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen**

Die Patienten, der durch die Zulassung festgelegten Zielpopulation mit aggressivem und symptomatischem MTC und nicht resektabler, lokal fortgeschrittener oder metastasierter Erkrankung, versterben fast immer am MTC. Diesem steht bislang keine effektive Therapie gegenüber, die bestehenden Behandlungsmöglichkeiten dienen lediglich einer Linderung der Symptome der Erkrankung. Der langfristige Erhalt einer akzeptablen Lebensqualität steht für die betroffenen Patienten an erster Stelle. Eine neue Therapie sollte in der Lage sein, das Fortschreiten der Erkrankung und damit ihrer Begleitsymptome relevant hinauszuzögern, ein Tumorsprechen zu erreichen, und sie sollte darüber hinaus auch bei Einsatz über einen langen Zeitraum hinweg gut verträglich sein.

Vandetanib ist ein Multi-Tyrosinkinase-Inhibitor, der in der pivotalen Studie 58 (D4200C00058) in der Indikation „nicht resektables, lokal fortgeschrittenes oder

metastasiertes MTC“ verglichen mit Placebo eine signifikante Verbesserung im primären Endpunkt „progressionsfreies Überleben“ ebenso zeigen konnte wie im sekundären Endpunkt „Zeit bis zur Schmerzprogression“.

Vandetanib 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 (6, 7), des bei der Entstehung aller erblichen Varianten von MTC und bis zu 80% aller Fälle von sporadischem MTC ursächlich beteiligten Protoonkogens. 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.

Im Rahmen der randomisierten, kontrollierten Studie D4200C00058 konnte für die patientenrelevanten Endpunkte „Progressionsfreies Überleben (*Progression-Free Survival*, PFS)“, „Objektive Ansprechrates (*Objective Response Rate*, ORR)“ und „Zeit bis zur Schmerzprogression (*Time to Worsening of Pain*, TWP)“ ein statistisch signifikanter Vorteil von Vandetanib gegenüber der mit Placebo behandelten Kontrollgruppe, welche ansonsten nur „*Best Supportive Care* (BSC, beide Behandlungsgruppen)“ erhielt, gezeigt werden. Zusätzlich wurden als Surrogate für den klinischen Endpunkt Progression die Parameter CTN und CEA für das biochemische Ansprechen untersucht.

### Gesamtüberleben (OS)

Aufgrund der geringen Ereignisraten zum Zeitpunkt des *Cut-off* und des möglichen Einflusses des *Cross-Over* – und einer damit möglicherweise vorliegenden Verzerrung der Ergebnisse – wurde die statistische Auswertung des Gesamtüberlebens nicht für die Nutzenbewertung herangezogen. Das Gesamtüberleben wurde dennoch in Abschnitt 4.3.1.3.1.1 dargestellt, um die Evidenz zu Vandetanib zu vervollständigen. Eine erste Analyse der Gesamtüberlebensrate zeigte ein *Hazard Ratio* von 0,89 zugunsten von Vandetanib. Statistische Signifikanz wurde nicht erreicht (KI: [0,28;2,85],  $p=0,713$ ).

### Verbesserung des Progressionsfreien Überlebens (PFS)

Bezüglich des PFS wurde in der Vandetanib-Gruppe ein statistisch signifikanter Vorteil gegenüber Placebo beobachtet (HR:0,46 95%-KI: [0,31; 0,69]). Die Wahrscheinlichkeit des PFS (Kaplan-Meier- Schätzer) war in der Vandetanib-Gruppe zu allen Zeitpunkten um ca. 20% höher: zu Monat 6 (91,4% vs. 74,5%), Monat 12 (83,3% vs. 62,7%) und zu Monat

24 (58,8% vs. 41,7%). Das mediane PFS war zum Analyse *Cut-off* in der Vandetanib-Gruppe noch nicht erreicht, wurde jedoch durch Anpassung eines Weibull-Modells auf 30,5 Monate geschätzt und lag um ca. 11 Monate höher als in der Vergleichsgruppe (19,3 Monate). Angesichts des in Abschnitt 4.3.1.3.1.2 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie und der Patientenrelevanz dieses Endpunktes ergibt sich für diesen Endpunkt der Beleg eines erheblichen Zusatznutzens (siehe Tabelle 4-90).

#### Objektive Ansprechrates (ORR)

Anhand der objektiven Ansprechrates (*Objective Response Rate*, ORR) konnte eine Verringerung der Tumorlast durch Vandetanib gezeigt werden. Eine partielle Remission (gemäß RECIST Kriterien) erreichten 45% der Patienten in der Vandetanib-Gruppe – im Vergleich zu 13% der Patienten unter BSC (wobei für 12 von 13 Patienten in der Placebo-Gruppe ein Ansprechen erst nach dem *Cross-Over* auf Vandetanib erreicht wurde). Keiner der Patienten in beiden Gruppen erreichte eine komplette Remission. Angesichts des statistisch signifikanten Ergebnisses und der Effektgröße für das objektive Ansprechen (OR für Nicht-Ansprechen: 0,18, 95%-KI: [0,09;0,33]) sowie der Aussagekraft des erbrachten Nachweises ergibt sich der Beleg eines erheblichen Zusatznutzens (siehe Tabelle 4-90).

#### Zeit bis zur Schmerzprogression (TWP)

Auch 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 opiathaltigen Schmerzmitteln, war unter Vandetanib im Median mehr als doppelt so lang wie in der Vergleichsgruppe (7,85 Monate gegenüber 3,25 Monate). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (HR: 0,61, 95%-KI: [0,43;0,87]). 62,5% der mit Vandetanib behandelten Patienten konnten ihren Verbrauch von opiathaltigen Schmerzmitteln senken; dies gelang 37,5% der Patienten auf Placebo, welche als aktive Therapie nur BSC erhielten. Angesichts des statistisch signifikanten Ergebnisses sowie der Aussagekraft des erbrachten Nachweises kann hieraus der Beleg eines beträchtlichen Zusatznutzens abgeleitet werden (siehe Tabelle 4-90).

#### Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G

Daten zur gesundheitsbezogenen Lebensqualität wurden rein exploratorisch analysiert. Die Angaben der Patienten bezüglich ihrer Lebensqualität im Hinblick auf ihre Gesundheit in den FACT-G Fragebögen zeigten jedoch in den beiden untersuchten Behandlungsgruppen Vandetanib und BSC keine statistisch signifikanten Unterschiede. Es ist anzunehmen, dass das objektive Ansprechen durch 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 Analyse nicht. Ein Zusatznutzen bezüglich des Endpunktes „gesundheitsbezogene Lebensqualität“ ist für Vandetanib nicht belegt (siehe Tabelle 4-90).

### Unerwünschte Ereignisse

Bei nahezu allen Patienten beider Gruppen trat mindestens ein unerwünschtes Ereignis auf. Die Gesamtrate unerwünschter Ereignisse in der Vandetanib- bzw. Placebo-Gruppe betrug 99,6% bzw. 90,9%. Das Ergebnis für schwerwiegende unerwünschte Ereignisse, Abbrüche aufgrund unerwünschter Ereignisse und für unerwünschte Ereignisse vom Schweregrad CTCAE Grad 3 oder höher zeigte signifikant höhere Raten in der Vandetanib-Gruppe (CTCAE Grad 3 oder höher 55,4% vs. 24,2%; SUE 30,7% vs. 13,1%) Der Anteil der Patienten mit unerwünschten Ereignissen, die zum Absetzen der Studienbehandlung führten (12,1% vs. 3,0%, OR: 4,41, 95%-KI: [1,31; 14,8]), war unter Vandetanib im Vergleich zu Placebo größer. Für diese Endpunkte zeigte sich somit ein Beleg für einen größeren Schaden (siehe Tabelle 4-90). 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). Zusammenfassend ist ein größerer Schaden durch Vandetanib somit belegt; sein Ausmaß kann in Betracht der Gesamtraten unerwünschter Ereignisse als eher zu vernachlässigen bezeichnet werden.

### 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 eine progrediente Erkrankung als Surrogat für den patientenrelevanten Endpunkt Morbidität (krankheitsbedingte Symptome). Signifikante Ergebnisse des biochemischen Ansprechens stehen in Zusammenhang mit diesem Endpunkt. Durchfall und Flush-Syndrom treten häufig infolge humoraler Faktoren auf (krankheitsbedingt hohe CTN-Werte). Das biochemische Ansprechen für CTN und CEA zeigt sich in einer Spiegelreduzierung für 69,3% bzw. 51,5% der mit Vandetanib behandelten Patienten im Vergleich zu 3,0% bzw. 2,0% bei Behandlung gemäß BSC (OR für Ansprechen: 72,86, 95%-KI: [26,22;303,2] bzw. 52,03, 95%-KI: [15,95;320,3] / OR für Nicht-Ansprechen: 0,01, 95%-KI: [0,00;0,04] bzw. 0,02, 95%-KI: [0,00;0,08]).

### **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 Patienten auf, deren Morbidität durch das um mehr als 11 Monate verlängerte progressionsfreie Überleben, eine Verringerung der Tumorlast und eine Verbesserung der Schmerzsituation im Vergleich mit der derzeit einzig möglichen Behandlungsalternative, BSC, relevant reduziert wird. 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. Angesichts der Schwere der Erkrankung und des erheblichen *medical need* in der Zielpopulation ist der Zusatznutzen von Vandetanib auch nach Abwägung der Vorteile (erheblicher Zusatznutzen in fast jedem der betrachteten Endpunkte) und des Schadens in der Gesamtschau zum gegenwärtigen Zeitpunkt als **erheblich** einzustufen, da dem gegenüber zwar eine höhere als unter BSC gemessene Rate unerwünschter Ereignisse steht, es aber nicht angemessen erscheint, in der Gesamtschau den aufgezeigten positiven Nutzen deshalb abzuwerten:

1. auch aus klinisch-therapeutischer Sicht dürfte unter allen direkt gemessenen Endpunkten die mindestens 11-monatige Verlängerung des für diese unheilbaren Patienten hochrelevanten Endpunktes PFS von vorrangigem Wert sein und
2. waren die unerwünschten Ereignisse unter Vandetanib in der Regel hinreichend gut beherrschbar und nur 12% der Patienten brachen die Behandlung wegen unerwünschter Ereignisse ab.

Zum Zeitpunkt der primären Analyse der pivotalen Studie sind erst 15% der Patienten verstorben gewesen, so dass eine abschließende Bewertung des Gesamtüberlebens erst nach der finalen Analyse durchgeführt werden kann.

Zudem basieren die Feststellungen der Zulassungsbehörde hinsichtlich des positiven Nutzen-Risiko-Verhältnisses von Vandetanib – welchen gemäß AM-NutzenV diese Nutzenbewertung nicht widersprechen darf – auf einem Vergleich mit BSC. Folglich muss Vandetanib allein schon auf dieser Grundlage ein Zusatznutzen bescheinigt werden, da gemäß Fachinformation ein unterschiedliches Nutzenausmaß bei verschiedenen Subgruppen als denkbar angesehen werden kann, dies andererseits aber nicht zum Ausschluss der Subgruppen mit möglicherweise geringerem Nutzen von der Zulassung geführt hat.

Unter Berücksichtigung obiger Erwägungen und in Anbetracht der zusätzlichen Tatsache, dass es sich bei Vandetanib um die erste 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 vorliegt:

- 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 hervorzugehen scheint- mit dem Kriterium der Nachhaltigkeit grundsätzlich den Anspruch an langfristige symptomfreie Zeiträume von mehreren Jahren verbindet. Die Patienten aus der Zielpopulation von Vandetanib haben nach derzeitigem Kenntnisstand eine mediane verbleibende Lebenserwartung von insgesamt nur noch 2 bis 3 Jahren; vor diesem Hintergrund müssen eine PFS-Verlängerung um mindestens 11 Monate und eine Verlängerung der Zeitdauer bis zur Schmerzprogression um ca. 5 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 (8). 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 (9).

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% (10). Bei Patienten mit Fernmetastasen liegt sie jedoch nur bei etwa 40% mit einer medianen Gesamtüberlebensrate von 2-3 Jahren (11, 12). 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 (10-12). Bei etwa 50% aller MTC-Patienten flammt die Erkrankung nach Behandlung (kurative Operation/Thyreoidektomie) erneut auf (10). 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 (11, 13).

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 (14).

Zum gegenwärtigen Zeitpunkt ist 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 nach einer 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 ist 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 (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 einer randomisierten kontrollierten Studie (RCT).

#### Patientenpopulation

Die Zielpopulation der zu bewertenden Intervention sind erwachsene Patienten mit aggressivem und symptomatischem MTC in nicht resektablem, lokal fortgeschrittenem oder metastasiertem Stadium.

#### Intervention

Das zu bewertende Arzneimittel ist Vandetanib 300 mg Tablette (Handelsname Caprelsa®) in der zugelassenen Dosierung von 300 mg einmal täglich oral.

### Vergleichstherapie

Die Standardtherapie beim MTC ist die Operation in kurativer Intention. Für das fortgeschrittene bzw. metastasierte MTC existiert zu diesem Zeitpunkt keine zugelassene bzw. allgemein empfohlene Therapie.

Die aktuellen Behandlungsempfehlungen und Guidelines der *European Society for Medical Oncology* (ESMO; (15)) und der *American Thyroid Association* (ATA; (16)) raten vom Einsatz zytotoxischer Chemotherapie zur Behandlung von Patienten mit fortgeschrittenem MTC aufgrund der unbefriedigenden Resultate ab. Als Erstlinientherapie empfehlen beide Fachgesellschaften stattdessen die Behandlung der Patienten mit TKI im Rahmen klinischer Studien; die aktuelle Leitlinie (Version 3.2011) des *National Comprehensive Cancer Network* (NCCN) zum Schilddrüsenkarzinom nennt bei MTC bereits explizit Vandetanib als in Betracht zu ziehende Option (17).

Entsprechend erhalten MTC-Patienten der Zielpopulation anstatt einer systemischen Therapie unterstützende Behandlungen im Rahmen der *Best Supportive Care* (BSC). Diese umfasst die nach dem Stand der medizinischen Wissenschaften bestmöglichen unterstützenden Behandlungsmaßnahmen, also die individuelle Behandlung physischer sowie psychischer Beschwerden mit den Zielen (18):

- 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 sicherte 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 (19). Sowohl die Patienten in der Vandetanib-Gruppe als auch in der Placebogruppe erhielten eine als BSC eingestufte Begleitbehandlung. Somit bildet die Placebogruppe BSC im Versorgungsalltag der Zielpopulation von Vandetanib weitestgehend ab.

### 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:

- Gesamtüberleben (*Overall Survival, OS*)
- Progressionsfreies Überleben (*Progression-Free Survival, PFS*)
- Objektive Ansprechrate (*Objective Response Rate, ORR*)
- Zeit bis zur Schmerzprogression (*Time to Worsening of Pain, TWP*)
- Gesundheitsbezogene Lebensqualität (*Health Related Quality of Life, HRQoL*, anhand des FACT-G Fragebogens)
- Unerwünschte Ereignisse (UE)

### 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 RCT hinausgehend werden unterstützende Informationen aus zwei nicht randomisierten Studien herangezogen, um die Evidenz zu Vandetanib im MTC zu vervollständigen. Diese Studien sind gemäß der in Abschnitt 4.2.2 angegebenen 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. Auch die in diesen Studien erhobenen Endpunkte wurden der Vollständigkeit halber in Anhang 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 4-102) und Studie D4200C00068 (Tabelle 4-103) patientenrelevante Endpunkte im Sinne von Tabelle 4-2 (siehe unten) nur erfasst, 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 (20).

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 mit nicht resektablem, lokal fortgeschrittenen oder metastasiertem medullärem Schilddrüsenkarzinom (*Medullary Thyroid Carcinoma*, MTC) eingeschlossen und anhand patientenrelevanter Endpunkte ausgewertet. Aufgrund der Seltenheit der Erkrankung ist die in die Nutzenbewertung eingeschlossene Studie weltweit die erste Phase-III-Studie, 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 4-2.

Tabelle 4-2: Prädefinierte Ein- und Ausschlusskriterien zur Selektion relevanter klinischer Studien zum medizinischen Nutzen des zu bewertenden Arzneimittels Vandetanib

Kategorie	Prädefinierte Einschlusskriterien	Prädefinierte Ausschlusskriterien
Patienten (P)	Erwachsene Patienten beiderlei Geschlechts mit aggressivem und symptomatischem medullärem Schilddrüsenkarzinom (MTC) in nicht resektablem lokal fortgeschrittenem oder metastatischem Stadium	Patienten unter 18 Jahren; Indikationen außerhalb des MTC
Intervention (I)	Behandlung mit Vandetanib 300 mg einmal täglich oral	Behandlung mit abweichenden Dosierungen von Vandetanib
Kontrollgruppe (C)	Behandlung mit Placebo/ <i>Best Supportive Care</i>	Vom Einschlusskriterium abweichende/keine Kontrollgruppe
Endpunkte (O)	Mindestens <b>einer</b> der patientenrelevanten Endpunkte wird berichtet: <ul style="list-style-type: none"> <li>• Gesamtüberleben (OS)</li> <li>• Progressionsfreies Überleben (PFS)</li> <li>• Objektive Ansprechrates (ORR)</li> <li>• Zeit bis zur Schmerzprogression (TWP)</li> <li>• Gesundheitsbezogene Lebensqualität /HRQoL) gemessen anhand des FACT-G</li> <li>• Unerwünschte Ereignisse (UE)</li> </ul>	
Studientyp	Randomisierte kontrollierte Studie (RCT)	Vom Einschlusskriterium abweichendes Studiendesign
Studiendauer	Keine Einschränkungen; ein medianes <i>Follow-Up</i> für primäre Endpunktanalyse von möglichst etwa 24 Monaten (Hinweis: <i>Follow-Up</i> 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 (21)). Studien mit signifikant davon abweichender Patientenpopulation wurden ausgeschlossen.

Das Einschlusskriterium bezüglich Population spiegelt den Zulassungsstatus von Vandetanib wider. 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 (22). 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. Der Rate an diesbezüglichen UE und SUE in beiden Armen der Studie nach zu urteilen (je >90%), ist von einer relevanten Inkongruenz der Population der Zulassungsstudien und der Population gemäß Zulassungsstatus nicht auszugehen. Für das Einschlusskriterium bezüglich der Population reicht es für die Nutzenbewertung aus, wenn mindestens 80% der in die Studie eingeschlossenen Patientinnen und Patienten dieses Kriterium erfüllen (23). Diese Forderung ist durch die Zulassungsstudie erfüllt.

Die Wahl des Kriteriums zur 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 Placebogruppe ist als Kontrollgruppe geeignet, solange in dieser Gruppe begleitende Maßnahmen zur Symptomkontrolle erlaubt waren. 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 Fall von Vandetanib nicht erforderlich, da keine Zulassungen vorliegen, die mehr als 12 Monate vor Einreichung des Dossiers erteilt wurden. Zwei Zulassungen von Vandetanib wurden in den USA am 06.04.2011 (24) und in Kanada am 12.01.2012 erteilt (*Approval Letter* zum Zeitpunkt der Einreichung des Dossiers noch nicht veröffentlicht). Literaturrecherchen für etwaige indirekte Vergleiche entfielen, da die Nutzenbewertung anhand eines direkten Vergleichs durchgeführt und keine indirekten Vergleiche zur Bewertung des medizinischen Zusatznutzens von Vandetanib herangezogen wurden.

#### **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](http://clinicaltrials.gov), [clinicalstudyresults.org](http://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.

Das Datum der Suche, sowie die einzelnen Suchstrategien für jede Datenbankabfrage sind in Anhang 4-B dokumentiert.

#### **4.2.3.4 Selektion relevanter Studien**

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

Die durch die gemäß Abschnitt 4.2.3.3 durchgeführte Studienregister-Abfrage identifizierten Studien wurden gemäß den in Abschnitt 4.2.1 prädefinierten Kriterien zum Ein- bzw. Ausschluss relevanter Studien (Tabelle 4-2) von zwei unabhängigen Personen selektiert. Auf Basis aller verfügbaren Informationen aus 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 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 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

*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“ Verzerrungspotenzial 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 nichtkontrollierten 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.

## 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)<sup>1</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-<sup>2</sup> bzw. STROBE-Statements<sup>3</sup> 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

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

<sup>2</sup> 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.

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

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

*Benennen Sie die Patientencharakteristika und 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 Studienpopulation der in die Nutzenbewertung eingehenden RCT wurde anhand folgender Charakteristika beschrieben:

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

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

- Gesamtüberleben (*Overall Survival, OS*)
- Progressionsfreies Überleben (*Progression-Free Survival, PFS*)
- Objektive Ansprechrate (*Objective Response Rate, ORR*)
- Zeit bis zur Schmerzprogression (*Time to Worsening of Pain, TWP*)
- Gesundheitsbezogene Lebensqualität (*Health Related Quality of Life, HRQoL*, anhand des FACT-G Fragebogens)
- Unerwünschte Ereignisse (UE)

Alle aufgeführten Zielgrößen stellen patientenrelevante Endpunkte dar. Die Patientenrelevanz begründet sich wie folgt:

### 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 (25). 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 (26). 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 die Zielpopulation von Vandetanib, eine onkologische Population, die trotz des aggressiven Verlaufs der Erkrankung eine vergleichsweise lange Überlebensdauer hat (27-30).

Für klinische Zulassungsstudien bei onkologischen Therapien schreiben die Zulassungsbehörden und Ethikkommissionen zudem häufig ein *Cross-Over* im Studiendesign vor, welches teilweise schon bei der Studienplanung vorgesehen werden muss, oder nachträglich für bereits laufende Zulassungsstudien gefordert wird, wenn sich etwa zeigt, dass die Patienten in der Verumgruppe stärker profitieren als die Patienten in der Placebogruppe. *Cross-over* Studiendesign bedeutet, dass Patienten 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 nachgewiesene 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 *Cross-Over* Design sichtbar wären, während der inkrementelle Effekt sinkt (26).

### Progressionsfreies Überleben

In der palliativen onkologischen Therapie stellt ein Herausögern des Fortschreitens der Erkrankung und der damit verbundenen Symptomatik eines der wichtigsten Therapieziele dar, insbesondere bei einem zu erwartenden langen Gesamtüberleben (31-33). 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 (34), 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

schwer belasteten Patienten wird zudem der Erfolg der Behandlung durch den Stillstand der Erkrankung vor Augen geführt, was umso mehr von Bedeutung sein kann, da die Erfolgsaussichten einer medikamentösen Intervention mit Antidepressiva für psychische Anpassungsstörungen bei fortgeschrittener Tumorerkrankung bzw. damit einhergehenden Komplikationen gering sind (35). 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 (36):

1. Die Reduzierung tumorbedingter Symptome aufgrund einer Verringerung der Tumorgroße oder Unterdrückung des fortschreitenden Wachstums
2. Die Hinauszögerung 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 Med. 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 (33). Hier steht nicht nur die Hinauszögerung 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 zur Auseinandersetzung und Akzeptanz der neuen Lebenssituation sind für den Patienten von großer Bedeutung. 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 der RECIST Kriterien (37, 38). PFS eignet sich besonders für fortgeschrittene Krankheitsstadien (28) und bietet den Vorteil kleinerer Patientenzahlen in den klinischen Studien und einer kürzeren Beobachtungszeit im Vergleich zum Gesamtüberleben (32), welches insbesondere im Falle einer langen Überlebenszeit nach der Progression oft keine statistisch signifikanten Aussagen mehr über die untersuchte Therapie erlaubt (39). 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 (36). Da in onkologischen Studiendesigns ein Wechsel der Patienten in den Behandlungsarm mit aktiver Medikation (*Cross-Over*) meist erst nach Progression vorgesehen ist, hat das *Cross-Over* keinen verzerrenden Einfluss auf die Auswertung des PFS, wie dies hingegen in der Regel beim Gesamtüberleben gegeben ist (26, 27, 40-42). Die Deutsche Gesellschaft für Hämatologie und Onkologie 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 (32). Bei der Beurteilung von PFS sollte allerdings berücksichtigt werden, dass die Verzögerung der Progression und der damit einhergehenden zunehmenden Morbidität nur dann einen Vorteil für die Patienten darstellt, wenn gleichzeitig das Nebenwirkungsprofil der Therapie und die Lebensqualität gut oder zumindest akzeptabel ist (41, 42). FDA und EMA haben bereits mehrere neue Medikamente auf Basis einer PFS-Verbesserung zugelassen (41-45).

### 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 der standardisierten und etablierten RECIST-Kriterien bewertet, um die Veränderungen der Tumorgröße im zeitlichen Verlauf darzustellen. Die ORR misst dabei die Verringerung der Tumorlast und umfasst sowohl komplettes als auch partielles Ansprechen, allerdings nicht den Stillstand der Erkrankung (25, 28, 41) 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 dieser Endpunkt in vielen Studien als primärer Endpunkt für die Wirksamkeit als Zulassungsgrundlage bei der FDA (25, 41), 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 (27, 29).

### Zeit bis zur Schmerzprogression

Schmerz stellt eine bedeutsame Komponente der Symptomatik des aggressiven, nicht resektablen, lokal fortgeschrittenen oder metastasierten MTC dar. 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 opiathaltiger Schmerzmittel abgeleitet. Für die Patienten ist eine möglichst langfristige Verzögerung der Schmerzprogression von großer Bedeutung, da Schmerzen die Lebensqualität stark beeinträchtigen und auch der Einsatz von Schmerzmitteln von entsprechenden Nebenwirkungen begleitet ist.

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 (46). 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 (47).

### 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 (48) 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 wurden. 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 (49).

### Unerwünschte Arzneimittelwirkungen

Ergänzend zum Nachweis des Nutzens dient die Erfassung unerwünschter Arzneimittelwirkungen zur Nutzen-Risiko-Bewertung. 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<sup>4</sup> 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 Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet<sup>5</sup> 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<sup>6</sup> 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.

---

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

<sup>5</sup> 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.

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

#### 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.*

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. die 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 aufgrund der Überrekrutierung (Ergebnis nur auf der Basis der gemäß Fallzahlplanung zu rekrutierenden Patienten) (4.3.1.2.2)
- Adjustierte Analysen für PFS, OS, TWP und objektive Ansprechrate 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)
- In einer Sensitivitätsanalyse für den Endpunkt „Zeit bis zur Schmerzprogression“ (TWP) wurden Modifikationen der Definition des TWP ausgewertet, um eine valide Beurteilung der Schmerzprogression unter Berücksichtigung beider Komponenten, des Schmerzes (gemessen anhand valdierter Skalen) und des Schmerzmittelgebrauchs, abzugeben. Hierbei wurden unterschiedliche Strategien für die Ersetzung fehlender Werte verwendet.

#### 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.*

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

- Geschlecht (männlich, weiblich)

- Rasse (kaukasisch, andere)
- WHO *Performance Status* (WHO-PS), (0,  $\geq 1$ )
- RET Mutationssatus (positiv, negativ, unbekannt)
- CTN Verdopplungszeit ( $\leq 24$  Monate,  $> 24$  Monate, unbekannt)
- CEA Verdopplungszeit ( $\leq 24$  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 ( $< 65$ ,  $\geq 65$ )

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 Effektmodifikation wurden mithilfe von Interaktionstest abgeleitet. Ein p-Wert kleiner als 0,2 wurde als Hinweis auf eine Effektmodifikation gewertet.

Für die Endpunkte „Objektive Ansprechrates“, „gesundheitsbezogene Lebensqualität“ und „unerwünschte Ereignisse“ waren keine Subgruppenanalysen geplant.

Für den Endpunkt „Zeit bis zur Schmerzprogression“ wurde eine explorative Subgruppenanalyse, getrennt nach *Baseline*-Opioidanalgetika-Gebrauch, durchgeführt.

#### 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<sup>7</sup>. 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“<sup>8</sup>, „Multiple-Treatment-Meta-Analysen“<sup>9</sup> oder auch

<sup>7</sup> 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.

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

„Netzwerk-Meta-Analysen“<sup>10</sup>, 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<sup>11</sup>.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen<sup>12</sup>.

*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. BSC

---

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

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

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

<sup>12</sup> 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.

wurde in einem Beratungsgespräch vom G-BA (29.09.2011, siehe Modul 3A, Abschnitt 3.1.2) als zweckmäßige Vergleichstherapie festgelegt (1).

