Dokumentvorlage, Version vom 20.01.2011

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Enzalutamid (Xtandi®)

Astellas Pharma GmbH

Modul 4 A

Behandlung des metastasierten kastrationsresistenten Prostatakarzinoms bei erwachsenen Männern, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 29.08.2013

Inhaltsverzeichnis

Seite

Tabellenverzeichnis ... 4
Abbildungsverzeichnis ... 12
Abkürzungsverzeichnis ... 13
4 Modul 4 – allgemeine Informationen ... 15
4.1 Zusammenfassung der Inhalte von Modul 4 ... 16
4.2 Methodik ... 32
4.2.1 Fragestellung ... 32
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung ... 33
4.2.3 Informationsbeschaffung ... 35
4.2.3.1 Studien des pharmazeutischen Unternehmers ... 35
4.2.3.2 Bibliografische Literaturrecherche ... 35
4.2.3.3 Suche in Studienregistern ... 36
4.2.3.4 Selektion relevanter Studien ... 37
4.2.4 Bewertung der Aussagekraft der Nachweise ... 37
4.2.5 Informationssynthese und -analyse ... 40
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien ... 40
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien ... 41
4.2.5.3 Meta-Analysen ... 45
4.2.5.4 Sensitivitätsanalysen ... 46
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren ... 47
4.2.5.6 Indirekte Vergleiche ... 51
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen ... 53
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel ... 54
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel ... 54
4.3.1.1.1 Studien des pharmazeutischen Unternehmers ... 54
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche ... 55
4.3.1.1.3 Studien aus der Suche in Studienregistern ... 58
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel ... 59
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 60
4.3.1.2.1 Studiendesign und Studienpopulationen ... 60
4.3.1.2.2 Verzerrungspotenzial auf Studienebene ... 74
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien ... 75
4.3.1.3.1 Endpunkte – RCT ... 75
4.3.1.3.1.1 Endpunkt „Gesamtüberleben“ - RCT ... 75
4.3.1.3.1.2 Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ – RCT ... 79
4.3.1.3.1.3 Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ - RCT ... 82

4.3.1.3.1.4 Endpunkt „Schmerz“ - RCT 87
4.3.1.3.1.5 Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ - RCT 95
4.3.1.3.1.6 Endpunkt „Gesundheitsbezogene Lebensqualität“ - RCT 99
4.3.1.3.1.7 Endpunkt „Sicherheit“ - RCT 111
4.3.1.3.2 Subgruppenanalysen – RCT 124
4.3.1.3.2.1 Subgruppenanalysen zur Studienpopulation-RCT 124
4.3.1.3.2.2 *A priori* definierte Subgruppenanalysen zur Studienpopulation 157
4.3.1.3.2.3 Zusammenfassung der Ergebnisse der Subgruppenanalysen aus RCT 163
4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien 173
4.3.2 Weitere Unterlagen 183
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 183
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 183
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 183
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 184
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 184
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 186
4.3.2.2 Nicht randomisierte vergleichende Studien 187
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 187
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 187
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 188
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien 188
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 189
4.3.2.3 Weitere Untersuchungen 189
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 189
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 190
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen 190
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 190
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 191
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen 191
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens 191
4.4.1 Beurteilung der Aussagekraft der Nachweise 191
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß 200
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 210
4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens 210
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 211
4.5.1 Begründung für die Vorlage indirekter Vergleiche 211
4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen 211

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 211
4.5.4 Verwendung von Surrogatendpunkten 212
4.6 Liste der eingeschlossenen Studien 213
4.7 Referenzliste 215
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 219
Anhang 4-B : Suchstrategien – Suche in Studienregistern 223
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund 224
Anhang 4-D : Liste der abgebrochenen Studien 226
Anhang 4-E : Liste der laufenden Studien 227
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT 228
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 248
Endpunkt: Gesamtüberleben 257
Endpunkt: Radiographisches progressionsfreies Überleben 260
Endpunkt: Zeit bis zur ersten skelettbezogenen Komplikation 263
Endpunkt: Schmerzlinderung 266
Endpunkt: Schmerzprogressionsrate 269
Endpunkt: Zeit bis zur Schmerzprogression 272
Endpunkt: Gesundheitsbezogene Lebensqualität nach FACT-P 275
Endpunkt: Gesundheitsbezogene Lebensqualität nach EQ-5D 278
Endpunkt: Unerwünschte Ereignisse 281
Endpunkt: Unerwünschte Ereignisse nach CTCAE Grad ≥ 3 284
Endpunkt: Schwerwiegende unerwünschte Ereignisse 287
Endpunkt: Studienabbruch aufgrund unerwünschter Ereignisse 290
Endpunkt: Tod aufgrund unerwünschter Ereignisse 293

Tabellenverzeichnis

Seite

Tabelle 4-1: Ein- und Ausschlusskriterien 17
Tabelle 4-2: Enzalutamid vs. Placebo: Zusammenfassung des Zusatznutzens auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM 22
Tabelle 4-3: Zusammenfassung der Effektschätzer für Subgruppen, bei denen ein Beleg für oder ein Hinweis auf eine Interaktion vorlagen, aus der randomisiert kontrollierten Studie AFFIRM 26
Tabelle 4-4: Ein- und Ausschlusskriterien. 33
Tabelle 4-5: Gegenüberstellung der Unterschiede in Domäne 5 der Versionen 3 und 4 des FACT-P. 44
Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewerteten Arzneimittel 54
Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 55
Tabelle 4-8: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel 58
Tabelle 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel 59
Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 61
Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel 62
Tabelle 4-12: Charakterisierung der Interventionen: Auflistung der Begleitmedikamente, welche mehr als 10 % der Patienten des jeweiligen Arms während der Behandlung erhielten (sowie ausgewählte Andere) 63
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel 65
Tabelle 4-14: Charakterisierung der Subgruppen „Gabe von Biphosphonaten während der Studie (nein vs. ja)“ – RCT mit dem zu bewertenden Arzneimittel 66
Tabelle 4-15: Charakterisierung der Subgruppen „Gabe von Steroiden während der Studie (nein vs. ja)“ – RCT mit dem zu bewertenden Arzneimittel 67
Tabelle 4-16: Charakterisierung der Subgruppen „ECOG-*Performance*-Status (0 oder 1 vs. 2) zu Studienbeginn“ – RCT mit dem zu bewertenden Arzneimittel 68
Tabelle 4-17: Charakterisierung der Subgruppen „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ – RCT mit dem zu bewertenden Arzneimittel 69
Tabelle 4-18: Charakterisierung der Subgruppen „Alter (< 65 vs. ≥ 65)“ – RCT mit dem zu bewertenden Arzneimittel 70
Tabelle 4-19: Charakterisierung der Subgruppen “Anzahl vorheriger Chemotherapien 71

Tabelle 4-20: Charakterisierung der Subgruppen “Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein)“ – RCT mit dem zu bewertenden Arzneimittel 72
Tabelle 4-21: Charakterisierung der Auswertungspopulationen – RCT mit dem zu bewertenden Arzneimittel 73
Tabelle 4-22: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel 74
Tabelle 4-23: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel 75
Tabelle 4-24: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel (Fortsetzung von Tabelle 4-23) 75
Tabelle 4-25: Operationalisierung von Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 76
Tabelle 4-26: Bewertung des Verzerrungspotenzials für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 77
Tabelle 4-27: Ergebnisse für Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 77
Tabelle 4-28: Operationalisierung von Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ in RCT mit dem zu bewertenden Arzneimittel 79
Tabelle 4-29: Bewertung des Verzerrungspotenzials für Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ in RCT mit dem zu bewertenden Arzneimittel 79
Tabelle 4-30: Ergebnisse für Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ in RCT mit dem zu bewertenden Arzneimittel 80
Tabelle 4-31: Operationalisierung von Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel 82
Tabelle 4-32: Bewertung des Verzerrungspotenzials für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel 82
Tabelle 4-33: Ergebnisse für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ aus RCT mit dem zu bewertenden Arzneimittel 84
Tabelle 4-34: Ergebnisse der Sensitivitätsanalyse für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ aus RCT mit dem zu bewertenden Arzneimittel 86
Tabelle 4-35: Operationalisierung von Endpunkt „Schmerzlinderung“ in RCT mit dem zu bewertenden Arzneimittel 87
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Endpunkt „Schmerzlinderung“ in RCT mit dem zu bewertenden Arzneimittel 87
Tabelle 4-37: Ergebnisse für den Endpunkt „Schmerzlinderung“ in RCT mit dem zu bewertenden Arzneimittel 88
Tabelle 4-38: Operationalisierung von Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel 89
Tabelle 4-39: Bewertung des Verzerrungspotenzials für Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel 89

Tabelle 4-40: Ergebnisse für den Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel (*Responder*-Analyse) 90
Tabelle 4-41: Ergebnisse für Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel (Veränderung gegenüber Studienbeginn) 91
Tabelle 4-42: Operationalisierung von Endpunkt „Zeit bis zur Schmerzprogression“ in RCT mit dem zu bewertenden Arzneimittel 93
Tabelle 4-43: Bewertung des Verzerrungspotenzials für Endpunkt „Zeit bis zur Schmerzprogression“ in RCT mit dem zu bewertenden Arzneimittel 93
Tabelle 4-44: Ergebnisse für Endpunkt „Zeit bis zur Schmerzprogression“ in RCT mit dem zu bewertenden Arzneimittel 94
Tabelle 4-45: Operationalisierung von Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ in RCT mit dem zu bewertenden Arzneimittel 95
Tabelle 4-46: Bewertung des Verzerrungspotenzials für Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ in RCT mit dem zu bewertenden Arzneimittel 96
Tabelle 4-47: Aufstellung von mit Lähmungserscheinungen assoziierten unerwünschten Ereignissen, unterschieden nach Schweregrad des Ereignisses (nach NCI-CTCAE, Version 4.0) 97
Tabelle 4-48: Operationalisierung von Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel 99
Tabelle 4-49: Bewertung des Verzerrungspotenzials für Endpunkt „gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel 99
Tabelle 4-50: Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel (*Responder*-Analyse) 100
Tabelle 4-51: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel (Veränderung gegenüber Studienbeginn) 101
Tabelle 4-52: Operationalisierung von Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel 104
Tabelle 4-53: Bewertung des Verzerrungspotenzials für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel 104
Tabelle 4-54: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Gesamtpunktzahl basierend auf der VAS) 106
Tabelle 4-55: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Shift-Tabelle für die Veränderung von Studienbeginn zu Woche 49) 107
Tabelle 4-56: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Shift-Tabelle für die Veränderung von Studienbeginn zu letzte verfügbare Erhebung) 109
Tabelle 4-57: Operationalisierung von Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 111

Tabelle 4-58: Bewertung des Verzerrungspotenzials für Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 111
Tabelle 4-59: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 112
Tabelle 4-60: Operationalisierung von Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 113
Tabelle 4-61: Bewertung des Verzerrungspotenzials für Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 113
Tabelle 4-62: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“, Enzalutamid vs. Placebo in RCT mit dem zu bewertenden Arzneimittel 114
Tabelle 4-63: Operationalisierung von Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 115
Tabelle 4-64: Bewertung des Verzerrungspotenzials für Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 115
Tabelle 4-65: Ergebnisse für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“, Enzalutamid vs Placebo in RCT mit dem zu bewertenden Arzneimittel 116
Tabelle 4-66: Operationalisierung von Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 117
Tabelle 4-67: Bewertung des Verzerrungspotenzials für Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 117
Tabelle 4-68: Ergebnisse für den Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 118
Tabelle 4-69: Operationalisierung von Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 119
Tabelle 4-70: Bewertung des Verzerrungspotenzials für Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 119
Tabelle 4-71: Ergebnisse für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 120
Tabelle 4-72: Operationalisierung von Endpunkt „Krampfanfälle“ in RCT mit dem zu bewertenden Arzneimittel 121
Tabelle 4-73: Bewertung des Verzerrungspotenzials für Endpunkt „Krampfanfälle“ in RCT mit dem zu bewertenden Arzneimittel 121
Tabelle 4-74: Übersicht über die im Enzalutamid-Arm auftretenden Krampfanfälle. [Astellas Pharma GmbH 2013b] 123
Tabelle 4-75: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests (p-Werte) der randomisiert kontrollierten Studie AFFIRM 126
Tabelle 4-76: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 129

Tabelle 4-77: Subgruppenanalyse zur Studienpopulation „Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des *Screenings* (ja vs. nein)“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 130

Tabelle 4-78: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel 131

Tabelle 4-79: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel 132

Tabelle 4-80: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel 133

Tabelle 4-81: Subgruppenanalyse zur Studienpopulation „Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel 133

Tabelle 4-82: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel 134

Tabelle 4-83: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel 135

Tabelle 4-84: Subgruppenanalysen zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel 136

Tabelle 4-85: Subgruppenanalysen zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel 137

Tabelle 4-86: Subgruppenanalysen zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel 138

Tabelle 4-87: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel .. 139

Tabelle 4-88: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel 140

Tabelle 4-89: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ (*Responder*-Analyse) in RCT mit dem zu bewertenden Arzneimittel..... 141

Tabelle 4-90: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ (Veränderung gegenüber Studienbeginn) in RCT mit dem zu bewertenden Arzneimittel 142

Tabelle 4-91: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt Gesundheitsbezogene Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel 144

Tabelle 4-92: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel 145

Tabelle 4-93: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel 146

Tabelle 4-94: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel .. 147

Tabelle 4-95: Subgruppenanalyse zur Studienpopulation „Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel 148

Tabelle 4-96: Subgruppenanalyse zur Studienpopulation „Gabe von Bisphosphonaten während der Studie (ja vs. nein)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 149

Tabelle 4-97: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 150

Tabelle 4-98: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 151

Tabelle 4-99: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel 152

Tabelle 4-100: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 153

Tabelle 4-101: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 154

Tabelle 4-102: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs.

≥ 4)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 155

Tabelle 4-103: Subgruppenanalyse zur Studienpopulation „Gabe von Bisphosphonaten während der Studie (ja vs. nein)“ für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel 156

Tabelle 4-104: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests für prädefinierte Subgruppen-Analysen zum Endpunkt „Gesamtüberleben“ der randomisiert kontrollierten Studie AFFIRM 159

Tabelle 4-105: Subgruppenanalyse zur Studienpopulation „Hämoglobin zu Studienbeginn“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel 160

Tabelle 4-106: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests für prädefinierte Subgruppen-Analysen zu sicherheitsbezogenen Ereignissen, die unter der Behandlung auftraten, aus der randomisiert kontrollierten Studie AFFIRM 161

Tabelle 4-107 Subgruppenanalyse zur Studienpopulation nach „Alter (< 75 vs. ≥ 75)“ für sicherheitsbezogene Ereignisse in RCT mit dem zu bewertenden Arzneimittel 162

Tabelle 4-108: Subgruppenanalyse zur Studienpopulation nach „Anamnese einer Hypertonie (ja vs. nein)“ für sicherheitsbezogene Ereignisse in RCT mit dem zu bewertenden Arzneimittel 162

Tabelle 4-109: Zusammenfassung der Effektschätzer für Subgruppen, bei denen ein Beleg oder ein Hinweis auf eine Interaktion vorlagen, aus der randomisiert kontrollierten Studie AFFIRM 168

Tabelle 4-110: Enzalutamid vs. Placebo: Zusammenfassung der Effekte auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM 180

Tabelle 4-111: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche 184

Tabelle 4-112: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 184

Tabelle 4-113: Operationalisierung von <Endpunkt xxx> 185

Tabelle 4-114: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche 185

Tabelle 4-115: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche 186

Tabelle 4-116: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien 188

Tabelle 4-117: Operationalisierung von <Endpunkt xxx> 188

Tabelle 4-118: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien 189

Tabelle 4-119: Operationalisierung von <Endpunkt xxx> 190

Tabelle 4-120: Enzalutamid vs. Placebo: Zusammenfassung des Zusatznutzens auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM 206

Tabelle 4-121: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 210
Tabelle 4-122: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens 211
Tabelle 4-123 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel 226
Tabelle 4-124 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel 227
Tabelle 4-125 (Anhang): Studiendesign und -methodik für die Studie AFFIRM (CRPC2) . 229
Tabelle 4-126: Übersicht über den zeitlichen Ablauf der Erhebungen und Untersuchungen innerhalb des Studienprogramms 246
Tabelle 4-127 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie AFFIRM 249

Abbildungsverzeichnis

Seite

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche, Stand: 8.7.2013 – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel 56

Abbildung 2: Ergebnisse der im CSR dargestellten Subgruppenanalysen für den Endpunkt „Gesamtüberleben“. 158

Abbildung 3: “*Points to consider on application with: 1. meta-analyses; 2. one pivotal study*” [European Medicines Agency (EMA) 2001] 194

Abbildung 4: Flussdiagramm Patientenfluss in der Studie AFFIRM 245

Abkürzungsverzeichnis

Abkürzung	**Bedeutung**
ALT	Alanin Aminotransferase
ASCO	*American Society of Clinical Oncology*
AST	Aspartat Aminotransferase
BPI	*Brief Pain Inventory*
CMC	*Clinically Meaningful Changes*
CONSORT	*Consolidated Standards of Reporting Trials*
CSR	*Clinical Study Report*
CT	Computertomographie
CTCAE	*Common Toxicity Criteria for Adverse Events*
DFS	*Disease-free survival*
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
ECOG	*Eastern Cooperative Oncology Group*
EG	Europäische Gemeinschaft
EMA	*European Medicines Agency*
EQ-5D	*European Quality of Life Five-Domain Scale*
FACT-G	*Functional Assessment of Cancer Therapy - General*
FACT-P	*Functional Assessment of Cancer Therapy - Prostate*
G-BA	Gemeinsamer Bundesausschuss
HR	*Hazard Ratio*
ICTRP	*International Clinical Trials Registry Platform*
(I) DMC	*(Independent) Data Monitoring Committee*
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	*Intention to treat*
IVRS	*Interactive Voice Response System*
IWRS	*Interactive Web Response System*
KI	Konfidenzintervall
LDH	Laktat-Dehydrogenase
MedDRA	*Medical Dictionary for Regulatory Activities*
MRI	Magnetresonanztomographie
n.e.	nicht erreicht

NA	Nicht Anwendbar
NCI	*National Cancer Institute*
NYHA	*New York Heart Association*
PCWG2	*Prostate Cancer Clinical Trials Working Group 2*
PSA	Prostata-spezifisches Antigen
QTcF	*QT correction formula by Fridericia*
RCT	*Randomized Controlled Trial*
RECIST v1.1	*Response Evaluation Criteria in Solid Tumors version 1.1*
rPFS	Radiographisches Progressionsfreies Überleben
RR	Relatives Risiko
SD	*Standard Deviation*
SGB	Sozialgesetzbuch
SMD	*Standardized Mean Difference*
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VAS	*Visual Analogue Scale*
VerfO	Verfahrensordnung des Gemeinsamen Bundesausschusses
WHO	*World Health Organization*
ZNS	Zentrales Nervensystem

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

Enzalutamid ist ein Arzneimittel zur medikamentösen Behandlung des metastasierten kastrationsresistenten Prostatakarzinoms bei erwachsenen Männern, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet.

Ziel des vorliegenden Dossiers ist es, den medizinischen Nutzen und Zusatznutzen von Enzalutamid (Xtandi®) + *Best Supportive Care* anhand der Ergebnisse aus randomisierten, kontrollierten klinischen Studien in Bezug auf verschiedene patientenrelevante Endpunkte zu bewerten. Die Bewertung erfolgt für die Zulassungspopulation unter Berücksichtigung der Ergebnisse verschiedener Subgruppenanalysen. Als Vergleichstherapie wird dabei die vom G-BA festgelegte zweckmäßige Vergleichstherapie *Best Supportive Care* herangezogen, als welche die Therapie verstanden wird, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität gewährleistet.

Datenquellen

Zur Identifizierung relevanter Studien für die Bewertung des medizinischen Nutzens und Zusatznutzens von Enzalutamid wurde eine umfassende systematische Suche in Studienregistern und in bibliografischen Datenbanken durchgeführt. Des Weiteren wurden alle von Astellas durchgeführten Studien sowie Studien, an welchen Astellas finanziell beteiligt war, hinsichtlich ihrer Relevanz für die Nutzenbewertung überprüft. Im Rahmen dieser Suche wurde eine randomisierte, placebokontrollierte, doppelt-verblindete Phase-III-Studie zu Wirksamkeit und Sicherheit von Enzalutamid bei Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet, identifiziert. Als Datenquellen für das Dossier wurde der vollständige Studienbericht herangezogen. Außerdem wurden ergänzende Analysen durchgeführt, um den Anforderungen des G-BA gemäß Beratungsgespräch und Verfahrensordnung Rechnung zu tragen.

Ein-/Ausschlusskriterien für Studien

Die Ein-/Ausschlusskriterien für Studien sind in Tabelle 4-1 dargestellt.