### **4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen**

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

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

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

#### **4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel**

##### **4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel**

###### **4.3.1.1.1 Studien des pharmazeutischen Unternehmers**

*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 4-3: 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

*Geben Sie an, welchen Stand die Information in Tabelle 4-3 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 untersucht (Vgl. Anhang 4-B). Der Stand der Information in Tabelle 4-3 ist vom 05. Januar 2012.

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

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

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
-	-

*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.*

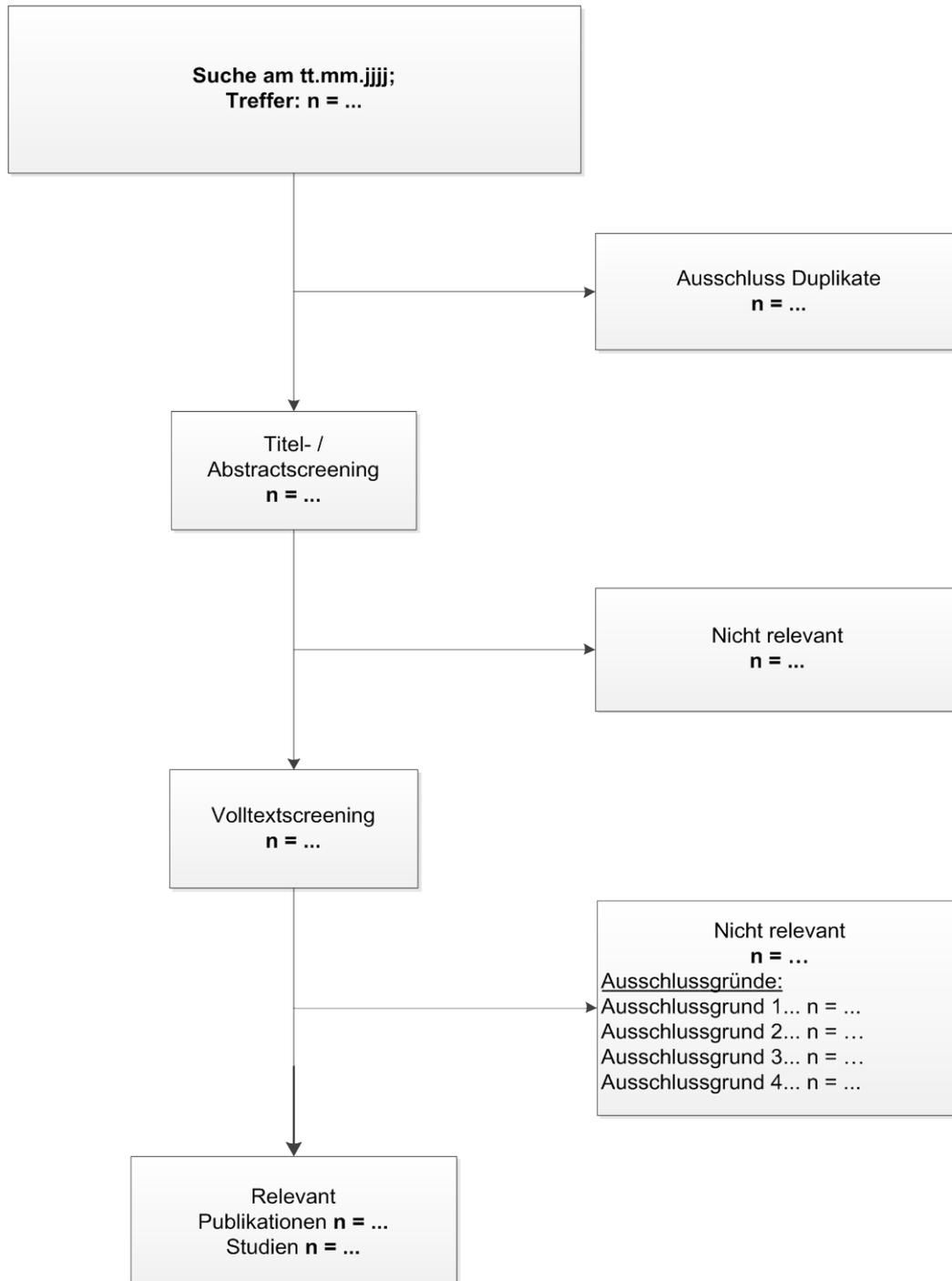


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die Voraussetzungen für eine erforderliche Literaturrecherche sind nicht gegeben. Zum einen wurde für das zu bewertende Arzneimittel eine Zulassung zum ersten Mal am 06.04.2011, also noch innerhalb der unter Abschnitt 4.2.3.2 benannten 12-Monats-Frist, in den USA

erteilt, zum anderen wurden keine indirekten Vergleiche durchgeführt. Deshalb entfällt eine bibliografische Literaturrecherche.

#### 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 4-3) 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 <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. <sup>b</sup> )
D4200C00058	International Clinical Trials Registry Platform Search Portal [ <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00410761">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00410761</a> ], clinicaltrials.gov [ <a href="http://www.clinicaltrials.gov/ct2/show/NCT00410761">http://www.clinicaltrials.gov/ct2/show/NCT00410761</a> ]	ja	n. d.

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 <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
D4200 C0005 8	ja	ja	nein	ja (19)	ja (50) ( <a href="#">NCT00410761</a> )	Ja (51)
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
--						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

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

##### 4.3.1.2.1 Studiendesign und Studienpopulationen

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

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-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 ( <i>Cross-Over</i> )	Erwachsene Patienten mit messbarem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom	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	<u>Primärer Endpunkt:</u> Progressionsfreies Überleben (PFS); <u>Patientenrelevante sekundäre Endpunkte:</u> <ul style="list-style-type: none"> <li>• Gesamtüberleben (OS)</li> <li>• Zeit bis zur Schmerzprogression (TWP)</li> <li>• objektive Ansprechrate (ORR)</li> <li>• gesundheitsbezogene Lebensqualität (HRQoL) anhand des FACT-G</li> <li>• unerwünschte Ereignisse (UE)</li> </ul>

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

Studie	Vandetanib	Placebo	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.</i>
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									
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

S=Schwarz, K=Kaukasisch, O=Orientalisch, A=Andere

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				
Vandetanib	21,6	57,1	18,2	3,0
Placebo	10,0	70,0	17,0	3,0

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	N	Bestes Ansprechen bei vorangegangener Chemotherapie gegen Medulläres Schilddrüsenkarzinom (MTC)						
		Komplettes Ansprechen	Partielles Ansprechen	Stabile Erkrankun g	Progression	Nicht auswertbar	Nicht anwendbar	Unbekannt
D4200C00058								
Vandetanib	50	0,0	0,9	6,1	7,8	0,4	6,5	0,0
Placebo	18	1,0	0,0	11	1,0	1,0	3,0	1,0

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		
Vandetanib	39,0	61,0
Placebo	42,0	58,0

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						
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

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						
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

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							
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

Tabelle 4-16: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 7: Familiengeschichte mit MTC/ Begleitende Syndrome) (Alle Angaben in %)

Studie Gruppe	Familiengeschichte mit MTC				Begleitende Syndrome			
	Ja	Nein	Unbekannt	FMTC	MEN2a	MEN2b	Keins	Unbekannt
D4200C00058								
Vandetanib	42,9	42,9	14,3	14,3	50,0	25,0	7,1	3,6
Placebo	80,0	20,0	0,0	20,0	60,0	0,0	20,0	0,0

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

Studie Gruppe	RET Mutationsstatus positiv (Vandetanib: 59,3; Placebo: 50,0)										
	Sporadisch (Vandetanib: 45,0; Placebo: 43,0)										
	<i>Mutationsposition</i>										
	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 9: RET-Mutationsstatus hereditär) (Alle Angaben in %)

Studie Gruppe	RET Mutationsstatus positiv (Vandetanib: 59,3; Placebo: 50,0)											
	Hereditär (Vandetanib: 11,7; Placebo: 5,0)										Unbekannt (Vandetanib: 2,6; Placebo: 2,0)	
	<i>Mutationsposition</i>											
	364	609	611	618	620	634	781	791	804	918	634	918
D4200C00058												
Vandetanib	0,4	0,4	0,4	0,9	0,9	4,8	0,0	0,4	0,4	3,5	0,4	2,2
Placebo	0,0	0,0	0,0	1,0	1,0	2,0	1,0	0,0	0,0	0,0	0,0	2,0

Tabelle 4-19: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 10: RET-Mutationsstatus negativ/unbekannt) (Alle Angaben in %)

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

Tabelle 4-20: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Teil 11: WHO-PS) (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					
Vandetanib	67,0	29,0	4,0	0,0	0,0
Placebo	58,0	38,0	4,0	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 D4200C00058 ist die einzige randomisierte Studie, die gemäß der Selektionskriterien relevant ist für die Nutzenbewertung von Vandetanib zur Behandlung des medullären Schilddrüsenkarzinoms bei erwachsenen Patienten mit aggressivem, symptomatischem, nicht resektablem, lokal fortgeschrittenem oder metastasiertem MTC. 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). 231 Patienten wurden der Behandlung mit Vandetanib zugeteilt und 100 Patienten der Placebogruppe. Es erhielten 231 Patienten die Behandlung mit Vandetanib und 99 Patienten Placebo, nachdem ein Patient in der Placebogruppe 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 Einnahmeregelerletzungen, *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 Studie entlassen und hatten die Option, in die offene Phase der Studie überzutreten und mit Vandetanib behandelt zu werden (*Cross-Over*). Alle Patienten werden für das Gesamtüberleben nachbeobachtet. Hinsichtlich Alter, Geschlecht, Größe und Rasse gab es keine größeren Unterschiede zwischen der Vandetanib- und der Placebo-Gruppe. Mit 53,4 Jahren in der Placebo-Gruppe und 50,7 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. In der Vandetanib-Gruppe sind 21,6% in der jüngsten Patientengruppe ( $\geq 18$  -  $< 40$ ), hingegen befinden sich in der vergleichbaren Placebogruppe nur 10% der Patienten. 70% der Placebo-Patienten entstammen der Altersklasse ( $\geq 40$  -  $< 65$ ), deren Anteil im Behandlungsarm mit Vandetanib liegt hingegen bei nur 57,1%. Der Anteil der kaukasischen Patienten mit 97% in der Placebo-Gruppe war etwas größer als in der Vandetanib-Gruppe mit 94,4%.

Hinsichtlich des besten Ansprechens bei vorheriger MTC Chemotherapie liegen für „Stabile Erkrankung“ (Vandetanib vs. Placebo: 6,1% vs. 11,0%) und „Progression“ (Vandetanib vs. Placebo 7,8% vs. 1,0%) deutliche Unterschiede zwischen den Behandlungsarmen vor. 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 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 4-21: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
D4200C00058	ja	ja	ja	ja	nein	ja*	niedrig

\* Abweichung von der geplanten Fallzahlplanung: Anstelle von 232 Patienten wurden 331 Patienten randomisiert.

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Für die Studie D4200C00058 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, der Therapie des nicht resektablen, lokal fortgeschrittenen oder metastasierten medullären Schilddrüsenkarzinoms, bei Patienten mit einem symptomatisch-aggressiven Verlauf der Erkrankung, 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 beinhaltet. Die Auswertung der Wirksamkeits-Endpunkte erfolgte zentral.

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 Prüfbüros 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 (52).

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.

#### 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 4-22: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	OS	PFS	ORR	Biochemisches Ansprechen*	TWP	Unerwünschte Ereignisse	FACT-G
D4200C00058	ja	ja	ja	ja	ja	ja	ja

\* Surrogatendpunkt

#### 4.3.1.3.1 Endpunkte – RCT

##### 4.3.1.3.1.1 Gesamtüberleben (OS) – 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 4-23: 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 4-24: 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-Modell enthielt folgende Kovariaten: RET Mutationsstatus, Zeitraum zur Verdopplung des CTN-Wertes, Zeitraum 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 Placebogruppe zum *Cross-Over* nach Krankheitsprogression stark eingeschränkt. Als Resultat des *Cross-Over* wird der Effekt im Placeboarm 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.

Die Gesamtüberlebensrate soll erneut untersucht werden, wenn 50% der Patienten verstorben sind. Diese Ergebnisse müssen dann, unter Einbezug der zu Vandetanib übergegangenen Placebo-Patienten (*Cross-Over*), erneut interpretiert werden.

Unter Berücksichtigung aller oben genannter Aspekte und durch den Umstand, dass die Ergebnisse der Gesamtüberlebensrate durch die geringe Zahl an beobachteten Todesfällen sowie das *Cross-Over* verzerrt sein könnten, wird im Rahmen des Dossiers auf die weitere Darstellung 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 4-25: Ergebnisse für Gesamtüberleben (OS) aus RCT mit dem zu bewertenden Arzneimittel

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]	0,7675

\* 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 zur Verdopplung des Calcitoninwertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär, sporadisch, unbekannt).

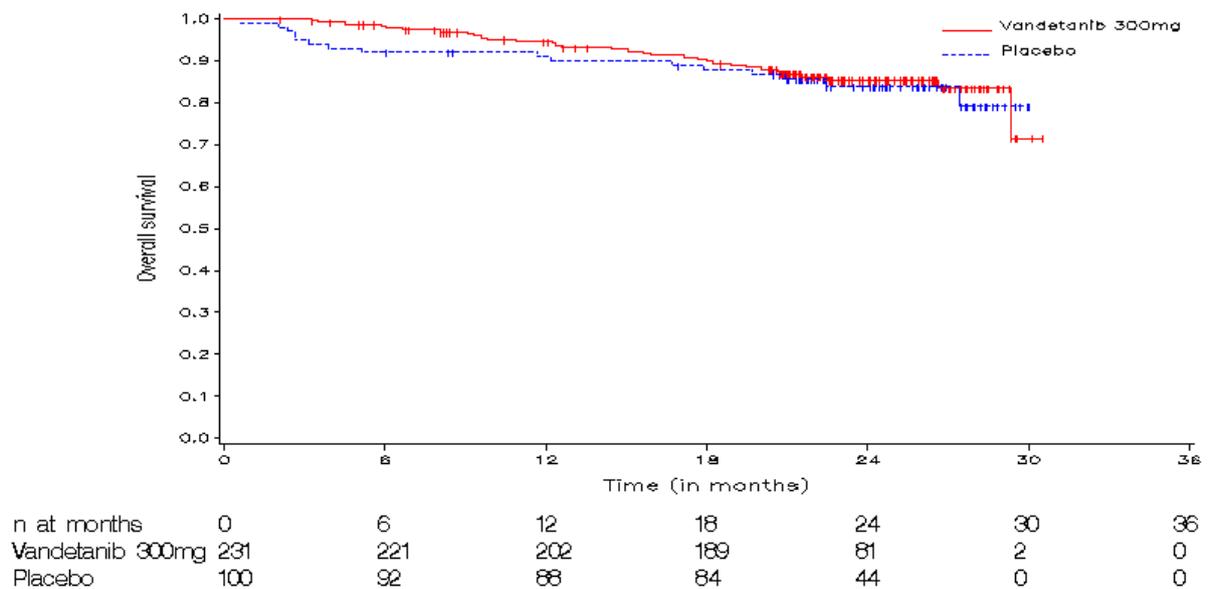


Abbildung 2: Kaplan-Meier-Kurve für Gesamtüberleben (ITT-Population)

Eine erste Analyse der Gesamtüberlebensrate (Tabelle 4-25) zeigte ein *Hazard Ratio* von 0,89 zugunsten von Vandetanib. Statistische Signifikanz wurde nicht erreicht (KI: [0,28; 2,85],  $p=0,713$ ). Diese Ergebnisse sind unter Vorbehalt zu interpretieren, da wie bereits erwähnt, zum Zeitpunkt der Analyse (Daten *Cut-off*) nur 15% der Patienten verstorben waren. Zusätzlich – durch das Protokoll vorab definiert – bestand die Möglichkeit für Patienten aus der Placebogruppe, während der offenen Studienphase Vandetanib zu erhalten (*Cross-Over*). Infolgedessen sind die Ergebnisse der Placebo-Gruppe verschleiert und der geschätzte Behandlungseffekt kaum interpretierbar. Abbildung 2 spiegelt obige Ergebnisse wider. Eine endgültige Aussage zum Gesamtüberleben kann erst nach der finalen Auswertung nach Eintreten von  $\geq 50\%$  Ereignissen getroffen werden.

*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 3: 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 Progressionsfreies Überleben (PFS) – 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 4-26: Operationalisierung von Progressionsfreies Überleben (PFS)

Studie	Operationalisierung
D4200C00058	Zeit bis zum Auftreten der ersten objektiven Krankheitsprogression nach RECIST v1.0 (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 4-27: 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 RECIST v1.0 (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 zu Monat 6, 12 und 24 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 zur Verdopplung des CTN-Wertes, Zeitraum 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 D4200C00058 eine Analyse der ITT-Population 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.

*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 4-28: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Vandetanib		Placebo		Behandlungseffekt (Vandetanib:Placebo)	
	N	Anzahl Ereignisse (%)	N	Anzahl Ereignisse (%)	<i>Hazard Ratio</i> [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 zur Verdopplung des Calcitoninwertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie MTC-Status (hereditär,sporadisch, unbekannt).

Tabelle 4-29: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel

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

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

Tabelle 4-30: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel (Kaplan-Meier-Methode)

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 4-31: Ergebnisse für Progressionsfreies Überleben (PFS) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Progressionsstatus	Ereignistyp	Anzahl Patienten (%)		
			Vandetanib (N=231)	Placebo (N=100)	Total (N=331)
D4200C00058	Progression	Total	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	Total	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 (53,4)
		Vor Progression ausgeschieden	16 (16,9)	5 (5,0)	21 (6,3)
		Sicherheitsgründe	1 (0,4)	0 (0,0)	1 (0,3)
		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

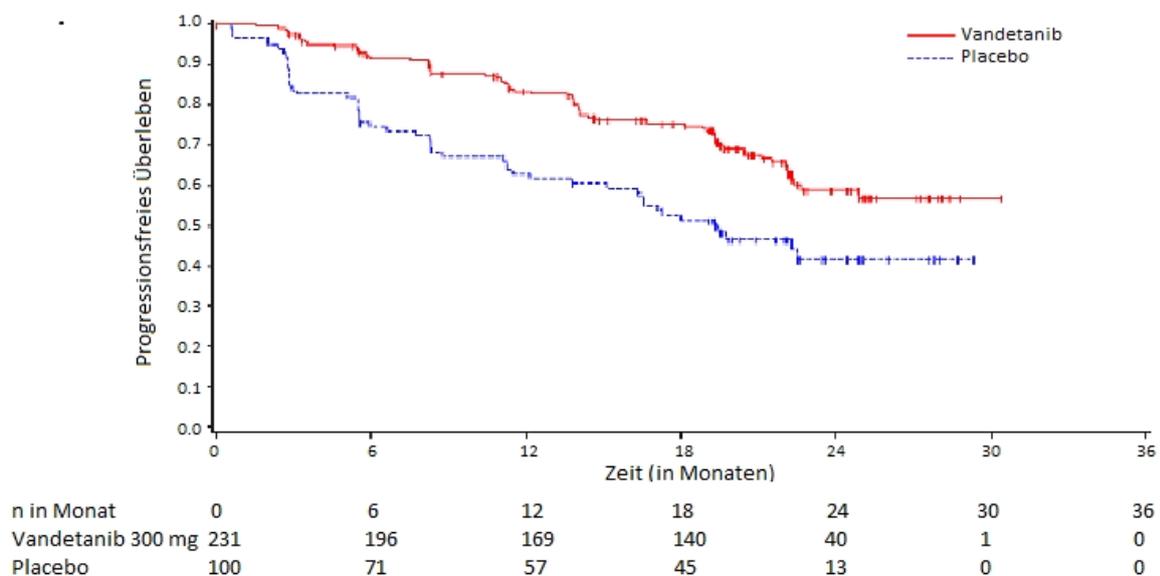


Abbildung 4: Kaplan-Meier-Kurve für Progressionsfreies Überleben (ITT-Population)

Bezüglich des PFS (Tabelle 4-28) zeigt sich 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 4-30).

In Tabelle 4-29 zeigt sich, dass der Anteil an progressionsfreien Patienten (Kaplan-Meier-Schätzer) in der Vandetanib-Gruppe zu jedem der Zeitpunkte höher war als in der Vergleichsgruppe. Zu Monat 6 betrug der Anteil der Patienten ohne Progression in der Vandetanib-Gruppe 91,4% gegenüber 74,5% in der Placebo-Gruppe. Zu Monat 12 lagen die Anteile der progressionsfreien Patienten bei 83,3% (Vandetanib), bzw. 62,7% (Placebo) und zu Monat 24 bei 58,5% bzw. 41,7%.

Abbildung 4 unterstreicht die oben genannten Ergebnisse anhand der grafischen Darstellung der Kaplan-Meier-Kurve für das PFS.

In Tabelle 4-31 wird der PFS-Status am Tag des Daten *Cut-off* angegeben. Es hat sich gezeigt, dass der Anteil der progressiven Patienten in der Vandetanib-Gruppe mit 31,6% um knapp 20% geringer ist als in der Placebogruppe, wo etwas über die Hälfte der Patienten (51%) eine Krankheitsprogression erlitten haben. Die Aufteilung der Progressionsfälle auf RECIST Progression und Tod ist in beiden Gruppen verhältnismäßig gleich.

Somit ist der Anteil der Patienten, 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 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%) oder Ausscheiden, bevor eine Progression auftreten konnte (16,9% vs. 5,0%). Gründe für das vorzeitige Ausscheiden waren Sicherheitsgründe (0,4% vs. 0,0%), Protokollabweichung (0,4% vs. 1,0%) und Beendigung auf Wunsch des Patienten (6,1% vs. 4,0%).

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]). Auch bei Adjustierung für mögliche prognostische Faktoren (RET Mutationsstatus, Zeitraum zur Verdopplung des CTN-Wertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien und den MTC-Status) ergab sich ein statistisch signifikanter Vorteil von Vandetanib gegenüber Placebo hinsichtlich der PFS (HR<sub>adj</sub>:0,46, 95%-KI: [0,32;0,68]).

*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 5: 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.

#### **4.3.1.3.1.3 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 4-32: Operationalisierung der objektiven Ansprechrates (ORR)

Studie	Operationalisierung
D4200C00058	Anteil der Patienten, deren bestes objektives Ansprechen im Zeitraum vom Behandlungsbeginn bis zur Progression (BOR) 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 4-33: Bewertung des Verzerrungspotenzials für die objektive Ansprechrates (ORR) 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 Objektive Ansprechrates 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 Ansprechrates 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 aller 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.

Neben der objektiven Ansprechrates wurde entsprechend den RECIST Kriterien (zentraler Review) die Krankheitskontrollrate aus dem besten objektiven Ansprechen abgeleitet. Die Krankheitskontrollrate stellt eine Erweiterung der objektiven Ansprechrates dar. Neben dem partiellen oder kompletten Ansprechen berücksichtigt sie zudem den stabilen Krankheitsstatus ( $\geq 24$  Wochen) seit Randomisierung. Durch die Implementierung der ORR in die Krankheitskontrollrate spiegeln sich statistisch signifikante Ergebnisse der ORR auch in der Krankheitskontrollrate wieder. Im Umkehrschluss bedeutet dies, dass statistisch signifikante Ergebnisse der Krankheitskontrollrate die Ergebnisse der objektiven Ansprechrates bestätigen.

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 zur Verdopplung des CTN-Wertes, Zeitraum 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 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 von *Cross-Over* wird eher zuungunsten von Vandetanib angenommen. Mögliche Abweichungen werden aufgrund der kleinen Patientenzahl (von 13 Patienten mit Ansprechen in der Placebogruppe) als „gering“ eingestuft.

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 4-34: Ergebnisse für die objektive Ansprechrate (ORR) aus RCT mit dem zu bewertenden Arzneimittel

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 zur Verdopplung des CTN-Wertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien sowie den MTC-Status.

Tabelle 4-35: Ergebnisse für die objektive Ansprechrate (ORR) ohne Berücksichtigung der Bewertungen der offenen Behandlungsphase (Sensitivitätsanalyse) aus RCT mit dem zu bewertenden Arzneimittel

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 4-36: Ergebnisse für die objektive Ansprechrate (ORR) nach RECIST aus RCT mit dem zu bewertenden Arzneimittel

Studie	Ansprech-Status	Bestes objektives Ansprechen (BOR)	Anzahl (%) der Patienten		
			Vandetanib 300 mg (N=231)	Placebo (N=100)	Total (N=331)
D4200C00058	Ansprechen	Total	104 (45,0)	13 (13,0) <sup>a</sup>	117 (35,3)
		Komplettes Ansprechen	0 (0,0)	0 (0,0)	0 (0,0)
		Partielles Ansprechen	104 (45,0)	13 (13,0) <sup>a</sup>	117 (35,3)
	Kein Ansprechen	Total	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)		

<sup>a</sup> Ansprechbeginn war in 12 von 13 Fällen während der offenen (*open label*) Behandlungsphase mit Vandetanib.

Tabelle 4-34 zeigt, dass es eine statistisch signifikante Verbesserung des ORR für Vandetanib verglichen mit Placebo gibt. Das *Odds Ratio* liegt bei 5,48 mit einem 95%igen Konfidenzintervall von [2,99;10,79] und einem p-Wert von <0,0001. Dieses Ergebnis spricht deutlich für eine bessere ORR innerhalb der Vandetanib-Gruppe im Vergleich zu der Placebo-Gruppe. Insgesamt betrug die ORR 45% in der Vandetanib-Gruppe und 13% in der Placebo-Gruppe. Von den 13 Patienten mit Ansprechen in der Placebo-Gruppe hatte lediglich 1 Patient bereits unter Placebo ein beginnendes Ansprechen. Für die verbleibenden 12 Patienten der Placebo-Gruppe setzte die Remission erst während der offenen Behandlungsphase unter Vandetanib ein. Grundlage für den Wechsel (*Cross-Over*) war ein positiver Review durch den Prüfarzt, welcher nicht immer durch den zentralen Review bestätigt werden konnte (Tabelle 4-36). Diese Vorgehensweise birgt einen Vorteil zugunsten von Placebo. Das Resultat der Sensitivitätsanalyse ohne Berücksichtigung der Daten nach *Cross-Over* zeigt folglich einen noch deutlicheren Effekt für Vandetanib (Tabelle 4-35): Objektives Ansprechen erlebten 101/231 Patienten unter Vandetanib und 1/100 Patienten unter Placebo (OR [95%-KI]: 76,91 [16,68; 1366]).

*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 6: 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 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 4-37: Operationalisierung des biochemischen Ansprechens (CTN; CEA)

Studie	Operationalisierung
--------	---------------------

D4200C00058	<p>Die biochemische Ansprechrate für CTN bzw. CEA ist der Anteil der Patienten, deren bestes biochemisches Ansprechen einem kompletten Ansprechen (CR, <i>Complete Response</i>) oder partiellen Ansprechen (PR, <i>Partial Response</i>) entspricht. Es waren nur Erhebungen während der randomisierten Behandlung zu berücksichtigen. Die Analyse der Spiegel erfolgte durch ein Zentrallabor.</p> <p><u>CR</u>: Vollständige Normalisierung des CTN- bzw. CEA-Spiegels (CTN: <math>\leq 10</math> pg/ml für Männer und <math>\leq 5</math> pg/ml für Frauen; CEA: <math>\leq 2,5</math> pg/ml), die durch eine Wiederholungsmessung <math>&gt;4</math> Wochen später bestätigt sein muss</p> <p><u>PR</u>: Eine Abnahme des CTN- bzw. CEA-Spiegels um mindestens 50% vom CTN- bzw. CEA-Ausgangswert (<i>Baseline</i>), die durch eine Wiederholungsmessung <math>&gt;4</math> Wochen später bestätigt sein muss</p>
-------------	--

*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 4-38: 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:  $\leq 10$  pg/ml für Männer und  $\leq 5$  pg/ml für Frauen; CEA:  $\leq 2,5$  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 zu Visiten (außer zu Woche 1 und 2) unter randomisierter Behandlung gewonnen wurden. Ein fehlender Wert führte zu einem nicht auswertbaren biochemischen Ansprechen. 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 wurde durch ein unabhängiges, verblindetes, Zentrallabor durchgeführt, 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 wurden, liegt hier kein Verzerrungspotential auf Grund eines vorzeitigen Wechsels (*Cross-Over*) von Placebo zu 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 Berichterstattung noch sonstige das Verzerrungspotential beeinflussende Punkte gefunden werden.

*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 4-39: Ergebnisse für die biochemische Ansprechrate (CTN) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Vandetanib 300 mg		Placebo		Behandlungseffekt (Vandetanib:Placebo)	
	N	Anzahl Patienten mit Ansprechen (%)	N	Anzahl Patienten mit Ansprechen (%)	<i>Odds Ratio</i> [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 4-40: Ergebnisse für das beste objektive biochemische Ansprechen (CTN) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Ansprech- Status	Bestes objektives Ansprechen	Anzahl (%) der Patienten		
			Vandetanib 300 mg (N=231)	Placebo (N=100)	Total (N=331)
D4200C00058	Ansprechen	Total	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	Total	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 4-41: Ergebnisse für die biochemische Ansprechrate (CEA) aus RCT mit dem zu bewertenden Arzneimittel

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 4-42: Ergebnisse für das beste objektive biochemische Ansprechen (CEA) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Ansprech-status	Bestes objektives Ansprechen (BOR)	Anzahl (%) der Patienten		
			Vandetanib 300 mg (N=231)	Placebo (N=100)	Total (N=331)
D4200C00058	Ansprechen	Total	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	Total	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 4-39 und Tabelle 4-41 zeigen eine statistisch signifikante Verbesserung der biochemischen Ansprechrate für Vandetanib verglichen mit Placebo. Für CTN liegt das *Odds Ratio* bei 72,86 mit einem 95%igen Konfidenzintervall von [26,33;303,2] und einem p-Wert von <0,0001. Die CTN-Ansprechraten betragen 69,3% für Vandetanib und 3,0% für Placebo. Für CEA liegt das *Odds Ratio* bei 52,03 mit einem 95%igen Konfidenzintervall von [15,95;320,3] und einem p-Wert von <0,0001. Die CEA-Ansprechraten betragen 51,1% für Vandetanib und 2,0% für Placebo.

*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.

#### 4.3.1.3.1.5 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 4-43: Operationalisierung von TWP

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

*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 4-44: Bewertung des Verzerrungspotenzials für Verzögerung der TWP 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.

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 Opiatanalgetika abgeleitet. Der kombinierte Endpunkt für den TWP wurde erreicht, wenn der Patient eine Progression des „schwersten Schmerzes“ ( $\geq 2$  Punkte im *Worst Pain Score* des BPI-SF) berichtete oder Opiatanalgetika 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 „Schlimmster 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, sowohl ihren Schmerz zum Zeitpunkt der Beantwortung des Fragebogens (über die letzten 24 Stunden), als auch den stärksten, geringsten und durchschnittlichen Schmerz der letzten vorangegangenen Wochen zu bewerten. 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 mittels Patienten-Telefoninterviews gemäß Studienplan berichtet.

Patientenberichteter 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.

Für „Schlimmster Schmerz“ war „Progression“ definiert als eine Erhöhung um  $\geq 2$  Punkte zum Ausgangswert. „Verbesserung“ war definiert als eine Reduzierung um  $\geq 2$  Punkte zum Ausgangswert, „Schlimmster Schmerz“ ohne eine Erhöhung der Schmerzmittel im Vergleich zum Ausgangswert um  $\geq 10$  mg/Tag Morphinsulfat-Äquivalent. „Sonstige“ war definiert als keine Veränderung.

Für den Schmerzmittelgebrauch war „Verschlechterung“ definiert als eine Erhöhung um  $\geq 10$  mg/Tag Morphinsulfat Äquivalent zum Ausgangswert. „Verbesserung“ war definiert als Reduzierung um  $>50\%$  des Gebrauchs von Schmerzmitteln gegenüber dem letzten Besuch, ohne Erhöhung des „Schlimmster Schmerz“ um  $\geq 2$  zum 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 wurde 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 erfolgte 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 zur Verdopplung des CTN-Wertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien, MTC-Status (hereditär, sporadisch, unbekannt). Weitere Sensitivitätsanalysen umfassten Modifikationen der Definition des TWP: Erstens wurde die TWP allein anhand des „schwersten Schmerzes“ auf dem BPI-SF ohne Berücksichtigung des Schmerzmittelgebrauchs hergeleitet, zweitens wurden dabei fehlende Werte von Patienten ersetzt, die aufgrund des im Präzenträum lokalen *reads* der CT/MRT-Aufnahmen wegen Progression zensiert wurden.

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.

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 4-45: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel

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 zur Verdopplung des CTN-Wertes, Zeitraum 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 4-46: Sensitivitätsanalyse für TWP ohne Berücksichtigung des Schmerzmittelgebrauchs aus RCT mit dem zu bewertenden Arzneimittel

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	107 (46,3)	100	52 (52,0)	0,72 [0,50;1,02]	0,0665

\* 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 wurden ersetzt, die aufgrund des im Präzentrum lokalen reads der CT/MRT scans wegen Progression zensiert wurden.

Tabelle 4-47: Ergebnisse für TWP aus RCT mit dem zu bewertenden Arzneimittel

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

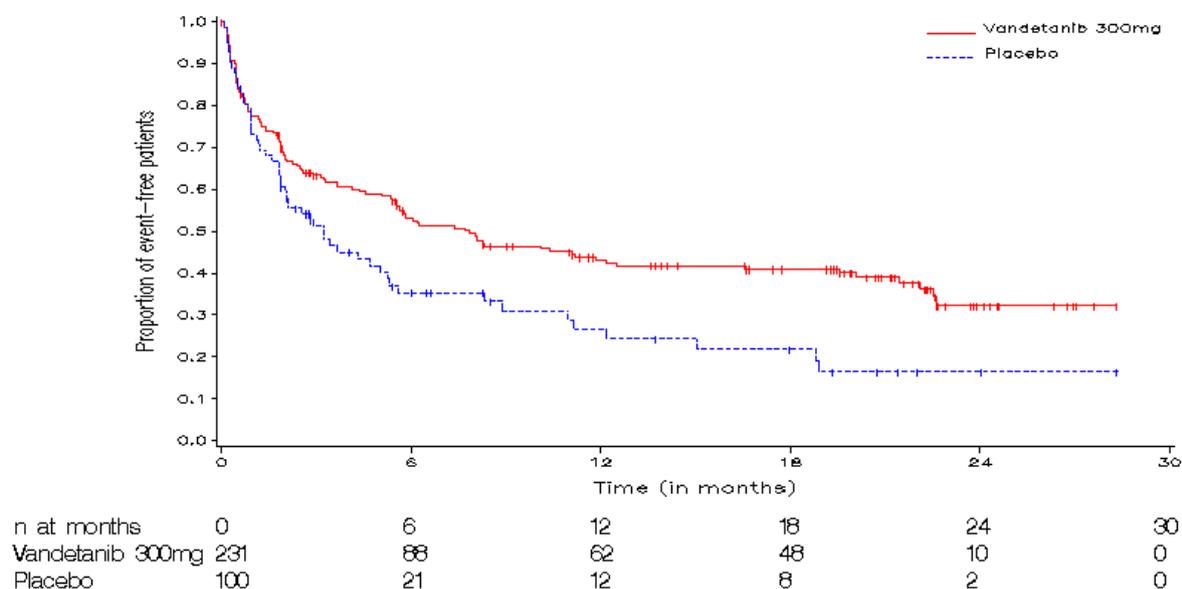


Abbildung 8: Kaplan-Meier-Kurve für Zeit bis zur Schmerzprogression (ITT-Population)

Es zeigte sich eine statistisch signifikante Verbesserung der TWP für Vandetanib im Vergleich zu Placebo mit einem *Hazard Ratio* von 0,61 (95%-KI: [0,43;0,87], p-Wert: 0,0062). Die mediane Zeit (Tabelle 4-47) bis zur Schmerzprogression betrug in der Vandetanib-Gruppe 7,85 Monate und in der Placebogruppe 3,25 Monate, was einer Verbesserung von knapp 5 Monaten entspricht. Abbildung 8 unterstreicht die oben genannten Ergebnisse anhand der grafischen Darstellung der Kaplan-Meier-Kurve, welche die Zeit bis zur Schmerzprogression für beide Behandlungsgruppen zeigt. Aus der Abbildung und den Ergebnissen wird ersichtlich, dass die Patienten der Vandetanib-Gruppe einen längeren Zeitraum erwarten konnten, bis sich Schmerzen entwickeln, eine Schmerzprogression eintritt oder eine Erhöhung der Opiatdosis nötig ist, um die Schmerzen zu kontrollieren.

Beide Sensitivitätsanalysen (Tabelle 4-46), die den Gebrauch von Schmerzmittel für die Definition des Endpunktes nicht berücksichtigten, zeigten zwar einen numerischen Vorteil für

Vandetanib, der Unterschied zwischen den Behandlungsgruppen war aber nicht signifikant. Das Hazard Ratio lag bei 0,76 (95%-KI: [0,53;1,09]), bzw. bei 0,72 (95%-KI: [0,50;1,02]) nach Ersetzung fehlender Werte. 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.

*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 9: 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.

#### **4.3.1.3.1.6 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 4-48: Operationalisierung des Endpunktes Gesundheitsbezogene Lebensqualität

Studie	Operationalisierung
D4200C00058	Der totale 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 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 4-49: 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) (53) 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 (54, 55).

Dieser Endpunkt wurde 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 (bereits zu 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 mangelhaften 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.

*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 4-50: Ergebnisse für gesundheitsbezogene Lebensqualität(HRQoL) gemessen anhand des FACT-G - aus RCT mit dem zu bewertenden Arzneimittel

Studie	Erhebungszeitpunkt	n <sup>a</sup>	Mittelwert	SD	Median	Min	Max
D4200C00058							
<b>Vandetanib</b>	<i>Baseline</i>	218	79,11	16,716	81,45	34,1	106,0
	<b>Woche 12</b>	174	77,06	17,323	77,00	25,4	108,0
	<b>Woche 24</b>	177	77,22	17,410	79,50	37,2	108,0
	<b>Woche 36</b>	154	79,68	16,808	81,33	29,5	108,0
	<b>Woche 48</b>	149	78,70	17,114	80,00	30,3	108,0
	<b>Woche 60</b>	141	78,84	17,421	81,00	22,0	108,0
	<b>Woche 72</b>	117	78,93	16,871	79,00	31,5	108,0
	<b>Woche 84</b>	116	79,82	15,629	80,00	36,0	108,0
	<b>Woche 96</b>	62	79,51	15,843	81,00	44,0	108,0
	<b>Woche 108</b>	31	73,50	16,477	70,00	39,0	108,0
	<b>Woche 120</b>	15	72,53	17,230	70,00	42,7	95,0
	<b>Woche 132</b>	1	84,67	NB	84,67	84,7	84,7
<b>Placebo</b>	<i>Baseline</i>	92	77,53	14,180	76,40	42,7	105,0
	<b>Woche 12</b>	73	76,97	16,709	78,17	36,0	104,0
	<b>Woche 24</b>	63	78,89	15,478	81,00	32,8	104,0
	<b>Woche 36</b>	51	80,52	15,037	81,83	36,0	106,0
	<b>Woche 48</b>	40	77,82	15,940	78,58	36,0	101,0
	<b>Woche 60</b>	38	77,69	14,194	78,08	39,0	102,0
	<b>Woche 72</b>	32	77,81	17,026	75,00	32,0	105,0
	<b>Woche 84</b>	25	74,67	14,791	72,00	52,8	107,0
	<b>Woche 96</b>	14	75,21	12,085	74,42	49,2	95,0
	<b>Woche 108</b>	9	69,93	10,411	71,50	52,2	84,0
	<b>Woche 120</b>	3	67,22	19,687	60,00	52,2	89,5

<sup>a</sup> Anzahl an Patienten mit vorliegenden Daten

SD = Standardabweichung (*Standard Deviation*)

NB = nicht berechenbar

Tabelle 4-51: Veränderung der gesundheitsbezogenen Lebensqualität (HRQoL) zum Ausgangswert gemessen anhand des FACT-G – aus RCT mit dem zu bewertenden Arzneimittel

Studie	Erhebungszeitpunkt	n <sup>a</sup>	Mittelwert	SD	Median	Min	Max
D4200C00058							
<b>Vandetanib</b>	<b>Woche 12</b>	168	-2,01	11,171	-1,00	-38,4	24,7
	<b>Woche 24</b>	168	-3,58	11,678	-2,80	-34,7	31,9
	<b>Woche 36</b>	147	-2,32	11,782	-2,10	-34,8	29,0
	<b>Woche 48</b>	142	-2,25	12,743	-1,40	-38,6	40,0
	<b>Woche 60</b>	136	-2,23	13,316	-1,40	-35,0	54,0
	<b>Woche 72</b>	112	-2,07	13,490	-1,30	-44,1	38,0
	<b>Woche 84</b>	111	-1,74	12,693	-1,30	-40,0	44,0
	<b>Woche 96</b>	57	-2,28	12,089	-1,50	-24,8	26,2
	<b>Woche 108</b>	28	-3,27	13,499	-0,90	-46,8	19,0
	<b>Woche 120</b>	14	-6,09	15,155	-5,15	-43,1	20,0
	<b>Woche 132</b>	1	11,20	NB	11,20	11,2	11,2
<b>Placebo</b>	<b>Woche 12</b>	69	-1,87	9,107	-1,30	-30,5	16,4
	<b>Woche 24</b>	62	-0,13	9,167	0,20	-21,5	20,4
	<b>Woche 36</b>	49	-0,65	9,701	-0,60	-25,8	29,2
	<b>Woche 48</b>	39	-2,82	11,314	-1,10	-26,0	30,2
	<b>Woche 60</b>	37	-4,09	9,859	-3,00	-25,5	17,0
	<b>Woche 72</b>	31	-2,80	12,739	-3,00	-29,8	29,0
	<b>Woche 84</b>	25	-4,36	10,567	-5,00	-28,0	14,0
	<b>Woche 96</b>	14	-3,10	7,735	-3,45	-19,8	9,9
	<b>Woche 108</b>	9	-4,80	11,078	-3,00	-24,5	9,2
	<b>Woche 120</b>	3	-1,90	3,804	-2,10	-5,6	2,0

<sup>a</sup> Anzahl an Patienten mit vorliegenden DatenSD = Standardabweichung (*Standard Deviation*)

NB = nicht berechenbar

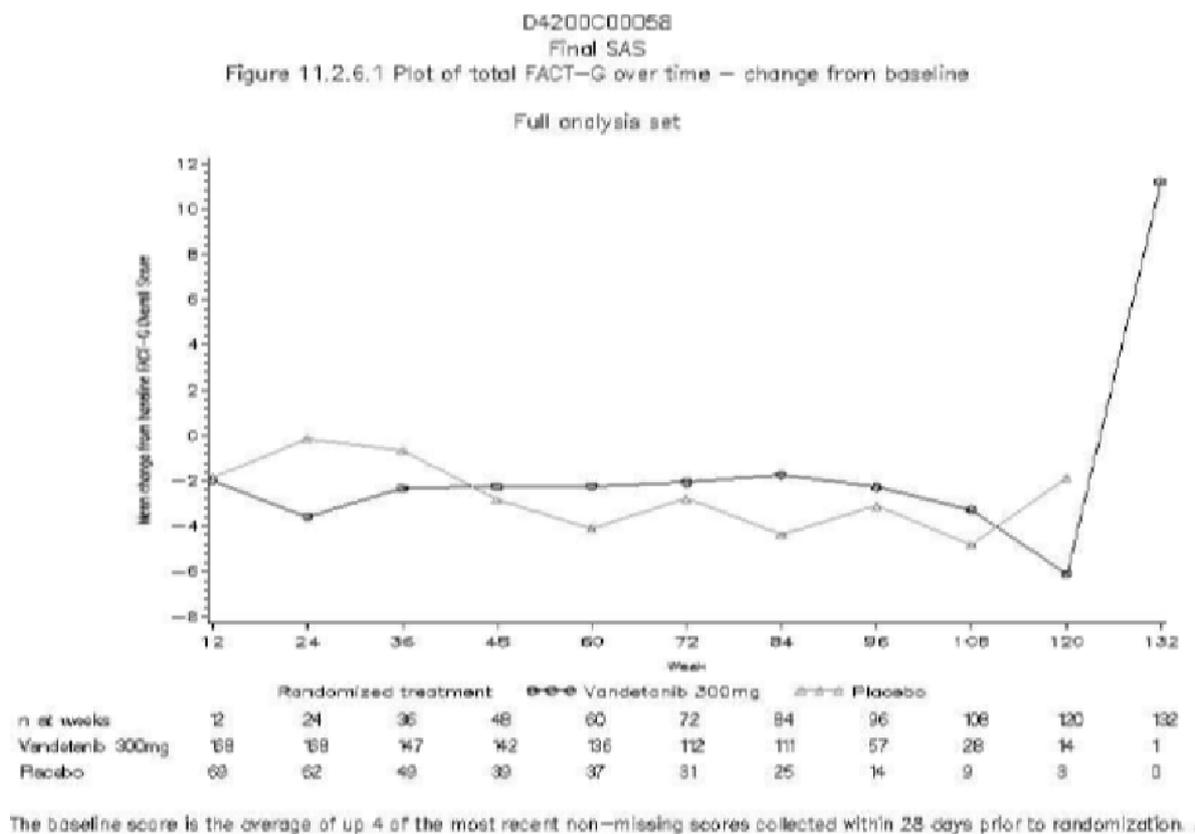


Abbildung 10: FACT-G über die Zeit – Veränderung der gesundheitsbezogenen Lebensqualität zum Ausgangswert

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 4-50 und Tabelle 4-51 der beiden Behandlungsgruppen wird deutlich, dass keine bedeutenden Unterschiede zwischen den Behandlungsgruppen bestehen. Die grafische Darstellung (Abbildung 10) verdeutlicht die Ergebnisse. Es ergab sich jedoch keine Evidenz dafür, dass die gesundheitsbezogene Lebensqualität durch Vandetanib nachteilig beeinflusst wurde.

*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.

#### **4.3.1.3.1.7 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 4-52: Operationalisierung von Unerwünschte Ereignisse (UE)

Studie	Operationalisierung
D4200C00058	Gesamtrate unerwünschter Ereignisse Schwerwiegende unerwünschte Ereignisse Studienabbruch wegen unerwünschter Ereignisse  Jeweils operationalisiert als Anteil der Patienten mit mindesten einem Ereignis

*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 D4200C00058 gemäß GCP Standards erfolgte und die in die Analyse eingeschlossenen Patienten für alle Operationalisierungen identisch ist.

Tabelle 4-53: 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 Randomisierung und Analyse-Cut-off. Die Analyse wurde anhand der Sicherheits (Safety) -Population durchgeführt, d. h. anhand aller Patienten, die mindestens eine Dosis der Studienmedikation bekommen haben. 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.

Tabelle 4-54 fasst die Ergebnisse zu den Gesamtraten unerwünschter bzw. therapiebezogener unerwünschter Ereignisse, schwerwiegender unerwünschter Ereignisse und die Studienabbruchraten aufgrund von unerwünschten Ereignissen 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 4-54: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Studie	Vandetanib (N=231)		Placebo (N=99)		Total (N=330)
D4200C00058					
AE Kategorie	Anzahl Patienten <sup>1</sup> (%)	Ereignisrate (pro 1000 Patientenjahre)	Anzahl an Patienten <sup>1</sup> (%)	Ereignisrate (pro 1000 Patientenjahre)	Anzahl Patienten (%)
Jegliche UEs	230 (99,6)	21729,8	90 (90,9)	4374,3	320 (97,0)
Vandetanib zugeschriebene <sup>2</sup> UEs	222 (96,1)	9136,4	59 (59,6)	1154,7	281 (85,2)
UEs vom CTCAE Grad 3 oder höher	128 (55,4)	663,9	24 (24,2)	270,7	152 (46,1)
Jegliche SUEs (einschließlich Ereignisse mit Ausgang = Tod)	71 (30,7)	258,5	13 (13,1)	133,9	84 (25,5)
Jegliche SUEs mit Ausgang = Tod	5 (2,2)	15,1	2 (2,0)	19,6	7 (2,1)
Jegliche UEs die zur Absetzung Vandetanibs führten	28 (12,2)	85,6	3 (3,0)	29,5	31 (9,4)
Alle anderen signifikanten UEs <sup>3</sup>	0 (0,0)	0,0	0 (0,0)	0,0	0 (0,0)

<sup>1</sup> Patienten mit mehreren Ereignissen in derselben Kategorie werden nur einfach in dieser Kategorie gezählt

<sup>2</sup> Durch den Prüfarzt eingeschätzt

<sup>3</sup> Jegliche UEs, die vom Prüfarzt als signifikant eingeschätzt werden.

Ereignisrate = (Anzahl Patienten mit UEs / Totale Dauer des *Follow-up* über alle Patienten in einer gegebenen Gruppe) x 1000

SUE =Schwere unerwünschte Ereignisse

Tabelle 4-55: Ergebnisse für Unerwünschte Ereignisse (UE) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

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 (12,2)	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;14,8]	23,00 [2,87;184,16]
	RR [95% -KI]		2,34 [1,36;4,03]	4,00 [1,25;12,85]	1,10 [1,03;1,17]

Tabelle 4-56: 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)

Studie	Vandetanib 300 mg (N=99)	Placebo (N=99)	Total (N=330)
D4200C00058			
UE Kategorie	Anzahl Patienten <sup>1</sup> (%)	Anzahl Patienten <sup>1</sup> (%)	Anzahl Patienten (%)
Gesamtes Auftreten von QTc bedingten Ereignissen	36 (15,6)	4 (4,0)	40 (19,6)
Jegliche SUEs mit Ausgang = Tod	0	0	0
Schwerwiegende UEs	3 (1,3)	0	3 (1,3)
Jegliche UEs die zur Absetzung Vandetanibs führten	2 (0,9)	1 (1,0)	3 (1,9)
UEs vom CTCAE Grad 3 oder höher	20 (8,7)	3 (3,0)	23 (11,7)
Gesamtes Auftreten von UEs der bevorzugten Bezeichnung „ <i>electrocardiogram QT prolonged</i> “ (nach MedDRA)	33 (14,3)	1 (1,0)	7 (15,3)
Jegliche SUEs mit Ausgang = Tod	0	0	0
Schwerwiegende UEs	1 (0,4)	0	1 (0,4)
Jegliche UEs die zur Absetzung Vandetanibs führten	2 (0,9)	0	2 (0,9)
UEs vom CTCAE Grad 3 oder höher	18 (7,8)	1 (1,0)	19 (8,8)

Wie in Tabelle 4-54 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). Der Anteil der Patienten mit unerwünschten Ereignissen vom *Common Terminology Criteria for Adverse Events* (CTCAE) Grad 3 oder höher (55,4% vs. 24,2%), mit SAEs (30,7% vs. 13,1%) und mit unerwünschten Ereignissen, die zur Absetzung der Studienbehandlung führten (12,1% vs. 3,0%), 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%,).

Im Vergleich zu Placebo zeigte sich für unerwünschte Ereignisse, schwerwiegende unerwünschte Ereignisse und Studienabbrüche wegen unerwünschter Ereignisse ein statistisch signifikanter Unterschied zuungunsten von Vandetanib. Daraus ergibt sich ein Beleg für einen größeren Schaden von Vandetanib gegenüber Placebo für UE, SUE und Studienabbruch wegen unerwünschter Ereignisse (Tabelle 4-55).

Die fünf am häufigsten berichteten unerwünschten Ereignisse in der Vandetanib-Gruppe vs. Placebo waren Diarrhö, Hautausschlag, Übelkeit, Bluthochdruck und Kopfschmerzen.

- Diarrhö: 56,3% vs. 26,3%
- Hautausschlag: 45,0% vs. 11,1%
- Übelkeit: 33,3% vs. 16,2%
- Bluthochdruck: 31,6% vs. 5,1%
- Kopfschmerzen: 25,5% vs. 9,1%

Unerwünschte Ereignisse (Vandetanib vs. Placebo) wie Rückenschmerzen (9,1% vs. 20,2%), Gelenkschmerzen (7,8% vs. 10,1%) und Schmerzen in den Extremitäten (6,9% vs. 13,1%) wurden weniger häufig in der Vandetanib-Gruppe berichtet als in der Placebogruppe.

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 Modul3A, Abschnitt 3.4).

Im Vergleich zu Placebo zeigte sich 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 4-56). Die QTc-bedingten unerwünschten Ereignisse wurden im Risk-Management Plan für Vandetanib berücksichtigt (siehe Modul 3A, 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 12: 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 4.3.1.3.1.2.

*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 Alter, Geschlecht und Schwere der Krankheit wurde nur 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.

Für den Endpunkt „Objektive Ansprechrate (ORR)“ wurden keine Subgruppenanalysen durchgeführt.

Für den Endpunkt „Zeit bis zur Schmerzprogression (TWP)“ wurde eine explorative Subgruppenanalyse- nach *Baseline*-Opioidanalgetika-Gebrauch durchgeführt, Subgruppenresultate wurden in einem *Forest-Plot* dargestellt.

Generell ist anzumerken, dass die Resultate von Subgruppenanalysen, insbesondere auf Studienebene, nicht zu hoch bewertet, sondern kritisch beurteilt werden sollten. Kleine Stichproben innerhalb der Subgruppen, ein Ungleichgewicht zwischen den Subgruppen und die steigende Anzahl statistischer Tests erhöhen die Wahrscheinlichkeit für zufällig statistisch signifikante Ergebnisse. Diese Aspekte sollten bei der Bewertung der im Folgenden dargestellten Subgruppenresultate Berücksichtigung finden.

Ein im Protokoll vorgesehener globaler Interaktionstest (zum Niveau 0,1), der die 6 als prognostisch eingestuften *Baseline*-Kovariaten (RET Mutationsstatus, Zeitraum zur Verdopplung des Calcitoninwertes, Zeitraum zur Verdopplung des CEA-Wertes, Anzahl und Ansprechen vorangegangener Therapien und den MTC-Status) im Cox-PH-Modell einschloss, ergab keinen Hinweis auf eine Interaktion ( $p=0,177$ ). Post hoc wurden individuelle Interaktionstests für alle pre-spezifizierten Subgruppenfaktoren durchgeführt und zeigten 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 Hinweise auf eine Interaktion, die in beiden Fällen qualitativer Natur war (Tabelle 4-57). 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 keine Änderung des aus der Bewertung der Gesamtpopulation abgeleiteten Fazits.