Tabelle 4-1: Ein- und Ausschlusskriterien

E1	Patientenpopulation (Indikation)	Erwachsene Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet.
E2	Intervention	Enzalutamid plus *Best Supportive Care*
E3	Vergleichstherapie	*Best Supportive Care*
E4	Endpunkte (patientenrelevant)	**Mortalität** • Gesamtüberleben **Morbidität** • Radiographisches progressionsfreies Überleben • Zeit bis zur ersten skelettbezogenen Komplikation • Schmerz • Lähmungen **Lebensqualität** **Sicherheit**: • Unerwünschte Ereignisse (gesamt) • Unerwünschte Ereignisse mit CTCAE Grad ≥ 3 • Schwerwiegende unerwünschte Ereignisse • Studienabbruch aufgrund unerwünschter Ereignisse • Unerwünschte Ereignisse, die zum Tod führten
E5	Studientyp (Design)	Randomisierte kontrollierte Studie (RCT)
E6	Vollpublikation verfügbar	Ja (Studienbericht)
E7	Publikationssprache	Deutsch oder Englisch
E8	Studiendauer	Keine Einschränkung
A1	Studienart	Tierexperimentelle Studien
A2	Doppelpublikation	Doppelpublikationen ohne relevante Zusatzinformation

Alle den Einschlusskriterien genügenden und nicht unter die Ausschlusskriterien fallenden Studien wurden in die Bewertung einbezogen.

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

Für die in die Nutzenbewertung eingeschlossene Phase-III-Studie wurde das Verzerrungspotenzial der Ergebnisse sowohl auf Studienebene als auch auf Ebene der einzelnen Endpunkte anhand der Vorgaben des IQWiG und der Verfahrensordnung beurteilt. Dazu wurden endpunktübergreifende und endpunktspezifische Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet.

Die Beurteilung des Verzerrungspotenzials wurde durch einen zweiten Reviewer überprüft. Diskrepanzen in der Beurteilung durch die Reviewer wurden durch Diskussion aufgelöst.

Für dichotome Endpunkte wurde der Therapieeffekt in der Regel anhand des Relativen Risikos (RR) zusammengefasst. *Time-to-event* Variablen wurden anhand der medianen Zeit bis zum Eintreten des Ereignisses und der Hazard Ratio (HR) dargestellt. Für kontinuierliche Variablen wurden standardisierte Mittelwertdifferenzen (*Standardized Mean Difference* (SMD)) berechnet.

Neben der Beurteilung der patientenrelevanten Endpunkte auf der Ebene der Zulassungspopulation erfolgte eine Bewertung aller Endpunkte für die *post hoc* definierten Subgruppen „Gabe von Bisphosphonaten während der Studie", „Gabe von Steroiden während der Studie", „ECOG-*Performance*-Status bei Studienbeginn", „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn", „Alter", „Anzahl vorheriger Chemotherapien" und „viszerale Erkrankungen (Lunge und/ oder Leber) zum Zeitpunkt des Screenings". Möglicherweise variierende Effekte zwischen verschiedenen Subgruppen wurden mittels Interaktionstests bewertet. Im Fall eines Hinweises auf ($p \leq 0,2$) oder eines Beleges ($p \leq 0,05$) für eine Interaktion erfolgte eine Darstellung der Ergebnisse der Subgruppenanalysen.

Die Bewertung des Ausmaßes des Zusatznutzens erfolgte in Anlehnung an den Entwurf des IQWiG zur Aktualisierung der Allgemeinen Methoden vom 18.4.2013. Wie dort beschrieben wurden für die verschiedenen Zielgrößenkategorien in Abhängigkeit vom Schweregrad des Ereignisses/der Nebenwirkung verschiedene Schwellenwerte zur Bewertung des Ausmaßes des Zusatznutzens herangezogen.

Da zur Bewertung des medizinischen Zusatznutzens des zu bewertenden Arzneimittels Enzalutamid plus *Best Supportive Care* gegenüber der zweckmäßigen Vergleichstherapie *Best Supportive Care* eine direkt vergleichende placebokontrollierte Studie vorliegt, war die Durchführung eines indirekten Vergleiches nicht erforderlich. Da nur eine Studie für die Nutzenbewertung zur Verfügung stand, wurden keine Meta-Analysen durchgeführt.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die AFFIRM-Studie ist eine randomisierte, placebokontrollierte, multinationale, multizentrische, Phase-III-Doppelblindstudie. Ziel der Studie war die Untersuchung der Wirksamkeit und Sicherheit von Enzalutamid im Vergleich zu Placebo bei Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet. Patienten in beiden Studienarmen erhielten neben der Therapie mittels Enzalutamid bzw. Placebo eine individualisierte, unterstützende Behandlung nach *Best Supportive Care*, z.B. eine adäquate Schmerztherapie. Das Zuteilungsverhältnis auf die Enzalutamid- bzw. Placebogruppe betrug 2:1. Für die Abschlussanalyse wurde eine Anzahl von 650 Todesereignissen festgelegt. Mit Erreichen eines signifikanten Ergebnisses zur Interimanalyse, welche nach 520 Todesereignissen durchgeführt wurde, wurde die Studie jedoch wie geplant beendet, die Interimanalyse stellt somit die Abschlussanalyse dar. Die Studie wurde in 15 Ländern duchgeführt. Primäres Zielkriterium war das „Gesamtüberleben". Daneben wurden u.a. das „radiographische

progressionsfreie Überleben“, die „Zeit bis zum Auftreten der ersten skelettbezogenen Komplikation“, „Schmerz“, „Lebensqualität“ und „Sicherheit“ erfasst.

Das Verzerrungspotential der Studie wurde mit „niedrig“ bewertet, sie wies eine hohe Ergebnissicherheit auf. Aufgrund der Größe der Studie, ihres randomisierten, kontrollierten, doppelblinden Designs, der vorliegenden internen wie externen Validität sowie der klinischen Relevanz, statistischen Signifikanz, Datenqualität und Konsistenz der Ergebnisse entsprach die Studie den Anforderungen der EMA an eine pivotale Studie, was die EMA durch die Erteilung der Zulassung bestätigte. Auf die Anwendung dieser Kriterien wird auch vom IQWiG im Entwurf zur Aktualisierung der Methoden 4.0 vom 18.4.2013 verwiesen. Im vorliegenden Dossier wurde die AFFIRM-Studie deshalb zur Ableitung eines Belegs für einen Zusatznutzen herangezogen.

Im Vergleich zu einer Placebo-Behandlung erwies sich die Behandlung mit Enzalutamid auf Studienpopulationsebene bezüglich der Endpunkte „Gesamtüberleben“, „radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“, „Schmerzprogressionsrate“, „Zeit bis zur Schmerzprogression“, „Gesundheitsbezogene Lebensqualität nach FACT-P“ und hinsichtlich des Auftretens „Unerwünschter Ereignisse nach CTCAE Grad $\geq$ 3“ als die statistisch siginifikant überlegene Therapie:

Für das „Gesamtüberleben“ konnte für Patienten unter Enzalutamid-Behandlung ein Überlebensvorteil von 4,8 Monaten gegenüber den Patienten in der Placebogruppe erzielt werden. Das mediane Gesamtüberleben verlängerte sich dabei von 13,6 Monaten (95 %-KI [11,3; 15,8]) auf 18,4 Monate (95 %-KI [17,3; n.e.]), wobei das Hazard Ratio 0,63 (95 %-KI [0,53; 0,75], $p < 0,0001$) zugunsten von Enzalutamid betrug.

Für das „radiographische progressionsfreie Überleben“ erwies sich Enzalutamid mit einem um 5,4 Monate länger andauernden radiographischen progressionsfreien Überleben (Enzalutamid: 8,3 Monate (95 %-KI [8,2; 9,4]; Placebo: 2,9 Monate (95 %-KI [2,8; 3,4]) und einem Hazard Ratio von 0,40 (95 %-KI [0,35; 0,47], $p < 0,0001$) als die signifikant überlegene Therapie.

Die „Zeit bis zur ersten skelettbezogenen Komplikation“ betrug für Patienten im Enzalutamid-Arm 16,7 Monate (95 %-KI [14,6; 19,1]). Damit trat das erste skelettale Ereignis 3,4 Monate später als bei Patienten im Placebo-Arm auf (13,3 Monate (95 %-KI [9,9; n.e.]). Das Hazard Ratio lag bei einem Wert von 0,69 (95 %-KI [0,57; 0,84], $p = 0,0001$).

Bezüglich der „Schmerzprogressionsrate“ zeigte sich für Enzalutamid im Vergleich zu Placebo im Rahmen der *Responder*-Analyse mit einem Wert von 0,71 (95 %-KI [0,59; 0,87], $p = 0,0008$) ein signifikant geringeres relatives Risiko, eine Schmerzprogression zu erleiden. Die Betrachtung der Veränderung der Schmerzpunktzahl zwischen Studienbeginn und Woche 13 bestätigte den Vorteil von Enzalutamid. Hier betrug die SMD für den Unterschied -0,48 (95 %-KI [-0,62; -0,34], $p < 0,0001$).

Auch hinsichtlich der „Zeit bis zur Schmerzprogression“ profitierten die Patienten von einer Enzalutamid-Therapie. Während der Median für diesen Endpunkt in der Enzalutamid-Gruppe zum Zeitpunkt der Datenerhebung noch nicht erreicht war, betrug er in der Placebogruppe 13,8 Monate (95 %-KI [13,8; n.e.]). Das Hazard Ratio lag hier bei einem Wert von 0,56 (95 %-KI [0,41; 0,78], p = 0,0004).

Für die Veränderung der „Gesundheitsbezogenen Lebensqualität nach FACT-P“ ergab sich im Rahmen der *Responder*-Analyse ein relatives Risiko bezüglich der Verbesserung der Lebensqualität von 2,81 (95 %-KI [2,05; 3,84], p < 0,0001) (Berechnung für das „Gegenereignis“: RR = 0,68 [0,63; 0,74]) zugunsten von Enzalutamid. Bei der Betrachtung der Veränderung der FACT-P-Gesamtpunktzahl der letzten verfügbaren Erhebung gegenüber Studienbeginn zeigte sich ebenfalls ein Vorteil für Enzalutamid, wobei die SMD für den Unterschied zwischen den Therapiegruppen 0,42 (95 %-KI [0,27; 0,56], p < 0,0001) betrug.

Die Auswertung des Sicherheitsendpunkts „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ ergab für Patienten unter Enzalutamid-Therapie im Vergleich zu Placebo mit einem Wert von 0,85 (95 %-KI [0,76; 0,96], p = 0,009) ein signifikant geringeres relatives Risiko ein unerwünschtes Ereignis nach CTCAE Grad ≥ 3 zu erleiden.

Für alle anderen Endpunkte („Schmerzlinderung“ „Gesundheitsbezogene Lebensqualität nach EQ-5D“, „Unerwünschte Ereignisse“, „Schwerwiegende unerwünschte Ereignisse“ „Studienabbruch aufgrund unerwünschter Ereignisse“ und „ Tod aufgrund unerwünschter Ereignisse“) gab es keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen.

Alle berichteten patientenrelevanten Endpunkte wurden an Hand *post hoc* definierter Subgruppenanalysen näher beleuchtet. Mögliche Effektmodifikatoren wurden dazu mittels Interaktionstests bewertet, und beim Auftreten von Hinweisen auf oder Belegen für eine Interaktion wurden die Subgruppenergebnisse im Detail berichtet.

Für den primären Endpunkt „Gesamtüberleben“ zeigte die Behandlung mit Enzalutamid für alle Patienten der Studienpopulation, mit Ausnahme derjenigen Patienten, welche zum Zeitpunkt des *Screenings* unter einer viszeralen Erkrankung der Lunge und/oder Leber litten, eine statistisch signifikante Verlängerung der Überlebenszeit gegenüber der Placebo-Behandlung. Während der signifikante Vorteil von Enzalutamid auf Studienpopulationsebene für das „radiographische progressionsfreie Überleben“ über alle Subgruppen hinweg gezeigt werden konnte, profitierten bezüglich der „Zeit bis zur ersten skelettbezogenen Komplikation“ vor allem Patienten, welche keine begleitenden Steroide bekamen, von einer Enzalutamid-Behandlung. Für Patienten welche Steroide erhielten, war ein starker Trend in diese Richtung ebenfalls vorhanden, jedoch nicht signifikant nachweisbar. Innerhalb der Kategorie „Schmerzen“ ergaben sich sowohl für die „Schmerzlinderung“, für die „Zeit bis zur Schmerzprogression“ und für die Auswertung der „Schmerzprogressionsrate“ als Veränderung gegenüber Studienbeginn durch die Subgruppenanalysen keine Veränderungen hinsichtlich der Aussagen der Hauptanalysen. Im Rahmen der *Responder*-Analyse der „Schmerzprogressionsrate“ zeigte sich jedoch Enzalutamid vor allem für Patienten mit einer

höheren Schmerzbelastung, für Patienten ohne begleitende Steroidgabe und für Patienten unter 65 Jahren als die signifikant überlegene Therapie. Die Subgruppenanalysen zu den Lebensqualitätsendpunkten ergaben, dass die auf Studienpopulationsebene erzielten Ergebnisse für die FACT-P *Responder*-Analyse, die Betrachtung der Veränderung der Gesamtpunktzahl in Woche 49 gegenüber Studienbeginn und die Auswertung der Lebensqualität nach EQ-5D durch die betrachteten Subgruppenmerkmale nicht wesentlich beeinflusst wurden. Bezüglich der Auswertung des FACT-P als Veränderung der letzten verfügbaren Erhebung gegenüber Studienbeginn zeigte sich jedoch ein Einfluss der begleitenden Steroidgabe. Hier erwies sich Enzalutamid bei Patienten ohne begleitende Steroidgabe im Vergleich zu Placebo als signifikant überlegen, während die Signifikanzschwelle für Patienten mit begleitender Steroidgabe knapp überschritten wurde. Bei der Analyse der Sicherheitsparameter ergaben sich für die Endpunkte „unerwünschte Ereignisse“ und „Studienabbruch aufgrund unerwünschter Ereignisse“ keine Veränderungen bezüglich der Ergebnisse auf Studienpopulationsebene. Für die Endpunkte „unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ und „schwerwiegende unerwünschte Ereignisse“ zeigte sich in der Enzalutamid-Gruppe ein geringeres relatives Risiko für das Auftreten des Ereignisses für Patienten ohne begleitende Steroidmedikation, Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 und Patienten mit einer Schmerzpunktzahl < 4 in der Enzalutamid-Gruppe. Für die „unerwünschten Ereignisse nach CTCAE Grad ≥ 3“ trat dieser Effekt zusätzlich bei Patienten mit begleitender Bisphosphonatgabe zutage. Auch das Relative Risiko aufgrund eines unerwünschten Ereignisses zu sterben, war bei Patienten, welche studienbegleitend Bisphosphonate erhielten, im Enzalutamid-Arm signifikant erniedrigt.

Tabelle 4-2: Enzalutamid vs. Placebo: Zusammenfassung des Zusatznutzens auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
Mortalität		
Gesamtüberleben	HR = 0,63 [0,53; 0,75] absolute Differenz: 4,8 Monate absoluter Zeitraum: 18,4 vs. 13,6 Monate p < 0,0001	Endpunktkategorie: Überlebensdauer KIs < 0,85 Beleg für Zusatznutzen Ausmaß: erheblich
Morbidität		
Radiographisches progressionsfreies Überleben (rPFS)	HR = 0,40 [0,35; 0,47] absolute Differenz: 5,4 Monate absoluter Zeitraum: 8,3 vs. 2,9 Monate p < 0,0001	Endpunktkategorie: schwerwiegende /schwere Symptome / Folgekomplikationen KIs < 0,75 Beleg für einen Zusatznutzen Ausmaß: erheblich
Zeit bis zur ersten skelettbezogenen Komplikation	HR = 0,69 [0,57; 0,84] absolute Differenz: 3,4 Monate absoluter Zeitraum: 16,7 vs. 13,3 Monate p = 0,0001	Endpunktkategorie: schwerwiegende /schwere Symptome / Folgekomplikationen KIs < 0,90 Beleg für einen Zusatznutzen Ausmaß: beträchtlich
- Zeit bis zur Bestrahlung des Knochens	HR = 0,71 [0,55; 0,91] absolute Differenz: n.e. p = 0,0058	[a]
- Zeit bis zur Knochenoperation	HR = 2,49 [0,30; 20,79] absolute Differenz: n.e. p = 0,38	[a]
- Zeit bis zur pathologischen Knochenfraktur	HR = 0,85 [0,47; 1,56] absolute Differenz: n.e. p = 0,60	[a]
- Zeit bis zur Rückenmarkkompression	HR = 1,00 [0,64; 1,55] absolute Differenz: n.e. p = 0,99	[a]
- Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen	HR = 0,54 [0,28; 1,04] absolute Differenz: n.e. p = 0,06	[a]
Schmerzlinderung	RR = 6,73 [0,99; 45,88] p > 0,05	Hinweis auf keinen Zusatznutzen oder –schaden

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
Schmerzprogressionsrate (*Responder*-Analyse)	RR = 0,71 [0,59; 0,87] p = 0,0008	Endpunktkategorie: Nicht schwerwiegende Symptome und Nebenwirkungen KIs < 0,90 Hinweis[c] auf einen Zusatznutzen Ausmaß: gering
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)	SMD = -0,48 [-0,62; -0,34] p < 0,0001	Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
Zeit bis zur Schmerzprogression	HR = 0,56 [0,41; 0,78] absolute Differenz: n.e. p = 0,0004	Endpunktkategorie: Nicht schwerwiegende Symptome und Nebenwirkungen KIs < 0,80 Beleg für einen Zusatznutzen Ausmaß: beträchtlich
Gesundheitsbezogene Lebensqualität		
FACT-P (*Responder*-Analyse)	RR = 2,81 [2,05; 3,84] Gegenereignis: RR = 0,68 [0,63; 0,74][b] p < 0,0001	Endpunktkategorie: Lebensqualität KIs < 0,75 Hinweis[c] auf einen Zusatznutzen Ausmaß: erheblich
FACT-P (Veränderung Woche 49 gegenüber Studienbeginn)		Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-P Gesamtpunktzahl	SMD = 0,22 [-0,23, 0,67] p = 0,33	
- FACT-G Gesamtpunktzahl	SMD = 0,19 [-0,25; 0,63] p = 0,40	
- FACT-P Körperliches Wohlbefinden	SMD = -0,07 [-0,50; 0,36] p = 0,76	[a]
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,23 [-0,20; 0,66] p = 0,29	[a]
- FACT-P Seelisches Wohlbefinden	SMD = 0,19 [-0,25; 0,63] p = 0,40	[a]
- FACT-P Funktionelles Wohlbefinden	SMD = 0,09 [-0,35; 0,53] p = 0,70	[a]
- FACT-P Prostatakarzinom-Subskala	SMD = 0,05 [-0,39; 0,49] p = 0,83	[a]
FACT-P (Veränderung letzte Erhebung		

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
gegenüber Studienbeginn)		
- FACT-P Gesamtpunktzahl	SMD = 0,42 [0,27; 0,56] p < 0,0001	Endpunktkategorie: Lebensqualität Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-G Gesamtpunktzahl	SMD = 0,36 [0,21; 0,50] p < 0,0001	Endpunktkategorie: Lebensqualität Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-P Körperliches Wohlbefinden	SMD = 0,26 [0,11; 0,40] p = 0,0006	[a]
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,06 [-0,08; 0,21] p = 0,41	[a]
- FACT-P Seelisches Wohlbefinden	SMD = 0,33 [0,18; 0,47] p < 0,0001	[a]
- FACT-P Funktionelles Wohlbefinden	SMD = 0,30 [0,15; 0,45] p < 0,0001	[a]
- FACT-P Prostatakarzinom-Subskala	SMD = 0,34 [0,20; 0,49] p < 0,0001	[a]
EQ-5D VAS Gesamtpunktzahl (Veränderung Woche 49 gegenüber Studienbeginn)	SMD = 0,85 [-0,69; 2,40] p = 0,26	Endpunktkategorie: Lebensqualität Hinweis auf keinen Zusatznutzen oder –schaden
EQ-5D VAS Gesamtpunktzahl (Veränderung last assessment gegenüber Studienbeginn)	SMD = -0,03 [-0,36; 0,30] p = 0,86	Endpunktkategorie: Lebensqualität Hinweis auf keinen Zusatznutzen oder –schaden
Sicherheit		
Unerwünschte Ereignisse	RR = 1,00 [0,99; 1,02] p = 0,67	Endpunktkategorie: Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden
Unerwünschte Ereignisse CTCAE Grad ≥ 3	RR = 0,85 [0,76; 0,96] p = 0,009	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen KIs < 1,00 Beleg für einen Zusatznutzen Ausmaß: gering
Schwerwiegende unerwünschte Ereignisse	RR = 0,87 [0,74; 1,02] p = 0,08	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden
Studienabbruch aufgrund unerwünschter Ereignisse	RR = 0,78 [0,53; 1,15] p = 0,20	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
		schaden
Tod aufgrund unerwünschter Ereignisse	RR = 0,82 [0,43; 1,58] p = 0,55	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden

a: Die Subskalen werden nicht zur separaten Ableitung eines Zusatznutzens herangezogen, sondern dienen der Absicherung des Zusatznutzens auf Ebene der Gesamtskala.

b: Berechnung des RR für das Gegenereignis

c: Liegt ein hohes Verzerrungspotenzial vor, wurde die Aussagekraft um eine Kategorie gemindert, d.h. ein Beleg wurde in einem solchen Fall als Hinweis bewertet.

d: Wie in den Methoden 4.0 [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011] vorgeschlagen, wurde zur Ableitung des Zusatznutens eine *Responder*-Analyse durchgeführt.