Tabelle 4-57: Individuelle Interaktionstests (post-hoc) - PFS

Subgruppenbildender Faktor (Subgruppen Kategorien)	p-Wert Interaktionstest	Art der Interaktion
<b>RET Mutationsstatus</b> (positiv, negativ, unbekannt)	0,8028	
<b>Zeitraum zur Verdopplung des Calcitoninwertes</b> (≤24 Monate, >24 Monate, unbekannt)	0,2784	
<b>Zeitraum zur Verdopplung des CEA-Wertes</b> (≤24 Monate, >24 Monate, unbekannt)	0,0062	quantitativ
<b>Anzahl vorangegangener Therapien</b> (≥1 Vorbehandlung, keine Vorbehandlung)	0,6806	
<b>MTC-Status</b> (hereditär, sporadisch, unbekannt)	0,8693	
<b>Geschlecht</b> (Männer, Frauen)	0,9974	
<b>WHO Performance Status</b> (≥1, 0)	0,9948	
<b>p-VEGF</b> (hoch, niedrig, unbekannt)	0,4544	
<b>p-VEGFR2</b> (hoch, niedrig, unbekannt)	0,0560	quantitativ
<b>p-bFGF</b> (hoch, niedrig, unbekannt)	0,6365	

#### 4.3.1.3.2.1 Progressionsfreies Überleben (PFS)

Tabelle 4-58: PFS Subgruppenanalysen nach *Baseline*-Charakteristika

Studie D4200C00058			Behandlungseffekt (Vandetanib:Placebo)			
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>Männlich</b>	Vandetanib	134	47 (35,1)	0,46	[0,28;0,77]	0,0028
	Placebo	56	32 (57,1)			
<b>Weiblich</b>	Vandetanib	97	26 (26,8)	0,44	[0,22;0,85]	0,0147
	Placebo	44	19 (43,2)			
<b>Kaukasisch</b>	Vandetanib	218	70 (32,1)	0,46	[0,31;0,70]	0,0002
	Placebo	97	50 (51,5)			
<b>Andere Rasse</b>	Vandetanib	13	3 (23,1)	NB	NB	NB
	Placebo	3	1 (33,3)			
<b>WHO PS ≥ 1</b>	Vandetanib	77	28 (36,4)	0,49	[0,27;0,90]	0,0204
	Placebo	42	23 (54,8)			
<b>WHO PS = 0</b>	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

NB = nicht berechenbar

Tabelle 4-59: PFS Subgruppenanalysen nach Biomarkern

Studie		Behandlungseffekt (Vandetanib:Placebo)				
D4200C00058						
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>RET- Mutationsstatus positiv</b>	Vandetanib	137	47 (34,3)	0,45	[0,26;0,78]	0,0041
	Placebo	50	27 (54,0)			
<b>RET- Mutationsstatus negativ</b>	Vandetanib	2	1 (50,0)	0,87	[0,10;7,66]	0,8967
	Placebo	6	5 (83,3)			
<b>RET- Mutationsstatus unbekannt</b>	Vandetanib	92	25 (27,2)	0,49	[0,26;0,95]	0,0341
	Placebo	44	19 (43,2)			
<b>CTN Verdopplungszeit ≤ 24 Monate</b>	Vandetanib	124	39 (31,5)	0,31	[0,17;0,56]	0,0001
	Placebo	46	27 (58,7)			
<b>CTN Verdopplungszeit &gt; 24 Monate</b>	Vandetanib	83	23 (27,7)	0,56	[0,29;1,06]	0,0754
	Placebo	43	19 (44,2)			
<b>CTN Verdopplungszeit unbekannt</b>	Vandetanib	24	11 (45,8)	0,57	[0,17;1,88]	0,3548
	Placebo	11	5 (45,5)			
<b>CEA Verdopplungszeit ≤ 24 Monate</b>	Vandetanib	69	25 (36,2)	0,19	[0,10;0,36]	<0,0001
	Placebo	33	26 (78,8)			
<b>CEA Verdopplungszeit &gt; 24 Monate</b>	Vandetanib	119	28 (23,5)	0,81	[0,42;1,58]	0,5349
	Placebo	48	14 (29,2)			
<b>CEA Verdopplungszeit unbekannt</b>	Vandetanib	43	20 (46,5)	0,55	[0,24;1,25]	0,1551
	Placebo	19	11 (57,9)			
<p><i>Hazard Ratios</i> basieren auf einem Cox-PH-Modell mit Behandlung als Faktor, p-Werte auf einem <i>Log-Rank</i> Test</p> <p>CTN/CEA Verdopplungszeit wurde abgeleitet aus der natürlichen Logarithmierung der einzelnen CTN/CEA Messungen (<i>pre-dose</i>). Zur Ermittlung der Steigung <i>m</i> wird an diese Daten wird eine lineare Regressionsgerade angepasst. Die CTN/CEA Verdopplungszeit ergibt sich als: <math>m / \ln 2</math>.</p> <p>Negative Verdopplungszeiten werden der &gt;24 Monate Gruppe zugewiesen</p>						

Tabelle 4-60: PFS Subgruppenanalyse nach Krankheitsstatus (MTC-Status und therapeutische Vorbehandlung)

Studie D4200C00058		Behandlungseffekt (Vandetanib:Placebo)				
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>Hereditäre oder Keimbahn-Mutation</b>	Vandetanib	28	7 (25,0)	0,30	[0,04;2,50]	0,2657
	Placebo	5	2 (40,0)			
<b>Sporadische oder unbekannte Mutation</b>	Vandetanib	203	66 (32,5)	0,48	[0,32;0,72]	0,0004
	Placebo	95	49 (51,6)			
<b>Metastatisch</b>	Vandetanib	217	67 (30,9)	0,42	[0,28;0,63]	<0,0001
	Placebo	97	51 (52,6)			
<b>Lokal fortgeschritten</b>	Vandetanib	14	6 (42,9)	NB	NB	NB
	Placebo	3	0 (0)			
<b>≥ 1 Vorbehandlung</b>	Vandetanib	90	31 (34,4)	0,43	[0,24;0,78]	0,0056
	Placebo	42	24 (57,1)			
<b>Keine Vorbehandlung</b>	Vandetanib	141	42 (29,8)	0,49	[0,29;0,85]	0,0106
	Placebo	58	27 (46,6)			
<b>Ansprechen nach Vorbehandlung</b>	Vandetanib	5	1 (20,0)	0	NB	NB
	Placebo	2	(0,0)			
<b>Kein Ansprechen nach Vorbehandlung</b>	Vandetanib	42	19 (45,2)	0,73	[0,33;1,60]	0,4291
	Placebo	23	11 (47,8)			
<b>NB/ unbekannter BOR zu Vorbehandlung</b>	Vandetanib	184	53 (28,8)	0,38	[0,23;0,60]	<0,0001
	Placebo	75	40 (53,3)			

*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*)  
NB = nicht berechenbar

Tabelle 4-61: PFS Subgruppenanalysen nach Plasma Biomarkern

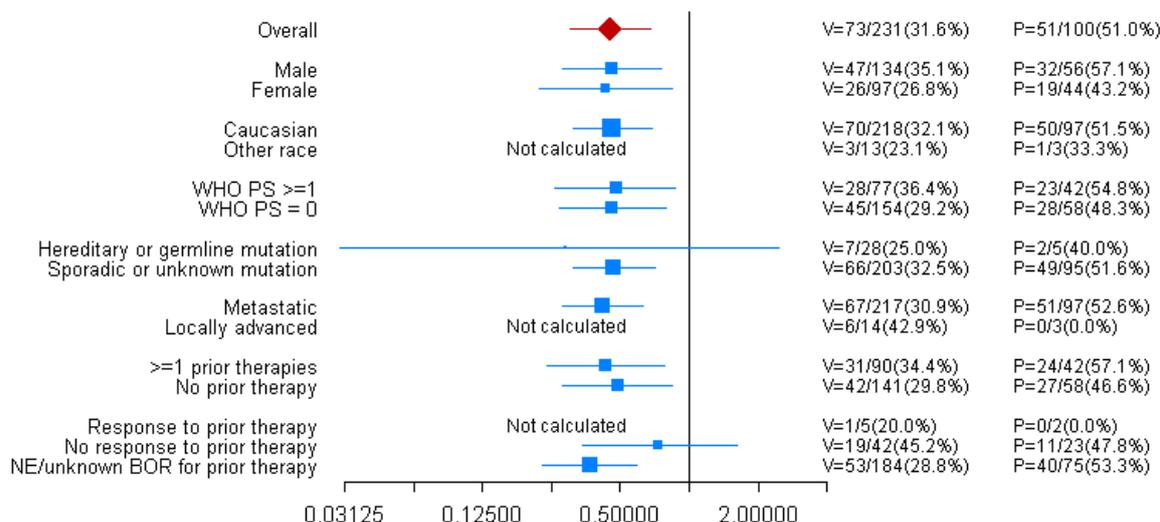
Studie		Behandlungseffekt (Vandetanib:Placebo)				
D4200C00058						
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>Hoher Baseline p-VEGF</b>	Vandetanib	115	41 (35,7)	0,56	[0,32;0,97]	0,0382
	Placebo	51	25 (49,0)			
<b>Niedriger Baseline p-VEGF</b>	Vandetanib	101	25 (24,8)	0,37	[0,19;0,73]	0,0039
	Placebo	42	20 (47,6)			
<b>Unbekannter Baseline p-VEGF</b>	Vandetanib	15	7 (46,7)	0,27	[0,07;1,01]	0,0517
	Placebo	7	6 (85,7)			
<b>Hoher Baseline p-VEGFR2</b>	Vandetanib	155	40 (25,8)	0,60	[0,35;1,03]	0,0626
	Placebo	69	26 (37,7)			
<b>Niedriger Baseline p-VEGFR2</b>	Vandetanib	61	26 (42,6)	0,20	[0,09;0,42]	<0,0001
	Placebo	24	19 (79,2)			
<b>Unbekannter Baseline p-VEGFR2</b>	Vandetanib	15	7 (46,7)	0,27	[0,07;1,01]	0,0517
	Placebo	7	6 (85,7)			
<b>Hoher Baseline p-bFGF</b>	Vandetanib	107	39 (36,4)	0,54	[0,32;0,94]	0,0278
	Placebo	49	26 (53,1)			
<b>Niedriger Baseline p-bFGF</b>	Vandetanib	108	27 (25,0)	0,41	[0,21;0,80]	0,0090
	Placebo	43	19 (44,2)			
<b>Unbekannter Baseline p-bFGF</b>	Vandetanib	16	7 (43,8)	0,38	[0,11;1,31]	0,1252
	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 und vor Randomisierung

Tabelle 4-62: PFS Subgruppenanalyse nach Altersgruppe

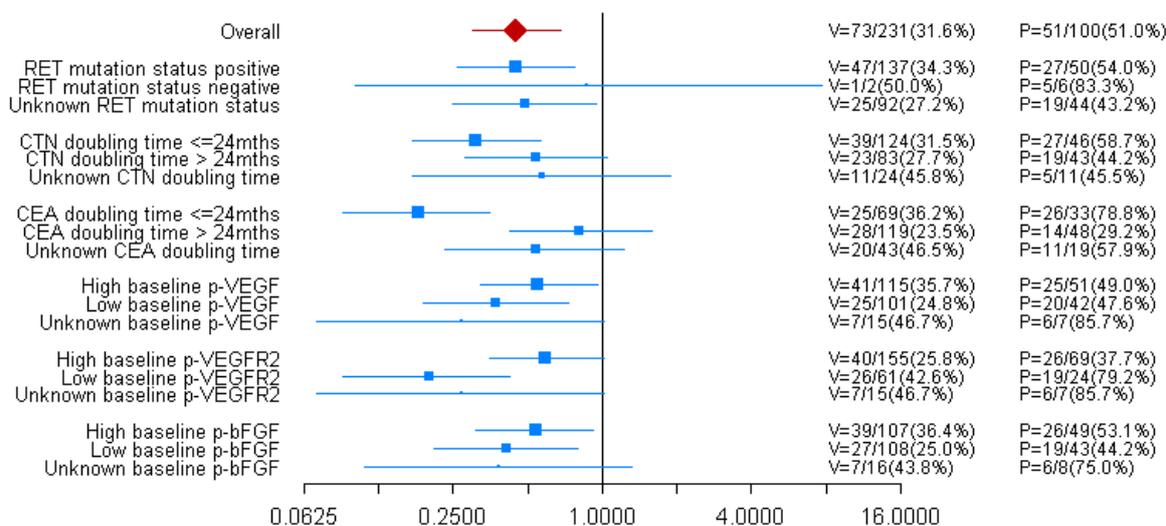
Studie		Behandlungseffekt (Vandetanib:Placebo)				
D4200C00058						
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>Jünger als 65 Jahre</b>	Vandetanib	182	55 (30,2)	0,50	[0,32;0,80]	0,0036
	Placebo	80	37 (46,3)			
<b>65 Jahre und älter</b>	Vandetanib	49	18 (36,7)	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.



V = Vandetanib, P = Placebo

Abbildung 13: Forest-Plot der Subgruppenanalyse für PFS – Baseline- Charakteristiken und Krankheitsstatus



V = Vandetanib, P = Placebo

Abbildung 14: Forest-Plot der Subgruppenanalyse für PFS – Biomarker

Eine explorative Subgruppenanalyse nach Alter wurde 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.

Subgruppenanalysen – insbesondere bei nur einer vorliegenden Studie – sind allgemein kritisch zu sehen und vorsichtig zu interpretieren. Zusammenfassend existieren innerhalb der Subgruppen bzgl. *Baseline*-Charakteristika, Biomarker, Krankheitsstatus und Plasma Biomarker keine relevanten Unterschiede bezüglich des progressionsfreien Überlebens.

#### 4.3.1.3.2.2 Objektive Ansprechrate (ORR)

Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt.

#### 4.3.1.3.2.3 Biochemisches Ansprechen (CTN; CEA)

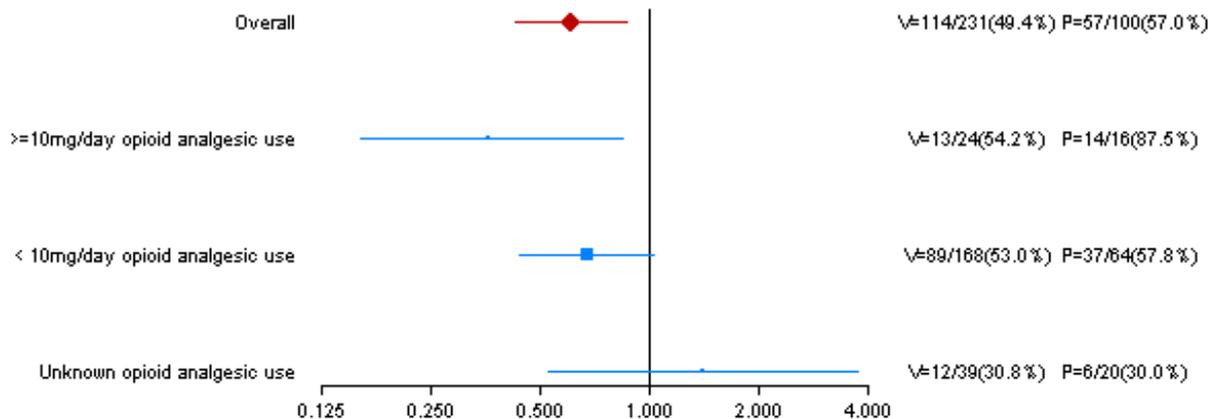
Für diesen Endpunkt wurden keine Subgruppenanalysen durchgeführt.

#### 4.3.1.3.2.4 Zeit bis zur Schmerzprogression (TWP)

Tabelle 4-63: TWP Subgruppenanalyse nach Schmerzmittelgebrauch

Studie		Behandlungseffekt (Vandetanib:Placebo)				
D4200C00058						
Subgruppe	Behandlung	N	Anzahl Ereignisse (%)	Hazard Ratio	95%- KI	2-seitiger p-Wert
<b>≥ 10 mg/Tag Opioidanalgetika Gebrauch</b>	Vandetanib	24	13 (54,2)	0,36	[0,16;0,84]	0,0176
	Placebo	16	14 (87,5)			
<b>&lt; 10 mg/Tag Opioidanalgetika Gebrauch</b>	Vandetanib	168	89 (53,0)	0,68	[0,44;1,03]	0,0698
	Placebo	64	37 (57,8)			
<b>Opioidanalgetika Gebrauch unbekannt</b>	Vandetanib	39	12 (30,8)	1,41	[0,53;3,74]	0,4936
	Placebo	20	6 (30,0)			

Analyse wurde mittels Log-Rank Test, mit Behandlung als einzigem Faktor, durchgeführt. Schmerzprogression wird als „Verschlechterung“ ohne Verbesserung innerhalb der nächsten 14 Tage erachtet; Der Gebrauch von Opioidanalgetika zum Ausgangswert wurde zu Morphinsulfat in mg/Tag konvertiert



V = Vandetanib, P = Placebo

Abbildung 15: *Forest-Plot* der Subgruppenanalyse für TWP

Die visuelle Prüfung des *Forest-Plots* spiegelt die Konsistenz des TWP über die zwei Gruppen mit bekanntem Opioidanalgetika Gebrauch hinweg wider. In der Gruppe „≥10 mg/Tag Opioidanalgetika Gebrauch“ liegt das *Hazard Ratio* bei 0,36 (95%-KI: [0,16;0,84],  $p=0,0176$ ) und ist signifikant. In der komplementären Gruppe, „<10 mg/Tag Opioidanalgetika Gebrauch“, liegt das *Hazard Ratio* bei 0,68 und ist mit einem 95%-KI [0,44;1,03] und einem  $p$ -Wert von 0,0698 nicht signifikant.

In dieser Subgruppenanalyse zeigt sich in der „≥10 mg/Tag Opioidanalgetika Gebrauch“-Gruppe eine signifikante Überlegenheit der Vandetanib-Gruppe bezüglich des TWP. Aufgrund nur einer vorliegenden Studie sowie kleiner Populations- und Ereigniszahlen sind diese Ergebnisse allerdings mit Vorsicht zu interpretieren.

#### 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.*

Die Ergebnisse der Studie 58 zeigen für den primären Endpunkt „Progressionsfreies Überleben (PFS)“ mit einer Reduktion des Risikos um 54% eine statistisch signifikante Verbesserung für Vandetanib im Vergleich zu Placebo (HR:0,46, 95%-KI: [0,31;0,69]). Die Wahrscheinlichkeit, 6, 12 bzw. 24 Monate progressionsfrei zu leben, liegt in der Vandetanib-Gruppe im Schnitt um fast 20% über der in der Placebo-Gruppe und am Daten *Cut-off* wies

ein um 20% kleinerer Anteil der Vandetanib-Gruppe eine Krankheitsprogression auf als in der Placebo-Gruppe. Das geschätzte mediane progressionsfreie Überleben für die Vandetanib-Gruppe entspricht mit 30,5 Monaten verglichen zu 19,3 Monaten in der Placebo-Gruppe, einer Verlängerung des medianen PFS um ca. 11 Monate. Die bezüglich des PFS durchgeführten Subgruppenanalysen identifizierten keine Subgruppe, welche nicht von Vandetanib profitiert.

Ebenfalls konnte eine statistisch signifikante Verbesserung zugunsten von Vandetanib gegenüber Placebo bezogen auf die objektive Ansprechrate gezeigt werden. Insgesamt betrug die ORR 45% in der Vandetanib-Gruppe und 13% in der Placebogruppe (OR:5,48 95%-KI: [2,99;10,79],  $p < 0,0001$ ). Der Effekt von Vandetanib bestätigte sich für 12 von 13 Patienten in der Placebogruppe. Nur ein Patient hatte bereits unter Placebo ein beginnendes Ansprechen des Tumors. Für die verbleibenden 12 Patienten der Placebogruppe setzte das Ansprechen erst während der offenen Behandlungsphase unter Vandetanib ein.

Auch für den Surrogatendpunkt „biochemische Ansprechen (CTN und CEA)“ konnte eine signifikante Verbesserung zugunsten von Vandetanib gegenüber Placebo gezeigt werden: CTN: 69,3% versus 3,0% (OR: 72,86, 95%-KI: [26,22;303,2],  $p < 0,0001$ ); und CEA: 51,5% versus 2,0% (OR: 52,03, 95%-KI: [15,95;320,3],  $p < 0,0001$ ), wobei hohe CTN-Werte mit sekretorischer Diarrhö verknüpft sind.

Eines der für die Patienten 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 dem Schmerzmittelgebrauch und den Antworten des BPI-SF. Es zeigte sich eine statistisch signifikante Verbesserung der TWP für Vandetanib (114 Ereignisse, 49,4%) im Vergleich zu Placebo (57 Ereignisse, 57,0%), mit einem Hazard Ratio von 0,61 (95%-KI: [0,43;0,87],  $p$ -Wert: 0,0062.). Die mediane Zeit bis zur Schmerzprogression betrug in der Vandetanib-Gruppe 7,85 Monate und in der Placebo-Gruppe 3,25 Monate, was einer Verbesserung von knapp 5 Monaten entspricht.

Bei nahezu allen Patienten beider Gruppen trat mindestens ein unerwünschtes Ereignis auf. Das Ergebnis bezüglich der Gesamtrate der unerwünschten Ereignisse zeigte statistisch signifikante Unterschiede zwischen den Behandlungsgruppen, wies jedoch eine geringe Effektgröße auf (RR: 1,1, 95%-KI: [1,01;1,17]). Die Anteile der Patienten mit unerwünschten Ereignissen vom CTCAE Grad 3 oder höher (RR: 2,29, 95%-KI: [1,58;3,30]), SUEs (RR: 2,34, 95%-KI: [1,36;4,03]) und unerwünschten Ereignissen die zum Studienabbruch führten (RR: 4,00 [1,25;12,85]) 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. Der Anteil aufgrund unerwünschter Ereignisse verstorbener Patienten war in beiden Gruppen gleich (2,2% vs. 2,0%).

## 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 4-64: 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 4.3.1.3.1.2.

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

Tabelle 4-65: 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 4-66: Operationalisierung von &lt;Endpunkt xxx&gt;

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 4-67: Bewertung des Verzerrungspotenzials für &lt;Endpunkt xxx&gt; 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 4-68: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1.2)
<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 4.3.1.3.1.3.

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 4-69: 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 4-70: Operationalisierung von &lt;Endpunkt xxx&gt;

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 4-71: Verzerrungsaspekte für &lt;Endpunkt xxx&gt; – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende 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 der in Abschnitt 4.2.2 angegebenen 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 4-72: 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 4-72 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 4-59 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.

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-73 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-73: 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.*

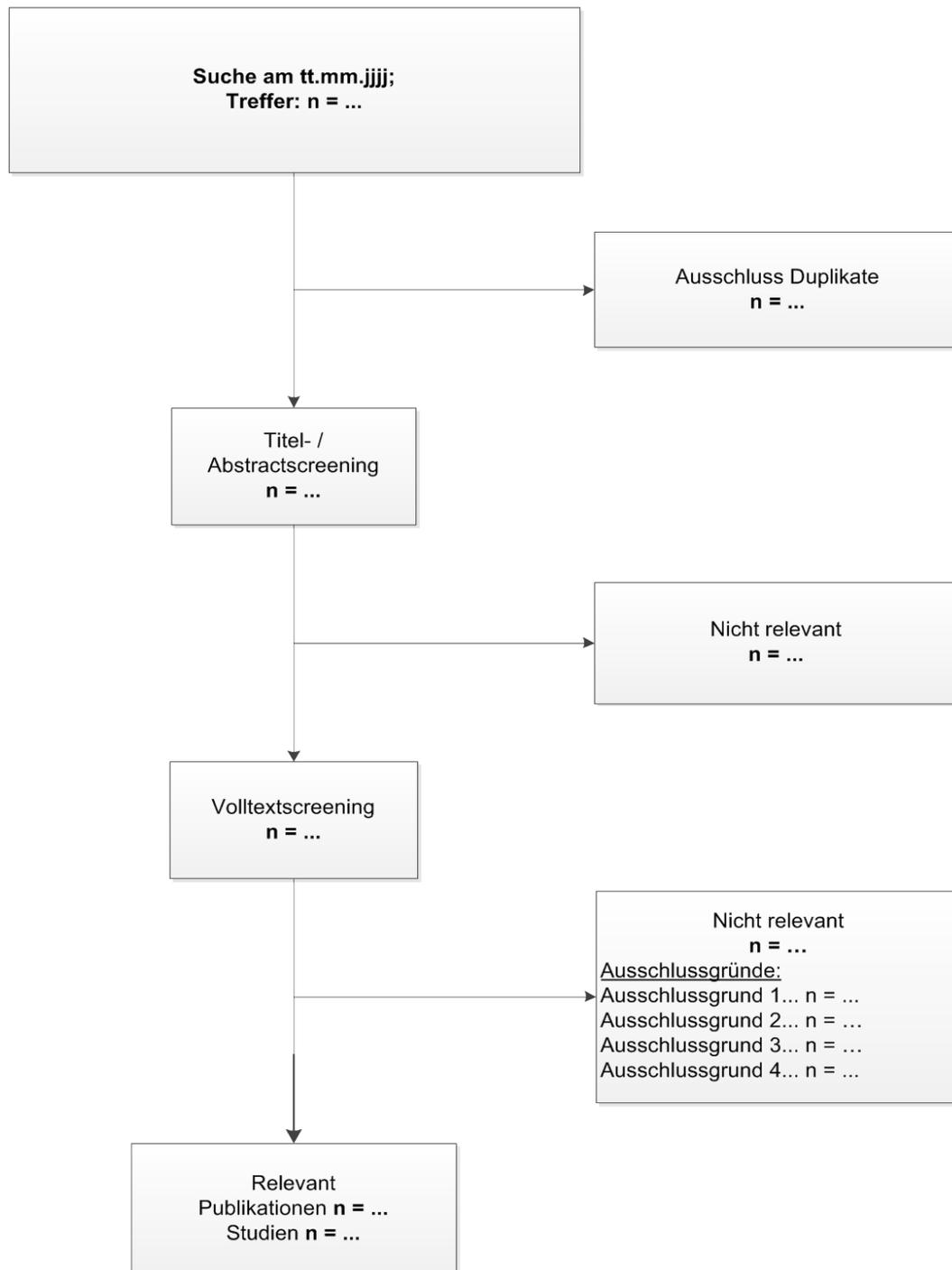


Abbildung 16: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen

Die Voraussetzungen für eine erforderliche Literaturrecherche sind nicht gegeben (siehe Abschnitt 4.3.1.1.2). Aufgrund dessen wurde keine bibliographische Literaturrecherche durchgeführt.

#### 4.3.2.3.1.3 Studien aus der Suche in Studienregistern

Tabelle 4-74: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. <sup>b</sup> )
D4200C00008	Clinicaltrials.gov: <a href="#">NCT00098345</a> (56)	ja	nein
D4200C00068	Clinicaltrials.gov: <a href="#">NCT00358956</a> (57)	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 4-75: Charakterisierung der eingeschlossenen Studien – weitere Untersuchungen

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten*)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	<RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<relevante Charakteristika, z. B. Schweregrad>		<ggf. Run-in, Behandlung, Nachbeobachtung>		
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: <ul style="list-style-type: none"> <li>Objektive Ansprechrates (ORR)</li> </ul> Patientenrelevante sekundäre Endpunkte: <ul style="list-style-type: none"> <li>Progressionsfreies Überleben (PFS)</li> <li>WHO Performance Status (WHO PS)</li> <li>Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACIT-D</li> <li>Unerwünschte Ereignisse (UE)</li> </ul>
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: <ul style="list-style-type: none"> <li>Objektive Ansprechrates (ORR)</li> </ul> Patientenrelevante sekundäre Endpunkte: <ul style="list-style-type: none"> <li>Progressionsfreies Überleben (PFS)</li> <li>WHO Performance Status (WHO PS)</li> <li>Gesundheitsbezogene Lebensqualität (HRQoL) gemessen anhand des FACT-G</li> <li>Unerwünschte Ereignisse (UE)</li> </ul>

\* 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 Ansprechrates (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 4-76: 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 4-77: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 1)

Studie	N	Alter (Jahre)	Geschlecht t w / m (%)	ethnische Zugehörig keit (%)		Gewicht (kg) Mittelwert (SD) [Spannweite]	Größe (cm) Mittelwert (SD) [Spannweite]
		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 4-78: Charakterisierung der Studienpopulationen – Weitere Untersuchungen (Teil 2)

Studie	N	Zeit seit Diagnose (Jahre)	Vorangegangene Theraphien für das MTC (%)				WHO Performance Status (%)		
		Mittelwert (SD) [Spannweite]	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 verheilte 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 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 4-79: Verzerrungspotenzial auf Studienebene – Weitere Untersuchungen mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial 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 existierte keine Vergleichsgruppe. Folglich waren in beiden Studien sowohl die Patienten als auch die Prüfarzte unverblindet. Die Aussagekraft nicht vergleichender Studien durch das Fehlen einer Vergleichsgruppe, die etwa hinsichtlich der natürlichen Krankheitsentwicklung aber auch hinsichtlich sogenannter Placeboeffekte adjustiert, ist als gering einzustufen.

Im Hinblick auf den Zugewinn an Informationen, z. B. die Nachhaltigkeit des Behandlungseffekts oder Nutzensaussagen einer begrenzten 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 4-80: 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 (partielltes Ansprechen) erreichten. Das beste objektive Ansprechen wurde entsprechend der 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 (partielltes Ansprechen) erreichten. Das beste objektive Ansprechen wurde entsprechend der 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 4-81 dargestellt.

Tabelle 4-81: Ergebnisse für die objektive Ansprechrate (ORR) – Weitere Untersuchungen

Studie	N	ORR (%)	Ansprechen				
		Wert [95%- KI]	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.