KI: Konfidenzintervall

KIs: obere Grenze Konfidenzintervall

RR: Relatives Risiko

HR: Hazard Ratio

SMD: *Standardized Mean Difference*

CTCAE: *Common Toxicity Criteria for Adverse Events*

n.e.: nicht erreicht

Tabelle 4-3: Zusammenfassung der Effektschätzer für Subgruppen, bei denen ein Beleg für oder ein Hinweis auf eine Interaktion vorlagen, aus der randomisiert kontrollierten Studie AFFIRM

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, 95 %-KI, p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Gesamtüberleben			
Gabe von Steroiden während der Studie	Ja	HR = 0,70 [0,56; 0,88], p = 0,002	p = 0,06
	Nein	HR = 0,49 [0,37; 0,65], p < 0,0001	Hinweis
Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings	Ja	HR = 0,78 [0,56; 1,09], p = 0,15	p = 0,15
	Nein	HR = 0,57 [0,46; 0,70], p < 0,0001	Hinweis
Hämoglobinwert zu Studienbeginn	≤ Median	HR = 0,68 [0,55; 0,85], p = 0,0005	p = 0,14
	> Median	HR = 0,50 [0,37; 0,68], p < 0,0001	Hinweis
Radiographisches progressionsfreies Überleben			
Gabe von Steroiden während der Studie	Ja	HR = 0,60 [0,49; 0,73], p < 0,0001	p < 0,0001
	Nein	HR = 0,26 [0,21; 0,33], p < 0,0001	Beleg
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	HR = 0,34 [0,28; 0,40], p < 0,0001	p < 0,0015
	≥ 4	HR = 0,58 [0,45; 0,75], p < 0,0001	Beleg
Alter	< 65	HR = 0,46 [0,36; 0,59], p < 0,0001	p = 0,18
	≥ 65	HR = 0,38 [0,32; 0,45], p < 0,0001	Hinweis
Anzahl vorheriger Chemotherapien	1	HR = 0,38 [0,32; 0,44], p < 0,0001	p = 0,10
	≥ 2	HR = 0,48 [0,36; 0,63], p < 0,0001	Hinweis
Zeit bis zur ersten skelettbezogenen Komplikation			
Gabe von Steroiden während der Studie	Ja	HR = 0,78 [0,61; 1,01], p = 0,06	p = 0,12
	Nein	HR = 0,56 [0,41; 0,75], p = 0,0001	Hinweis
Alter	< 65	HR = 0,60 [0,44; 0,83], p = 0,002	p = 0,20
	≥ 65	HR = 0,77 [0,60; 0,99], p = 0,04	Hinweis

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, 95 %-KI, p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Schmerzprogressionsrate (Responder-Analyse)			
Gabe von Steroiden während der Studie	Ja	RR = 0,84 [0,64; 1,10], p = 0,20	p = 0,15
	Nein	RR = 0,59 [0,44; 0,79], p = 0,0004	Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,85 [0,67; 1,06], p = 0,14	p = 0,003
	≥ 4	RR = 0,39 [0,26; 0,60], p < 0,0001	Beleg
Alter	< 65	RR = 0,59 [0,42; 0,83], p = 0,002	p = 0,16
	≥ 65	RR = 0,79 [0,61;1,01], p = 0,06	Hinweis
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)			
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	SMD = -0,46 [-0,61; -0,31], p < 0,0001	p = 0,14
	2	SMD = -0,68 [-1,22; -0,14], p = 0,02	Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	SMD = -0,46 [-0,62; -0,30], p < 0,0001	p = 0,10
	≥ 4	SMD = -0,58 [-0,88; -0,28], p < 0,001	Hinweis
Lebensqualität nach FACT-P (Responder-Analyse)			
Gabe von Steroiden während der Studie	Ja	RR = 2,10 [1,38; 3,19], p < 0,0005 Gegenereignis: 0,65 [0,58; 0,72]	p = 0,09
	Nein	RR = 3,69 [2,29; 5,95], p < 0,0001 Gegenereignis: 0,73 [0,64; 0,84]	Hinweis
Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl –Woche 49			
Gabe von Steroiden während der Studie	Ja	SMD = -0,09 [-0,77; 0,59], p = 0,79	p = 0,16
	Nein	SMD = 0,57[-0,08; 1,22], p = 0,08	Hinweis
Alter	< 65	SMD = -0,86 [-2,08; 0,37], p = 0,17	p = 0,02
	≥ 65	SMD = 0,47 [-0,03; 0,96], p = 0,06	Beleg

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, 95 %-KI, p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			
Gabe von Steroiden während der Studie	Ja Nein	SMD = 0,22 [-0,01; 0,45], p = 0,06 SMD = 0,61 [0,41; 0,81], p < 0,0001	p = 0,03 Beleg
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1 2	SMD = 0,38 [0,23; 0,54], p < 0,0001 SMD = 0,86 [0,25; 1,47], p = 0,006	p = 0,10 Hinweis
Lebensqualität nach FACT-G (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl –Woche 49			
Gabe von Steroiden während der Studie	Ja Nein	SMD = -0,13 [-0,77; 0,52], p = 0,70 SMD = 0,49 [-0,16; 1,13], p = 0,14	p = 0,17 Hinweis
Alter	< 65 ≥ 65	SMD = -0,67 [-1,72; 0,38], p = 0,21 SMD = 0,41 [-0,08; 0,90], p = 0,10	p = 0,04 Beleg
Lebensqualität nach FACT-G (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			
Gabe von Steroiden während der Studie	Ja Nein	SMD = 0,19 [-0,04; 0,41], p = 0,10 SMD = 0,53 [0,33; 0,72], p < 0,0001	p = 0,05 Beleg
Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl –Woche 49			
Gabe von Steroiden während der Studie	Ja Nein	SMD = NA (NA; NA), p = NA SMD = -0,75 [-2,92; 1,43], p = 0,48	p = 0,06 Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4 ≥ 4	SMD = -0,85 [-3,03; 1,33], p = 0,42 SMD = 2,61 [-0,16; 5,39], p = 0,06	p = 0,02 Beleg
Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			
Durchschnittliche Schmerzpunktzahl	< 4	SMD = -0,14 [-0,51; 0,23], p = 0,46	p = 0,20

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, 95 %-KI, p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	≥ 4	SMD = 0,35 [-0,42; 1,11], p = 0,36	Hinweis
Anzahl vorheriger Chemotherapien	1	SMD = -0,22 [-0,61; 0,16], p = 0,25	p = 0,03
	≥ 2	SMD = 0,67 [-0,01; 1,36], p = 0,05	Beleg
Unerwünschte Ereignisse nach* CTCAE *Grad ≥ 3			
Gabe von Bisphosphonaten während der Studie	Ja	RR = 0,76 [0,64; 0,91], p = 0,002	p = 0,10[c]
	Nein	RR = 0,94 [0,80; 1,10], p = 0,45	Hinweis
Gabe von Steroiden während der Studie	Ja	RR = 0,95 [0,83; 1,08], p = 0,43	p = 0,10 bzw. p = 0,02[a]
	Nein	RR = 0,70 [0,56; 0,86], p = 0,0009	Hinweis bzw. Beleg
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	RR = 0,83 [0,73; 0,94], p = 0,004	p = 0,31 bzw. p = 0,14[a]
	2	RR = 1,04 [0,79; 1,37], p = 0,79	Keine Interaktion bzw. Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,81 [0,69; 0,94], p = 0,006	p = 0,37 bzw. p = 0,18[a]
	≥ 4	RR = 0,95 [0,79; 1,13], p = 0,55	Keine Interaktion bzw. Hinweis
Schwerwiegende Unerwünschte Ereignisse			
Gabe von Steroiden während der Studie	Ja	RR = 0,94 [0,78; 1,13], p = 0,53	p = 0,25 bzw. p = 0,14[a]
	Nein	RR = 0,74 [0,56; 0,96], p = 0,03	Keine Interaktion bzw. Hinweis
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	RR = 0,82 [0,69; 0,97], p = 0,02	p = 0,04
	2	RR = 1,34 [0,86; 2,08], p = 0,19	Beleg
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,81 [0,66; 0,99], p = 0,04	p = 0,28 bzw. p = 0,19[a]
	≥ 4	RR = 0,99 [0,78; 1,25], p = 0,93	Keine Interaktion bzw. Hinweis
Tod aufgrund unerwünschter Ereignisse			
Gabe von Bisphosphonaten während der Studie	Ja	RR = 0,27 [0,08; 0,92], p = 0,04	p = 0,03
	Nein	RR = 1,41 [0,60; 3,29], p = 0,43	Beleg

a: Berechnet nach Deeks 2010 [Deeks 2010]

HR: Hazard Ratio
RR: Relatives Risiko
SMD: *Standardized Mean Difference*
ECOG: *Eastern Cooperative Oncology Group*
BPI: *Brief Pain Inventory*
FACT-P: *Functional Assessment of Cancer Therapy- Prostate*
FACT-G: *Functionel Assessment of Cancer Therapy- General*
EQ-5D: *European Quality of Life Five-Domain Scale*
CTCAE*: Common Toxicity Criteria for Adverse Events*
NA: nicht anwendbar (Grund: Nur ein Patient in der Placebogruppe, Schätzer nicht berechenbar)

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Im vorliegenden Bericht wurde der Zusatznutzen von Enzalutamid + *Best Supportive Care* im Vergleich zur Behandlung mit *Best Supportive Care* alleine als zweckmäßige Vergleichstherapie bei der Behandlung erwachsener Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet, bewertet. Die Therapie mit Enzalutamid in dieser Patientengruppe zeigte in der Zulassungsstudie AFFIRM einen erheblichen, therapeutisch bedeutsamen Zusatznutzen. Die Studie weist ein niedriges Verzerrungspotential auf und wurde u.a. aufgrund ihrer Größe und der Ergebnissicherheit als Beleg für den Zusatznutzen gewertet. Die Effektrichtung der untersuchten Endpunkte war gleichgerichtet, die Verzerrung auf Endpunktebene – mit Ausnahme der Lebensqualität – war niedrig, für Subgruppen ergaben sich keine abweichenden Ergebnisse im Vergleich zur Ebene der Zulassungspopulation. Der erhebliche Zusatznutzen von Enzalutamid + *Best Supportive Care* im Vergleich zur Behandlung mit *Best Supportive Care* zeigt sich sowohl in einer signifikanten Verbesserung des Überlebens, wie auch in einer Reduktion der Schmerzen sowie der Zeit bis zur ersten skelettbezogenen Komplikation, einer Verbesserung der Lebensqualität und einem vergleichbaren, teils besserem Sicherheitsprofil. (Ausnahme: Der Zusatznutzen bzw. verminderte Schaden weist für die Subgruppe „studienbegleitende Gabe von Bisphosphonaten“ für den „Tod aufgrund eines unerwünschten Ereignisses“ ein beträchtliches Ausmaß (Hinweis) auf, wofür auf Ebene der Studienpopulation kein statistisch signifikanter Unterschied bestand.)

Die Zusammenschau der signifikant positiven Effekte in den Bereichen Mortalität, Morbidität und Lebensqualität sowie der Sicherheit ergibt einen Beleg für einen erheblichen Zusatznutzen für Enzalutamid im Vergleich zu Best Supportive Care.

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.

Das Ziel des vorliegenden Berichts ist die Bewertung des Zusatznutzens von Enzalutamid + *Best Supportive Care* im Vergleich zur Behandlung mit *Best Supportive Care* alleine als zweckmäßige Vergleichstherapie bei der Behandlung erwachsener Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet.

Die Bewertung erfolgte bezüglich der patientenrelevanten Endpunkte Mortalität („Gesamtüberleben“), Morbidität („Radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“, „Schmerz“, „Lähmungen“), Lebensqualität und Sicherheit („unerwünschte Ereignisse“, „unerwünschte Ereignisse mit CTCAE Grad ≥ 3“, „schwerwiegende unerwünschte Ereignisse“, „Studienabbruch aufgrund unerwünschter Ereignisse“, „unerwünschte Ereignisse, die zum Tod führten“).

In die Bewertung gingen ausschließlich direkt vergleichende, randomisierte und kontrollierte Studien (RTCs) ein.

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.

Tabelle 4-4: Ein- und Ausschlusskriterien.

E1	Patientenpopulation (Indikation)	Erwachsene Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet
E2	Intervention	Enzalutamid plus *Best Supportive Care*
E3	Vergleichstherapie	*Best Supportive Care*
E4	Endpunkte (patientenrelevant)	**Mortalität** • Gesamtüberleben **Morbidität** • Radiographisches progressionsfreies Überleben • Zeit bis zur ersten skelettbezogenen Komplikation • Schmerz • Lähmungen **Lebensqualität** **Sicherheit**: • Unerwünschte Ereignisse (gesamt) • Unerwünschte Ereignisse mit CTCAE Grad ≥ 3 • Schwerwiegende unerwünschte Ereignisse • Studienabbruch aufgrund unerwünschter Ereignisse • Unerwünschte Ereignisse, die zum Tod führten
E5	Studientyp (Design)	Randomisierte kontrollierte Studie (RCT)
E6	Vollpublikation verfügbar	Ja (Studienbericht)
E7	Publikationssprache	Deutsch oder Englisch
E8	Studiendauer	Keine Einschränkung
A1	Studienart	Tierexperimentelle Studien
A2	Doppelpublikation	Doppelpublikationen ohne relevante Zusatzinformation

Alle den Einschlusskriterien genügenden und nicht unter die Ausschlusskriterien fallenden Studien wurden in die Bewertung einbezogen.

Patientenpopulation:

Enzalutamid ist indiziert bei erwachsenen Männern mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet. [Astellas Pharma GmbH 2013a]

Intervention:

Die zu prüfende Substanz ist Enzalutamid, verabreicht als orale Therapie gemäß Zulassung in einer täglichen Dosis von 160 mg, eingenommen als 4 Kapseln mit je 40 mg, unabhängig von den Mahlzeiten. [Astellas Pharma GmbH 2013a] *+ Best Supportive Care*

Vergleichstherapie:

Als Vergleichsintervention wurde nach Maßgabe des G-BA (Nachgang zum Beratungsgespräch vom 12.08.2012 [Gemeinsamer Bundesausschuss (G-BA) 2012] *Best Supportive Care* herangezogen: *„Die zweckmäßige Vergleichstherapie für Enzalutamide […] ist* Best Supportive Care *(z.B. eine adäquate Schmerztherapie). Als* Best Supportive Care *wird die Therapie verstanden, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität gewährleistet"*. [Gemeinsamer Bundesausschuss (G-BA) 2012]

Endpunkte:

Im vorliegenden Dossier wurden die folgenden Endpunkte bewertet: „Gesamtüberleben", „Zeit bis zur ersten skelettbezogenen Komplikation", „Schmerz", „Lähmungen" und „gesundheitsbezogene Lebensqualität", sowie die Sicherheitsendpunkte „unerwünschte Ereignisse" (gesamt), „unerwünschte Ereignisse mit CTCAE Grad ≥ 3", „schwerwiegende unerwünschte Ereignisse", „Studienabbruch aufgrund unerwünschter Ereignisse" und „unerwünschte Ereignisse, die zum Tod führten". Als ein wichtiger Surrogatparameter in der Behandlung des metastasierten kastrationsresistenten Prostatakarzinoms wurde außerdem das „radiographische progressionsfreie Überleben" erfasst.

Die Studien wurden auf das Vorhandensein mindestens eines dieser Endpunkte überprüft.

Studientyp:

Für die hier untersuchte Indikation und Vergleichsintervention sind randomisierte, kontrollierte Studien (RCTs) möglich und durchführbar. Da RCTs bei adäquater Durchführung die höchste Ergebnissicherheit aufweisen, wurden für die Bewertung des medizinischen Nutzens und Zusatznutzens ausschließlich Studien mit diesem Design berücksichtigt.

Vollpublikation verfügbar:

Als Vollpublikation gelten Studienberichte oder andere Berichte über die Studie, welche eine ausreichende Informationsbasis im Sinne der VerfO und damit eine Bewertung der Studie ermöglichen.

Studiendauer:

Bezüglich der Studiendauer wurde keine Einschränkung vorgenommen.

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

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

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist 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.

Die bibliografische Literaturrecherche wurde am 08.07.2013 in MEDLINE®, der Cochrane-, der Embase®- und Embase®-Alert-Datenbank durchgeführt. Für jede Datenbank wurde eine eigene, adaptierte Suchstrategie verwendet. Zunächst wurde in MEDLINE®, Embase®- und Embase®-Alert eine systematische Suche nach randomisierten Studien zu Enzalutamid durchgeführt.

Die Suche in der Cochrane-Datenbank erfolgte aufgrund der geringeren Variabilität und Differenziertheit der potenziellen Suchalgorithmen ohne Einschränkung nach klinischen Studien.

Es erfolgte keine zeitliche Einschränkung bei der Suche in der Cochrane- und der Embase®-Alert-Datenbank. Die MEDLINE®-Datenbank wurde von 1950 bis 08.07.2013 durchsucht und die Embase®-Datenbank von 1947 bis 08.07.2013.

4.2.3.3 Suche in Studienregistern

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

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

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für

einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

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

Durchsucht wurden die Studienregister *clinicaltrials.gov* (http://clinicaltrials.gov/) und das *International Clinical Trials Registry Platform Search Portal* (http://apps.who.int/trialsearch/) der *World Health Organization* (WHO). Gemäß den Angaben auf der Website des G-BA steht die Informationsplatform *clinicalstudyresults.org* (http://www.clinicalstudyresults.org) seit Ende 2011 nicht mehr zur Verfügung. Daher ist gemäß dem G-BA eine Abfrage in dieser nicht mehr erforderlich. Die Suchstrategie richtete sich nach der Syntax und den Optionen der jeweiligen Suchoberfläche. Pro Register wurde eine Suche zu klinischen Studien mit Enzalutamid durchgeführt. Die Suchen im Register von *clinicaltrials.gov* und im Studienregister des *International Clinical Trials Registry Platform Search Portal* erfolgten am 13.02.2013. Es wurde keine zeitliche Einschränkung der Suchen vorgenommen. Ein Update der Suchen erfolgte am 25.06.2013. Zudem wurde firmenintern nach entsprechenden laufenden Studien recherchiert.

4.2.3.4 Selektion relevanter Studien

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

Die Liste der Treffer wurde anhand der prädefinierten Selektionskriterien von zwei Reviewern unabhängig voneinander durchgesehen. Dabei erfolgte zunächst eine Selektion aufgrund der Titel und, falls vorhanden, der Abstracts. Treffer bzw. Publikationen, die dabei bereits ersichtlich eines der Einschlusskriterien nicht erfüllten, wurden ausgeschlossen. Die Volltexte der verbliebenen, potenziell relevanten Studien wurden erneut anhand der prädefinierten Kriterien hinsichtlich ihrer Relevanz bewertet. Aufgrund des Volltextes auszuschließende Publikationen wurden mit Angabe des Ausschlussgrundes aufgelistet. Im Fall von Diskrepanzen zwischen den beiden Reviewern wurden diese diskutiert und im Konsens gelöst.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert

werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-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.

Das Verzerrungspotenzial der Ergebnisse wurde auf Basis der Vorgaben des IQWiG [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011] und der Verfahrensordnung [Gemeinsamer Bundesausschuss (G-BA) 2013] für jede in die Nutzenbewertung eingeschlossene Studie und separat für jeden patientenrelevanten Endpunkt bewertet. Dazu wurden folgende, endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet:

A. Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene:

- Adäquate Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie des Behandlers
- Hinweis auf ergebnisgesteuerte Berichterstattung
- Sonstige, das Verzerrungspotenzial beeinflussende Punkte

B. Aspekte des Verzerrungspotentials der Ergebnisse von Endpunkten:

- Adäquate Umsetzung des ITT-Prinzips
- Verblindung der Endpunkterheber
- Hinweise auf ergebnisgesteuerte Berichterstattung
- Sonstige, das Verzerrungspotenzial beeinflussende Punkte

Die Beurteilung des Verzerrungspotenzials wurde durch einen zweiten Reviewer überprüft. Diskrepanzen in der Beurteilung durch die Reviewer wurden durch Diskussion aufgelöst.

Die Kriterien zur Beurteilung wurden entsprechend der Vorgaben im Anhang 4-G gewählt.