#### 4.3.2.3.2 Progressionsfreies Überleben (PFS) – weitere Untersuchungen

Tabelle 4-82: Operationalisierung des progressionsfreien Überleben

Studie	Operationalisierung
D4200C00008	Der Zeitraum zwischen erster Vandetanibgabe und einer objektiven Progression oder Tod. Eine Progression wurde entsprechend der 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 der 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 4-83: Ergebnisse für das Progressionsfreies Überleben (PFS) – Weitere Untersuchungen

Studie	N	PFS (Monate)		Status* n (%)	
		Wert [95% -KI]	Progression	Tod	Ereignisfrei
D4200C00008	30	27,9 [19,4;NB]	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)

NB=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 4-83 dargestellt.

#### 4.3.2.3.3 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 4-84: 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 4-85: Ergebnisse des WHO Performance Status (WHO PS) – Weitere Untersuchungen

Studie	N	Veränderung zwischen dem Ausgangswert (Baseline) und der 24 Wochen (Follow-up) n (%)				
		Verbesserung	Gleichbleibend	Verschlechterung		Kein Wert in Woche 24
		-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 4-85 dargestellt

#### 4.3.2.3.4 Gesundheitsbezogene Lebensqualität (HRQoL) – Weitere Untersuchungen

Tabelle 4-86: 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 4-87: 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 4-87 dargestellt.

**4.3.2.3.3.5 Unerwünschte Ereignisse (UE) – weitere Untersuchungen**

Tabelle 4-88: 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 $\geq 3$
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 $\geq 3$

CTCAE = *Common Terminology Criteria for Adverse Events*

\* Ein UE wurde als medikamentenbedingt eingestuft, falls der Arzt zumindest einen begründeten Verdacht hatte, dass die Medikamentengabe 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 4-89: Unerwünschte Ereignisse – Weitere Untersuchungen

Studie	N	Gesamtrate					
		UE	Schwerwiegendes UE	zum Tode führender UE	UE, die zum Abbruch der Studienmedikation führten	Medikamentenbedingter* UE	UE mit CTCAE Grad $\geq 3$ #
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 4-89 dargestellt.

Die häufigsten UEs in den Studien D4200C00008 und D4200C00068 waren Diarrhö, Hautausschlag, Müdigkeit und Übelkeit (nur in Studie D4200C00008).

#### 4.3.2.3.6 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.1.3.*

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 schwerwiegenden unerwünschten Ereignisse endeten noch während des Beobachtungszeitraums. In Studie D4200C00068 führte ein schwerwiegendes unerwünschtes

Ereignis zum Tod. Drei weitere schwerwiegende unerwünschte Ereignisse traten während des Beobachtungszeitraums 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 entsprechenden 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 signifikanten Effekt vorliegen, deren Ergebnis nicht durch weitere vergleichbare, endpunktbezogene ergebnissichere Studien infrage gestellt wird (23). 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 *CHMP Points to consider on application with: 1. meta-analyses; 2. one pivotal study* (58) und können im Fall von Vandetanib als erfüllt betrachtet werden. Im Folgenden wird dargelegt, warum diese Kriterien für die Studie D4200C00058 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. Das endpunktspezifische Verzerrungspotential ist mit Ausnahme der Endpunkte „Gesamtüberleben“ und „gesundheitsbezogene Lebensqualität“ sowohl für den primären Endpunkt „Progressionsfreies Überleben“, als auch für die sekundären Endpunkte „objektive Ansprechrates“, „Zeit bis zur Schmerzprogression“ und „unerwünschte Ereignisse“ niedrig.

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 in seltenen Erkrankungen ergibt (59) – wurde ein vergleichsweise umfassendes Patientenkollektiv eingeschlossen. Studie 58 ist folglich 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 Vergleichstherapie sind statistisch signifikant, die entsprechenden Effektschätzer sind präzise und klinisch relevant. Die beobachteten Effekte gelten über präspezifizierte Subgruppen hinweg. Damit sind die wesentlichen Kriterien hinsichtlich interner und externer Validität, statistischer und klinischer Relevanz der Ergebnisse, Konsistenz des Effekts über Subgruppen hinweg erfüllt.

Die RCT mit Vandetanib beinhaltet einen direkten Vergleich mit Placebo. Die Kriterien hinsichtlich Population und Behandlung richten sich nach der durch die Zulassung festgelegten Zielpopulation des zu bewertenden Arzneimittels Vandetanib (siehe Fachinformation (21)). Die Wahl der Kontrollgruppe erfolgte gemäß Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA am 29.09.2011 (1) 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. Die Placebogruppe 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 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. Gesamtüberleben zählt aufgrund seiner absolut eindeutigen Erhebung und klaren Definition (nur Tod zählt als Ereignis) zu den harten und damit validen Endpunkten. Durch den Umstand, dass die Ergebnisse der Gesamtüberlebensrate durch die geringe Zahl an beobachteten Todesfällen sowie das *Cross-Over* verzerrt sein könnten, wurde, wie bereits an anderer Stelle im Dossier diskutiert, auf eine weitere Darstellung verzichtet.

Das progressionsfreie Überleben (PFS) ist ein patientenrelevanter Endpunkt, der sich direkt in das Hinauszögern zunehmender progressionsbedingter 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 v1.0 (standardisiertes Kriterium) erhoben. Die objektive Ansprechrates (ORR) misst die Verringerung der Tumorlast und umfasst sowohl komplettes (CR) als auch partielles

Ansprechen (PR). Die Kriterien für CR oder PR entsprechen den RECIST Kriterien und basieren auf standardisierten reproduzierbaren bildgebenden Verfahren. Die Validität der Endpunkte „progressionsfreies Überleben“ und „objektive Ansprechrare“ ist somit gegeben.

Ebenfalls stellt das biochemische Ansprechen, einen Endpunkt dar, der als Surrogat für die krankheitsbedingten Symptome die Ergebnisse zu PFS und ORR auf wertvolle Weise ergänzt, da die Bewertung des CTN/CEA Status auf harten und objektiv messbaren Kriterien (Laborwerten; in der ausgewerteten Studie durch ein zentrales Labor durchgeführt) beruht.

Als Instrument zur Erhebung einer Veränderung der Zeit bis zur Schmerzprogression unter Studienmedikation im Vergleich zu Baseline 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 (46). 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 (47).

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 über den Krankheitsverlauf. (53, 54)

Die Meldung und Berichterstattung unerwünschter Ereignisse erfolgte in der RCT mit Vandetanib gemäß den Standards der *Good Clinical Practice* (GCP).

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 z. B. 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.

### **Statistisch signifikanter Zusatznutzen von Vandetanib gegenüber Best Supportive Care bezüglich PFS, ORR, biochemischem Ansprechen und TWP**

Im Rahmen der randomisierten, kontrollierten Studie D4200C00058 konnte für die patientenrelevanten morbiditätsbezogenen Endpunkte „Progressionsfreies Überleben (Progression Free Survival, PFS)“, „Objektive Ansprechrate (Objective Response Rate, ORR)“ und „Zeit bis zur Schmerzprogression (Time to Worsening of Pain, TWP)“ sowie auch für den bezüglich Progression und Morbidität aussagekräftigen Surrogatparameter „biochemisches Ansprechen (CTN; CEA)“ ein statistisch signifikanter Vorteil von Vandetanib gegenüber der mit Placebo behandelten Kontrollgruppe, welche ansonsten als aktive Therapie nur Best Supportive Care (BSC) erhielt, gezeigt werden.

In Anlehnung an die durch das IQWiG operationalisierten Kriterien der AM-NutzenV (60) wird das Ausmaß des Zusatznutzens auf Endpunktebene wie folgt bestimmt:

Tabelle 4-90: Ausmaß des Zusatznutzens auf Endpunktebene: Vandetanib vs. Placebo

Endpunktkategorie	Effektschätzer [95%-KI]	Ableitung des Ausmaßes <sup>b</sup>
<b>Endpunkt</b>	<b>Vandetanib vs. Placebo (BSC)</b>	
	<b>Wahrscheinlichkeit<sup>a</sup></b>	
<b>Mortalität</b>		
Gesamtüberleben (OS)	Aufgrund geringer Anzahl an	Zusatznutzen nicht belegt

Endpunktkategorie Endpunkt	Effektschätzer [95%-KI] Vandetanib vs. Placebo (BSC) Wahrscheinlichkeit <sup>a</sup>	Ableitung des Ausmaßes <sup>b</sup>
	<i>Todesfällen und der potenziellen Verzerrung durch das Cross-Over nicht für die Nutzenbewertung herangezogen (HR: 0,89 [0,28;2,85])</i>	
<b>Morbidität</b>		
Progressionsfreies Überleben (PFS)	HR: 0,46 [0,31;0,69] Verlängerung der medianen PFS-Zeit um ca. 11 Monate (30,5 Monate vs. 19,3 Monate) Wahrscheinlichkeit: <i>Beleg</i>	Zusatznutzen, Ausmaß: erheblich
Objektive Ansprechrate (ORR)	OR: 5,48 [2,99;10,79] Für Nicht-Ansprechen: OR: 0,18 [0,09;0,33] RR: 0,63 [0,55;0,73] Wahrscheinlichkeit: <i>Beleg</i>	Zusatznutzen, Ausmaß: erheblich
Biochemisches Ansprechen (CTN)	OR: 72,86 [26,22;303,2] Für Nicht-Ansprechen: OR: 0,01 [0,00;0,04] RR: 0,32 [0,26;0,39] Wahrscheinlichkeit: <i>Beleg</i>	Zusatznutzen, Ausmaß: erheblich
Biochemisches Ansprechen (CEA)	OR: 52,03 [15,95;320,3] Für Nicht-Ansprechen: OR: 0,02 [0,00;0,08] RR: 0,49 [0,43;0,57] Wahrscheinlichkeit: <i>Beleg</i>	Zusatznutzen, Ausmaß: erheblich
Zeit bis zur Schmerzprogression (TWP)	HR: 0,61 [0,43;0,87] Wahrscheinlichkeit: <i>Beleg</i>	Zusatznutzen, Ausmaß: beträchtlich
<b>Gesundheitsbezogene Lebensqualität</b>		
HRQoL, gemessen anhand des FACT-G		Zusatznutzen nicht belegt
<b>Nebenwirkungen</b>		
Unerwünschte Ereignisse (UE)	OR: 23,00 [2,87;184,16] RR: 1,10 [1,03;1,17] 99,6% vs. 90,9% Wahrscheinlichkeit: <i>Beleg</i>	Größerer Schaden
Schwerwiegende unerwünschte Ereignisse (SUE)	OR: 2,94 [1,54;5,60] RR: 2,34 [1,36;4,03] 30,7% vs. 13,1% Wahrscheinlichkeit: <i>Beleg</i>	Größerer Schaden
Studienabbrüche aufgrund unerwünschter Ereignisse	OR: 4,41 [1,31;14,8] RR: 4,00 [1,25;12,85] 12,2% vs. 3,0% Wahrscheinlichkeit: <i>Beleg</i>	Größerer Schaden

<sup>a</sup>: Angabe der Wahrscheinlichkeit, sofern statistisch signifikante Unterschiede vorliegen.

<sup>b</sup>: Einschätzungen zur Effektgröße erfolgen je nach Endpunktkategorie mit unterschiedlichen Grenzen anhand der oberen Grenze des Konfidenzintervalls, gemäß der vom IQWiG vorgeschlagenen Operationalisierung.

### Gesamtüberleben (OS)

Aufgrund der geringen Ereignisraten zum Zeitpunkt des *Cut-off* und des möglichen Einflusses des *Cross-Over* - und einer damit möglicherweise vorliegenden Verzerrung der Ergebnisse - wurde die statistische Auswertung des Gesamtüberlebens nicht für die Nutzenbewertung herangezogen. Eine erste Analyse der Gesamtüberlebensrate zeigte ein *Hazard Ratio* von 0,89 zugunsten von Vandetanib. Statistische Signifikanz wurde nicht erreicht (KI: [0,28; 2,85],  $p=0,713$ ).

### Verbesserung des Progressionsfreien Überlebens (PFS)

Bezüglich des PFS wurde in der Vandetanib-Gruppe ein statistisch signifikanter Vorteil gegenüber Placebo beobachtet (HR: 0,46 95%-KI: [0,31;0,69]). Die Wahrscheinlichkeit des PFS (Kaplan-Meier Schätzer) war in der Vandetanib-Gruppe zu allen Zeitpunkten um ca. 20% höher: zu Monat 6 (91,4% versus 74,5%), Monat 12 (83,3% versus 62,7%) und zu Monat 24 (58,8% versus 41,7%). Das mediane PFS war zum Analyse-*Cut-off* in der Vandetanib-Gruppe noch nicht erreicht, wurde jedoch mithilfe der Anpassung eines Weibull-Modells auf 30,5 Monate geschätzt und lag um ca. 11 Monate höher als in der Vergleichsgruppe (19,3 Monate). Angesichts des in Abschnitt 4.3.1.3.1.2 beschriebenen niedrigen Verzerrungspotentials sowie der hohen Aussagekraft der Studie und der schwerwiegenden Auswirkungen eines Tumorprogresses, insbesondere im Hinblick auf die Morbidität und Lebensqualität der Patienten, ergibt sich für diesen patientenrelevanten Endpunkt der Beleg eines erheblichen Zusatznutzens (siehe Tabelle 4-90).

### Objektive Ansprechrate (ORR)

Objektives Ansprechen (*Objective Response Rate*, ORR), in Form einer partiellen Remission gemäß RECIST Kriterien, wurde von 45% der Patienten in der Vandetanib-Gruppe erreicht, im Vergleich zu 13% der Patienten unter BSC. Angesichts des statistisch signifikanten Ergebnisses und der Effektgröße (OR für Nicht-Ansprechen: 0,18, 95%-KI: [0,09;0,33]) sowie der Aussagekraft des erbrachten Nachweises ergibt sich der Beleg eines erheblichen Zusatznutzens (siehe Tabelle 4-90).

### Surrogatendpunkt: Biochemisches Ansprechen (CTN; CEA)

Das biochemische Ansprechen (Calcitonin (CTN); Carcinoembryonales Antigen (CEA)) zeigt sich in einer Spiegelreduzierung für 69,3% bzw. 51,5% der mit Vandetanib behandelten Patienten im Vergleich zu nur 3,0% bzw. 2,0% bei Behandlung gemäß BSC (OR für

Ansprechen: 72,86, 95%-KI: [26,22;303,2] bzw. 52,03, 95%-KI: [15,95;320,3] / OR für Nicht-Ansprechen: 0,01, 95%-KI: [0,00;0,04] bzw. 0,02, 95%-KI: [0,00;0,08]). Angesichts des statistisch signifikanten Ergebnisses und der Größe des Effekts sowie der Aussagekraft des erbrachten Nachweises kann hieraus der Beleg eines erheblichen Zusatznutzens abgeleitet werden, da das biochemische Ansprechen ein geeigneter Surrogatmarker für den klinischen Progress ebenso wie relevante Nebenwirkungen der Erkrankung (sekretorische Diarrhö, Flush) ist und damit als valider Marker für Morbidität zu gewertet werden kann (siehe Tabelle 4-90).

#### Zeit bis zur Schmerzprogression (TWP)

Auch 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 opiathaltigen Schmerzmitteln, war unter Vandetanib im Median mehr als doppelt so lang wie in der Vergleichsgruppe (7,85 Monate gegenüber 3,25 Monate). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (HR: 0,61, 95%-KI: [0,43;0,87]). 62,5% der mit Vandetanib behandelten Patienten konnten ihren Verbrauch von opiathaltigen Schmerzmitteln senken; dies gelang nur 37,5% der Patienten auf Placebo, welche nur BSC erhielten. Angesichts des statistisch signifikanten Ergebnisses sowie der Aussagekraft des erbrachten Nachweises und unter Berücksichtigung der Schwere der Schmerzen kann hieraus der Beleg eines beträchtlichen Zusatznutzens abgeleitet werden (siehe Tabelle 4-90).

#### Gesundheitsbezogene Lebensqualität (FACT-G)

Obwohl die Angaben der Patienten bezüglich ihrer gesundheitsbezogenen Lebensqualität im Hinblick auf ihre Gesundheit in den FACT-G Fragebögen in den beiden untersuchten Behandlungsgruppen Vandetanib und BSC keine statistisch signifikanten Unterschiede gezeigt haben, ist doch davon auszugehen, dass allein schon das objektive Ansprechen der Behandlung in Verbindung mit einer Schmerzverminderung den Patienten in der palliativen Situation ein gewisses Maß an Zuversicht und Lebensmut geben. Bezüglich der Lebensqualität läßt sich kein Zusatznutzen belegen (siehe Tabelle 4-90). Daten zur Lebensqualität waren nur exploratorisch analysiert worden und konnten nicht aussagekräftig interpretiert werden. Ein Zusatznutzen bezüglich des Endpunkts „gesundheitsbezogene Lebensqualität“ ist für Vandetanib nicht belegt. Es ergab sich jedoch auch kein Anhaltspunkt für eine Verminderung der gesundheitsbezogenen Lebensqualität durch Vandetanib.

#### Unerwünschte Ereignisse

Bei nahezu allen Patienten beider Gruppen trat mindestens ein unerwünschtes Ereignis auf. Die Gesamtrate unerwünschter Ereignisse in der Vandetanib- bzw. Placebogruppe betrug 99,6% bzw. 90,9%. Das Ergebnis für schwerwiegende unerwünschte Ereignisse, Abbrüche aufgrund unerwünschter Ereignisse und für unerwünschte Ereignisse vom Schweregrad CTCAE Grad 3 oder höher zeigte signifikant höhere Raten in der Vandetanib-Gruppe (CTCAE Grad 3 oder höher 55,4% vs. 24,2%; SUE 30,7% vs. 13,1%). Der Anteil der

Patienten mit unerwünschten Ereignissen, die zum Absetzen der Studienbehandlung führten (12,1% vs. 3,0%, OR: 4,41, 95%-KI: [1,31;14,8]), war unter Vandetanib im Vergleich zu Placebo größer. Für diese Endpunkte zeigte sich somit ein Beleg für einen größeren Schaden (siehe Tabelle 4-90). 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 3A, Abschnitt 3.4.3). Die unter Vandetanib beobachtete Abbruchrate von 12% zeigt, dass die Nebenwirkungen für die Patienten vor dem Hintergrund ihrer Krankheitssituation insgesamt akzeptabel waren.

Zusammenfassend ist ein größerer Schaden durch Vandetanib somit belegt; sein Ausmaß kann in Betracht der Gesamtraten unerwünschter Ereignisse als eher zu vernachlässigen bezeichnet werden.

### **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 geschätzten 4000-5350 Schilddrüsenkarzinom-Neuerkrankungen in Deutschland im Jahr 2010 (61) haben nur ungefähr 5-10% MTC (also 200 bis maximal 535 Patienten (62)), wovon sich grob geschätzt gut 32% in einem fortgeschrittenen Stadium gemäß Indikation von Vandetanib befinden (vgl. Modul 3A, Abschnitt 3.2.3).

Zum gegenwärtigen Zeitpunkt ist keine effektive oder anerkannte Behandlungsoption für fortgeschrittenes MTC verfügbar. Die Patienten leiden körperlich (z.B. Schmerzen, Durchfälle) 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 herausgezögert 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 nach einer Behandlungsoption. Eine solche Therapie sollte in der Lage sein, das Fortschreiten der Erkrankung und ihrer Begleitsymptome hinauszuzögern, ein Tumorsprechen zu erreichen und sie sollte darüber hinaus auch bei Einsatz über einen langen Zeitraum hinweg gut verträglich sein.

Die Standardtherapie beim MTC ist die Operation in kurativer Intention. Für das fortgeschrittene bzw. metastasierte MTC existiert zu diesem Zeitpunkt keine zugelassene bzw. allgemein empfohlene Therapie.

Die aktuellen Behandlungsempfehlungen und Guidelines der *European Society for Medical Oncology* (ESMO; (15)) und der *American Thyroid Association* (ATA; (16)) raten vom Einsatz zytotoxischer Chemotherapie zur Behandlung von Patienten mit fortgeschrittenem MTC aufgrund der unbefriedigenden Resultate ab. Als Erstlinientherapie empfehlen beide Fachgesellschaften stattdessen die Behandlung der Patienten mit TKI im Rahmen klinischer Studien; die aktuelle Leitlinie (Version 3.2011) des *National Comprehensive Cancer Network* (NCCN) zum Schilddrüsenkarzinom nennt bei MTC bereits explizit Vandetanib als in Betracht zu ziehende Option (17).

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, wie 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 Patienten auf, deren Morbidität durch das um mehr als 11 Monate verlängerte progressionsfreie Überleben, eine Verringerung der Tumorlast und eine Verbesserung der Schmerzsituation im Vergleich mit der derzeit einzig möglichen Behandlungsalternative, BSC, relevant reduziert wird. 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.

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.

Angesichts der Schwere der Erkrankung und des erheblichen *medical need* in der Zielpopulation ist der Zusatznutzen von Vandetanib auch nach Abwägung der Vorteile (erheblicher Zusatznutzen in fast jedem der betrachteten Endpunkte) und des Schadens in der Gesamtschau zum gegenwärtigen Zeitpunkt als erheblich einzustufen, da dem gegenüber zwar eine höhere als unter BSC gemessene Rate unerwünschter Ereignisse steht, es aber nicht angemessen erscheint, in der Gesamtschau den aufgezeigten positiven Nutzen deshalb abzuwerten:

1. Auch aus klinisch-therapeutischer Sicht dürfte unter allen direkt gemessenen Endpunkten die mindestens 11-monatige Verlängerung des für diese unheilbaren Patienten hochrelevanten Endpunktes PFS von vorrangigem Wert sein und
2. waren die unerwünschten Ereignisse unter Vandetanib in der Regel hinreichend gut beherrschbar und nur 12% der Patienten brachen die Behandlung wegen unerwünschter Ereignisse ab.

Zum Zeitpunkt der primären Analyse der pivotalen Studie waren erst 15% der Patienten verstorben. Eine abschließende Bewertung kann erst nach der finalen Analyse des Gesamtüberlebens in Abhängigkeit von den eingetroffenen Ereignissen durchgeführt werden.

Zudem basieren die Feststellungen der Zulassungsbehörde hinsichtlich des positiven Nutzen-Risiko-Verhältnisses von Vandetanib - welchen gemäß AM-NutzenV diese Nutzenbewertung nicht widersprechen darf - auf einem Vergleich mit BSC. Folglich muss Vandetanib allein schon auf dieser Grundlage ein Zusatznutzen bescheinigt werden, weil gemäß Fachinformation ein unterschiedliches Nutzensausmaß bei verschiedenen Subgruppen als denkbar angesehen werden kann, dies andererseits aber nicht zum Ausschluss derjenigen Subgruppen mit möglicherweise geringerem Nutzen von der Zulassung geführt hat.

Unter Berücksichtigung obiger Erwägungen und in Anbetracht der zusätzlichen Tatsache, dass es sich bei Vandetanib um die erste zugelassene zielgerichtete Behandlung für die betreffende MTC-Teilpopulation mit aggressiver und symptomatischer Erkrankung handelt („Solist“), 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 Caprelsa- 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 vorliegt:

- 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 hervorzugehen scheint- mit dem Kriterium der Nachhaltigkeit grundsätzlich den Anspruch an 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 mindestens 11 Monate und eine Verlängerung der Zeitdauer bis zur Schmerzprogression um ca. 5 Monate aus Sicht der Patienten und der sie behandelnden Ärzte als bereits sehr nachhaltig erscheinen.

#### 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 4-91: 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 metastatischem Stadium	erheblich

Laut Text der Zulassung vom 17.02.2012 sollte bei Patienten, deren RET-Mutationsstatus nicht bekannt oder negativ ist, vor 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 Placebogruppe (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 der Patientengruppe mit RET-negativem Status lässt sich mittels der in Studie 58 gewonnenen Daten aufgrund der geringen Zahl dieser Patienten nicht statistisch signifikant bewerten. Aufgrund des Wirkprofils von Vandetanib gegen weitere in die Entstehung solider Tumore sehr häufig ursächlich involvierte Tyrosinkinase (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. Um diese Fragestellung zu beantworten, wird

AstraZeneca eine offene Studie initiieren, 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 aufgrund der Zulassung des Arzneimittels unter „Besonderen Bedingungen“ gemäß Artikel 14(7) der Verordnung (EG) Nr. 726/2004 (63) von der EMA für Vandetanib festgelegt wurden (22) 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 4-92: 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 der in Abschnitt 4.2.2 angegebenen 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 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 4-102) und Studie D4200C00068 (Tabelle 4-103) patientenrelevante Endpunkte im Sinne von Tabelle 4-1 (siehe oben) nur erfasst, 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 Zulassungsdossier an 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<sup>13</sup>, Molenberghs 2010<sup>14</sup>). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006<sup>15</sup>) 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<sup>16</sup>) soll

<sup>13</sup> Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

<sup>14</sup> 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.

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

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

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 Auffassung von AstraZeneca sind die betrachteten Endpunkte PFS, ORR, TWP und HRQoL in dieser Indikation patientenrelevant und werden entsprechend in den Abschnitten 4.1 bis 4.4 beschrieben, obgleich AstraZeneca die Einschätzungen des IQWiG aus vorausgegangenen Bewertungen bekannt sind. Der Endpunkt „Biochemisches Ansprechen“ als prognostischer Marker für die Progression wird als Surrogat für den patientenrelevanten Endpunkt krankheitsbedingte Symptome (Morbidität) 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 zur Verfügung. Die MTC-Zellen sezernieren Calcitonin (CTN) und Carcinoembryonales Antigen (CEA). Hohe Blutkonzentrationen an CTN gehen mit einem Fortschreiten der Erkrankung sowie mit Durchfall und *Flush*-Syndrom einher. Die Durchfälle können bei Patienten mit CTN-Konzentrationen über 30.000 pg/ml besonders schwerwiegende Ausmaße annehmen und regelmäßige ärztliche Intervention erfordern, was die Lebensqualität für die Patienten drastisch reduziert. Eine Senkung des CTN-Spiegels ist also im unmittelbaren Interesse der Patienten, da sie erst durch eine Verminderung der Durchfälle ihre Mobilität wieder uneingeschränkt genießen können. 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 (64). Auch CEA ist für diesen Tumor ein geeigneter Marker; erhöhte Plasmakonzentrationen dieses Parameters deuten auf fortschreitende Erkrankung und die Bildung von Metastasen hin (65).

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 Placebo einbezogen.

Die Charakteristika und die Ergebnisse der RCT D4200C00058 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 zu Vandetanib zu vervollständigen. Die Charakteristika und die Ergebnisse dieser nicht-randomisierten, offenen Studien wurden im Abschnitt 4.3.2.3 dargestellt.

Tabelle 4-93: 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	(50), (19)
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	(56), (66)
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	(57), (67)

\*in clinicaltrials.gov Fehler in der Auszeichnung, in Studienprotokoll steht Phase III

#### 4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), 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) 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.
- (3) 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.
- (4) 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.
- (5) Gemeinsamer Bundesausschuss. 5. Kapitel der Verfahrensordnung. Bewertung des Nutzens von Arzneimitteln nach § 35a SGB V. Berlin: Gemeinsamer Bundesausschuss; 2011.
- (6) 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.
- (7) 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.
- (8) Dralle H. Maligne Schilddrüsentumoren. In: Deutsche Krebsgesellschaft, Hrsg. Kurzgefasste interdisziplinäre Leitlinien: W. Zuckschwerdt Verlag; 2008.
- (9) 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.
- (10) Sippel RS, Kunnimalaiyaan M, Chen H. Current management of medullary thyroid cancer. *Oncologist.* 2008;13(5):539-47. Epub 2008/06/03.
- (11) 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.
- (12) 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.
- (13) 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.
- (14) 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.