Die Bewertung des ITT-Prinzips wurde nach folgenden Regeln vorgenommen. Fehlten mehr als 15 % der Auswertungspopulation oder betrug die Differenz der zwischen den Behandlungsarmen in der Auswertungspopulation fehlenden Patientendaten mehr als 5 %, wurden die Studien bzw. die betroffenen Endpunkte als potenziell hoch verzerrt bewertet.

Das Verzerrungspotenzial wurde jeweils in zwei Kategorien als „niedrig“ oder „hoch“ eingestuft.

Bei einem niedrigen Verzerrungspotenzial kann davon ausgegangen werden, dass die Ergebnisse der Studie nicht relevant verzerrt sind. Ein hohes Verzerrungspotenzial wurde dann zugewiesen, wenn eine relevante Verzerrung nicht ausgeschlossen werden konnte. Bei einer relevanten Verzerrung auf Studienebene ist die Gesamtaussage der Studie in Frage zu

stellen, da eine Behebung der Mängel möglicherweise zu anderen Schlussfolgerungen führen würde.

In einem ersten Schritt wurde das Verzerrungspotenzial gesamthaft pro Studie an Hand der unter A aufgeführten Aspekte als „hoch“ oder „niedrig“ bewertet. Anschließend wurde das Verzerrungspotenzial auf Endpunktebene für jeden Endpunkt separat ebenfalls in den Kategorien „niedrig“ oder „hoch“ bestimmt.

Wurde das Verzerrrungspotential bereits auf Studienebene mit „hoch“ beurteilt, wurden die Endpunkte jeweils trotzdem neu bewertet.

Dabei hatte die Beurteilung des Verzerrungspotentials keinen Einfluss auf den Einschluss einer Studie für die Nutzenbewertung. So wurden „hoch“ verzerrte Studien nicht ausgeschlossen, die Einstufung dient jedoch als Diskussionsgrundlage für heterogene Ergebnisse und die Einschätzung der Sicherheit einer Aussage.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)[1]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[2] bzw. STROBE-Statements[3] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

Auf der Basis der Angaben im Studienbericht wurden alle Detailinformationen zum Studiendesign und den Ergebnisdaten extrahiert und in den vorgegebenen Tabellen und Modulen des Dossiers dokumentiert. Die vorgegebene Tabelle (Item 2b-14b) zu Design und

[1] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[2] Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.
[3] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Durchführung der Studie wurde im Anhang des Dossiers nach den Vorgaben von CONSORT für jede Studie separat ausgefüllt, alle Ergebnisdaten werden im Abschnitt 4.3 dargestellt.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

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

Benennen Sie die Patientencharakteristika und Endpunkte, zu denen Ergebnisse dargestellt werden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Im vorliegenden Bericht wird die Studie anhand von Designcharakteristika (Studiendauer, Zahl der randomisierten Patienten, Ort und Zeitraum der Durchführung, primäre und sekundäre Zielkriterien) beschrieben. Darüber hinaus wird die krankheitsspezifische Therapie sowie die Begleitmedikation in den eingeschlossenen Studien dargestellt.

Die Studienpopulation wurde durch die wesentlichen Ein- und Ausschlusskriterien, demografische Daten (Alter, ECOG-*Performance*-Status, PSA-Wert) sowie durch Charakteristika der Therapie (Intervention und Vergleichstherapie, Begleittherapie) beschrieben. Die Studienpopulation sowie die Behandlungsgruppen wurden deskriptiv verglichen. Dabei wurden die Ein- und Ausschlusskriterien für die Teilnahme an der Studie ebenfalls dokumentiert. Auch für die hier analysierten Subgruppen wurden die Baseline-Charakteristika dargestellt.

Die Ergebnisse zu den in der Studie berichteten patientenrelevanten Endpunkten wurden für die jeweiligen Behandlungsarme beschrieben und verglichen. Zunächst wurden die in der Studie verwendeten Endpunkte den verwendeten Zielgrößen zugeordnet (Operationalisierung des Endpunkts). Die detaillierten Daten pro Endpunkt inklusive aller Auswertungen werden pro Endpunkt in gesonderten Kapiteln (siehe Abschnitt 4.3) dargestellt.

Mortalität

- Gesamtüberleben

Morbidität

- Radiographisches progressionsfreies Überleben
- Zeit bis zur ersten skelettbezogenen Komplikation
- Schmerz
- Lähmungen

Lebensqualität

Sicherheit:

- Unerwünschte Ereignisse (gesamt)
- Unerwünschte Ereignisse mit CTCAE Grad ≥ 3
- Schwerwiegende unerwünschte Ereignisse
- Studienabbruch aufgrund unerwünschter Ereignisse
- Unerwünschte Ereignisse, die zum Tod führten

Auswahl der Endpunkte und Patientenrelevanz:

Im Beratungsgespräch [Gemeinsamer Bundesausschuss (G-BA) 2012] wurde bereits festgehalten: „*Die Endpunkte Gesamtüberleben, Zeit bis zur ersten skelettbezogenen Komplikation, Schmerzlinderung, gesundheitsbezogene Lebensqualität sowie Nebenwirkungen stellen patientenrelevante Endpunkte dar*."

„Gesamtüberleben": Das Gesamtüberleben ist aufgrund seiner klaren Definition und des Fehlens von methodisch bedingten Unschärfen ein direktes und zuverlässiges Maß zur Darstellung eines Therapieeffekts und damit als valider und patientenrelevanter Endpunkt anerkannt.

Auch die Patientenrelevanz des Endpunkts "Progressionsfreies Überleben" wurde im Beratungsgespräch thematisiert. [Gemeinsamer Bundesausschuss (G-BA) 2012] Weitere Angaben hierzu finden sich im Abschnitt 4.5.4 von Modul 4.

„Zeit bis zur ersten skelettbezogenen Komplikation": Eine skelettbezogene Komplikation ist definiert als pathologische Fraktur, Bestrahlung des Knochens, Rückenmarkkompression oder

operativer Eingriff am Knochen, sowie eine Änderung der antineoplastischen Therapie, um den Knochenschmerz zu behandeln. [Medivation Inc. 2009]

Skelettbezogene Komplikationen tragen wesentlich zum Beschwerdebild des mCRPC-Patienten bei und reduzieren dessen Lebensqualität massiv. Eine Verlängerung der Zeit bis zum Auftreten der ersten skelettbezogenen Komplikation bedeutet eine Verringerung der Morbidität und ist deshalb als patientenrelevanter Endpunkt einzustufen.

Im Beratungsgespräch [Gemeinsamer Bundesausschuss (G-BA) 2012] wurde hierzu weiterhin festgelegt: *„In Bezug auf den kombinierten Endpunkt "Zeit bis zur ersten skelettbezogenen Komplikation" sind im Dossier zusätzlich die Einzelkomponenten getrennt darzulegen, da sowohl die Bewertung als auch die Gewichtung der einzelnen patientenrelevanten Endpunkte unterschiedlich ausfallen kann. In der vorliegenden Therapiesituation stellen insbesondere Lähmungen und die Lähmungsbedingte Harninkontinenz - hervorgerufen durch die Metastasierung, nicht durch die Therapie - wichtige patientenrelevante Morbiditätsendpunkte dar.*“

„Lebensqualität“: Die Lebensqualität wird hier anhand des validierten Fragebogens FACT-P bestimmt. [Esper 1997] Zusätzlich kam der EQ-5D zum Einsatz.

Angaben zu den verwendeten Instrumenten sowie zur Validierung:

Beim *Functional Assessment of Cancer Therapy - Prostate* (FACT-P)-Fragebogen, bestehend aus fünf Domänen, handelt es sich um einen patientenrelevanten, international verwendeten Fragebogen zur Erhebung der gesundheitsbezogenen Lebensqualität bei Männern, welche sich einer Prostatakarzinom-Behandlung unterziehen. [Esper 1997] Der Fragebogen zeigt eine gute Korrelation zum klinischen Verlauf der Erkrankung. [Cella 2009] Er setzt sich zusammen aus dem FACT-G-Fragebogen (*Functional Assessment of Cancer Therapy – General*), welcher generell für Patienten mit Krebserkrankungen verwendet wird und die Domänen 1-4 (Körperliches Wohlbefinden, Verhältnis zu Freunden, Bekannten und Familie, seelisches Wohlbefinden, Funktionelles Wohlbefinden) umfasst, sowie der Prostatakarzinom-Subskala (Domäne 5). Sowohl die Version 3 des FACT-P als auch die Version 4 des FACT-G gelten als validiert. [Esper 1997; Webster 1999] Astellas verwendete zur Befragung der Patienten die Version 4 des FACT-P, welche aus den im Folgenden aufgeführten Gründen ebenfalls als valide betrachtet werden kann: Die Domänen 1-4 des verwendeten Fragebogens entsprechen der validierten Version 4 des FACT-G-Fragebogens, insofern kann dieser Teil des FACT-P als validiert angesehen werden. Die Prostatakarzinom-Subskala (Domäne 5) der verwendeten Version 4 unterscheidet sich von der validierten Version 3 lediglich hinsichtlich des exakten Wortlauts in zwei Fragen und einer in Version 3 zusätzlich gestellten, die Domäne 5 zusammenfassenden Frage (siehe Tabelle 4-5). Diese zusätzliche Frage wurde in der Version 3 nur zu Testzwecken integriert, nicht aber zur Berechnung der Punktzahl verwendet. [Esper 1997] Insofern bestehen in der Prostatakarzinom-Subskala (Domäne 5) hinsichtlich der zu vergebenden Punktzahl und der Ermittlung dieser Punktzahl keine Unterschiede zwischen den Versionen 3 und 4.

Da sich somit die Unterschiede zwischen den Versionen 3 und 4 in der Domäne 5 des FACT-P auf minimale Formulierungsänderungen in zwei Fragen beschränken, und die Domänen 1-4 der Version 4 des FACT-P identisch mit dem validierten FACT-G-Fragebogen sind, kann davon ausgegangen werden, das die von Astellas ermittelten Ergebnisse zum Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ valide sind.

Tabelle 4-5: Gegenüberstellung der Unterschiede in Domäne 5 der Versionen 3 und 4 des FACT-P.

FACT-P Version 3	FACT-P Version 4
Frage 37: *„I have certain areas of my body where I experience significant pain“*	Frage P2: *„I have certain parts of my body where I experience pain“*
Frage 46: *“I am able to have and keep an erection”*	Frage BL5: *“I am able to have and maintain an erection”*
Frage 47: *“Looking at the above 12 questions (Anmerkung: alle Fragen der Domäne 5), how much would you say these additional concerns affect your quality of life?” (Bewertung von 0 (not at all) bis 10 (very much so)*	Frage nicht enthalten

Clinically Meaningful Changes (CMC): Eine Verbesserung der Lebensqualität war definiert als eine Verbesserung von 10 Punkten in der globalen FACT-P Auswertung, welche bei 2 aufeinanderfolgenden Messungen, die mindestens 3 Wochen auseinander lagen, im Vergleich zum Studienbeginn erzielt werden musste. Da klinisch bedeutende Veränderungen (*Clinically Meaningful Changes*, CMC) durch einen Unterschied von 6 bis 10 Punkten in der FACT-P Gesamtpunktzahl abgebildet werden [Cella 2009], handelt es sich bei der im Dossier verwendeten Differenz von 10 Punkten um eine Differenz, die zur Erhebung der Verbesserung der Lebensqualität geeignet ist.

Der *European Quality of Life Five-Domain Scale* (EQ-5D)-Fragebogen, bestehend aus zwei Teilen, ist ein patientenrelevantes, innerhalb Europas verwendetes Instrument zur Erfassung des Gesundheitszustandes und der Auswirkungen auf die gesundheitsbezogene Lebensqualität. Da der EQ-5D ein unspezifisches Instrument ohne Indikationsbezug ist, werden krankheitsspezifische Änderungen nicht abgebildet.

„Schmerzlinderung“: Wie andere Tumorerkrankungen ist auch das metastasierte kastrationsresistente Prostatakarzinom häufig mit ausgeprägten Schmerzen verbunden. Eine Schmerzlinderung ist für den betroffenen Patienten eine erhebliche Reduktion der Krankheitslast. Das Ausmaß der Schmerzlinderung ist somit als patientenrelevanter Endpunkt anzusehen. Hier wurde zur Bestimmung der Schmerzlinderung der BPI-SF-Fragebogen (*Brief Pain Inventory - short form*) eingesetzt. [Atkinson 2010]

Wie im Beratungsgespräch [Gemeinsamer Bundesausschuss (G-BA) 2012] festgelegt, werden die beiden Endpunkte „Schmerz“ und „Lebensqualität“ separat dargestellt.

„Sicherheit“: Der patientenrelevante Endpunkt „Sicherheit“ beinhaltet die Parameter: „unerwünschte Ereignisse (UE)“, „UE nach CTCAE Grad ≥ 3“, „schwerwiegende UE (SUE)“, „Studienabbruch aufgrund UE“, und „Tod aufgrund UE“. Bei der Darstellung der unerwünschten Ereignisse orientiert sich Astellas an der Vorgehensweise des IQWiG bei der Nutzenbewertung von onkologischen Präparaten.

Die Patientenrelevanz rührt daher, dass unerwünschte Wirkungen einen individuell wahrgenommenen oder objektiv feststellbaren physischen oder psychischen Schaden darstellen, der zu einer mehr oder weniger starken kurz- oder langfristigen Verkürzung der Lebenserwartung, Erhöhung der Morbidität oder Beeinträchtigung der Lebensqualität führt oder führen kann. [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011]

4.2.5.3 Meta-Analysen

Sofern die vorliegenden Studien dazu geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden. Dabei sollen verschiedene Studientypen nicht in einer Meta-Analyse zusammengefasst werden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten[4] erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges’ g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet[5] werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen

[4] *DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-188.*
[5] *Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*

Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität[6] erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

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.

Folgende Angaben zu geplanten Sensitivitätsanalysen werden im Studienbericht gemacht:

Gesamtüberleben:

Es wurde das Hazard Ratio (Enzalutamid/Placebo) sowie das 95 %-Konfidenzintervall auf der Basis eines stratifizierten Cox-Regression-Modells berechnet. Im Rahmen der Sensitivitätsanalyse wurde ebenfalls der unstratifizierte log-rank-Test sowie ein unstratifiziertes Cox-Regression-Modell angewendet (Hazard Ratio für Enzalutamid/Placebo und 95 %-KI). [Medivation Inc. 2012a]

Radiographisches Progressionsfreies Überleben:

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

Folgende Sensitivitätsanalysen wurden für die Investigator-beurteilte Progression sowie die daraus abgeleitete Progression definiert: [Medivation Inc. 2012a]

1. Im Fall einer Krankheitsprogression, die zwischen zwei geplanten radiographischen Beurteilungen dokumentiert wurde, wurde das erste Datum einer radiographischen Beurteilung, an dem Evidenz für eine Krankheitsprogression vorlag, als Datum der Krankheitsprogression verwendet.

2. Um den Einfluss der Compliance für die radiographische Untersuchung zu beurteilen, wurden Patienten mit Krankheitsprogression oder im Todesfall nach zwei oder mehr aufeinanderfolgenden radiographischen Beurteilungen, so beurteilt, als hätten sie ein „Ereignis“ gehabt.

Zeit bis zur ersten skelettbezogenen Komplikation:

Im Rahmen einer Sensitivitätsanalyse wurdefür die erste skelettbezogene Komplikation eine Änderung der anti-neoplastischen Therapie, die aufgrund der Behandlung von Knochenschmerzen in Folge von Knochenmetastasen notwendig wurde, nicht als Teil der Definition für die „erste skelettbezogene Komplikation“ gewertet. [Medivation Inc. 2012a]

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.

Die Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, welche die Effekte beeinflussen könnten, untersucht. Effektmodifikatoren konnten direkte Patientencharakteristika (Subgruppenmerkmale) oder Spezifika der Behandlungen oder der Komedikation sein. Im Unterschied zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen bestand hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen mit unterschiedlichen Merkmalen und unterschiedlicher Behandlung aufzudecken.

Dafür wurden in der eingeschlossenen Studie AFFIRM die folgenden Subgruppen-Analysen *a priori* definiert:

Der Endpunkt „Gesamtüberleben“ wurde für folgende Subgrupppen berechnet:

- ECOG-*Performance Status* bei Studienbeginn (0 oder 1 vs. 2);
- Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs. ≥ 4);
- Alter (< 65 vs. ≥ 65);
- Geographische Region (Nordamerika vs. Europa vs. übrige Welt);
- Gleason Score bei Diagnose (≤ 7 vs. ≥ 8);
- Anzahl vorheriger Chemotherapien (1 vs. ≥ 2);
- Art der Progression zu Studienbeginn (nur PSA-Progression vs. radiographische Progression mit oder ohne PSA Progression);

- Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein);
- PSA-Wert zu Studienbeginn (≤ Median vs. > Median);
- Laktat-Dehydrogenase-Wert zu Studienbeginn (≤ Median vs. > Median);
- Hämoglobin-Wert zu Studienbeginn (≤ Median vs. > Median).

Sicherheitsbezogene Ereignisse (UE, unter der Behandlung auftretend; SUE, unter der Behandlung auftretend; Studienabbruch aufgrund UE, basierend auf dem Grund für den Behandlungsabbruch laut dem Dokumentationsbogen für Behandlungsabbruch; Studienabbruch aufgrund UE, basierend auf der bezüglich der Studienmedikation ergriffenen Maßnahme laut dem Dokumentationsbogen für UEs und Tod aufgrund UE, unter der Behandlung auftretend) wurden für die folgenden Subgruppen ausgewertet:

- Alter (< 65, ≥ 65, ≥ 75);
- Anamnese einer Hypertonie (ja vs. nein);
- Anamnese kardiovaskulärer Erkrankungen (ja vs. nein).

Um die Anforderungen der Verfahrensordnung bzgl. der Auswertung von Subgruppen zu berücksichtigen, wurden zusätzlich folgende Analysen *a posteriori* durchgeführt:

Für die Endpunkte „Gesamtüberleben“, „radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“, „Schmerz“, „Lebensqualität“ sowie die sicherheitsbezogenen Ereignisse wurden folgende Subgruppen analysiert (soweit sie nicht bereits durch die prädefinierten Analysen für das Gesamtüberleben und die sicherheitsbezogenen Ereignisse abgedeckt waren):

- Gabe von Bisphosphonaten während der Studie (ja vs. nein);
- Gabe von Steroiden während der Studie (ja vs. nein);
- ECOG-*Performance-Status* bei Studienbeginn (0 oder 1 vs. 2);
- Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs. ≥ 4);
- Alter (< 65 vs. ≥ 65);
- Anzahl vorheriger Chemotherapien (1 vs. ≥ 2);
- Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein).

Berechnungsmethoden Interaktionsterme bei Subgruppen:

Angaben zur Methodik finden sich im *Statistical Analysis Plan* des CSR bzw. im *Statistical Analysis Plan* zu den Zusatzauswertungen (beides bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert). Zusätzlich wurden für die *Responder*- und die dichotomen Ereignisse die Interaktionsterme mittels Review-Manager, Methode nach Deeks berechnet. Für alle Subgruppen wurden in einem ersten Schritt mittels Review-Manager, Version 5.1, Tests auf Interaktionen durchgeführt. [Deeks 2010] Hierfür wurde ein Test auf Heterogenität über Subgruppen hinweg durchgeführt. Dabei ist $\hat{\theta}_j$ die Summe der Effektgröße für die Subgruppe j, mit dem Standardfehler ($\hat{\theta}$jSE)$\hat{\Theta}$. Die Summe für die Effektgröße kann sowohl mit einer *fixed effect* wie auch mit einer *random effect* Analyse berechnet werden. Für das *random effect model* (das im vorliegenden Bericht angewendet wurde), gelten die folgenden Formeln, die dann jeweils auf eine einzelne Subgruppe angewendet werden:

Pro Studie erhält jeder Effektschätzer ein Gewicht:

$$w_i' = \frac{1}{\mathrm{SE}\left\{\hat{\theta}_i\right\}^2 + \hat{\tau}^2}.$$

Die Summe für die Effektgröße ergibt sich durch

$$\hat{\theta}_{DL} = \frac{\sum w_i' \hat{\theta}_i}{\sum w_i'},$$

und

$$\mathrm{SE}\left\{\hat{\theta}_{DL}\right\} = \frac{1}{\sqrt{\sum w_i'}}.$$

Für Ratios werden alle Berechnungen auf einer logarithmischen Skala durchgeführt.