- (15) Pacini F, Castagna MG, Brilli L, Pentheroudakis G. Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2010;21 Suppl 5:v214-9. Epub 2010/06/29.
- (16) 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.
- (17) Tuttle RM, Ball DW, Byrd D, Dilawari RA, Doherty GM, Duh Q-Y, et al. Thyroid Carcinoma. NCCN Guidelines(TM). 2011;Version 3.2011:1-87.
- (18) Salzwimmer M. Best Supportive Care bei Patienten mit Kopf-Hals-Tumoren. *Wien Med Wochenschr.* 2008;158(9-10):278-82. Epub 2008/06/19.
- (19) 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.
- (20) 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.
- (21) AstraZeneca GmbH. Fach- und Gebrauchsinformation (Zusammenfassung der Merkmale des Arzneimittels) Caprelsa® 100 mg/300 mg Filmtabletten. 2012.
- (22) Committee for Medicinal Products for Human Use (CHMP). Assessment Report Caprelsa. London: European Medicines Agency, 2011.
- (23) Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden - Version 4.0 vom 23.09.2011. In: IQWiG, Hrsg. Köln: IQWiG; 2011.
- (24) Center for Drug Evaluation and Research (CDER). Approval Package for Application Number 022405Orig1s000. In: Department of Health and Human Services, Hrsg. Silver Spring, MD: Food and Drug Administration; 2011.
- (25) Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist.* 2008;13 Suppl 2:19-21. Epub 2008/06/17.
- (26) Aidelsburger P, Wasem J. Kosten-Nutzen-Bewertungen von onkologischen Therapien. Deutsche Krebsgesellschaft, 2008.
- (27) 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.
- (28) 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>.
- (29) 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.
- (30) 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.
- (31) 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.
- (32) 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.
- (33) 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.
- (34) Committee for Medicinal Products for Human Use (CHMP). Guideline on the evaluation of anticancer medicinal products in man. London: European Medicines Agency, 2011.
- (35) Häuser W. Pharmakotherapie in der Psychoonkologie. *Onkologische Welt*. 2010;32-4.
- (36) Verband Forschender Arzneimittelhersteller e.V. (vfa). vfa-Beitrag zur G-BA Diskussionsrunde "Frühe Nutzenbewertung onkologischer Arzneimittel". Berlin: Verband Forschender Arzneimittelhersteller e.V.; 2010.
- (37) Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-47. Epub 2008/12/23.
- (38) 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.
- (39) 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.
- (40) 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.
- (41) Center for Drug Evaluation and Research (CDER). Guidance for Industry - Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. Food and Drug Administration (FDA). 2007:1-19.
- (42) Mayfield E. Progression-Free Survival: Patient Benefit or Lower Standard? *NCI Cancer Bulletin*. 2008.
- (43) 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.
- (44) 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.
- (45) 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>.
- (46) 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.
- (47) 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.
- (48) 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.
- (49) 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.
- (50) An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer [Online-Datenquelle]. 2011 [Zugriffsdatum: 13.01.2012]. Verfügbar unter: <http://www.clinicaltrials.gov/ct2/show/NCT00410761>.
- (51) 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.
- (52) AstraZeneca Pharmaceuticals LP. Vandetanib (ZD6474) for the Treatment of Medullary Thyroid Carcinoma (MTC). Clinical Overview. 2010.
- (53) FACIT.org. FACT-G Quality of Life Questionnaire (Version 4) - German. Elmhurst, IL, USA: FACIT.org; 2006.
- (54) FACIT.org. FACT-G Scoring Guidelines (Version 4). Elmhurst, IL, USA: FACIT.org; 2007.
- (55) FACIT.org. FACIT Administration and Scoring Guidelines - Scoring the FACT-G. Elmhurst, IL, USA: FACIT.org; 2010.
- (56) Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer [Online-Datenquelle]. 2011 [Zugriffsdatum: 16.01.2012]. Verfügbar unter: <http://www.clinicaltrials.gov/ct2/show/NCT00098345>.
- (57) A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer [Online-Datenquelle]. 2011 [Zugriffsdatum: 16.01.2012]. Verfügbar unter: <http://www.clinicaltrials.gov/ct2/show/NCT00358956>.
- (58) Committee for Proprietary Medicinal Products (CPMP). Points to Consider on Application with 1. Meta-Analyses; 2. One Pivotal Study. London: European Medicines Agency, 2001.
- (59) 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.
- (60) 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.
- (61) Robert-Koch-Institut. Krebs in Deutschland 2005/2006. Berlin: Robert-Koch-Institut, Gesundheitsberichterstattung des Bundes. 2010.

- (62) Reiners C, Dietlein M, Luster M. Struma maligna - Schilddrüsenkarzinome. Deutsche Medizinische Wochenschrift. 2008;133(43):2215-28.
- (63) 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.
- (64) 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.
- (65) 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.
- (66) 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.
- (67) 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.

**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:

<b>Datenbankname</b>	EMBASE	
<b>Suchoberfläche</b>	Ovid	
<b>Datum der Suche</b>	08.12.2010	
<b>Zeitsegment</b>	1980 to 2010 week 50	
<b>Suchfilter</b>	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle <sup>17</sup> ] – Strategy minimizing difference between sensitivity and specificity	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
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

Wie in Abschnitt 4.2.3.2 erläutert, war eine bibliografische Literaturrecherche nicht erforderlich.

<sup>17</sup> 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.

**Anhang 4-B: Suchstrategien – Suche in Studienregistern**

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 0 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:

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

<b>Studienregister</b>	<i>International Clinical Trials Registry Platform Search Portal</i>
<b>Internetadresse</b>	<a href="http://apps.who.int/trialsearch/">http://apps.who.int/trialsearch/</a>
<b>Datum der Suche</b>	05. Januar 2012
<b>Suchstrategie</b>	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]
<b>Treffer</b>	12

In der Datenbank *International Clinical Trials Registry Platform Search Portal* konnten zwölf Studien identifiziert werden (Tabelle 4-94).

Tabelle 4-94 (Anhang): Im *International Clinical Trials Registry Platform Search Portal* identifizierte Studien

<b>Recruitment status</b>	<b>Main ID</b>	<b>Public Title</b>	<b>Date of Registration</b>
<i>Not yet recruiting</i>	NCT01496313	<a href="#">To Compare The Effects Of Two Doses Of Vandetanib In Patients With Advanced Medullary Thyroid Cancer</a>	27.12.2011
<i>Recruiting</i>	ISRCTN68146831	<a href="#">Vandetanib in non-cisplatin fit patients with urothelial cancers</a>	20.07.2009
<i>Recruiting</i>	NCT00923247	<a href="#">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...</a>	17.06.2009
<i>Recruiting</i>	NCT00811369	<a href="#">Trial to Evaluate the Therapeutic Benefit of Fulvestrant in Combination With ZACTIMA in Postmenopausal Women With Bone Predominant, Hormone Receptor Positive Metastatic Breast Cancer</a>	15.12.2008
<i>Not recruiting</i>	NCT00537095	<a href="#">Efficacy and Safety of Zactima™ in Patients With Metastatic Papillary or Follicular Thyroid Cancer</a>	27.09.2007
<i>Not recruiting</i>	NCT00507091	<a href="#">Phase I Irinotecan, 5-Fluorouracil and Leucovorin Combination</a>	10.07.2007
<i>Not recruiting</i>	NCT00498797	<a href="#">E3-Hormone Refractory Prostrate Cancer Taxotere Combination</a>	09.07.2007
<i>Not recruiting</i>	NCT00496665	<a href="#">ZD6474 (Zactima) and Metronomic Chemotherapy in Advanced Breast Cancer</a>	03.07.2007
<i>Not recruiting</i>	NCT00410761	<a href="#">An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer</a>	06.12.2006
<i>Not recruiting</i>	NCT00358956	<a href="#">A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer</a>	28.07.2006
<i>Not recruiting</i>	NCT00098345	<a href="#">Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer</a>	07.12.2004
<i>Not recruiting</i>	NCT00047840	<a href="#">This Study is to Assess the Efficacy and Safety of ZD6474 in Subjects With Non-Small Cell Lung Cancer.</a>	18.10.2002

<b>Studienregister</b>	clinicaltrials.gov
<b>Internetadresse</b>	<a href="http://www.clinicaltrials.gov">http://www.clinicaltrials.gov</a>
<b>Datum der Suche</b>	05. Januar 2012
<b>Suchstrategie</b>	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]
<b>Treffer</b>	9

In der Datenbank clinicaltrials.gov konnten neun Studien identifiziert werden (Tabelle 4-95).

Tabelle 4-95 (Anhang): In clinicaltrials.gov identifizierte Studien

Status	Studie
<i>Active, not recruiting</i>	<a href="#">An Efficacy Study Comparing ZD6474 to Placebo in Medullary Thyroid Cancer</a> Condition: Thyroid Cancer Intervention: Drug: ZD6474 (Vandetanib) NCT ID: NCT00410761 Other IDs: D4200C00058
<i>Active, not recruiting</i>	<a href="#">A Study To Assess ZD6474 (ZACTIMA™) Monotherapy In Locally Advanced or Metastatic Hereditary Medullary Thyroid Cancer</a> Condition: Thyroid Cancer Intervention: Drug: ZD6474 (vandetanib) NCT ID: NCT00358956 Other IDs: D4200C00068
<i>Not yet recruiting</i>	<a href="#">To Compare The Effects Of Two Doses Of Vandetanib In Patients With Advanced Medullary Thyroid Cancer</a> Condition: Thyroid Cancer Intervention: Drug: 300mg vandetanib; Drug: 150mg vandetanib NCT ID: NCT01496313 Other IDs: D4200C00097, 2011-004701-24
<i>Active, not recruiting</i> <i>Has Results</i>	<a href="#">Efficacy and Tolerability of ZD6474 in Patients With Thyroid Cancer</a> Condition: Thyroid Cancer Intervention: Drug: ZD6474 (vandetanib) NCT ID: NCT00098345 Other IDs: D4200C00008
<i>Recruiting</i>	<a href="#">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</a>

Status	Studie
	<p>Conditions: Locally Advanced or Metastatic Medullary Thyroid Cancer; Medullary Thyroid Cancer</p> <p>Intervention: Behavioral: Patient outreach</p> <p>NCT ID: NCT01298323</p> <p>Other IDs: D4200C00088, 2010-023428-26</p>
<i>Active, not recruiting</i> <i>Has Results</i>	<p><a href="#">Efficacy and Safety of Vandetanib (ZD6474) in Patients With Metastatic Papillary or Follicular Thyroid Cancer</a></p> <p>Condition: Thyroid Neoplasms</p> <p>Interventions: Drug: Vandetanib; Other: Placebo</p> <p>NCT ID: NCT00537095</p> <p>Other IDs: D4200C00079</p>
<i>Recruiting</i>	<p><a href="#">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)</a></p> <p>Conditions: Breast Cancer; Colon Cancer; Prostate Cancer; Lung Cancer; Adrenal Cancer; Renal Cancer; Gastric Cancer; Ovarian Cancer; Cervical Cancer; Thyroid Cancer; Sarcoma</p> <p>Interventions: Drug: VELCADE, PS-341 (Bortezomib, NSC #681239); Drug: ZACTIMA (Vandetanib) (ZD6474)</p> <p>NCT ID: NCT00923247</p> <p>Other IDs: 090089, 09-C-0089</p>
<i>Recruiting</i>	<p><a href="#">Vandetanib to Treat Children and Adolescents With Medullary Thyroid Cancer</a></p> <p>Conditions: Multiple Endocrine Neoplasia Type 2A; Multiple Endocrine Neoplasia Type 2B</p> <p>Intervention: Drug: ZACTIMA (Vandetanib)</p> <p>NCT ID: NCT00514046</p> <p>Other IDs: 070189, 07-C-0189</p>
<i>Withdrawn</i>	<p><a href="#">Vandetanib and Docetaxel in Treating Patients With Advanced Solid Tumors</a></p> <p>Conditions: Breast Cancer; Head and Neck Cancer; Lung Cancer; Prostate Cancer; Unspecified Adult Solid Tumor, Protocol Specific</p> <p>Interventions: Drug: docetaxel; Drug: vandetanib; Genetic: proteomic profiling; Other: laboratory biomarker analysis; Other: pharmacological study</p> <p>NCT ID: NCT00937417</p> <p>Other IDs: CDR0000647578, SWOG-S0716</p>

**Studienregister** clinicalstudyresults.org

**Internetadresse** <http://www.clinicalstudyresults.org>

**Datum der Suche** 30. August 2011

**Suchstrategie** Thyroid disease [studied indication or disease] AND not yet determined for new product OR not marketed [drug name]

**Treffer** 0

Zusätzlich wurde nach Studien in der Datenbank [clinicalstudyresults.org](http://clinicalstudyresults.org) gesucht. Über die vorgegebene Eingabemaske konnte keine spezifische Auswahl zum zu bewertenden Arzneimittel getroffen werden. Durch die Suche nach Studien in der Indikation „thyroid disease“ mit der Auswahl Drug Name „*not marketed*“ oder „*not yet determined for new product*“ konnten keine Studien identifiziert werden.

Eine Aktualisierung der Suche in diesem Register zu einem späteren Zeitpunkt (Januar 2012) erwies sich aufgrund dauerhafter Nicht-Erreichbarkeit des Internetauftritts als nicht durchführbar.

**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 im Volltext gesichtet und ausgeschlossen.

**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 4-96 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
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 Studienpool enthält nur eine Studie; 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 4-97 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
D4200C00058	ja	ja	nein	ja (19)	ja (50) (NCT00410761)	ja (51)
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
<p>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.</p>						

**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 4-98 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 4-98 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-98 (Anhang): Studiendesign und -methodik für Studie D4200C00058

Item <sup>a</sup>	Charakteristikum	Studieninformation
-	<b>Studienziel</b>	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Bewertung der Wirksamkeit von Vandetanib versus Placebo
-	<b>Methoden</b>	
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	
4	Probanden / Patienten	331 randomisierte Patienten mit nicht resektablem, lokal fortgeschrittenem oder metastasiertem medullärem Schilddrüsenkarzinom
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> <li>• Bereitstellung einer schriftlichen Einverständniserklärung</li> <li>• Weiblich oder männlich im Alter von 18 Jahren und älter</li> <li>• Vorherig bestätigte histologische Diagnose von inoperablem, lokal fortgeschrittenem oder metastasiertem hereditärem oder sporadischem medullärem Schilddrüsenkarzinom. Nachweis musste in der Patientenakte bereitgestellt sein.</li> <li>• Lebenserwartung von 12 Wochen oder länger (<i>Diese Patienten wurden als am ehesten dazu fähig erachtet, die Studienprozeduren und die Behandlung zu vertragen</i>)</li> <li>• WHO PS 0-2 (<i>Diese Patienten wurden als am ehesten dazu fähig erachtet, die Studienprozeduren und die Behandlung zu vertragen</i>)</li> <li>• Fähigkeit, die Studienmedikation zu schlucken</li> <li>• Vorhandensein eines messbaren Tumors (nach vorab spezifizierten Kriterien) definiert (a) durch eine Einzelläsion, die <math>\geq 2</math>cm misst oder (b), durch folgende Kriterien bei mehreren Läsionen:</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Eine Technik mit <math>\leq 5</math> mm- Sektionen: Summe der Durchmesser <math>\geq 2</math> cm (keine Zielläsionen, die <math>&lt; 1</math> cm messen und keine Lymphknoten, die <math>&lt; 1,5</math> cm messen) ODER</p> <p>Eine Technik mit <math>&gt; 5</math> mm Sektionen: Summe der Durchmesser <math>\geq 4</math> cm (keine Zielläsion, die <math>&lt; 2</math> cm misst)</p> <ul style="list-style-type: none"> <li>• Calcitonin (<i>CTN, Calcitonin</i>) <math>\geq 500</math> pg/mL (gebräuchliche Einheit) oder <math>\geq 146,3</math> pmol/l (internationale Standardeinheit)</li> <li>• 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</li> <li>• Negativer Schwangerschaftstest für weibliche gebärfähige Patienten</li> </ul> <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> <li>• Hirnmetastasen oder Rückenmarkskompression, sofern nicht mindestens 4 Wochen vor der ersten Dosis behandelt und seit 10 Tagen stabil ohne Steroidbehandlung (<i>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 wären, in die Studienbewertungen und –prozeduren einzuwilligen</i>)</li> <li>• 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)</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>und/oder alle verbotenen Medikationen, die in dem revidierten Studienprotokoll (CSP, <i>Clinical Study Protocol</i>), Appendix E (siehe Anhang 12.1.1., Studienbericht), referenziert werden</p> <ul style="list-style-type: none"> <li>• Bedeutende Operation innerhalb 4 Wochen vor Randomisierung (<i>Dieses Ausschlusskriterium fand Anwendung, um das potenzielle Risiko der Wundheilungstoxizität für Patienten zu reduzieren</i>)</li> <li>• Die letzte Dosis der vorhergehenden Chemotherapie wurde weniger als 4 Wochen vor der Randomisierung erhalten (<i>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.</i>)</li> <li>• Strahlentherapie innerhalb der letzten 4 Wochen vor der Randomisierung (mit der Ausnahme von palliativer Strahlentherapie) (<i>Dieses Ausschlusskriterium fand Anwendung, um ein potenzielles Risiko von überschneidender Toxizität durch Strahlentherapie zu reduzieren</i>)</li> <li>• Serum Bilirubin &gt;1,5 x Obergrenze des Referenzbereichs (ULRR, <i>Upper Limit of Reference Range</i>) (<i>Dies dient als generelles Sicherheits-Ausschlusskriterium</i>)</li> <li>• Kreatinin-Clearance &lt;30 mL/min</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>(bestimmt durch die Cockcroft-Gault Formel) (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit für eine erhöhte Belastung durch Vandetanib bei Patienten mit schwerem Nierenversagen</i>)</p> <ul style="list-style-type: none"> <li>• Kalium &lt;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) (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT Verlängerung entstehen</i>)</li> <li>• Alanin-Aminotransferase, -Aspartat-Aminotransferase oder Alkalische Phosphatase &gt;2,5 x ULRR, oder &gt;5,0 x ULRR, wenn der Prüfarzt dieses mit Lebermetastasen zusammenhängend beurteilt. (<i>Dies dient als generelles Sicherheits-Ausschlusskriterium</i>)</li> <li>• 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 (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien, die durch Vandetanib-assoziierte QT-Verlängerung</i>)</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p><i>entstehen)</i></p> <ul style="list-style-type: none"> <li>• 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. (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien, die durch Vandetanib-assoziierte QTVerlängerung entstehen)</i></li> <li>• Angeborenes Long-QT-Syndrom oder Verwandter ersten Grades mit unerklärtem plötzlichen Tod mit unter 40 Jahren (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen)</i></li> <li>• QT-Verlängerung mit anderen Medikationen die eine Einstellung der Medikation erforderten (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen)</i></li> <li>• Vorhandensein eines Linksschenkelblocks (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund von Schwierigkeiten, das QT-Intervall bei Patienten mit Linksschenkelblock akkurat zu messen.</i>)</li> <li>• QTc mit Bazett's Formel unmessbar oder <math>\geq 480</math> ms beim Screening-EKG. Zu beachten: Wenn ein Patient ein QTc-Intervall <math>\geq 480</math> ms beim Screening-EKG hatte, konnte das Screening-EKG 2 Mal wiederholt werden (mindestens 24 Stunden verschoben) für insgesamt 3 EKGs.</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>(Die durchschnittliche QTc von den 3 <i>Screening</i>-EKGs musste &lt;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 <i>Screening</i>-QTc &lt; 460 ms sein. <i>(Dieses Ausschlusskriterium fand Anwendung aufgrund eines potenziell erhöhten Risikos für Arrhythmien die durch Vandetanib-assoziierte QT-Verlängerung entstehen)</i></p> <ul style="list-style-type: none"> <li>• Hypertonie, die nicht durch medikamentöse Therapie kontrolliert wird (systolischer Blutdruck &gt;160 Millimeter Quecksilbersäule (mmHg) oder diastolischer Blutdruck &gt;100 mmHg) <i>(Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit von Vandetanib, den Blutdruck zu erhöhen)</i></li> <li>• 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 <i>(Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit einer vorherigen oder aktuellen Malignität, die Studienendpunkte zu beeinflussen)</i></li> <li>• Jegliche ungeklärte chronische Toxizität durch eine vorherige Anti-Krebs-Therapie, größer als CTCAE Grad 2 <i>(Dies dient als generelles Sicherheitsausschlusskriterium)</i></li> <li>• 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) <i>(Dieses Sicherheitsausschlusskriterium fand Anwendung, um potenzielle</i></li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p><i>Interaktionen zwischen investigativen Wirkstoffen zu verhindern und um Komplikationen bei der Interpretation der Ergebnisse zu vermeiden)</i></p> <ul style="list-style-type: none"> <li>• Vorherige Exposition gegenüber Vandetanib (<i>Dieses Ausschlusskriterium fand Anwendung aufgrund der Möglichkeit, dass Vandetanib bei Patienten, die eine vorherige Behandlung mit dem Medikament erhalten haben, weniger wirksam sein könnte</i>)</li> <li>• Aktuell schwanger oder stillend (<i>Diese Ausschlusskriterium fand Anwendung aufgrund unbekannter Risiken, die während einer Schwangerschaft oder Stillzeit mit einer Vandetanib-Behandlung assoziiert sein könnten</i>)</li> <li>• Mitwirkung an der Planung und Durchführung der Studie (trifft auf AstraZeneca Mitarbeiter und den Mitarbeiterstab der Forschungseinrichtung zu) (<i>Generelles prozedurales Ausschlusskriterium fand Anwendung, um Interessenskonflikte zu vermeiden</i>)</li> <li>• Vorherige Randomisierung oder Behandlung in der gegenwärtigen Studie</li> </ul> <p>Restriktionen:</p> <ul style="list-style-type: none"> <li>• 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 Langzeitempfangnischutz, die</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Spirale oder Tubenligatur erlaubt. Orale Kontrazeption alleine war nicht akzeptabel; zusätzliche mechanische Verhütungsmethoden in Verbindung mit Spermizid mussten benutzt werden.</p> <ul style="list-style-type: none"> <li>• 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.</li> <li>• Patienten, die Blutspender waren, durften während der Studie und in den folgenden 12 Wochen nach der letzten Dosis des Studienwirkstoffs kein Blut spenden.</li> <li>• Jegliche Begleitmedikationen, die die QTc hätten beeinflussen oder die CYP3A4-Funktion hätten auslösen können (siehe revidierter CSP, [Appendix 12.1.1])</li> <li>• 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.</li> </ul>
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

Item <sup>a</sup>	Charakteristikum	Studieninformation
		Ereignissen: Möglichkeit der Dosisreduzierung auf 200 mg Vandetanib pro Tag und danach auf 100 mg pro Tag.
<b>6</b>	Zielkriterien	
<b>6a</b>	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Im Folgenden steht nach dem Zielkriterium fett gedruckt in Klammern, um welche Art von Zielkriterium es sich handelt.</p> <p>Primäres Zielkriterium:</p> <p>Nachweis einer Verbesserung des PFS mit Vandetanib verglichen mit Placebo bei Patienten mit nichtresektablem, lokal fortgeschrittenem oder metastasiertem Schilddrüsenkarzinom (<b>Wirksamkeit</b>)</p> <p>Sekundäre Zielkriterien:</p> <ol style="list-style-type: none"> <li>1. Nachweis einer Verbesserung der objektiven <i>Ansprechrate</i> (ORR, <i>Objective Response Rate</i>), Krankheitskontrollrate (DCR, <i>Disease Control Rate</i>) und Dauer des Ansprechens (DOR, <i>Duration of Response</i>) mit Vandetanib verglichen mit Placebo (<b>Wirksamkeit</b>)</li> <li>2. Nachweis einer Verbesserung des OS bei Patienten mit MTC die mit Vandetanib behandelt wurden verglichen mit Placebo (<b>Wirksamkeit</b>)</li> <li>3. Nachweis einer Verbesserung des biochemischen Ansprechens mit Vandetanib verglichen mit Placebo (gemessen durch CTN und Carcinoembryonales Antigen (CEA)) (<b>Wirksamkeit</b>)</li> <li>4. Nachweis einer Verzögerung der TWP unter Patienten mit MTC nach Behandlung mit Vandetanib verglichen mit Placebo (<b>Patienten berichteter Endpunkt (PRO, Patient Reported Outcome)</b>)</li> <li>5. Bestimmung der Pharmakokinetik (PK) von Vandetanib in dieser Patientenpopulation und Untersuchung jeglichen Einflusses von Demografie und Pathophysiologie des Patienten</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>auf die PK (<b>PK</b>)</p> <ol style="list-style-type: none"> <li>6. Bewertung der Beziehung zwischen PK und QTc, der Sicherheit, der Wirksamkeit und Biomarkern (<b>PK/Pharmakodynamik</b>)</li> <li>7. Bestimmung der Sicherheit und Verträglichkeit der Vandetanib-Behandlung bei MTC Patienten (<b>Sicherheit</b>)</li> <li>8. Bestimmung des mutationellen Status des RET-Protoonkogens in der DNA, entnommen aus Tumorproben (<b>Biomarker</b>)</li> </ol> <p>Explorative Zielkriterien:</p> <ol style="list-style-type: none"> <li>1. Untersuchung des Behandlungseffekts von Vandetanib verglichen mit Placebo auf Diarrhö bei Patienten mit MTC. (<b>PRO</b>)</li> <li>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 (<b>Biomarker</b>)</li> <li>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 (<b>Biomarker</b>)</li> <li>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 (<b>Biomarker</b>)</li> <li>5. Untersuchung von Änderungen bei Tumor-Biomarkern der Inhibition von RET, VEGFR und EGFR Signalbahnen (<b>Biomarker</b>)</li> <li>6. Nachweis einer Verzögerung des TWP bei Patienten mit</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>MTC, die an der <i>Baseline</i> keine Schmerzen haben, definiert durch das Benötigen von &lt;10 mg/Tag Morphinsulfat oder Äquivalentem, nach der Behandlung mit Vandetanib verglichen mit Placebo (<b>PRO</b>)</p> <p>7. Nachweis einer Verzögerung des TWP bei Patienten mit MTC, die an der <i>Baseline</i> Schmerzen haben, definiert durch das Benötigen von <math>\geq 10</math> mg/Tag Morphinsulfat oder Äquivalentem, nach der Behandlung mit Vandetanib verglichen mit Placebo (<b>PRO</b>)</p> <p>8. Nachweis einer Reduktion des Gebrauchs von opioidanalgetischer Medikation bei Patienten mit MTC, die an der <i>Baseline</i> Schmerzen haben, definiert durch das Benötigen von <math>\geq 10</math> mg/Tag Morphinsulfat oder Äquivalentem, nach der Behandlung mit Vandetanib verglichen mit Placebo (<b>PRO</b>)</p> <p>9. Nachweis einer Gewichtssteigerung bei Patienten mit MTC die mit Vandetanib behandelt werden verglichen mit Placebo (<b>Wirksamkeit</b>)</p> <p>10. Nachweis einer Verzögerung der Zeit bis zur Verschlechterung des WHO PS (TDPS, <i>Time to Decline in Patient WHO Performance Status</i>) bei Patienten die mit Vandetanib behandelt werden verglichen mit Placebo (<b>Wirksamkeit</b>)</p> <p>11. Untersuchung des Effekts von Vandetanib, -verglichen mit Placebo, auf die Lebensqualität (QoL, <i>Quality of life</i>) des Patienten, gemessen durch die <i>Functional Assessment of Cancer Therapy General Scale</i> (FACT-G) (<b>PRO</b>)</p> <ul style="list-style-type: none"> <li>• Radiologische Auswertung alle 12 Wochen mittels RECIST. Patienten wurden bewertet bis objektive Progression eintrat.</li> <li>• Alle medizinischen Bilder wurden durch</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>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)</p> <ul style="list-style-type: none"> <li>• 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.</li> <li>• Bei Patienten mit Tumoren, die 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.</li> <li>• 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.</li> </ul> <p>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.</p> <p>Tumor-Biopsieproben wurden per Standard-Core-Biopsietechniken oder per Feinnadelbiopsieproben in Paraffinblöcken entnommen.</p> <ol style="list-style-type: none"> <li>1. Biopsie: Vor der ersten Dosis des Studienwirkstoffs</li> <li>2. Biopsie: In Woche 12 nach der Randomisierung (solange der Patient sein Einverständnis nicht zurückgezogen hat)</li> </ol> <p>Alle Tumorproben wurden bei Raumtemperatur zu dem Zentrallabor geschickt, welches vertraglich von AstraZeneca bestimmt wurde.</p> <p><u>Erhebungszeitpunkte:</u>  <i>Screening</i> (Tag -21 bis 0 vor</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>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 <i>Follow-up</i>, Überleben-<i>Follow-up</i></p> <p><u>Erhobene Parameter:</u></p> <p>Einmalig erhoben:</p> <p>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</p> <p>Im Studienverlauf mehrmals erhoben (genaue Erhebungszeitpunkte siehe Studienbericht):</p> <p>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*</p> <p>*= Parameter, die auch in der offenen Phase nach Progression erhoben wurden (zusätzlich: Rückgabe der verblindeten Medikamente, offene Vandetanib / Placebo Ausgabe / Rückgabe)</p>
<b>6b</b>	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>18. Mai 2009</p> <ul style="list-style-type: none"> <li>• 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-</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Zielkriterien überarbeitet.</p> <p>Begründung: Um TWP als einziges sekundäres PRO-Zielkriterium und alle anderen PRO-Zielkriterien als explorative Zielkriterien zu identifizieren.</p> <ul style="list-style-type: none"> <li>• Patientengewicht wurde von sekundärem zu explorativem Zielkriterium geändert.</li> </ul> <p>Begründung: Konsistent mit oben erwähnter Änderung sekundäre Zielkriterien bis auf ORR, DCR, DOR, biochemischem Ansprechen und TWP einzuschränken</p>
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 des PFS bei 2-seitigem $\alpha = 0.05$ 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 <i>Follow-up</i> 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	Sobald Patienten für die Studie überprüft werden, müssen sie einen Einschreibungscode (E-code, <i>Enrollment Code</i> ) zugewiesen bekommen. Dieser E-code setzt sich zusammen aus der Zentrumsnummer und der