Zunächst wird für jede Subgruppe ein Gewicht berechnet:

$$w_j = \frac{1}{\mathrm{SE}\left\{\hat{\theta}_j\right\}^2},$$

Dann wird eine (*fixed-effect*) Analyse der Summe der Effektgrößen über die Subgruppen hinweg durchgeführt:

$$\hat{\theta}_{tot} = \frac{\sum w_j \hat{\theta}_j}{\sum w_j}.$$

Die Test-Statistik für den Unterschied über die Subgruppen hinweg ist gegeben:

$$Q_{int} = \sum w_j \left(\hat{\theta}_j - \hat{\theta}_{tot}\right)^2.$$

Unter Annahme der Null-Hypothese, dass keine Unterschiede bezüglich des Behandlungseffekts über die Subgruppen hinweg bestehen, besteht eine Chi-Quadrat Verteilung mit 1-S Freiheitsgraden (wobei S die Anzahl der Subgruppen mit Effektgrößensummen ist). Das I^2 für einen Unterschied über Subgruppen hinweg wird folgendermaßen berechnet:

$$I^2 = \max\left\{100\% \times \frac{Q_{int} - (S-1)}{Q_{int}}, 0\right\}.$$

Dieser Wert gibt das Ausmaß der Inkonsistenz über die Subgruppen-Ergebnisse hinweg an und wird als Annäherung an den Anteil an der totalen Variation in den Subgruppenschätzern interpretiert, der sich aus der genuinen Variation über die Subgruppen hinweg ergibt und nicht aus einem Stichprobenfehler (*sampling error*). [Deeks 2010]

Das Ergebnis für den jeweiligen Interaktionstest wurde als Beleg auf eine Interaktion zwischen den Subgruppen gewertet, wenn ein p-Wert von < 0,05 vorlag und als Hinweis auf eine Interaktion zwischen den Subgruppen, wenn ein p-Wert < 0,2 vorlag. Lag mindestens ein Hinweis oder aber ein Beleg auf unterschiedliche Effekte in den Subgruppen vor, so wurden neben dem Gesamteffekt auch die einzelnen Subgruppenergebnisse berichtet. [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011]

Unabhängig von den Ergebnissen der Interaktionstests wurden zusätzlich die Ergebnisse aller Subgruppenanalysen bei den Studienberichten in Modul 5 hinterlegt.

4.2.5.6 Indirekte Vergleiche

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

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

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

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

[7] *Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.*
[8] *Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.*
[9] *Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331(7521): 897-900.*
[10] *Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks ofrandomized trials. Stat Methods Med Res 2008;17(3): 279-301.*
[11] *B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*
[12] *Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009; 338: b1147.*

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

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*

- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*

- *Art und Umfang von Sensitivitätsanalysen.*

Indirekte Vergleiche wurden nicht durchgeführt.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

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 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewerteten Arzneimittel

Studienbezeichnung	**Zulassungsstudie (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer**	**Therapiearme**
AFFIRM	ja	abgeschlossen	Beginn: September 2009 Ende: *data cutoff date*: 25. September 2011	Enzalutamid Placebo
9785-CL-0003	Nein	laufend (Studienstart (*first subject in*): geplant für November 2013)	Behandlungsdauer: 57 Tage	Enzalutamid (4 Kapseln à 40 mg täglich) Enzalutamid (1 Tablette à 160 mg täglich)

Geben Sie an, welchen Stand die Information in Tabelle 4-6 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.

Die Informationen in Tabelle 4-6 spiegeln den Stand zum 1. Juli 2013 wider.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-6 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-7: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
9785-CL-0003	Es handelt sich um eine Bioverfügbarkeitsstudie. In der Studie wird eine Einmalgabe (1 x 160 mg Enzalutamid als Tablette) mit einer Vierfachgabe (4 x 40 mg Enzalutamid als Kapsel) verglichen. Die Vergleichstherapie in der Studie entspricht somit nicht der zweckmäßigen Vergleichstherapie.

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 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

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

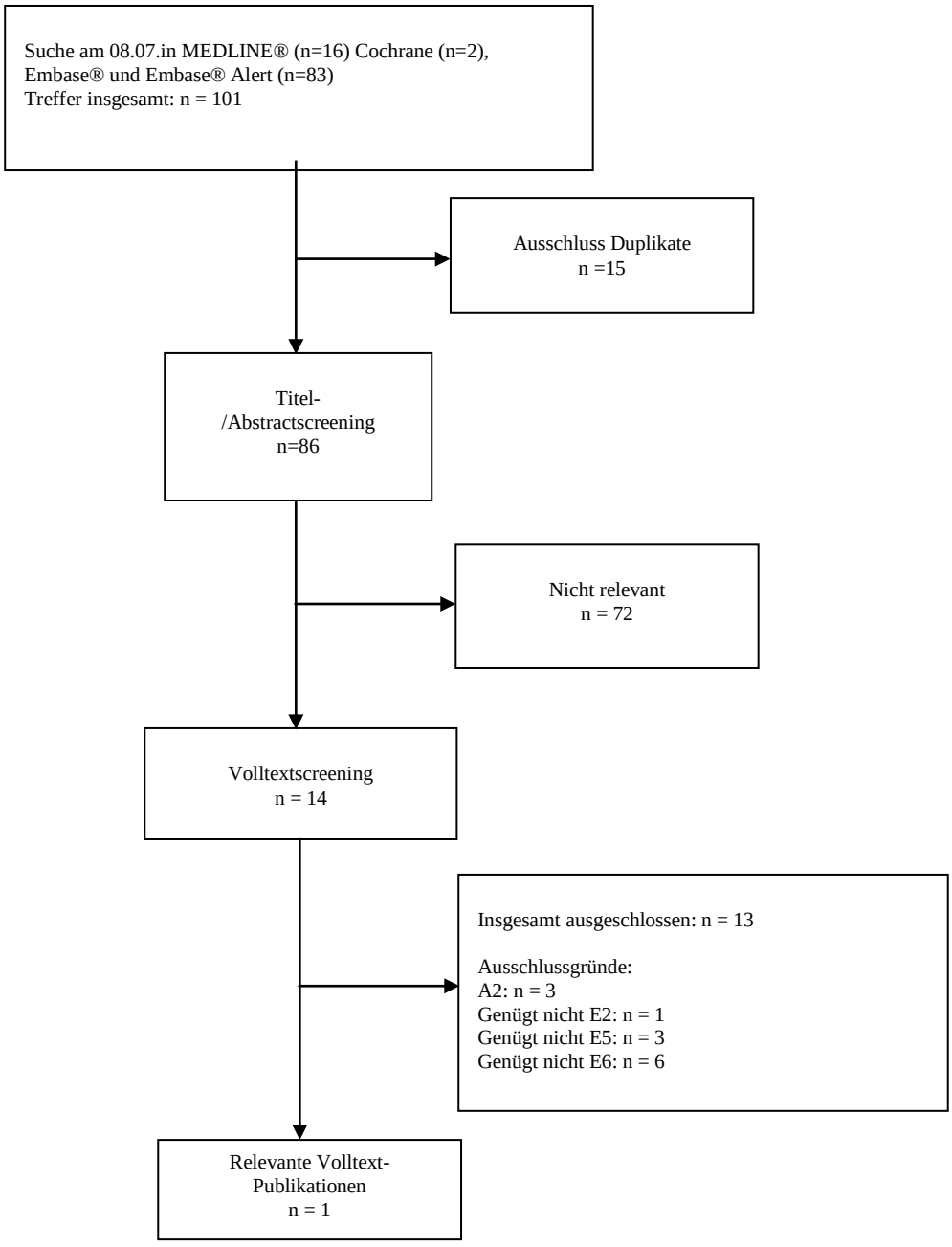


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche, Stand: 8.7.2013 – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die systematische Suche in bibliografischen Datenbanken ergab nach Ausschluss von Duplikaten, sowie der Selektion entsprechend der vordefinierten Ein- und Ausschlusskriterien auf Ebene von Titel und Abstract insgesamt 16 Publikationen, welche im Volltext gesichtet wurden. Davon stellten sich nach dem Volltextscreening 15 Publikationen als nicht relevant

(entsprechend der Ein- und Ausschlusskriterien, siehe Tabelle 4-4 und Anhang 4-C) heraus, sodass letztlich über die bibliografische Literaturrecherche eine relevante Volltextpublikation identifiziert werden konnte. Diese Publikation ist der Zulassungsstudie AFFIRM zuzuordnen. Daneben wurden 6 Abstrakts der Zulassungsstudie AFFIRM identifiziert, die jedoch nicht als Volltext publiziert vorliegen, so dass sie ausgeschlossen wurden. Um die publizierten Informationen zur Studie hier vollständig darzustellen, werden diese 6 Abtrakts hier jedoch mit aufgeführt.

Studie AFFIRM

Publikationen (alphabetisch geordnet):

Scher, H.I.; Fizazi, K.; Saad, F.; Taplin, M.E.; Sternberg, C.N.; Miller, K. et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. N.Engl.J.Med. 27.9.2012. 367, (13), 1187-1197. [Scher 2012]

Publikationen von Postern zur Studie AFFIRM (alphabetisch geordnet, ausgeschlossen, da keine Volltexte vorliegen):

Fleming, M.T.; Scher, H.I.; Fizazi, K.; Taplin, M.E.; Forer, D.; Hirmand, M. et al. Long-term responders to enzalutamide (ENZA) during the phase III AFFIRM trial: Baseline characteristics and efficacy outcomes. Journal of Clinical Oncology 2013; 31(6). [Fleming 2013]

Miller, K.; Scher, H.I.; Fizazi, K.; Basch, E.M.; Sternberg, C.N.; Hirmand, M. et al. Effect of enzalutamide on health-related quality of life (HRQoL) in men with metastatic castration-resistant prostate cancer (mCRPC) following docetaxel-based therapy: Results from the AFFIRM study. Journal of Clinical Oncology 2013; 31(6). [Miller 2013]

Mulders, P.; Scher, H.; Fizazi, K.; Saad, F.; Taplin, M.; Sternberg, C.; Miller, K.; De Wit, R.; Hirmand, M.; Selby, B.; De Bono, J., 2012. MDV3100, an androgen receptor signaling inhibitor, improves overall survival in patients with prostate cancer post docetaxel: Results from the Phase 3 AFFIRM study. Urology, 80, (3), -S30. [Mulders 2012]

Saad, F.; de Bono, J.S.; Shore, N.D.; Fizazi, K.; Hirmand, M.; Forer, D. et al. Efficacy outcomes by baseline prostate-specific antigen (PSA): Results from the phase 3 affirm trial. Journal of Urology 2013; 189(4): -e324. [Saad 2013]

Scher, H.I.; Fizazi, K.; Saad, F.; Chi, K.N.; Taplin, M.E.; Sternberg, C.N. et al. Impact of on-study corticosteroid use on efficacy and safety in the phase III AFFIRM study of enzalutamide (ENZA), an androgen receptor inhibitor. Journal of Clinical Oncology 2013; 31(6). [Scher 2013]

Sternberg, C.N.; De Bono, J.S.; Chi, K.N.; Fizazi, K.; Mulders, P.; Hirmand, M. et al. Outcomes in elderly patients with metastatic castration-resistant prostate cancer (mCRPC) treated with the androgen receptor inhibitor enzalutamide: Results from the phase III AFFIRM trial. Journal of Clinical Oncology 2013; 31(6). [Sternberg 2013]

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 4-6) 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 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

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

Tabelle 4-8: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Identifikationsorte (Name der Studienregister und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.)**
AFFIRM	***clinicaltrials.gov*:** NCT00974311 [Medivation Inc. 2012b] Hinterlegte Dokumente zu Studienergebnissen: Scher et al. 2012 [Scher 2012] **ICTRP *Search Portal*:** EUCTR2009-013174-41-BE [Medivation Inc. 2012c]	ja	ja
a: Zitat des Studienregistereintrags sowie, falls vorhanden, die im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. n. d.: bibliografische Literaturrecherche nicht durchgeführt.			

Bei der Suche in beiden Studienregistern wurde eine relevante Studie identifiziert. Ein Großteil der Treffer in der Suche des ICTPR *Search Portals* erwiesen sich als Duplikate zu den im Studienregister *clinicaltrials.gov* angezeigten Treffern. Bei den übrigen Treffern handelte es sich überwiegend nicht um randomisierte, kontrollierte Studien oder die Studienpopulation oder Intervention wich von der Zulassungspopulation für Enzalutamid bzw. der Intervention mittels Enzalutamid ab.

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 4-9: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen[a]		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie[b] (ja/nein)	Studie Dritter (ja/nein)	Studien-bericht (ja/nein [Zitat])	Registereintrag[c] (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
AFFIRM[d]	ja	ja	nein	Ja [27]	ja ***clinicaltrials.gov*:** NCT00974311 [28] Hinterlegte Dokumente zu Studienergebnissen: Scher et al. 2012 [39] **ICTRP *Search Portal:*** EUCTR2009-013174-41-BE [29]	Ja [39]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Keine Studien identifiziert						

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

Studie	Studienkategorie			verfügbare Datenquellen[a]		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie[b] (ja/nein)	Studie Dritter (ja/nein)	Studien-bericht (ja/nein [Zitat])	Registereintrag[c] (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Studiendesign und/oder -ergebnisse. d: Der SAP zur AFFIRM-Studie (Appendix 16.1.9) sowie die Zusatzauswertungen und der SAP zu den Zusatzauswertungen sind in den Unterlagen zu den Studienberichten im Modul 5 abgespeichert.						

Alle Unterlagen zu der eingeschlossenen Studie AFFIRM sowie die Dokumentation zu Methodik und Ergebnissen zusätzlicher Analysen, die für dieses Dossier durchgeführt wurden, sind im Abschnitt 4.6 aufgeführt und im Modul 5 hinterlegt (Dateien für Modul 4, Studienberichte).

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 4-10: 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**
AFFIRM	RCT, doppelblind, placebokontrolliert, multizentrisch,multi-national, Phase III Studie	Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet	Enzalutamid (n = 800) Placebo (n = 399)	9/2009 – abgeschlossen (*data cut off* Datum: 09/2011) Das *Screening* erfolgte von Woche -4 bis -1 (28 Tage). Die Medikation der Patienten erfolgte bis zum Wechsel der antineoplastischen Therapie (aufgrund von gesicherter Progression oder nicht akzeptabler Toxizität) bzw. dem Tod.[b] Bis 30 Tage nach Absetzen der Studienmedikation erfolgte ein *„safety follow up“* und anschließend ein *„long term follow up“* (60 Monate) Das Eintreten des 520. Todesfalls führte zur Durch-führung der Interimanalyse und diese zur Beendigung der Studie, die Studie lief somit 2 Jahre.	15 Länder in Nordamerika, Südamerika, Europa, Australien, Südafrika Zeitraum: 9/2009 –11/2010 (Zeit von Start bis Ende Patienteneinschluss) 9/2009 –9/2011 (Zeitraum der Durchführung) *data cutoff date* für die Interimanalyse: 25.09.2011. Die Interimanalyse führte zur Beendigung der Studie.	Primärer Endpunkt: Gesamtüberleben Sekundäre Endpunkte u.a.: Radiographisches progressionsfreies Überleben Zeit bis zum Auftreten des ersten skelettalen Ereignisses Schmerz Lebensqualität Sicherheit (Endpunkte ohne Patientenrelevanz siehe CONSORT Tabelle 4-125)[a]

a: Extrahierte primäre Endpunkte beinhalten Angaben ohne Berücksichtigung der Relevanz für diese Nutzenbewertung. Extrahierte sekundäre Endpunkte beinhalten ausschließlich Angaben zu relevanten verfügbaren Endpunkten für diese Nutzenbewertung.

b: Die mediane Behandlungsdauer für die Patienten unter Enzalutamid betrug 8,3 Monate und unter Placebo 3,0 Monate.

Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Enzalutamid	Placebo	Vorbehandlung, Begleitmedikation
AFFIRM	Enzalutamid 160 mg/täglich (4 Kapseln mit je 40 mg)	Placebo 4 Kapseln täglich	Vorangehende Chemotherapie: Alle Patienten hatten vor Studienbeginn eine oder zwei Chemotherapien erhalten (Zeitraum nicht näher definiert), wobei mindestens eine davon auf Docetaxel basierte. Begleitmedikationen: Die häufigsten Begleitmedikamente umfassten Analgetika, endokrine Therapeutika, Medikamente gegen säurebedingte Magen-Darm-Erkrankungen und Knochenerkrankungen, Entzündungshemmer und Antirheumatika (siehe Tabelle 4-12)

Tabelle 4-12: Charakterisierung der Interventionen: Auflistung der Begleitmedikamente, welche mehr als 10 % der Patienten des jeweiligen Arms während der Behandlung erhielten (sowie ausgewählte Andere)

Begleitmedikamente	**Enzalutamid (n = 800)**	**Placebo (n = 399)**	**Total (n = 1199)**
Anzahl der Patienten welche mindestens ein Begleitmedikament bekamen	800 (100 %)	399 (100 %)	1199 (100 %)
Analgetika	659 (82,4 %)	336 (84,2 %)	995 (83,0 %)
Paracetamol	414 (51,8 %)	199 (49,9 %)	613 (51,1 %)
Oxycodonhydrochlorid	138 (17,3 %)	86 (21,6 %)	224 (18,7 %)
Fenantyl	119 (14,9 %)	74 (18,5 %)	193 (16,1 %)
Morphinsulfat	125 (15,6 %)	66 (16,5 %)	191 (15,9 %)
Panadeine co (Codeinphosphat, Paracetamol)	77 (9,6 %)	49 (12,3 %)	126 (10,5 %)
Oxycodon	72 (9,0 %)	48 (12,0 %)	120 (10,0 %)
Antiepileptika	16 (2,0 %)	5 (1,3 %)	21 (1,8 %)
Carbamazepin	3 (0,4 %)	1 (0,3 %)	4 (0,3 %)
Levetiracetam	4 (0,5 %)	0 (0,0 %)	4 (0,3 %)
Clonazepam	3 (0,4 %)	0 (0,0 %)	3 (0,3 %)
Primidon	3 (0,4 %)	0 (0,0 %)	3 (0,3 %)
Natriumvalproat	2 (0,3 %)	1 (0,3 %)	3 (0,3 %)
Phenytoin	2 (0,3 %)	0 (0,0 %)	2 (0,2 %)
Ergenyl chrono (Natriumvalproat, Valproinsäure)	0 (0,0 %)	1 (0,3 %)	1 (< 0,1 %)
Oxcarbazepin	1 (0,1 %)	0 (0,0 %)	1 (< 0,1 %)
Phenobarbital	1 (0,1 %)	0 (0,0 %)	1 (< 0,1 %)
Pregabalin	0 (0,0 %)	1 (0,3 %)	1 (< 0,1 %)
Valproat seminatrium	0 (0,0 %)	1 (0,3 %)	1 (< 0,1 %)
Entzündungshemmer und Antirheumatika	397 (49,6 %)	201 (50,4 %)	598 (49,9 %)
Ibuprofen	170 (21,3 %)	96 (24,1 %)	266 (22,2 %)
Antithrombotika	310 (38,8 %)	172 (43,1 %)	482 (40,2 %)
Acetylsalicylsäure	137 (17,1 %)	79 (19,8 %)	216 (18,0 %)
Kortikosteroide für den systemischen Gebrauch	382 (47,8 %)	182 (45,6 %)	564 (47,0 %)
Prednison	190 (23,8 %)	102 (25,6 %)	292 (24,4 %)
Dexamethason	125 (15,6 %)	60 (15,0 %)	185 (15,4 %)
Diuretika	193 (24,1 %)	113 (28,3 %)	306 (25,5 %)
Furosemid	95 (11,9 %)	58 (14,5 %)	153 (12,8 %)
Medikamente gegen säurebedingte Magenstörungen	415 (51,9 %)	213 (53,4 %)	628 (52,4 %)
Omeprazol	160 (20,0 %)	95 (23,8 %)	255 (21,3 %)

Begleitmedikamente	**Enzalutamid (n = 800)**	**Placebo (n = 399)**	**Total (n = 1199)**
Medikamente gegen funktionale, gastrointestinale Störungen	198 (24,8 %)	101 (25,3 %)	299 (24,9 %)
Metoclopramid	77 (9,6 %)	42 (10,5 %)	119 (9,9 %)
Medikamente zur Behandlung von Knochenerkrankungen	383 (47,9 %)	183 (45,9 %)	566 (47,2 %)
Zoledronsäure	333 (41,6 %)	158 (39,6 %)	491 (41,0 %)
Endokrine Therapeutika	756 (94,5 %)	377 (94,5 %)	1133 (94,5 %)
Leuprorelinacetat	307 (38,4 %)	168 (42,1 %)	475 (39,6 %)
Goserelinacetat	203 (25,4 %)	109 (27,3 %)	312 (26,0 %)
Laxantien	303 (37,9 %)	158 (39,6 %)	461 (38,4 %)
Movicol (Macrogol 3350, Kaliumchlorid, Natriumbicarbonat, Natriumchlorid)	81 (10,1 %)	47 (11,8 %)	128 (10,7 %)
Macrogol	83 (10,4 %)	40 (10,0 %)	123 (10,3 %)
Lipidsenker	249 (31,1 %)	127 (31,8 %)	376 (31,4 %)
Simvastatin	90 (11,3 %)	36 (9,0 %)	126 (10,5 %)
Mineralstoffzusätze	285 (35,6 %)	151 (37,8 %)	436 (36,4 %)
Kalzium	82 (10,3 %)	39 (9,8 %)	121 (10,1 %)
Psychoanaleptika	122 (15,3 %)	64 (16,0 %)	186 (15,5 %)
Citalopram	20 (2,5 %)	7 (1,8 %)	27 (2,3 %)
Escitalopramoxalat	15 (1,9 %)	6 (1,5 %)	21 (1,8 %)
Paroxetin	12 (1,5 %)	5 (1,3 %)	17 (1,4 %)
Citalopramhydrobromid	9 (1,1 %)	7 (1,8 %)	16 (1,3 %)
Paroxetinhydrochlorid	10 (1,3 %)	5 (1,3 %)	15 (1,3 %)
Amitriptylin	6 (0,8 %)	7 (1,8 %)	13 (1,1 %)
Escitalopram	5 (0,6 %)	8 (2,0 %)	13 (1,1 %)
Amitriptylinhydrochlorid	7 (0,9 %)	5 (1,3 %)	12 (1,0 %)
Psycholeptika	288 (36,0 %)	122 (30,6 %)	410 (34,2 %)
Lorazepam	70 (8,8 %)	23 (5,8 %)	93 (7,8 %)
Zolpidemtartrat	45 (5,6 %)	19 (4,8 %)	64 (5,3 %)
Zopiclon	41 (5,1 %)	14 (3,5 %)	55 (4,6 %)
Alprazolam	31 (3,9 %)	18 (4,5 %)	49 (4,1 %)
Haloperidol	27 (3,4 %)	14 (3,5 %)	41 (3,4 %)
Bromazepam	25 (3,1 %)	15 (3,8 %)	40 (3,3 %)
Clonazepam	29 (3,6 %)	7 (1,8 %)	36 (3,0 %)
Zolpidem	22 (2,8 %)	9 (2,3 %)	31 (2,6 %)
Temazepam	16 (2,0 %)	14 (3,5 %)	30 (2,5 %)

Begleitmedikamente	**Enzalutamid (n = 800)**	**Placebo (n = 399)**	**Total (n = 1199)**
Midazolam	21 (2,6 %)	5 (1,3 %)	26 (2,2 %)
Oxazepam	12 (1,5 %)	7 (1,8 %)	19 (1,6 %)
Diazepam	12 (1,5 %)	5 (1,3 %)	17 (1,4 %)
Schilddrüsenmedikamente	42 (5,3 %)	7 (1,8 %)	49 (4,1 %)
Natriumlevothyroxin	21 (2,6 %)	4 (1,0 %)	25 (2,1 %)
Levothyroxin	14 (1,8 %)	2 (0,5 %)	16 (1,3 %)
Vitamine	237 (29,6 %)	105 (26,3 %)	342 (28,5 %)
Vitamin D NOS	97 (12,1 %)	41 (10,3 %)	138 (11,5 %)
Multivitamine (einfache)	73 (9,1 %)	43 (10,8 %)	116 (9,7 %)

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

Studie Gruppe	**N[a]**	**Alter Mittelwert (SD) Minimum Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn Mittelwert (SD) Median**
AFFIRM Enzalutamid	800	68,8 (7,96) 41,0 92,0	Status 0: 298 (37,3 %) Status 1: 432 (54,0 %) Status 2: 70 (8,8 %)	415,6 (930,76) 107,7
AFFIRM Placebo	399	68,6 (8,39) 49,0 89,0	Status 0: 156 (39,1 %) Status 1: 211 (52,9 %) Status 2: 32 (8,0 %)	389,4(1105,72) 128,3

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen

Tabelle 4-14: Charakterisierung der Subgruppen „Gabe von Biphosphonaten während der Studie (nein vs. ja)“ – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn:** **Mittelwert (SD)** **Median**
Gabe von Biphosphonaten: nein				
AFFIRM Enzalutamid	417 (52,1 %)	69,1 (8,34) 41 92	Status 0: 141 (33,8 %) Status 1: 238 (57,1 %) Status 2: 38 (9,1 %)	392,0 (925,96) 89,6
AFFIRM Placebo	216 (54,1 %)	68,7 (8,47) 49 89	Status 0: 82 (38,0 %) Status 1: 109 (50,5 %) Status 2: 25 (11,6 %)	438,6 (1419,17) 133,5
Gabe von Biphosphonaten: ja				
AFFIRM Enzalutamid	383 (47,9 %)	68,5 (7,53) 43 87	Status 0: 157 (41,0 %) Status 1: 194 (50,7 %) Status 2: 32 (8,4 %)	441,4 (936,48) 131,4
AFFIRM Placebo	183 (45,9 %)	68,6 (8,33) 49 89	Status 0: 74 (40,4 %) Status 1: 102 (55,7 %) Status 2: 7 (3,8 %)	331,2 (536,78) 123,1

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-15: Charakterisierung der Subgruppen „Gabe von Steroiden während der Studie (nein vs. ja)“ – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG–*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn:** **Mittelwert (SD)** **Median**
Gabe von Steroiden: nein				
AFFIRM Enzalutamid	418 (52,3 %)	69,0 (7,96) 43 88	Status 0: 196 (46,9 %) Status 1: 201 (48,1 %) Status 2: 21 (5,0 %)	337,9 (893,27) 77,5
AFFIRM Placebo	217 (54,4 %)	69,1 (8,56) 49 89	Status 0: 96 (44,2 %) Status 1: 106 (48,8 %) Status 2: 15 (6,9 %)	366,9 (1388,36) 99,4
Gabe von Steroiden: ja				
AFFIRM Enzalutamid	382 (47,8 %)	68,6 (7,98) 41 92	Status 0: 102 (26,7 %) Status 1: 231 (60,5 %) Status 2: 49 (12,8 %)	500,7 (964,12) 169,8
AFFIRM Placebo	182 (45,6 %)	68,1 (8,18) 49 89	Status 0: 60 (33,0 %) Status 1: 105 (57,7 %) Status 2: 17 (9,3 %)	416,1 (621,96) 177,1

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-16: Charakterisierung der Subgruppen „ECOG-*Performance*-Status (0 oder 1 vs. 2) zu Studienbeginn“ – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn** **Mittelwert (SD)** **Median**
ECOG-*Performance*-Status 0 oder 1				
AFFIRM Enzalutamid	730 (91,3 %)	68,7 (8,01) 41 92	Status 0: 298 (40,8 %) Status 1: 432 (59,2 %)	377,5 (801,39) 103,6
AFFIRM Placebo	367 (92,0 %)	68,5 (8,41) 49 89	Status 0: 156 (42,5 %) Status 1: 211 (57,5 %)	399,2 (1146,28) 124,6
ECOG-*Performance*-Status 2				
AFFIRM Enzalutamid	70 (8,8 %)	69,9 (7,43) 54 85	Status 2: 70 (100,0 %)	812,8 (1752,40) 238,1
AFFIRM Placebo	32 (8,0 %)	69,8 (8,26) 51 85	Status 2: 32 (100,0 %)	276,0 (411,39) 131,0

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-17: Charakterisierung der Subgruppen „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn** **Mittelwert (SD)** **Median**
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* < 4				
AFFIRM Enzalutamid	574 (71,8 %)	69,0 (7,80) 43 88	Status 0: 261 (45,5 %) Status 1: 283 (49,3 %) Status 2: 30 (5,2 %)	388,2 (889,88) 94,5
AFFIRM Placebo	284 (71,2 %)	69,2 (8,43) 49 89	Status 0: 133 (46,8 %) Status 1: 136 (47,9 %) Status 2: 15 (5,3 %)	394,6 (1246,03) 113,2
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* ≥ 4				
AFFIRM Enzalutamid	226 (28,3 %)	68,4 (8,37) 41 92	Status 0: 37 (16,4 %) Status 1: 149 (65,9 %) Status 2: 40 (17,7 %)	485,3 (1026,10) 165,2
AFFIRM Placebo	115 (28,8 %)	67,1 (8,17) 49 85	Status 0: 23 (20,0 %) Status 1: 75 (65,2 %) Status 2: 17 (14,8 %)	376,3 (643,39) 149,6

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-18: Charakterisierung der Subgruppen „Alter (< 65 vs. ≥ 65)“ – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn** **Mittelwert (SD)** **Median**
Alter < 65				
AFFIRM Enzalutamid	232 (29,0 %)	59,1 (4,60) 41 64	Status 0: 99 (42,7 %) Status 1: 111 (47,8 %) Status 2: 22 (9,5 %)	390,6 (819,05) 85,8
AFFIRM Placebo	130 (32,6 %)	58,9 (3,81) 49 64	Status 0: 53 (40,8 %) Status 1: 69 (53,1 %) Status 2: 8 (6,2 %)	325,6 (492,95) 106,9
Alter ≥ 65				
AFFIRM Enzalutamid	568 (71,0 %)	72,8 (5,13) 65 92	Status 0: 199 (35,0 %) Status 1: 321 (56,5 %) Status 2: 48 (8,5 %)	425,9 (973,19) 113,5
AFFIRM Placebo	269 (67,4 %)	73,3 (5,51) 65 89	Status 0: 103 (38,3 %) Status 1: 142 (52,8 %) Status 2: 24 (8,9 %)	420,2 (1302,23) 134,2

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (ei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-19: Charakterisierung der Subgruppen "Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)" – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn** **Mittelwert (SD)** **Median**
Anzahl vorheriger Chemotherapien = 1				
AFFIRM Enzalutamid	579 (72,4 %)	68,5 (7,94) 41 88	Status 0: 227 (39,2 %) Status 1: 307 (53,0 %) Status 2: 45 (7,8 %)	339,2 (723,44) 87,3
AFFIRM Placebo	296 (74,2 %)	68,4 (8,42) 49 89	Status 0: 119 40,2 %) Status 1: 149 (50,3 %) Status 2: 28 (9,5 %)	381,4 (1240,18) 111,5
Anzahl vorheriger Chemotherapien ≥ 2				
AFFIRM Enzalutamid	221 (27,6 %)	69,6 (7,99) 48 92	Status 0: 71 (32,1 %) Status 1: 125 (56,6 %) Status 2: 25 (11,3 %)	615,9 (1309,79) 223,4
AFFIRM Placebo	103 (25,8 %)	69,3 (8,32) 51 89	Status 0: 37 (35,9 %) Status 1: 62 (60,2 %) Status 2: 4 (3,9 %)	412,3 (567,15) 181,9

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-20: Charakterisierung der Subgruppen "Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein)" – RCT mit dem zu bewertenden Arzneimittel

Studie **Gruppe**	**N**[a]	**Alter** **Mittelwert (SD)** **Minimum** **Maximum**	**ECOG-*Performance*-Status bei Studienbeginn: (0, 1, 2)**	**PSA-Wert (ng/ml) zu Studienbeginn** **Mittelwert (SD)** **Median**
Viszerale Erkrankung: ja				
AFFIRM Enzalutamid	196 (24,5 %)	68,7 (8,40) 41 92	Status 0: 59 (30,1 %) Status 1: 117 (59,7 %) Status 2: 20 (10,2 %)	393,6 (708,41) 104,2
AFFIRM Placebo	82 (20,6 %)	68,9 (8,64) 51 89	Status 0: 32 (39,0 %) Status 1: 40 (48,8 %) Status 2: 10 (12,2 %)	309,0 (506,62) 126,9
Viszerale Erkrankung: **n**ein				
AFFIRM Enzalutamid	604 (75,5 %)	68,8 (7,82) 43 88	Status 0: 239 (39,6 %) Status 1: 315 (52,2 %) Status 2: 50 (8,3 %)	422,8 (992,67) 108,8
AFFIRM Placebo	317 (79,4 %)	68,5 (8,34) 49 89	Status 0: 124 (39,1 %) Status 1: 171 (53,9 %) Status 2: 22 (6,9 %)	410,1 (1213,25) 128,3

a: Anzahl Patienten in der Auswertung
SD: *Standard Deviation*
ECOG: *Eastern Cooperative Oncology Group*
PSA: Prostata-spezifisches Antigen
Quelle: Zusatzauswertungen (bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert)

Tabelle 4-21: Charakterisierung der Auswertungspopulationen – RCT mit dem zu bewertenden Arzneimittel

Auswertungspopulation	**Enzalutamid (n = 800)**	**Placebo (n = 399)**
Randomisiert (ITT-Population)	800 (100 %)	399 (100 %)
Auswertungspopulation für FACT-P[a]	630 (78,8 %)	241 (60,4 %)
Auswertungspopulation für EQ-5D	146 (18,3 %)	68 (17,0 %)
Auswertungspopulation für Schmerzlinderung	49 (6,1 %)	15 (3,8 %)
Auswertungspopulation für Schmerzprogressionsrate	625 (78,1 %)	259 (64,9 %)
Auswertungspopulation für die Zeit bis zur Schmerzprogression	800 (100 %)	399 (100 %)
Sicherheitspopulation	800 (100 %)	399 (100 %)

a: Die Angaben weichen von den im *Clinical Study Report* (CSR) berichteten Zahlen ab (Tabelle 11.1-1), da die Auswertung der Fragebögen zunächst nicht konform zu den Auswerteregeln vorgenommen worden war.
FACT-P: *Functional Assessment of Cancer Therapy – Prostate*
EQ-5D: *European Quality of Live Five Domain Scale*

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

Die AFFIRM-Studie ist eine randomisierte, placebokontrollierte, multinationale, multizentrische, Phase-III-Doppelblindstudie. Ziel der Studie war die Untersuchung der Wirksamkeit und Sicherheit von Enzalutamid im Vergleich zu Placebo bei Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet. Patienten in beiden Studienarmen erhielten neben der Therapie mittels Enzalutamid bzw. Placebo eine individualisierte, unterstützende Behandlung nach *Best Supportive Care*, z.B. eine adäquate Schmerztherapie. Das Zuteilungsverhältnis auf die Enzalutamid- und Placebogruppe betrug 2:1.

Die Studie begann im September 2009, der letzte Patient wurde im November 2010 eingeschlossen, der erste Datenschnitt fand im September 2011 statt, die Patienten wurden noch fortlaufend beobachtet. Für die Abschlussanalyse wurde eine Anzahl von 650 Todesereignissen festgelegt. Da jedoch bereits zum Zeitpunkt der Interimanalyse nach 520 Todesfällen ein statistisch signifikantes Ergebnis vorlag, wurde die Studie auf Empfehlung des unabhängigen *Data Monitoring Committees* vorzeitig beendet. Die Interimanalyse stellt somit die Abschlussanalyse dar. Die Studie wurde in 15 Ländern duchgeführt. Primäres Zielkriterium war das Gesamtüberleben. Daneben wurden u.a. das radiographische progressionsfreie Überleben, die Zeit bis zum Auftreten der ersten skelettbezogenen Komplikation, Schmerz, Lebensqualität und Sicherheit erfasst.

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 4-22: 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			
AFFIRM	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der AFFIRM-Studie wurde die Randomisierungssequenz mittels eines computergesteuerten Randomisierungsschemas (permutierte Blockrandomisierung, stratifiziert nach ECOG-*Performance*-Status und Punktzahl der Frage 3 des BPI – *short form*) erzeugt. Die Verdeckung der Gruppenzuteilung wurde durch Mitteilung über die Zuweisung der Patienten zu einer Therapiegruppe über ein *interactive voice response system / interactive web response system* (IVRS/IWRS) gewährleistet. Alle Patienten, Untersucher, das Personal der Studienzentren sowie die Mitarbeiter des Sponsors waren verblindet, und die Enzalutamid- und Placebo-Kapseln waren bezüglich ihres Aussehens und ihrer Verpackung identisch. Im Studienbericht gab es keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Ein Vergleich des Studienberichts mit der Veröffentlichung von Scher et al., in welcher die Ergebnisse der AFFIRM-Studie berichtet werden [Scher 2012], zeigte eine übereinstimmende Berichterstattung der Ergebnisse. Sonstige Aspekte, die auf ein Verzerrungspotenzial hinweisen, wurden nicht identifiziert. Das Verzerrungspotenzial der Studie wurde daher mit „niedrig“ bewertet, da sie die geforderten Kriterien für eine Minimierung des Verzerrungspotenzials erfüllt.

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 4-23: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Gesamtüberleben	Radiographisches progressionsfreies Überleben	Zeit bis zur ersten skelettbezogenen Komplikation	Schmerzlinderung	Schmerzprogressionsrate	Zeit bis zur Schmerzprogression	Lähmungen[a]	Lähmungsbedingte Harninkontinenz[a]
AFFIRM	ja	ja	ja	ja	ja	ja	ja	

a: Aufgrund der geringen Ereigniszahl wurde der Endpunkt gesondert (fallbezogen) ausgewertet.

Tabelle 4-24: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel (Fortsetzung von Tabelle 4-23)

Studie	Gesundheitsbezogene Lebensqualität nach FACT-P	Gesundheitsbezogene Lebensqualität nach EQ-5D	Unerwünschte Ereignisse	Unerwünschte Ereignisse nach CTCAE Grad ≥ 3	Schwerwiedende unerwünschte Ereignisse	Studienabbruch aufgrund unerwünschter Ereignisse	Tod aufgrund unerwünschter Ereignisse	Krampfanfälle[a]
AFFIRM	ja	ja	ja	ja	ja	ja	ja	ja

a: Aufgrund der geringen Ereigniszahl wurde der Endpunkt gesondert (fallbezogen) ausgewertet.

4.3.1.3.1 Endpunkte – RCT

4.3.1.3.1.1 Endpunkt „Gesamtüberleben“ - 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 4-25: Operationalisierung von Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Das „Gesamtüberleben“ war definiert als die Zeit zwischen Randomisierung und Tod jedweder Ursache. Der Überlebensstatus wurde durch den Untersucher und/oder den Studienkoordinator mittels Kontakt mit dem Patienten, der Familie des Patienten oder eines anderen Arztes erhoben. Die Untersuchung fand bei regulären Studienvisiten, bei der *Safety-Follow-Up* Visite und bei Protokoll-spezifischen Analysen statt. Nach Beendigung der Behandlung und Abschluss der *Safety-Follow-Up* Visite kontaktierte der Untersucher oder der Studienkoordinator alle 12 ± 1 Wochen den Patienten, die Familie des Patienten oder den Hausarzt des Patienten, um den Überlebensstatus zu beurteilen oder den Grund für den Tod zu erfahren.

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 4-26: Bewertung des Verzerrungspotenzials für den Endpunkt „Gesamtüberleben“ 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Gesamtüberleben“ wurde daher mit „niedrig“ bewertet.

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 4-27: Ergebnisse für Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	Enzalutamid			Placebo			Enzalutamid vs. Placebo
	N[a]	Patienten mit Ereignis n (%)	Median [95 %-KI] Monate	N[a]	Patienten mit Ereignis n (%)	Median [95 %-KI] Monate	Hazard Ratio [95 %-KI] p-Wert
Gesamtüberleben							
AFFIRM	800	308 (38,5)	18,4 [17,3; n.e.]	399	212 (53,1)	13,6 [11,3; 15,8]	0,63 [0,53; 0,75] p < 0,0001[b] p < 0,0001[c]

a: Anzahl Patienten in der Auswertung

b: p-Wert basiert auf einem log-rank Test stratifiziert nach ECOG-*Performance*-Status und BPI – *short form* (Frage #3)

c: unstratifizierter Test

n.e.: nicht erreicht

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der ITT-Population.

In der AFFIRM-Studie starben im Auswertezeitraum unter einer Enzalutamid-Therapie 38,5 % (308/800) der Patienten, wohingegen 53,1 % (212/399) der Patienten in der Placebo-Gruppe verstarben. Das mediane „Gesamtüberleben“ in der Enzalutamid-Gruppe betrug 18,4 Monate (95 %-KI [17,3; n.e.]), während Patienten in der Placebogruppe im Median nur 13,6 Monate (95 %-KI [11,3; 15,8]) überlebten. Patienten unter Enzalutamid-Therapie hatten somit einen Überlebensvorteil von 4,8 Monaten. Das Hazard Ratio (stratifizierte Analyse) betrug 0,63 (95 %-KI [0,53; 0,75], $p < 0,0001$) und zeigte damit einen statistisch signifikanten Vorteil für die Enzalutamid-Gruppe im Vergleich zur Behandlungsgruppe unter Placebo.