Item <sup>a</sup>	Charakteristikum	Studieninformation
	Zuteilung gewährleistet war	<p>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 (eCSF, <i>electronic Case Report Form</i>) benutzt. Bei Eintritt eines Patienten in die Überprüfung sollte der Prüfarzt die zentrale Registrierung / das Randomisierungszentrum per Telefon anrufen, um den Patienten zu registrieren.</p> <p>Ablauf des gesamten Prozesses:</p> <ul style="list-style-type: none"> <li>• Patient wird identifiziert zur zentralen Registrierung / Randomisationszentrum mittels Initialen, E-code und Geburtstag.</li> <li>• IVRS (<i>Interactive Voice Response Services</i>) Die zentrale Registrierung / Randomisierungszentrum informiert den Prüfarzt, die MedID-Nummer dem Patienten beim Randomisierungsbesuch zuzuweisen.</li> <li>• Der Prüfarzt wird für jede nachfolgende Ausgabe-Visite IVRS hinzuziehen, um eine neue MedID-Nummer zuzuweisen</li> </ul> <p>Bei den Visiten zugewiesene MedID-Nummer entspricht der Behandlung zu der der Patient ursprünglich randomisiert wurde oder er wird eine Dosisreduzierung erhalten, wenn nötig.</p>
<b>10</b>	<p>Randomisierung, Durchführung</p> <p>Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?</p>	<ul style="list-style-type: none"> <li>• Bei Eintritt eines Patienten in die Studie sollte der Prüfarzt die zentrale Registrierung / das Randomisierungszentrum per Telefon anrufen, um den Patienten zu registrieren.</li> <li>• Zuteilung zur Behandlung durch Randomisierungsschema, welches von einer Computersoftware produziert wird.</li> </ul>
<b>11</b>	Verblindung	
<b>11a</b>	<p>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?</p>	<p>Studienmedikation wird durch MedID-Nummer gekennzeichnet, die mit dem Randomisierungsschema verknüpft ist. Patienten und Personal des Prüfzentrums waren verblindet, sonst keine weitere Informationen</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
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	<p>Es wurden verschiedene Analysepopulationen festgelegt:</p> <ol style="list-style-type: none"> <li>1. ITT-Population – alle randomisierten Patienten</li> <li>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</li> <li>3. Sicherheitspopulation – alle randomisierten Patienten, die mindestens eine Dosis der randomisierten Behandlung erhalten haben (d. h. mindestens eine Dosis von Vandetanib/Placebo)</li> <li>4. PK-Population – alle randomisierten Patienten mit validen Plasma-Konzentrationen von Vandetanib, die zur Vandetanib-Gruppe randomisiert wurden</li> <li>5. Offene Population– alle randomisierten Patienten, die mindestens eine Dosis der offenen Behandlung bekommen haben</li> </ol> <p>Variablen zum Zielkriterium Wirksamkeit:</p> <ol style="list-style-type: none"> <li>a. PFS <ul style="list-style-type: none"> <li><u>Primäranalyse</u></li> <li>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.</li> <li>Folgende Hypothesen wurden getestet:</li> <li>H0: Kein Unterschied zwischen 300 mg und Placebo</li> <li>H1: Unterschied zwischen 300 mg</li> </ul> </li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>und Placebo</p> <ul style="list-style-type: none"> <li>• Basierend auf ITT-Population wurde ein Log-Rank-Test durchgeführt</li> <li>• Angabe der Punktschätzer für das mediane PFS für jede Behandlungsgruppe (grafische Darstellung mit Kaplan-Meier)</li> <li>• Zusammenfassung des Progressionsstatus von Patienten zur Zeit der Analyse</li> <li>• Zusammenfassung der Anzahl an Patienten, die eine Krebs-Folgetherapie nach der Beendigung der randomisierten Behandlung erhielten (Zusätzlich Zusammenfassung über diese ersten Folgetherapien)</li> <li>• Zusammenfassung der medianen Dauer des <i>Follow-Up</i></li> </ul> <p><u>Sensitivitätsanalysen</u></p> <p>Für PFS:</p> <ul style="list-style-type: none"> <li>• PP-Analyse – Log-Rank-Test basierend auf PP-Population</li> <li>• PFS von zentraler Beurteilungsstelle, Daten in offener Phase ausgeschlossen – Log-Rank-Test</li> <li>• PFS von einrichtungseigener Beurteilungsstelle – Log-Rank-Test</li> <li>• 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</li> <li>• PFS von zentraler Beurteilungsstelle, keine</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Adjustierung für verkalkte Läsionen – Log-Rank-Test</p> <ul style="list-style-type: none"> <li>• PFS von zentraler Beurteilungsstelle, keine Adjustierung für hypodense oder hypointensive Läsionen – Log-Rank-Test</li> <li>• Whitehead’s Methode (Methode zum gruppierten Überleben) um den Einfluss einer unterschiedlichen Häufigkeit von Bewertungen zwischen Behandlungsgruppen zu schätzen</li> <li>• Globaler Interaktionstest mit 1%-Signifikanzlevel für das Cox-Modell</li> </ul> <p><u>Subgruppenanalysen:</u> Mittels Log-Rank-Test, grafische Darstellung mittels <i>Forest Plot</i> Folgende Subgruppen wurden vor der Entblindung vorgegeben:</p> <ul style="list-style-type: none"> <li>• RET-Mutationsstatus</li> <li>• CTN-Verdopplungszeit</li> <li>• CEA-Verdopplungszeit</li> <li>• Anzahl an vorherigen Therapien</li> <li>• Ansprechen auf die jüngste vorherige Therapie</li> <li>• MTC-Status</li> <li>• Geschlecht</li> <li>• Phase/Ausmaß der Erkrankung bei Eintritt</li> <li>• <i>Baseline</i> WHO PS</li> <li>• ethnische Zugehörigkeit</li> <li>• <i>Baseline</i> Werte für p-VEGF, p-VEGFR2 und p-bFGF</li> </ul> <p>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 <i>Odds Ratio</i></p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p><u>Sensitivitätsanalyse</u>            Logistische Regression mit Behandlungseffekt, RET-Mutationsstatus, CTN-Verdopplungszeit, CEA-Verdopplungszeit, Anzahl vorheriger Therapien, Ansprechen auf jüngste vorherige Therapie und MTC-Status</p> <p>c. DCR</p> <p><u>Primäranalyse</u>            Logistische Regression mit Kovariable „Behandlung“ (als einzige Kovariable, die auf der ITT-Population basiert)            Resultate in Form einer Schätzung des <i>Odds Ratio</i></p> <p><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 <i>Odds Ratio</i></p> <p>d. DOR            2 Arten der Definition:</p> <ul style="list-style-type: none"> <li>• Zeit von der Randomisierung bis zum Tag der nachgewiesenen Krankheitsprogression laut zentraler Beurteilungsstelle oder Tod in Abwesenheit einer Krankheitsprogression</li> <li>• Zeit vom Tag des ersten nachgewiesenen Ansprechens bis zum Tag der nachgewiesenen Progression laut zentraler Beurteilungsstelle oder Tod in Abwesenheit einer Krankheitsprogression</li> </ul> <p>DOR wurde zusammengefasst als Median des DORs nach beiden Definitionen            Kaplan-Meier-Kurven für DOR wurden erstellt.</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>e. OS</p> <p>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</p> <p><u>Primäranalyse</u></p> <p>Log-Rank-Test basierend auf der ITT-Population, grafische Darstellung durch Kaplan-Meier-Kurven.</p> <p><u>Sensitivitätsanalysen</u></p> <ul style="list-style-type: none"> <li>• Log-Rank-Test basierend auf der PP-Population</li> <li>• Cox-PH-Modell basierend auf ITT-Population (mit den gleichen Termen wie Cox- PH-Modell für PFS)</li> </ul> <p>Globaler Interaktionstest wurde basierend auf dem Cox-Modell durchgeführt</p> <p>f. Biochemisches Ansprechen</p> <p>Bestes CEA- und CTN- Ansprechen wird an der <i>Baseline</i> und während der Behandlung bewertet. Therapieansprecher sind die Patienten, die das beste biochemische Ansprechen bei CR und PR haben</p> <p>Definitionen:</p> <ul style="list-style-type: none"> <li>• Komplettes Ansprechen (CR, <i>Complete Response</i>): Komplette Normalisierung der CEA/CTN-Level nach der Behandlung, bestätigt durch eine wiederholte Prüfung des CEA/CTN-Level</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Partielles Ansprechen (PR, <i>Partial Response</i>) : Mindestens 50% Reduzierung des CEA/CTN-Levels (Referenz = <i>Baseline</i> - Level)</li> <li>• Progressive Erkrankung (PD, <i>Progressive Disease</i>): Mindestens 50% Erhöhung des CEA/CTN-Levels (Referenz = <i>Baseline</i> - Level)</li> <li>• Stabile Erkrankung ( <i>Stable Disease</i>): Weder ausreichende Normalisierung, um als PR zu gelten, noch ausreichende Normalisierungserhöhung, um als PD zu gelten (Referenz = <i>Baseline</i> - Level)</li> </ul> <p>g. Klinische Gewichtszunahme Definiert als Gewichtszunahme um &gt;5% gegenüber des <i>Baseline</i>-Werts und nicht gefolgt von einer Gewichtsüberprüfung, die eine Gewichtsabnahme gegenüber dem <i>Baseline</i>-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.</p> <p>h. TDPS Zu gegebenem Zeitpunkt war eine Verschlechterung des WHO PS als <math>\geq 1</math> Punkt-Steigerung des <i>Baseline</i>-Werts definiert. TDPS = Intervall vom Tag der Randomisierung bis zur ersten</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Feststellung einer Verschlechterung  <u>Primäranalyse</u>            Log-Rank-Test basierend auf ITT-Population            Resultate als <i>Hazard Ratios</i>            Zusätzlich: Punktschätzer des medianen TDPS. Grafische Darstellung: Kaplan-Meier-Kurven</p> <p>Variablen zu PRO:</p> <p>a. Patientenberichtete Schmerz erfassung (BPI, <i>Brief Pain Inventory</i>)            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 <i>Baseline</i>-Wert zu erstellen, dann einmal wöchentlich bis Entblindung.</p> <p>Erhobene Bewertungen:</p> <ul style="list-style-type: none"> <li>• Einzelpostenbewertung des Postens „Schlimmster Schmerz“</li> <li>• Arithmetisches Mittel der Bewertungen des 4-Posten-Schmerzstärke-Index</li> <li>• Arithmetisches Mittel der Bewertungen des 7-Posten Schmerz-Beeinträchtigungs-Index</li> </ul> <p>b. Opioidanalgetika-Gebrauch</p> <ul style="list-style-type: none"> <li>• Selbst-berichtend</li> <li>• Konvertiert zum Äquivalent zu Morphinsulfat in mg/Tag</li> <li>• Für jeden Patienten individuell aufgelistet</li> <li>• Zusammenfassung des Gebrauchs und der</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Änderungen im Gebrauch ausgehend vom <i>Baseline</i>-Wert</p> <p>Liniendiagramm der mittleren Änderung vom <i>Baseline</i>-Wert</p> <p>c. TWP</p> <ul style="list-style-type: none"> <li>• Abgeleitet vom Posten „Schlimmster Schmerz“ des BPI-SF oder vom Opioidanalgetika-Gebrauch.</li> <li>• Grad der Schmerzprogression entsprach entweder einer Erhöhung von <math>\geq 2</math> Punkten vom <i>Baseline</i>-Wert des Postens „Schlimmster Schmerz“ oder einer Erhöhung des Opioidanalgetika Gebrauchs um <math>\geq 10</math> mg/Tag vom <i>Baseline</i>-Wert.</li> <li>• Grad der Verbesserung des Schmerzes entsprach (a) entweder einer Abnahme von <math>\geq 2</math> Punkten vom <i>Baseline</i>-Wert des Postens „Schlimmster Schmerz“, während der Gebrauch von Opioidanalgetika um nicht mehr als 10 mg/Tag gegenüber dem <i>Baseline</i>-Wert zugenommen hat oder (b) einer Abnahme im Opioidanalgetika-Gebrauch gegenüber der vorherigen Visite um <math>&gt;50\%</math> ohne eine Zunahme des Postens „Schlimmster Schmerz“ um <math>\geq 2</math> Punkte vom <i>Baseline</i>-Wert</li> <li>• 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.</li> </ul> <p><u>Primäranalyse</u></p> <ul style="list-style-type: none"> <li>• Log-Rank-Test für TWP</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Resultate als <i>Hazard Ratios</i></li> <li>• Punktschätzer des medianen TWP für jede Behandlungsgruppe und grafische Darstellung des TWP in Kaplan-Meier - Kurven</li> </ul> <p><u>Sensitivitätsanalysen</u></p> <ul style="list-style-type: none"> <li>• Log-Rank-Test (basierend nur auf „Schlimmster Schmerz“ und nicht auf Opioidanalgetika-Gebrauch)</li> <li>• Cox-PH-Modell basierend auf ITT-Population (mit den gleichen Termen wie Cox-PH-Modell für PFS)</li> <li>• 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.</li> </ul> <p>Die Subgruppenanalyse wurde in der ITT-Population mittels Log-Rank-Test für Opioidanalgetika-Gebrauch durchgeführt</p> <p>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 <i>Odds Ratios</i></p> <p>e. Dauer des Opioidanalgetika-Gebrauchs Als Liste und Zusammenfassung</p> <p>f. FACT-G Mittels FACT-G Fragebogen, der</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>alle 12 Wochen bis Beendigung der randomisierten Studie vom Patienten ausgefüllt wurde.</p> <p>Berechnet wurden:</p> <p>Totaler FACT-G Wert und die 4 Subskalen Physisches Wohlbefinden (PWB), soziales Wohlbefinden, emotionales Wohlbefinden (EWB) und funktionales Wohlbefinden (FWB)</p>
<b>12b</b>	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Siehe 12a
-	<b>Resultate</b>	
<b>13</b>	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
<b>13a</b>	<p>Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die</p> <p>a) randomisiert wurden,</p> <p>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,</p> <p>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden</p>	<p>a) Vandetanib: 231, Placebo: 100</p> <p>b) Vandetanib: 231, Placebo: 99</p> <p>c) Patienten in ITT-Population: Vandetanib: 231, Placebo: 100</p> <p>Patienten in PP-Population: Vandetanib: 215, Placebo: 91</p>
<b>13b</b>	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>1 Patient (1%, Placebogruppe) erhielt keine randomisierte Behandlung aufgrund Tod</p> <p>Protokoll Abweichungen (Vandetanib (V):16, Placebo (P):9):</p> <ul style="list-style-type: none"> <li>• Abweichungen von den Basis RECIST-Kriterien (V:2 (0,9%), P:2 (2%)): <ul style="list-style-type: none"> <li>○ Basis RECIST Überprüfung mehr als 28 Tage vor der ersten Dosis der randomisierten Behandlung: V:2 (0,9%), P:2 (2%)</li> </ul> </li> <li>• Abweichung von den Einschlusskriterien (V:10 (4,3%), P:5 (5%)): <ul style="list-style-type: none"> <li>○ Kein messbarer Tumor am <i>Baseline</i>: V:9 (3,9%), P:5 (5%)</li> <li>○ Keine vorherig bestätigte histologische Diagnose von inoperablem, lokal fortgeschrittenem oder metastasiertem, hereditärem oder sporadischem medullärem</li> </ul> </li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Schilddrüsenkarzinom: V:1 (0,4%), P:0 (0%)</p> <ul style="list-style-type: none"> <li>• Abweichungen von den Ausschlusskriterien (V:1 (0,4%), P:0(0%)): <ul style="list-style-type: none"> <li>○ Vorherige oder aktuelle Malignitäten anderer Histologien innerhalb der letzten 5 Jahre (mit Ausnahme der im Protokoll spezifizierten): V:1 (0,4%), P:0 (0%)</li> </ul> </li> <li>• Abweichungen von der Begleitmedikation (V:0 (0%), P:1(1%)): <ul style="list-style-type: none"> <li>○ Einnahme von CYP3A4 Induzierern für mindestens 14 Tage während der randomisierten Behandlung; V:0 (0%), P:1(1%)</li> </ul> </li> <li>• Abweichungen von der Randomisierung (V:4 (1,7%), P:1 (1%)): <ul style="list-style-type: none"> <li>○ Mindestens eine Dosis der falschen randomisierten Behandlung erhalten: V:4 (1,7%), P:0 (0%)</li> <li>○ Wurde randomisiert, hat aber keine randomisierte Behandlung erhalten: V:0 (0%), P:1 (1%)</li> </ul> </li> </ul>
<b>14</b>	Aufnahme / Rekrutierung	
<b>14a</b>	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Erster Patient eingeschrieben: 23. November 2006</p> <p>Letzter Patient eingeschrieben: 19. Oktober 2007</p> <p>Daten <i>Cut-off</i>: 31. Juli 2009</p> <p>Mediane Dauer des <i>Follow-up</i> nach Daten <i>Cut-off</i>: 103 Wochen</p>
<b>14b</b>	Informationen, warum die Studie endete oder beendet wurde	Studie endet: voraussichtlich Dezember 2016
<b>a: nach CONSORT 2010.</b>		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