Diese Ergebnisse wurden durch die Sensitivitätsanalyse bestätigt, bei der sich für das unstratifizierte Hazard Ratio ein Wert von 0,63 (95 %-KI [0,53; 0,75]) ergab, mit einem ebenfalls statistisch signifikantem Vorteil von $p < 0,0001$.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.2 Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ – RCT

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

Tabelle 4-28: Operationalisierung von Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Das „ radiographisch progressionsfreie Überleben (rPFS)“ wurde vom Zeitpunkt der Randomisierung bis zum frühesten Nachweis einer radiographischen Progression oder Todeseintritt jedweder Ursache erfasst. Radiographische Progression war dabei definiert als das Auftreten entweder einer Knochenerkrankung nach den Konsensus-Leitlinien der *Prostate Cancer Clinical Trial Working Group 2* (PCWG) und/oder einer Krankheitsprogression im Weichteilgewebe, welche durch die *Response Evaluation Criteria in Solid Tumors* (RECIST),Version 1.1 definiert war. Falls es keinen berechtigten früheren Anlass gab, wurde die erste Untersuchung zur Beurteilung der radiographischen Progression in Woche 13 durchgeführt. Eine Progression war definiert als das Auftreten von zwei oder mehr neuen Knochenläsionen im Knochenscan oder einer mittels Computertomographie (CT) oder Magnetresonanztomographie (MRI) festgestellten Krankheitsprogression im Weichteilgewebe nach RECIST Version 1.1. Wurde in Woche 13 eine Progression festgestellt, musste dieser Befund nach mindestens 6 Wochen mittels einer zweiten Untersuchung bestätigt werden. Weitere Untersuchungen zur radiographischen Progression fanden in Woche 25 und danach alle 12 Wochen bis zum Eintritt des Todes statt. **Sensitivitätsanalyse:** Sensitivitätsanalysen wurden auf Basis der abgeleiteten (*derived*) Definition der Progression des Weichteilgewebes basierend auf RECIST v1.1 Kriterien und den modifizierten Zensierungsregeln durchgeführt. (siehe auch Abschnitt 4.2.5.4).

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 4-29: Bewertung des Verzerrungspotenzials für Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial wurde daher für den Endpunkt „rPFS“ mit „niedrig“ bewertet.

Stellen Sie die Ergebnisse für den Endpunkt „radiographisches 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 4-30: Ergebnisse für Endpunkt „Radiographisches progressionsfreies Überleben (rPFS)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**			**Placebo**			**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**N[a]**	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**Hazard Ratio [95 %-KI] p-Wert**
Radiographisches progressionsfreies Überleben							
AFFIRM	800	524 (65,5)	8,3 [8,2; 9,4]	399	337 (84,5)	2,9 [2,8; 3,4]	0,40 [0,35; 0,47] p < 0,0001
a: Anzahl Patienten in der Auswertung KI: Konfidenzintervall Alle Angaben beziehen sich auf die Analyse der ITT-Population.							

Das mediane, „radiographisch progressionsfreie Überleben“ betrug bei Patienten der Enzalutamid-Gruppe 8,3 Monate (95 %-KI [8,2; 9,4]) im Vergleich zu 2,9 Monaten (95 %-KI [2,8; 3,4]) in der Placebogruppe. Das „radiographische progressionsfreie Überleben“ war somit für Patienten unter Enzalutamid-Therapie um 5,4 Monate länger als für Patienten unter Placebo-Therapie. Das Hazard Ratio lag mit einem Wert von 0,40 (95 %-KI [0,35; 0,47], p < 0,0001) statistisch signifikant zugunsten der Enzalutamid-Gruppe.

Die Sensitivitätsanalyse für das radiographisch progressionsfreie Überleben zeigte folgende Ergebnisse (siehe auch Abschnitt 4.2.5.4):

- Der Median in Monaten bis zum Auftreten einer radiographischen Progression betrug bei Beurteilung der Progression durch den Investigator und modifizierte Annahmen zur Zensierung: Patienten der Enzalutamid-Gruppe 8,3 Monate (95 %-KI [8,2; 9,0]) im Vergleich zu 2,8 Monaten (95 %-KI [2,8; 3,2]) in der Placebogruppe. Das Hazard Ratio lag bei einem Wert von HR = 0,41 (95 %-KI [0,35; 0,47], p < 0,0001) zugunsten der Enzalutamid-Gruppe.

- Beurteilung als *derived* Progression und primäre Annahmen für die Zensierung: HR = 0;40 (95 %-KI [0,35; 0,47]).

- Beurteilung als *derived* Progression und modifizierte Annahmen zur Zensierung: HR = 0,41 (95 % KI [0,35; 0,47])

Die Ergebnisse der Sensitivitätsanalysen bestätigten somit die Ergebnisse der Hauptanalysen.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.3 Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ - RCT

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

Tabelle 4-31: Operationalisierung von Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die „Zeit bis zur ersten skelettbezogenen Komplikation“ war definiert als Zeitraum vom Zeitpunkt der Randomisierung bis zum ersten Auftreten einer skelettbezogenen Komplikation. Eine skelettbezogene Komplikation war definiert als eine Bestrahlung des Knochens oder ein chirurgischer Eingriff am Knochen, eine pathologische Knochenfraktur, eine Rückenmarkkompression oder eine Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen. Das Auftreten von skelettbezogenen Komplikationen wurde regelmäßig bei den geplanten Visiten (Studienvisiten, *Long-Term Follow-Up* Visiten, *Safety Follow Up* Visiten) und *Survival Follow-Up* Telefonanrufen erfasst. Die Beurteilung erfolgte dabei durch den Untersucher oder Studienkoordinator, welcher den Patienten, die Familie des Patienten oder dessen Hausarzt nach dem Auftreten von skelettbezogenen Komplikationen befragte.

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 4-32: Bewertung des Verzerrungspotenzials für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial

beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial wurde daher für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ mit „niedrig“ bewertet.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-33: Ergebnisse für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ aus RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**			**Placebo**			**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**N[a]**	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**Hazard Ratio [95 %-KI] p-Wert**
Zeit bis zur ersten skelettbezogenen Komplikation							
AFFIRM	800	287 (35,9)	16,7 [14,6; 19,1]	399	161 (40,4)	13,3 [9,9; n.e.]	0,69 [0,57; 0,84] p = 0,0001
Zeit bis zur Bestrahlung des Knochens							
AFFIRM	800	176 (22,0)	n.e. (18,7; n.e.)	399	101 (25,3)	n.e. (18,2; n.e.)	0,71 [0,55; 0,91] p = 0,0058
Zeit bis zur Knochenoperation							
AFFIRM	800	6 (0,8)	n.e. [n.e.; n.e.]	399	1 (0,3)	n.e. [n.e.; n.e.]	2,49 [0,30; 20,79] p = 0,38
Zeit bis zur pathologischen Knochenfraktur							
AFFIRM	800	31 (3,9)	n.e. [n.e.; n.e.]	399	16 (4,0)	n.e. [n.e.; n.e.]	0,85 [0,47; 1,56] p = 0,60
Zeit bis zur Rückenmarkkompression							
AFFIRM	800	66 (8,3)	21,6 [21,6; n.e.]	399	29 (7,3)	n.e. [n.e.; n.e.]	1,00 [0,64; 1,55] p = 0,99
Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen							
AFFIRM	800	20 (2,5)	n.e. [n.e.; n.e.]	399	16 (4,0)	n.e. [n.e.; n.e.]	0,54 [0,28; 1,04] p = 0,06

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
n.e.: nicht erreicht
Alle Angaben beziehen sich auf die Auswertung der ITT-Population.

Die mediane „Zeit bis zur ersten skelettbezogenen Komplikation“ betrug bei Patienten der Enzalutamid-Gruppe 16,7 Monate (95 %-KI [14,6; 19,1]) im Vergleich zu 13,3 Monaten (95 %-KI [9,9; n.e.]) in der Placebogruppe. Bei Patienten unter Enzalutamid-Therapie trat die erste skelettbezogene Komplikation demnach 3,4 Monate später auf als bei Patienten, die ein Placebo erhielten. Das Hazard Ratio war mit einem Wert von 0,69 (95 %-KI [0,57; 0,84], p = 0,0001) statistisch signifikant zugunsten der Enzalutamid-Gruppe.

Die mediane Zeit bis zur Bestrahlung des Knochens als erstem skelettbezogenen Ereignis war in beiden Behandlungsgruppen zum Zeitpunkt der Analyse noch nicht erreicht. Das Hazard Ratio war mit einem Wert von 0,71 (95 %-KI [0,55; 0,91], p = 0,0058) statistisch signifikant zugunsten der Enzalutamid-Gruppe.

Für die skelettbezogenen Ereignisse Zeit bis zur Knochenoperation (Hazard Ratio = 2,49 (95 %-KI [0,30; 20,79], p = 0,38), Zeit bis zur pathologischen Knochenfraktur (Hazard Ratio = 0,85 (95 %-KI [0,47; 1,56], p = 0,60), Zeit bis zur Rückenmarkkompression (Hazard Ratio = 1,00 (95 %-KI [0,64; 1,55], p = 0,99) und Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen (Harard Ratio = 0,54 (95 %-KI [0,28; 1,04], p = 0,06) ergaben sich keine statistisch signifikanten Unterschiede zwischen den beiden Therapiearmen.

Ergebnisse der Sensitivitätsanalyse für diesen Endpunkt bestätigen die Ergebnisse der Hauptanalyse: Für diese Sensitivitätsanalyse wurden Patienten, deren Therapie sich aufgrund von Knochenschmerzen bei Knochenmetastasen änderte, nicht mitgezählt. Der Vergleich für die beiden Therapiegruppen wurde hier nur für das Ereignis „Zeit bis zur ersten skelettbezogenen Komplikation“ berechnet. Die mediane „Zeit bis zur ersten skelettbezogenen Komplikation“ betrug bei Patienten der Enzalutamid-Gruppe 17,4 Monate (95 %-KI [15,7, n.e.]) im Vergleich zu 15,0 Monaten (95 %-KI [12,4, n.e.]) in der Placebogruppe. Bei Patienten unter Enzalutamid-Therapie trat die erste skelettbezogene Komplikation demnach 2,4 Monate später auf als bei Patienten, die ein Placebo erhielten. Das Hazard Ratio war mit einem Wert von 0,74 (95 %-KI [0,61; 0,91], p = 0,004) zugunsten der Enzalutamid-Gruppe. (siehe auch Tabelle 4-34)

Tabelle 4-34: Ergebnisse der Sensitivitätsanalyse für Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ aus RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**			**Placebo**			**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**N[a]**	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**Hazard Ratio [95 %-KI] p-Wert**
Zeit bis zur ersten skelettbezogenen Komplikation[b]							
AFFIRM	800	274 (34,3)	17,4 [15,7; n.e.]	399	147 (36,8)	15,0 [12,4; n.e.]	0,74 [0,61; 0,91] p = 0,004

In der Sensitivitätsanalyse wurden Patienten mit einer Änderung der antineoplastischen Therapie, die aufgrund der Notwendigkeit der Behandlung von Knochenschmerzen in Folge von Knochenmetastasen erfolgte, nicht berücksichtigt.

a: Anzahl Patienten in der Auswertung

b: Basiert auf dem ersten stattfindenden Ereignis. Unterschiedliche skelettbezogene Komplikationen werden beim Patienten zusammengezählt, es wird aber pro Komplikation jeweils nur ein Ereignis ausgewertet. Patienten, von denen nicht bekannt war, dass sie ein skelettbezogenes Ereignis hatten, wurden am Datum der letzten Bewertung dahingehend zensiert, dass es keinen Hinweis auf ein skelettbezogenens Ereignis gab.

n.e.: nicht erreicht

KI: Konfidenzintervall

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.4 Endpunkt „Schmerz“ - RCT

4.3.1.3.1.4.1 Endpunkt „Schmerzlinderung“ - RCT

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

Tabelle 4-35: Operationalisierung von Endpunkt „Schmerzlinderung“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die Auswertung des Endpunkts „Schmerzlinderung“ erfolgte basierend auf Frage #3 des *Brief Pain Inventory – short form Questionnaire (BPI)*, welche bei Studienbeginn und in Woche 13 von den Patienten in einem Schmerztagebuch erfasst wurde. Schmerz konnte auf einer Skala von 0 bis 10 bewertet werden, wobei 0 Punkte Schmerzfreiheit und 10 Punkte die höchste Schmerzstufe bedeuteten, und sich die Angabe der Patienten nach ihren schlimmsten erfahrenen Schmerz der vergangenen 24 Stunden richtete. „Schmerzlinderung“ war definiert als eine ≥ 30 %ige Reduktion des Schmerzempfindens in Woche 13 verglichen mit Studienbeginn, wobei der Anstieg des Analgetika Verbrauchs maximal 30 % betragen durfte. Für die Auswertung des Endpunkts kamen nur Patienten infrage, welche zu Studienbeginn Knochenmetastasen aufwiesen, das Schmerztagebuch an mindestens 4 von 7 Tagen während der *run-in phase* zu Studienbeginn ausgefüllt hatten, und dabei einen konstanten Analgetika-Verbrauch (≤ 30 % Variation im Analgetika-Verbrauch) und eine konstante und ausreichend hohe Schmerzbelastung aufwiesen. D.h. die tägliche Angabe der Schmerzpunktzahl durfte um nicht mehr als 2 Punkte variieren und die durchschnittliche Schmerzpunktzahl musste mindestens 4 Punkte betragen.

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 4-36: Bewertung des Verzerrungspotenzials für Endpunkt „Schmerzlinderung“ 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
AFFIRM	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Da sich nur insgesamt 64 Patienten in der Auswertungspopulation befanden, wurde das ITT-Prinzip hier nicht adäquat umgesetzt und somit das Verzerrungspotenzial für den Endpunkt „Schmerzlinderung“ mit „hoch“ bewertet. Es wurden nur Patienten in die Auswertung aufgenommen, die bereits zu Studienbeginn eine ausreichend hohe Schmerzbelastung und damit ein entsprechendes Potential zur Schmerzlinderung aufwiesen (die durchschnittliche Schmerzpunktzahl zu Studienbeginn musste mindesten 4 Punkte betragen). Sonstige das Verzerrungspotenzial beeinflussende Faktoren wurden nicht identifiziert.

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

Tabelle 4-37: Ergebnisse für den Endpunkt „Schmerzlinderung“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI][b] p-Wert[b]**
Schmerzlinderung					
AFFIRM	49	22 (44,9)	15	1 (6,7)	6,73 [0,99; 45,88] p > 0,05

a: Anzahl Patienten in der Auswertung
b: Berechnet nach Deeks 2010 [Deeks 2010]
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzlinderung

In der AFFIRM-Studie erfuhren unter Enzalutamid-Therapie 44,9 % (22/49) der Patienten eine Schmerzlinderung, während dies bei nur 6,7 % (1/15) der Patienten in der Placebogruppe der Fall war. Das Relative Risiko betrug 6,73 (95 %-KI [0,99; 45,88], p > 0,05), und zeigte somit einen deutlichen, wenn auch statistisch nicht signifikanten Vorteil für Patienten innerhalb der Enzalutamid-Gruppe.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.4.2 Endpunkt „Schmerzprogressionsrate“ – RCT

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

Tabelle 4-38: Operationalisierung von Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die Erhebung des Endpunkts „Schmerzprogressionsrate“ erfolgte basierend auf der Frage #3 des BPI – *short form*, welche von den Patienten im Rahmen eines Schmerztagebuchs beantwortet wurde. Analog zum Endpunkt „Schmerzlinderung“ wurde Schmerz auf einer Skala von 0 bis 10 bewertet, wobei 0 Punkte Schmerzfreiheit und 10 Punkte die höchste Schmerzstufe bedeuteten, und sich die Angabe der Patienten nach ihren schlimmsten erfahrenen Schmerz der vergangenen 24 Stunden richtete. In die Auswertung gingen nur Patienten ein, welche an mindestens 4 von 7 Tagen während der *run-in phase* zu Studienbeginn und an den 7 Tagen vor der Woche 13 Visite gültige Eintragungen im Schmerztagebuch hatten. „Schmerzprogression“ war definiert als jeglicher Anstieg der mittleren Schmerzpunktzahl zwischen Studienbeginn und Woche 13. Die „Schmerzprogressionsrate“ war definiert als die Anzahl der Patienten, die in Woche 13 im Vergleich zu Studienbeginn eine Schmerzprogression erfahren hatten.

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 4-39: Bewertung des Verzerrungspotenzials für Endpunkt „Schmerzprogressionsrate“ 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
AFFIRM	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet.

Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Die Auswertungspopulation für die „Schmerzprogressionsrate“ umfasste 884 Patienten. Somit befinden sich lediglich 74 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt und das Verzerrungspotenzial für den Endpunkt „Schmerzprogressionsrate“ wurde mit „hoch“ bewertet. Sonstige das Verzerrungspotenzial beeinflussende Faktoren wurden nicht identifiziert.

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

Tabelle 4-40: Ergebnisse für den Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel (*Responder*-Analyse)

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Schmerzprogressionsrate					
AFFIRM	625	174 (27,8)	259	101 (39,0)	0,71 [0,59; 0,87] p = 0,0008

a: Anzahl Patienten in der Auswertung

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die „Schmerzprogressionsrate“.

Die Anzahl Patienten in dieser Analyse unterscheidet sich von der Anzahl Patienten, die für die Veränderung der „Schmerzprogressionsrate“ gegenüber Studienbeginn zur Verfügung stand. Für die *Responder*-Analyse wurden Patienten mit fehlenden Werten zu Studienbeginn als „nicht-*Responder*“ gewertet, für die Analyse zur Veränderung gegenüber Studienbeginn wurden sie als „fehlend“ bewertet und nicht in die Analyse einbezogen.

In der AFFIRM-Studie trat bei 27,8 % der Patienten (174/625), die Enzalutamid bekamen, eine Schmerzprogression auf, wohingegen 39,0 % der Patienten (101/259), welche Placebo bekamen, eine Schmerzprogression erlitten. Der Unterschied zwischen den beiden Therapiegruppen lag somit bei 11,2 Prozentpunkten, wobei das Relative Risiko mit einem Wert von 0,71 (95 %-KI [0,59; 0,87], p = 0,0008) statistisch signifikant zugunsten der Enzalutamid-Gruppe ausfiel.

Tabelle 4-41: Ergebnisse für Endpunkt „Schmerzprogressionsrate“ in RCT mit dem zu bewertenden Arzneimittel (Veränderung gegenüber Studienbeginn)

Studie	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende[b]**	**Delta**	**N[a]**	**Studienbeginn**	**Ende[b]**	**Delta**		
Schmerzprogressionsrate										
AFFIRM	620	2,41 (2,15)	2,05 (2,16)	-0,35 (2,02)	259	2,23 (2,15)	2,87 (2,51)	0,64 (2,16)	-0,48 [-0,62; -0,34]	p < 0,0001

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

b: Woche 13

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die „Schmerzprogressionsrate“.

Die Anzahl Patienten in dieser Analyse unterscheidet sich von der Anzahl Patienten, die für die *Responder*-Analyse zur „Schmerzprogressionsrate“ zur Verfügung stand. Für die *Responder*-Analyse wurden Patienten mit fehlenden Werten zu Studienbeginn als „nicht-*Responder*“ gewertet, für die Analyse zur Veränderung gegenüber Studienbeginn wurden sie als „fehlend“ bewertet und nicht in die Analyse einbezogen.

Während die Schmerzpunktzahl bei Patienten unter Enzalutamid-Therapie zwischen Studienbeginn und Woche 13 im Mittel um 0,35 (SD: 2,02) Punkte abnahm, erhöhte sie sich bei Patienten in der Placebogruppe im Mittel um 0,64 (SD: 2,16) Punkte. Die SMD zwischen den beiden Therapiegruppen betrug -0,48 [-0,62; -0,34] und fiel mit einem p-Wert von $p < 0,0001$ statistisch signifikant zugunsten der Enzalutamid-Gruppe aus.

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

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.4.3 Endpunkt „Zeit bis zur Schmerzprogression“ – RCT

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

Tabelle 4-42: Operationalisierung von Endpunkt „Zeit bis zur Schmerzprogression“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Der Endpunkt „Zeit bis zur Schmerzprogression“ wurde bei allen randomisierten Patienten mittels des *Functional Assessment of Cancer Therapy-Prostate* (FACT-P)-Fragebogens erhoben. Die Aussage „Ich habe Schmerzen“ wurde von den Patienten mittels einer Skala von 0 bis 4 Punkten (0 = keine Schmerzen bis 4 = sehr große Schmerzen) bewertet. Die letzte verfügbare Erhebung von Schmerzen mittels des FACT-P-Fragebogens fand zu Studienbeginn und anschließend bei jeder regulären Visite beginnend mit Woche 13 statt. Die Patienten bekamen dazu den Fragebogen zum Selbstausfüllen mit nach Hause, und wurden angewiesen Angaben zu den jeweils 7 Tagen vor der nächsten klinischen Visite zu machen. „Schmerzprogression“ war definiert als Anstieg der Schmerzbewertung gegenüber Studienbeginn, wobei dieser Anstieg durch eine Folgebewertung mindestens 3 Wochen später bestätigt werden musste. Die „Zeit bis zur Schmerzprogression“ war definiert als die Zeit in Monaten von Studienbeginn bis zu einem Anstieg in der Schmerzbewertung.

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 4-43: Bewertung des Verzerrungspotenzials für Endpunkt „Zeit bis zur Schmerzprogression“ 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte

Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Zeit bis zur Schmerzprogression“ wurde daher mit „niedrig“ bewertet.

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 4-44: Ergebnisse für Endpunkt „Zeit bis zur Schmerzprogression“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**			**Placebo**			**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**N[a]**	**Patienten mit Ereignis n (%)**	**Median [95 %-KI] Monate**	**Hazard Ratio [95 %-KI] p-Wert**
Zeit bis zur Schmerzprogression							
AFFIRM	800	141 (17,6)	n.e. [n.e; n.e.]	399	57 (14,3)	13,8 [13,8; n.e.]	0,56 [0,41; 0,78] p = 0,0004
a: Anzahl Patienten in der Auswertung KI: Konfidenzintervall n.e.: nicht erreicht Alle Angaben beziehen sich auf die Analyse Auswertungspopulation für die „Zeit bis zur Schmerzprogression“.							