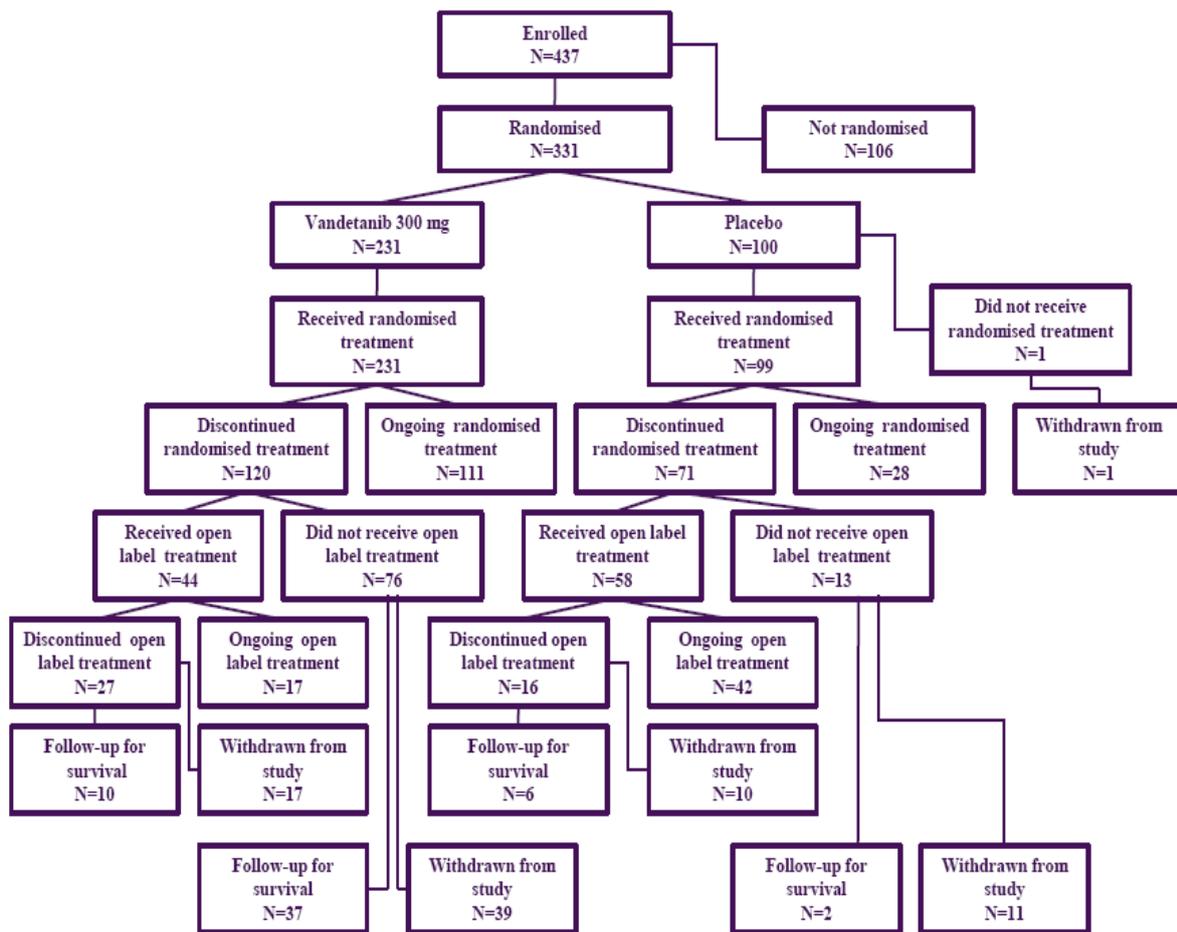


Abbildung 17: Flussdiagramm Patientenverteilung Studie D4200C00058

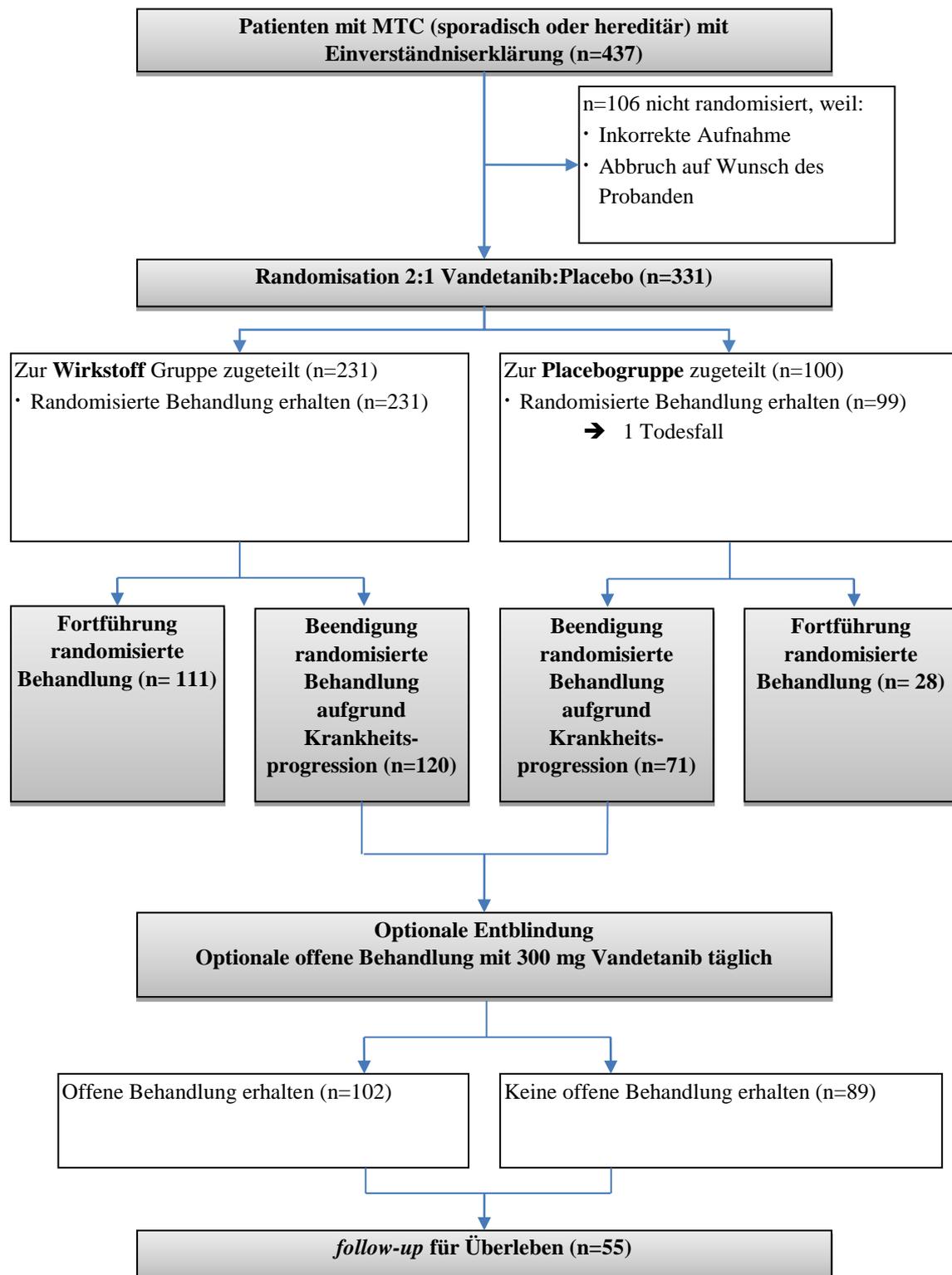


Abbildung 18: Flussdiagramm Studiendesign Studie D4200C00058

Tabelle 4-99 (Anhang): Studiendesign und -methodik für Studie D4200C00008

Item <sup>a</sup>	Charakteristikum	Studieninformation
-	<b>Studienziel</b>	
3	Zielsetzungen	<p>Primäres Ziel:</p> <p>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.</p> <p>Sekundäre Ziele:</p> <ol style="list-style-type: none"> <li>1. Untersuchung des biochemischen Ansprechens von Calcitonin (CTN) bei Behandlung mit Vandetanib</li> <li>2. Untersuchung, ob Vandetanib zu einem Rückgang der Stuhlgangshäufigkeit und Verbesserung der Stuhlgangkonsistenz bei Patienten mit symptomatischem Durchfall in Verbindung mit vererbbaeren hereditärem MTC vorliegt</li> <li>3. Untersuchung des progressionsfreien Überlebens (<i>Progression-Free Survival</i>, PFS)</li> <li>4. Untersuchung der Krankheitskontrollrate (<i>Disease Control Rate</i>, DCR), der Dauer des objektiven Ansprechens (<i>Duration of Objective Response</i>, DOR) und der Dauer der Krankheitskontrolle (<i>Duration of Disease Control</i>)</li> <li>5. Bewertung des WHO <i>Performance Status</i> (WHO PS)</li> <li>6. Beschreibung der Pharmakokinetik in dieser Population</li> <li>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</li> <li>8. Untersuchung der Sicherheit und Verträglichkeit von Vandetanib in dieser Patientenpopulation</li> </ol> <p>Explorative Ziele:</p> <ol style="list-style-type: none"> <li>1. Untersuchung des basalen CTN Levels nach einem <i>calcium-pentagastrin stimulation test</i> in ausgewählten Patienten mit einem kompletten biochemischen Ansprechen.</li> <li>2. Untersuchung, ob eine Veränderung des Baseline Plasma Level des <i>Vascular Endothelial Growth Factor</i> (VEGF), <i>Basic Fibroblast Growth Factor</i> (bFGF) und CEA prognostischen Wert bei Patienten mit MTC das Ansprechen von Vandetanib verändert.</li> <li>3. Effekt von Vandetanib auf DNA Mikroarray Muster</li> <li>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)</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
-	<b>Methoden</b>	
4	Studiendesign	Internationale, einarmige, multizentrische, offene, zweistufige 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	<p><i>Kohortenstudie</i> (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl an; Methoden der Nachbeobachtung)</p> <p><i>Fallkontrollstudie</i> (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)</p> <p><i>Querschnittsstudie</i> (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl)</p>	<p>Einschlusskriterien:</p> <ol style="list-style-type: none"> <li>1. Einverständniserklärung unterschrieben</li> <li>2. Männer oder Frauen im Alter <math>\geq 18</math></li> <li>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.</li> <li>4. WHO <i>Performance Status</i> 0-2</li> <li>5. Fähig, die Studienmedikation zu schlucken</li> <li>6. Mindestens eine messbare Läsion von mindestens 1 cm Durchmesser im CT-Aufnahme oder 2 cm mit konventionellen Techniken</li> <li>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</li> <li>8. Schlucken der Medikation als ganze Tablette ohne Kauen, Zerkleinern oder Teilen</li> </ol> <p>Ausschlusskriterien:</p> <ol style="list-style-type: none"> <li>1. Gehirnmetastasen oder Rückenmarkskompression, sofern nicht mind. 4 Wochen vor der ersten Dosis bestrahlt und stabil ohne Behandlung mit Steroiden für mind. eine Woche</li> <li>2. Nachweis eines Phäochromozytoms, basierend auf einem Vorkommen von Katecholaminen im Urin in</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>der 24h Untersuchung</p> <ol style="list-style-type: none"> <li>3. Weniger als <math>1.5 \times 10^9/L</math> Neutrophile oder weniger als <math>100,000/mm^3</math> Platelets</li> <li>4. Serum Bilirubin größer als <math>1.5 \times</math> oberes Limit des Referenzwertes (ULRR)</li> <li>5. Serum Kreatinin größer als <math>1.5 \times</math> ULRR</li> <li>6. Kaliumkonzentration kleiner <math>3.5 \text{ mEq/l}</math>, Kalzium (ionisiertes Kalzium oder für Albumin adjustiertes), oder Magnesium Konzentrationen außerhalb der Normalwerte, sofern nicht von AstraZeneca genehmigt. Zugabe von Elektrolyten ist möglich.</li> <li>7. ALT oder Aspartate Aminotransferase (AST) größer als <math>2.5 \times</math> ULRR, bzw. ALT oder AST größer als <math>5.0 \times</math> ULRR, falls der Prüfarzt einen Zusammenhang mit Lebermetastasen diagnostiziert oder alkalische Phosphatase (ALP) größer als <math>2.5 \times</math> ULRR</li> <li>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.</li> <li>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</li> <li>10. Vorgeschichte von Arrhythmien (multifokale ventrikuläre Extrasystolen, Bigeminus, Trigemini, 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.</li> <li>11. Chronisches Vorhofflimmern</li> <li>12. Vorgeschichte von QTc Intervall Verlängerungen bei anderen Medikamenten</li> <li>13. Angeborenes langes QT Syndrom</li> <li>14. QTc nach Bazett-Korrektur nicht messbar oder <math>\geq 480</math> ms bei der ECG Screening-Untersuchung. Zu beachten: Wenn ein Patient ein QTc Intervall <math>\geq 460</math> ms beim Screening EKG hatte, konnte der Screening EKG nach mindestens 24h wiederholt werden. Die durchschnittliche QTc beider EKGs musste dann unter <math>460</math> 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.</li> <li>15. Verbotene Begleitmedikation in dieser Studie</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>einschließlich 5HT3 Antagonisten: starke und mittlere CYP3A4 Hemmer, starke und mittlere CYP3A4Induktoren, Medikationen, die das QTc Intervall verlängern oder <i>Torsades de Pointes</i> hervorrufen, sofern nicht von AstraZeneca zugelassen. Siehe auch Kapitel 3.4.9.1 „<i>Cancer and other concomitant treatment</i>“ im Studienprotokoll.</p> <p>16. Vorgeschichte von linksventrikulärer Ejektionsfraktion (LVEF) weniger als 45%, gemessen durch <i>multi-gated acquisition scan</i> (MUGA) oder Echokardiogramm (ECHO) für Patienten mit früherer Anthrazyklin Therapy (Gesamtdosis &gt; 450 mg/m<sup>2</sup>) oder signifikante kardiovaskuläre Krankheit oder Brustbestrahlung, wie vom Prüfarzt bestimmt.</p> <p>17. Unkontrollierte Hypertension (systolischer Blutdruck &gt; 160 mm Hg oder diastolischer Blutdruck &gt; 100 mm Hg)</p> <p>18. Jegliche andere Malignität oder Vorgeschichte einer Malignität in den letzten 5 Jahren, sofern nicht von AstraZeneca zugelassen (mit Ausnahme von Nichtmelanom-Hautkrebs oder <i>Carcinoma in situ</i> des Zervix)</p> <p>19. Chemotherapie und / oder Bestrahlungstherapie, falls diese nicht in den letzten 4 Wochen vor erster Studienmedikation endeten</p> <p>20. Jegliche ungelöste chronische Toxizität mit CTCAE Grad größer 2 von vorherigen Antitumor Therapie, sofern nicht von AstraZeneca zugelassen.</p> <p>21. Teilnahme an einer klinischen Studie in den letzten 30 Tagen (abgesehen von Studien zur Lebensqualität)</p> <p>22. Aktuell schwanger oder stillend</p> <p>Herkunft: Patienten aus der USA und Frankreich</p>
<b>6b</b>	<p><i>Kohortenstudie</i> (Matching (Paarbildung) und Matchingkriterien, Anzahl der exponierten und der nicht exponierten Teilnehmer)</p> <p><i>Fallkontrollstudie</i> (Matching (Paarbildung) und Matchingkriterien, Anzahl der Kontrollen pro Fall)</p>	Nicht zutreffend
<b>7</b>	Variablen	<p>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.</p> <p>Sekundäre Zielkriterien: 1. Biochemisches Ansprechen: Änderung gegenüber des Baselinewertes bei Calcitonin (CTN) nach</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Behandlung mit Vandetanib</p> <ol style="list-style-type: none"> <li>2. Symptomatisches Ansprechen: Änderung gegenüber des Baselinewertes der Stuhlgangshäufigkeit und -konsistenz bei Patienten mit symptomatischem Durchfall in Verbindung mit vererbbaarem hereditärem MTC nach Behandlung mit Vandetanib</li> <li>3. Progressionsfreies Überleben (<i>Progression Free Survival</i>, PFS) gemessen von der ersten Dosisgabe bis Progression oder Tod</li> <li>4. <i>Disease Control Rate</i> (DCR; Rate der Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 24 Wochen)</li> </ol> <p>Dauer des objektiven Ansprechens (<i>Duration of Objective Response</i>, 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.</p> <p>Dauer der Krankheitskontrolle (<i>Duration of Disease Control</i>) gemessen von der ersten Dosisgabe bis zur Progression.</p> <ol style="list-style-type: none"> <li>5. Bewertung des WHO <i>Performance Status</i> (WHO-PS)</li> </ol> <p>Explorative Zielkriterien:</p> <ol style="list-style-type: none"> <li>1. FDG-PET (2-[F-18] fluor-2-deoxy-D-glucose positron emission tomography)</li> <li>2. Calcium-pentagastrin stimulation test</li> <li>3. CEA levels (<i>Carcinoembryonic antigen</i>)</li> <li>4. DNA (<i>Deoxyribonucleic acid</i>) Mikroarray Muster</li> <li>5. FACIT-D (<i>Functional Assessment of Chronic Illness Therapy – Diarrhea</i>)</li> <li>6. RET (<i>Rearranged during Transfection (proto-oncogene)</i>), phospho-RET, EGFR (<i>Epidermal growth factor receptor</i>), phospho-EGFR, VEGFR-2 (<i>Vascular endothelial growth factor receptor</i>), phospho-VEGFR-2, MAPK (<i>Mitogen-activated protein kinase</i>), phospho-MAPK, SHC protein (<i>Src homology 2 domain containing</i>), phospho-SHC protein</li> </ol>
8	Datenquellen/Messmethoden	Für das primäre Zielkriterium wurden radiologische Untersuchungen durchgeführt und entsprechend den RECIST Kriterien bewertet.
9	Bias	<p>Ein zentraler Review der CT/MRI <i>scans</i> für den primären Endpunkt wurde durchgeführt um einen Zentrumsbias zu vermeiden.</p> <p>Beim FACIT-D wurden entsprechende Maßnahmen getroffen, falls sich Hinweise auf eine Symptomatik bei den fehlenden Werten ergaben, um potentiellen Bias zu vermeiden.</p>
10	Studiengröße	Die Studie hatte ein zweistufiges Design mit jeweils 15 geplanten Patienten pro Stufe. Falls nach den ersten 15

Item <sup>a</sup>	Charakteristikum	Studieninformation
		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 <i>Confounding</i>	Grundlage: <i>Intention-To-Treat</i> -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“ „ <i>Duration of Disease Control</i> “ und „DOR“, einschließlich 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	<i>Kohortenstudie</i> (Umgang mit vorzeitigem Ausscheiden („loss to Follow-up“)) <i>Fallkontrollstudie</i> (Berücksichtigung des <i>Matchings</i> (Paarbildung) von Fällen und Kontrollen bei der Auswertung) <i>Querschnittsstudie</i> (Auswertungsmethoden, die die gewählte Strategie zur Stichprobenauswahl ( <i>Sampling strategy</i> ) berücksichtigen)	Patienten mit vorzeitigem Ausscheiden wurden als non-Responder betrachtet.
12e	Sensitivitätsanalysen	Nicht zutreffend
-	<b>Ergebnisse</b>	
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	35 Patienten wurden gescreent. 30 Patienten wurden in die Studie eingeschlossen und behandelt. 13 Patienten haben die Studie vorzeitig abgebrochen.

<b>Item<sup>a</sup></b>	<b>Charakteristikum</b>	<b>Studieninformation</b>
	Nachbeobachtung abgeschlossen wurde und deren Daten ausgewertet wurden)	
<b>13b</b>	Gründe für die Nicht-Teilnahme in jeder Studienphase	Nicht zutreffend
<b>13c</b>	Flussdiagramm	Siehe Flow Chart unten.
<b>17</b>	Weitere Auswertungen	Nicht zutreffend
<b>a: nach STROBE 2007</b>		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

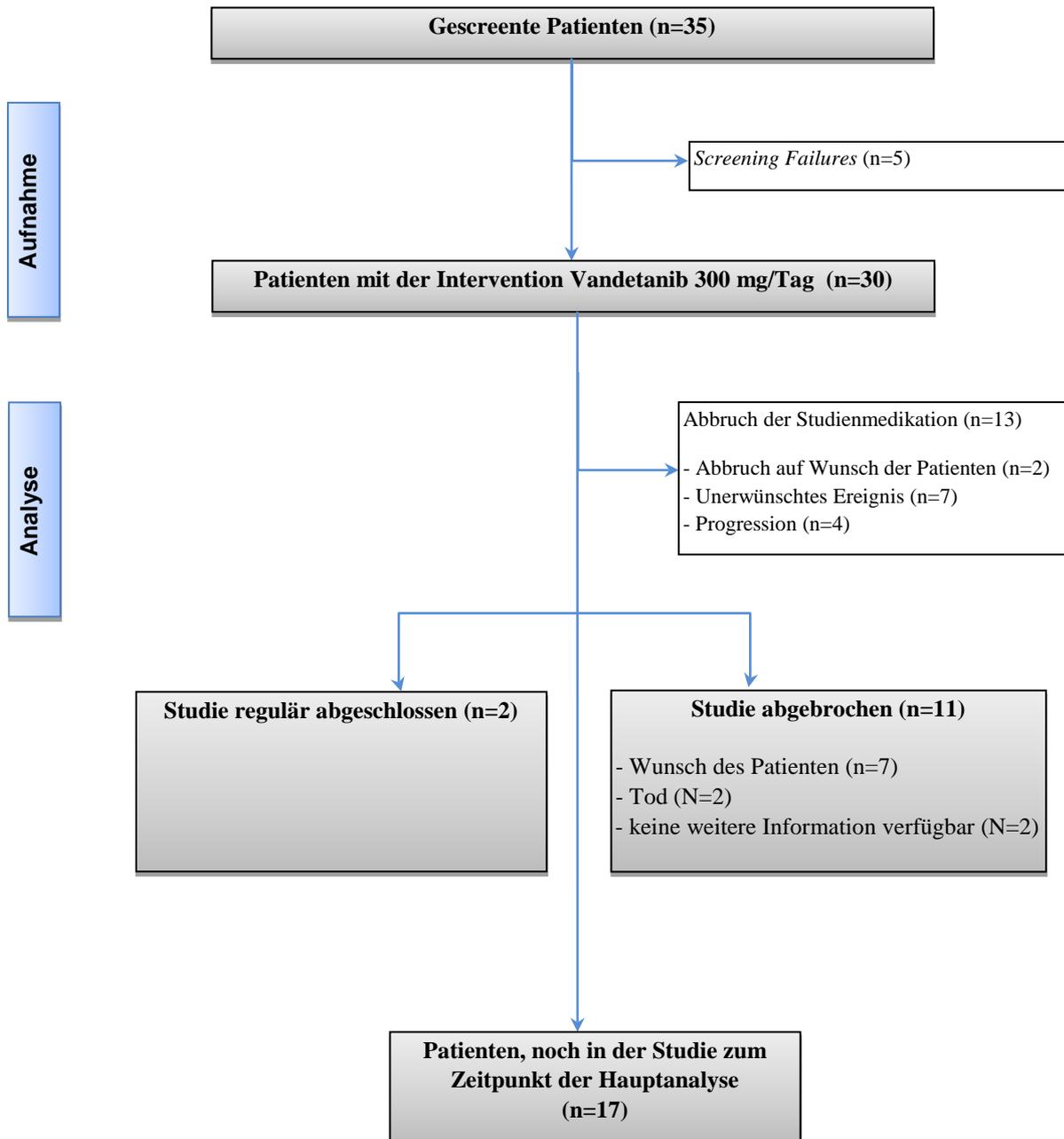


Abbildung 19: Flussdiagramm Studiendesign Studie D4200C00008

Tabelle 4-100 (Anhang): Studiendesign und -methodik für Studie D4200C00068

Item <sup>a</sup>	Charakteristikum	Studieninformation
-	<b>Studienziel</b>	
3	Zielsetzungen	<p>Primäres Ziel:</p> <p>Bewertung der objektiven Ansprechrate von Patienten, die mit 100 mg Vandetanib als Monotherapie behandelt wurden. Eine beobachtete Ansprechrate von 20% wurde als Studienziel definiert.</p> <p>Sekundäre Ziele:</p> <ol style="list-style-type: none"> <li>1. Untersuchung der Sicherheit und Verträglichkeit der Behandlung in dieser Patientenpopulation</li> <li>2. Untersuchung des progressionsfreien Überlebens (<i>Progression-Free Survival</i>, PFS)</li> <li>3. Untersuchung der Krankheitskontrollrate (<i>Disease Control Rate</i>, DCR), der Dauer bis zum objektiven Ansprechen (<i>Duration of Objective Response</i>, DOR) und der Dauer der Krankheitskontrolle (<i>Duration of Disease Control</i>)</li> <li>4. Bewertung der Veränderung des WHO Performance Status (WHO (PS)) gegenüber dem Ausgangswert unter Benutzung des World Health Organisation (WHO) PS</li> <li>5. Untersuchung, ob ein Anstieg der Stuhlgangshäufigkeit bei Patienten mit symptomatischem Durchfall in Verbindung mit hereditärem MTC vorliegt</li> <li>6. Bewertung des Effekts der Dosis 100 mg Vandetanib auf die Konzentration von Calcitonin (CTN) und des carcinoembryonalen Antigens (CEA)</li> <li>7. Beschreibung der Pharmakokinetik in dieser Population</li> <li>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</li> </ol> <p>Explorative Ziele:</p> <ol style="list-style-type: none"> <li>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</li> <li>2. Erforschung des Effekts von Vandetanib auf die Lebensqualität unter Benutzung der <i>Functional Assessment of Cancer Therapy General Scale</i> (FACIT-G)</li> <li>3. Einschätzung des Effekts von Vandetanib auf die Aktivität des RET Onkoproteins, des EGFRs und des VEGFR-2 im Tumorgewebe</li> <li>4. Bestätigung der Keimbahn-RET-Mutation im Tumorgewebe</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		5. Einschätzung des Effekts von Vandetanib auf zirkulierende Angiogenese-Biomarker, auf die Tumorlast und - auf urinbasierte Biomarker
-	<b>Methoden</b>	
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	<p><i>Kohortenstudie</i> (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl an; Methoden der Nachbeobachtung)</p> <p><i>Fallkontrollstudie</i> (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)</p> <p><i>Querschnittsstudie</i> (Einschlusskriterien, Herkunft der Teilnehmer, Methoden ihrer Auswahl)</p>	<p>Einschlusskriterien:</p> <ol style="list-style-type: none"> <li>1. Bereitstellung einer schriftlichen Einverständniserklärung</li> <li>2. Weiblich oder männlich im Alter von 18 Jahren und älter</li> <li>3. Histologisch bestätigte Diagnose: lokal fortgeschrittenes oder metastasierender hereditärer MTC, für den keine Standardtherapie mehr verfügbar ist</li> <li>4. Lebenserwartung <math>\geq</math> 12 Wochen</li> <li>5. WHO Performance Status 0-2</li> <li>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 (<math>&gt;</math> 5 mm Schichtdicke) gemäß des modifizierten RECIST-Kriteriums</li> <li>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</li> <li>8. Schlucken der Medikation als ganze Tablette ohne Kauen, Zerkleinern oder Teilen</li> </ol> <p>Ausschlusskriterien:</p> <ol style="list-style-type: none"> <li>1. Gehirnmastasen oder Rückenmarkskompression, sofern nicht mind. 4 Wochen vor der ersten Dosis bestrahlt und stabil ohne Behandlung mit Steroiden für mind. eine Woche</li> <li>2. Letzte Dosis der vorherigen Chemotherapie in den letzten 4 Wochen vor Beginn der Studientherapie</li> <li>3. Bestrahlung in den letzten 4 Wochen vor Beginn der Studientherapie</li> <li>4. Größere Operationen in den letzten 4 Wochen oder</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>unzureichend geheilter chirurgischer Eingriff vor Beginn der Studientherapie</p> <ol style="list-style-type: none"> <li>5. Ungeklärte Toxizitäten mit CTCAE-Grad &gt; 2 (Common Terminology Criteria for Adverse Events, National Institutes of Health, National Cancer Institute) von vorheriger Antitumorthherapie</li> <li>6. Serum-Bilirubin-Konzentration &gt; 1.5 mal dem oberen Grenzwert des Referenzbereichs (ULRR)</li> <li>7. Serum-Kreatinin-Konzentration &gt; 1.5 mal ULRR oder Kreatinin-Clearance ≤ 50 ml/min (Berechnet durch die Cockcroft-Gault-Formel)</li> <li>8. Alaninaminotransferase (ALT), Aspartataminotransferase (AST) oder alkalische Phosphatase (ALP) &gt; 2.5 mal ULRR ohne nachweisbare Lebermetastasen, sonst &gt; 5 mal ULRR falls der Prüfarzt einen Zusammenhang mit Lebermetastasen diagnostiziert</li> <li>9. Bedeutendes kardiales Ereignis (z.B. Herzinfarkt, Vena-cava-superior-Syndrom, New York Heart Association classification of heart disease &gt; 2) in den letzten 3 Monaten vor Studieneintritt oder Vorhandensein einer Herzerkrankung, die laut Meinung des Prüfarztes, das Risiko ventrikulärer Arrhythmien erhöht</li> <li>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</li> <li>11. Angeborenes long-QT-Syndrom oder Verwandtschaft ersten Grades mit einem ungeklärten plötzlichen Todesfall unter 40 Jahren</li> <li>12. QT-Verlängerung mit anderen Medikamenten, die einen Abbruch dieser Therapie erfordern</li> <li>13. Linksschenkelblock</li> <li>14. QTc nach Bazett-Korrektur nicht messbar oder ≥ 480 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.</li> <li>15. Kaliumkonzentration &lt; 4.0 mmol/l trotz Ergänzung, Kalzium (ionisiert oder Albumin-korrigiert) oder Magnesium außerhalb des Referenzbereichs trotz Ergänzung</li> <li>16. Schwangerschaft oder stillend Frauen</li> <li>17. Begleitende Medikation, welche die QTc-</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Verlängerung auslösen könnte (mit Ausnahme von Somatostatin oder Somatostatin-Analogen) und / oder verbotene Medikamente laut Studienprotokoll</p> <p>18. Nichtkontrollierte Hypertonie (systolischer Blutdruck &gt; 160 mm Hg oder diastolischer Blutdruck &gt; 100 mm Hg)</p> <p>19. Vorgegangene oder aktuelle bösartige Tumore an anderen Stellen in den letzten 5 Jahren mit Ausnahme eines Carcinoma in situ des Zervix oder angemessene behandelte Basalzellen- oder Plattenepithelkarzinom der Haut</p> <p>20. Anzeichen für schwere oder unkontrollierte systemische Erkrankungen; oder der aktuelle Gesundheitszustand ist für eine Studienteilnahme/Einhaltung des Protokolls nach Meinung des Prüfarztes unvorteilhaft.</p> <p>21. Vorgegangene Teilnahme an der vorliegenden Studie oder vorgegangene Teilnahme an einer Studie mit einer Behandlung durch Vandetanib</p> <p>22. Mitwirkung an der Planung und Durchführung der Studie (angewandt auf Personal von AstraZeneca oder Personal des Prüfzentrums)</p> <p>23. Erhalt jeglicher Prüfmedikation innerhalb 30 Tage vor Beginn dieser Studie</p>
<b>6b</b>	<p><i>Kohortenstudie</i> (Matching (Paarbildung) und Matchingkriterien, Anzahl der exponierten und der nicht exponierten Teilnehmer)</p> <p><i>Fallkontrollstudie</i> (Matching (Paarbildung) und Matchingkriterien, Anzahl der Kontrollen pro Fall)</p>	Nicht zutreffend
<b>7</b>	Variablen	<p>Variable für primäres Ziel:</p> <p>Objektive Ansprechrate (ORR) nach dem RECIST-Kriterium. ORR ist definiert als der Anteil (%) an Patienten, die eine komplette oder partielle Remission hatten.</p> <p>Variablen für sekundäre Ziele:</p> <ul style="list-style-type: none"> <li>- Unerwünschter Ereignissen (UE)</li> <li>- Klinisch signifikante abnormale Laborwerte, EKG-Veränderungen und Vitaldaten</li> <li>- Progressionsfreies Überleben (PFS): gemessen von der ersten Medikamentengabe bis zur Progression</li> <li>- <i>Disease Control Rate</i> (DCR; Rate der Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 12 Wochen)</li> <li>- Dauer des objektiven Ansprechen (DOR)</li> <li>- <i>Duration of Disease Control</i> (Dauer von erster Dosis bis zur Progression für Patienten mit Remission oder stabilem Krankheitsverlauf über mindestens 12 Wochen)</li> <li>- <i>Performance Status</i> (PS) basierend auf dem WHO-Kriterium</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>- Häufigkeit des Stuhlgangs bei Patienten mit symptomatischem Durchfall</li> <li>- Änderungen des CTN und CEA Levels</li> <li>- PK und PD von Vandetanib</li> <li>- QTc Intervall</li> </ul> Variablen für explorative Ziele: <ul style="list-style-type: none"> <li>- Lebensqualität (FACT-G)</li> <li>- RET Onkoprotein, RET Genmutation</li> <li>- EGFR und die VEGFR-2 Signalbahnen</li> <li>- Serum Protein</li> <li>- Plasma VEGF, VEGFR2 und bFGF</li> <li>- Urinäres Katecholamin, Metanephrin und Normetanephrin</li> </ul>
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 $\geq$ 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 <i>Confounding</i>	Grundlage: <i>Intention-to-Treat</i> -Prinzip Berechnung der Zielkriterien Berechnung von exakten 95%-Konfidenzintervallen zu den Endpunkten „ORR“, „CTN-Response“, „DCR“, Darstellung der Kaplan-Meier-Kurve zum Endpunkt „PFS“ „ <i>Duration of Disease Control</i> “ 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 <sup>a</sup>	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	<i>Kohortenstudie</i> (Umgang mit vorzeitigem Ausscheiden („loss to Follow-up“)) <i>Fallkontrollstudie</i> (Berücksichtigung des <i>Matchings</i> (Paarbildung) von Fällen und Kontrollen bei der Auswertung) <i>Querschnittsstudie</i> (Auswertungsmethoden, die die gewählte Strategie zur Stichprobenauswahl ( <i>Sampling strategy</i> ) berücksichtigen)	Patienten mit vorzeitigem Ausscheiden wurden als non-Responder betrachtet.
12e	Sensitivitätsanalysen	Nicht zutreffend
-	<b>Ergebnisse</b>	
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
<b>a: nach STROBE 2007</b>		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

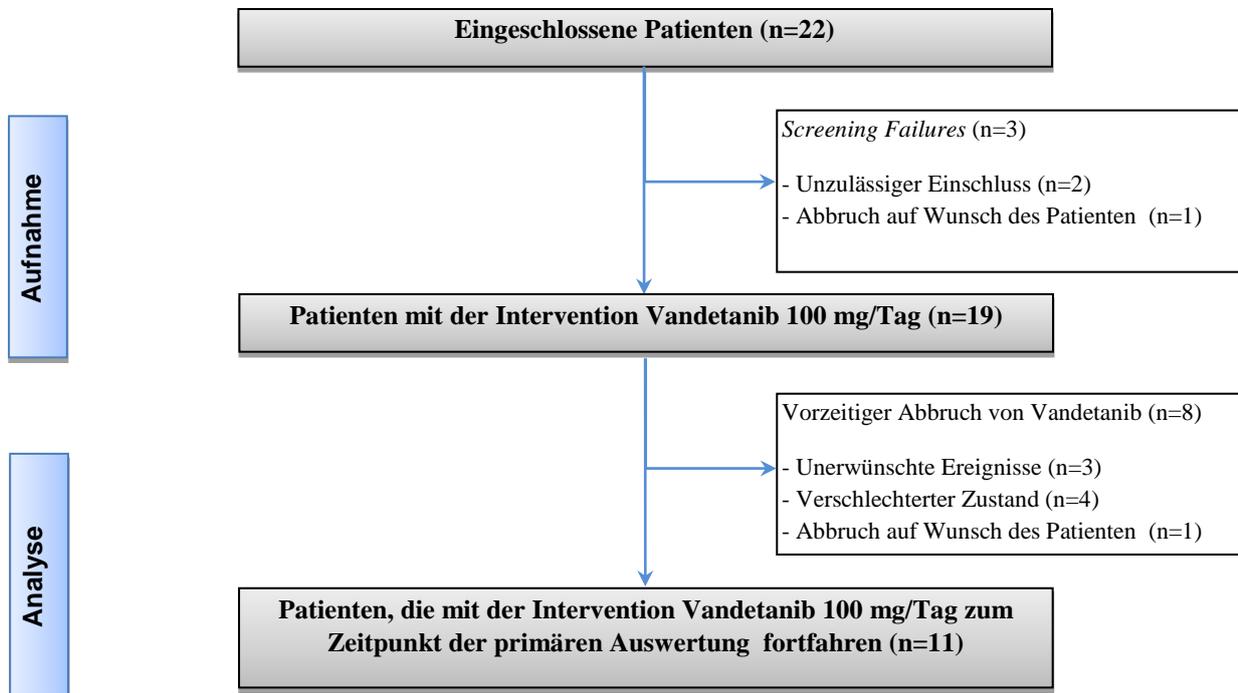


Abbildung 20: Flussdiagramm Studiendesign Studie D4200C00068

**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 4-101 (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:

---

---

**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 kodierte 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:

---

---

### 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:

---



---

## 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:** 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:

---

---

### 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 stark eingeschränkt

---

**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 stark eingeschränkt

---

**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:

---

---

### 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: \_\_ 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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen 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:** \_\_ 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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen 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:** \_\_ 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 Gruppeneinteilung 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 (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 Gruppeneinteilung 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:

**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:

---

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:

---

---

### 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 4-102 (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

Genauere 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 kodierte 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 Performance 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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen 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 4-103 (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 kodierte 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 Ansprechrage (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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen 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 Performance 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 Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen 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:

---

---