Die mediane „Zeit bis zur Schmerzprogression“ betrug in der Placebogruppe 13,8 Monate (95 %-KI [13,8; n.e.]). Da in der Enzalutamid-Gruppe zum Zeitpunkt der Datenerhebung weniger als 50 % der betrachteten Patienten ein Ereignis hatten, konnte hier noch kein Median berechnet werden. Das Hazard Ratio lag bei 0,56 (95 %-KI [0,41; 0,78], p = 0,0004) und fiel statistisch signifikant zugunsten der Enzalutamid-Gruppe aus.

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

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.5 Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ - RCT

In der vorliegenden Therapiesituation stellten insbesondere durch Metastasierung hervorgerufene „Lähmungen und die lähmungsbedingte Harninkontinenz“ patientenrelevante Morbiditätsendpunkte dar. Die Aufnahme dieser Endpunkte wurde vom G-BA im Beratungsgespräch explizit gefordert [Gemeinsamer Bundesausschuss (G-BA) 2012], deshalb wird im Rahmen des Dossiers gesondert auf diese Punkte eingegangen. Da die separate Erhebung dieser Endpunkte im Rahmen der AFFIRM-Studie allerdings nicht vorgesehen war, basiert die folgende Darstellung lediglich auf den zur Verfügung stehenden Angaben von Auswertungen von Einzelfällen.

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

Tabelle 4-45: Operationalisierung von Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die separate Erhebung des Endpunkts „Lähmungen und lähmungsbedingte Harninkontinenz“ war im Rahmen der AFFIRM-Studie nicht vorgesehen, weshalb auch keine allgemeine Operationalisierung vorgenommen werden kann. Die Einzelereignisse sind in den folgenden Abschnitten separat gelistet. Sie wurden auf der Basis der Einträge für unerwünschte Ereignisse (durch den Arzt auf dem CRF) und deren Kodierung mittels MedrA identifiziert. Harninkontinenz (*urinary incontinence)* : Als mögliche Ursachen für die Harninkontinenz wurden folgende Ereignisse bei jedem einzelnen Patienten analysiert*:* • Prostatektomie *(Prostatectomy)* • Andere operative Eingriffe *(Other surgery)* • Strahlentherapie *(Radiotherapy)* • Rückenmarkkompression *(Spinal Cord Compression)* • Urologische Krankengeschichte *(Urological medical history)* • Urologische unerwünschte Ereignisse *(Urological AEs)* Der Zeitpunkt des Auftretens dieser Ereignisse wurde in Zusammenhang mit dem Auftreten einer Harninkontinenz gesetzt und für jeden Einzelfall ausgewertet. Lähmungen: Um mögliche Lähmungen zu identifizieren, wurden die UEs, die mittels MedrA kodiert wurden, nach „*paralysis, parese, palsy*“ in verschiedenen Kombinationen und Schreibweisen, durchsucht. Alle Einzelereignisse, sowie deren Schweregrad werden hier im Dossier dargestellt.

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 4-46: Bewertung des Verzerrungspotenzials für Endpunkt „Lähmungen und lähmungsbedingte Harninkontinenz“ 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
AFFIRM	niedrig	ja	a)	nein	ja	„hoch“

a) Es erfolgte keine statistische Auswertung, bei der deskriptiven Darstellung wurden alle randomisierten Patienten berücksichtigt.

Begründen Sie für jede Studie die abschließende Einschätzung.

Da für die Ereignisse „Lähmungen und lähmungsbedingte Harninkontinenz“ keine statistische Auswertung durchgeführt werden konnte, wurde das Verzerrungspotential nur theoretisch bewertet. Aufgrund der geringen Ereigniszahl sowie der Heterogenität der Einzelereignisse liegen verschiedene weitere Einflussfaktoren mit Verzerrungspotential vor. Im Fall einer Auswertung müsste von einem hohen Verzerrungspotential ausgegangen werden. Dementsprechend wurden aus diesem Endpunkt auch keine Aussagen zum Zusatznutzen abgeleitet.

„Lähmungen“

Da die Anzahl der Patienten, welche von Lähmungserscheinungen betroffen waren, sehr klein war, und zudem unter medizinischen Aspekten in verschiedene Kategorien fielen, wird hier auf eine statistische Auswertung zugunsten einer tabellarischen bzw. deskriptiven Darstellung verzichtet. Da die jeweilige Ursache der Lähmungserscheinung nicht eindeutig eruiert werden kann, d.h. in der Regel unklar ist, ob die Lähmung durch eine Metastasierung hervorgerufen wurde, werden in Tabelle 4-47 alle während der Behandlungsperiode aufgetretenen unerwünschten Ereignisse, welche mit Lähmungen assoziiert sind, aufgelistet. Die Erhebung der Ereignisse erfolgte wie in der Operationalisierung für unerwünschte Ereignisse, siehe Tabelle 4-57, beschrieben.

Tabelle 4-47: Aufstellung von mit Lähmungserscheinungen assoziierten unerwünschten Ereignissen, unterschieden nach Schweregrad des Ereignisses (nach NCI-CTCAE, Version 4.0)

Lähmungserscheinung	**Enzalutamid (N[a] = 800)**		**Placebo (N[a] = 399)**	
	Grad ≥ 3	**Alle**	**Grad ≥ 3**	**Alle**
Lähmung der Zunge (*Tongue paralysis*)	0	1	0	1
Hirnnervenlähmung (*Cranial nerve paralysis*)	0	1	0	0
Trochlearislähmung (*IVth nerve paralysis*)	0	1	0	0
Abduzenslähmung (*VIth nerve paralysis*)	1	1	0	0
Paraplegie (*Paraplegia*)	0	0	1	2
Monoplegie (*Monoplegia*)	1	1	0	0
Fazialislähmung / -parese (*Facial paresis, Facial palsy*)	0	2	1	3
Poreneuslähmung (*Peroneal nerve palsy*)	0	1	0	1
Monoparese (*Monoparesis*)	1	4	0	0
Hypoglossusparese (*Hypoglossal nerve paresis*)	0	1	0	0
Paraparese (*Paraparesis*)	0	1	0	0
Summe	3 (0,38 %)	14 (1,75 %)	2 (0,50 %)	7 (1,75 %)
a: Anzahl Patienten in der Auswertung				

In Summe traten in der Enzalutamid-Gruppe 14 Fälle von Lähmungserscheinungen auf, von welchen 3 als ein Ereignis mit einem Schweregrad von ≥ 3 klassifiziert wurden. In der Placebogruppe wurden insgesamt 7 Fälle von Lähmungserscheinungen registriert, von welchen 2 einen Schweregrad von ≥ 3 aufwiesen. In beiden Gruppen traten demnach Lähmungserscheinungen gleich häufig auf (bei 1,75 % der Patienten). Da die Erhebung dieses Endpunkts vom G-BA gewünscht [Gemeinsamer Bundesausschuss (G-BA) 2012], aber im Rahmen der AFFIRM-Studie nicht vorgesehen war, kann basierend auf der zur Verfügung stehenden Datengrundlage nicht beurteilt werden, inwieweit diese Lähmungserscheinungen auf eine Metastasierung zurückzuführen sind oder andere Gründe haben.

„Lähmungsbedingte Harninkontinenz“

Im Rahmen der Erhebung von unerwünschten Ereignissen, welche in Tabelle 4-57 beschrieben ist, wurden in der Studie AFFIRM insgesamt 40 Fälle von Harninkontinenz berichtet. Von diesen Fällen traten 29 (3,6 %) in der Enzalutamid-Gruppe und 11 (2,8 %) in der Placebogruppe auf. Allerdings wurde bei keinem dieser Ereignisse eine Lähmungserscheinung als Ursache genannt. Sechs Patienten, die sich alle in der Enzalutamid-Gruppe befanden, erlitten ebenfalls eine Rückenmarkkompression, welche grundsätzlich aufgrund einer fortschreitenden Metastasierung eintreten kann. Bei zwei dieser sechs Patienten bestand die Harninkontinenz allerdings bereits vor der Rückenmarkkompression, insofern kommt dieses Ereignis nicht als Ursache infrage. Von den verbleibenden vier Patienten traten bei zwei die beiden Ereignisse Rückenmarkkompression und Harninkontinenz am selben Tag auf, bei den beiden anderen stellte sich die Harninkontinenz etwa einen bzw. zwei Monate nach der Rückenmarkkompression ein und steht somit nicht in einem unmittelbaren Zusammenhang. Insofern erscheint bei zwei Patienten eine fortschreitende Metastasierung als Ursache für die Harninkontinenz möglich, bei zwei weiteren Patienten ist dies nicht auszuschließen.

4.3.1.3.1.6 Endpunkt „Gesundheitsbezogene Lebensqualität“ - RCT

4.3.1.3.1.6.1 Endpunkt „Gesundheitsbezogene Lebensqualität“ nach FACT-P - RCT

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

Tabelle 4-48: Operationalisierung von Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die Auswertung der „gesundheitsbezogenen Lebensqualität“ erfolgte basierend auf der Version 4 des FACT-P-Fragebogens. Der FACT-P ist ein multidimensionaler, selbsterklärender Fragebogen, welcher 39 Fragen, aufgeteilt in 5 Domänen enthält. Auf einer Skala von 0 bis 4 (0 = trifft überhaupt nicht zu, 4 = trifft sehr zu) bewerten die Patienten dabei, inwieweit die jeweilige Aussagen auf sie zutreffen. Die Fragen der ersten vier Domänen richten sich generell an Patienten mit einer Krebserkrankung und beinhalten Fragen zu den Bereichen körperliches Wohlbefinden, Verhältnis zu Freunden, Bekannten und Familie, seelisches Wohlbefinden und Funktionsfähigkeit. Zusammengenommen ergeben diese vier Domänen den FACT-G-Fragebogen (*Functional Assessment of Cancer Therapy - General*) und die Summe der vergebenen Punkte ergibt die Gesamtpunktzahl für den FACT-G. Die Fragen der fünften Domäne, der Prostatakarzinom-Subskala, richten sich speziell an Prostatakarzinom-Patienten. Die Summe der vergebenen Punkte aller fünf Domänen ergibt die Gesamtpunktzahl für den FACT-P. Die Erhebung der Lebensqualität mittels FACT-P wurde zu Studienbeginn und im Folgenden bei jeder regulären Visite beginnend mit Woche 13 durchgeführt. Eine Verbesserung der Lebensqualität (*Responder*) war definiert als eine Verbesserung von 10 Punkten in der FACT-P Gesamtpunktzahl, welche bei 2 aufeinanderfolgenden Messungen, die mindestens 3 Wochen auseinander lagen, im Vergleich zu Studienbeginn erzielt werden musste. Angaben zur Validität des Fragebogens und der *Clinically Meaningful Changes* (CMC) finden sich in Abschnitt 4.2.5.2.

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 Endpunkt „gesundheitsbezogene Lebensqualität nach FACT-P“ 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
AFFIRM	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Die Auswertungspopulation für FACT-P umfasste 871 Patienten, somit befinden sich, unabhängig von der Auswertung des Endpunkts als *Responder*-Analyse oder als Veränderung gegenüber Studienbeginn maximal 73 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt und das Verzerrungspotenzial für den Endpunkt „gesundheitsbezogene Lebensqualität nach FACT-P" wurde mit „hoch" bewertet. Sonstige das Verzerrungspotenzial beeinflussende Faktoren wurden nicht identifiziert.

Stellen Sie die Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P" 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 den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P" in RCT mit dem zu bewertenden Arzneimittel (*Responder*-Analyse)

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Lebensqualität					
AFFIRM	630	264 (41,9)	241	36 (14,9)	2,81 [2,05; 3,84] p < 0,0001

a: Anzahl Patienten in der Auswertung

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für FACT-P

Anmerkung: Die Angaben weichen von den im CSR berichteten Ergebnissen ab (Tabelle 11.4.1.4.1-1), da die Auswertung der Fragebögen zunächst nicht konform zu den Auswerteregeln vorgenommen worden war.

Die „gesundheitsbezogene Lebensqualität nach FACT-P" verbesserte sich bei 41,9 % der Patienten (264/630) unter Enzalutamid-Therapie, wohingegen sie sich bei nur 14,9 % der Patienten (36/241), die Placebo bekamen, verbesserte. Der Unterschied in der Ansprechrate zwischen den Behandlungsgruppen betrug 27 Prozentpunkte und lag mit einem Relativen Risiko von 2,81 (95 %-KI [2,05; 3,84], p < 0,0001) statistisch signifikant zugunsten der Enzalutamid-Gruppe.

Tabelle 4-51: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ in RCT mit dem zu bewertenden Arzneimittel (Veränderung gegenüber Studienbeginn)

Studie	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studien beginn**	**Ende**	**Delta**	**N[a]**	**Studien beginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-P Gesamtpunktzahl - Woche 49										
AFFIRM	221	113,28 (19,18)	111,34 (22,56)	-1,57 (16,78)	21	117,66 (21,17)	113,80 (21,70)	-4,69 (16,52)	0,22 [-0,23; 0,67]	p = 0,33
Gesundheitsbezogene Lebensqualität nach FACT-P Gesamtpunktzahl - letzte verfügbare Erhebung										
AFFIRM	613	108,88 (21,17)	103,75 (23,84)	-5,07 (17,64)	241	110,21 (20,95)	97,89 (23,45)	-12,29 (17,48)	0,42 [0,27; 0,56]	p < 0,0001
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl - Woche 49										
AFFIRM	224	81,37 (13,94)	79,45 (16,53)	-1,53 (12,66)	22	85,04 (13,96)	81,19 (14,65)	-3,80 (10,36)	0,19 [-0,25; 0,63]	p = 0,40
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl - letzte verfügbare Erhebung										
AFFIRM	623	78,40 (15,35)	74,42 (17,26)	-4,00 (12,82)	249	79,14 (14,74)	70,53 (16,93)	-8,56 (12,98)	0,36 [0,21; 0,50]	p < 0,0001
Gesundheitsbezogene Lebensqualität nach FACT-P Körperliches Wohlbefinden - Woche 49										
AFFIRM	227	22,58 (4,47)	21,43 (5,41)	-1,09 (4,72)	23	22,23 (5,50)	21,60 (5,00)	-0,63 (4,64)	-0,07 [-0,50; 0,36]	p = 0,76
Gesundheitsbezogene Lebensqualität nach FACT-P Körperliches Wohlbefinden - letzte verfügbare Erhebung										
AFFIRM	643	21,39 (5,34)	19,17 (6,21)	-2,21 (5,26)	256	21,35 (5,43)	17,77 (6,49)	-3,57 (5,70)	0,26 [0,11; 0,40]	p = 0,0006
Gesundheitsbezogene Lebensqualität nach FACT-P Verhältnis zu Freunden, Bekannten und Familie - Woche 49										
AFFIRM	227	22,15 (5,20)	22,01 (4,85)	-0,14 (5,03)	23	23,80 (3,54)	22,41 (3,35)	-1,38 (2,36)	0,23 [-0,20; 0,66]	p = 0,29
Gesundheitsbezogene Lebensqualität nach FACT-P Verhältnis zu Freunden, Bekannten und Familie - letzte verfügbare Erhebung										

Studie	Enzalutamid Mittelwert (SD)				Placebo Mittelwert (SD)				SMD [95 %-KI]	p-Wert
	N[a]	Studien beginn	Ende	Delta	N[a]	Studien beginn	Ende	Delta		
AFFIRM	645	21,99 (4,78)	21,97 (4,60)	-0,08 (4,23)	254	21,97 (4,68)	21,55 (4,68)	-0,34 (3,84)	0,06 [-0,08; 0,21]	p = 0,41
Gesundheitsbezogene Lebensqualität nach FACT-P Seelisches Wohlbefinden - Woche 49										
AFFIRM	227	18,02 (4,14)	18,45 (4,27)	0,45 (3,35)	22	19,18 (4,73)	18,83 (4,33)	-0,27 (5,06)	0,19 [-0,25; 0,63]	p = 0,40
Gesundheitsbezogene Lebensqualität nach FACT-P Seelisches Wohlbefinden – letzte verfügbare Erhebung										
AFFIRM	642	17,39 (4,51)	17,13 (4,75)	-0,20 (3,68)	252	17,39 (4,43)	15,84 (4,98)	-1,48 (4,54)	0,33 [0,18; 0,47]	p < 0,0001
Gesundheitsbezogene Lebensqualität nach FACT-P Funktionelles Wohlbefinden - Woche 49										
AFFIRM	227	18,46 (5,63)	17,57 (6,10)	-0,80 (5,46)	22	19,46 (5,79)	18,35 (5,30)	-1,14 (5,09)	0,09 [-0,35; 0,53]	p = 0,70
Gesundheitsbezogene Lebensqualität nach FACT-P Funktionelles Wohlbefinden– letzte verfügbare Erhebung										
AFFIRM	641	17,64 (5,80)	16,09 (6,44)	-1,52 (5,32)	250	18,30 (5,39)	15,30 (6,25)	-3,07 (4,80)	0,30 [0,15; 0,45]	p < 0,0001
Gesundheitsbezogene Lebensqualität nach FACT-P Prostatakarzinom-Subskala - Woche 49										
AFFIRM	226	32,06 (7,17)	31,92 (7,44)	-0,05 (6,30)	22	32,27 (7,84)	32,61 (8,42)	0,01 (8,04)	0,05 [-0,39; 0,49]	p = 0,83
Gesundheitsbezogene Lebensqualität nach FACT-P Prostatakarzinom-Subskala – letzte verfügbare Erhebung										
AFFIRM	642	30,55 (7,43)	29,28 (7,94)	-1,16 (6,67)	248	30,88 (7,60)	27,43 (8,02)	-3,41 (6,33)	0,34 [0,20; 0,49]	p < 0,0001

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Die Angaben beziehen sich auf die für die jeweilige Analyse der innerhalb der Auswertungspopulation für FACT-P infrage kommenden Patienten.

Für die Veränderung der „gesundheitsbezogenen Lebensqualität nach der FACT-P“ sowie der FACT-G Gesamtpunktzahl ergaben sich bei Betrachtung der letzten verfügbaren Erhebung gegenüber Studienbeginn die folgenden Unterschiede zwischen den Therapiearmen: Die mittlere Gesamtpunktzahl nahm mit einer SMD von 0,42 (95 %-KI [0,27; 0,56]) $p < 0,0001$ (FACT-P) bzw. 0,36 (95 %-KI [0,21; 0,50]) $p < 0,0001$ (FACT-G) unter Enzalutamid-Therapie signifikant weniger stark ab als unter Placebo-Therapie. Bei Betrachtung der Woche 49 gegenüber Studienbeginn ergaben sich keine signifikanten Unterschiede zwischen den Therapiegruppen.

Für die einzelnen Subskalen Körperliches Wohlbefinden (SMD = 0,26; 95 %-KI [0,11; 0,40,], $p = 0,0006$), Verhältnis zu Freunden, Bekannten und Familie (SMD = 0,06; 95 %-KI [-0,08; 0,21], $p = 0,41$), Seelisches Wohlbefinden (SMD = 0,33; 95 %-KI [0,18; 0,47], $p < 0,0001$), Funktionelles Wohlbefinden (SMD = 0,30; 95 %-KI [0,15; 0,45], $p < 0,0001$) sowie die Prostatakarzinom-Subskala (SMD = 0,34; 95 %-KI [0,20; 0,49], $p < 0,0001$) fanden sich bei Betrachtung der letzten verfügbaren Erhebung gegenüber Studienbeginn bis auf die Subskala Verhältnis zu Freunden, Bekannten und Familie jeweils signifikante Unterschiede zugunsten der Enzalutamid-Therapie. Bei Betrachtung der Woche 49 gegenüber Studienbeginn ergaben sich keine signifikanten Unterschiede zwischen den Therapiegruppen.

Anzumerken ist, dass sich bei Betrachtung der Woche 49 mit einer Anzahl von maximal 250 Patienten deutlich weniger Patienten in der Auswertung befanden als bei Betrachtung der letzten verfügbaren Erhebung, für die mindestens 854 Patienten zur Verfügung standen.

Eine Auswertung der Veränderung der „gesundheitsbezogenen Lebensqualität nach FACT-P“ in Woche 25 gegenüber Studienbeginn wurde im Rahmen des *Genitourinary Cancers Symposiums* 2013 der ASCO (*American Society of Clinical Oncology*) als Posterpräsentation veröffentlicht. [Miller 2013]

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.6.2 Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ - RCT

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

Tabelle 4-52: Operationalisierung von Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die Erfassung der Lebensqualität mittels des Fragebogens *European Quality of Life Five-Domain Scale* (EQ-5D) wurde im Rahmen einer Protokolländerung nach Studienbeginn in die Studie integriert (*protocol amendment 2* am 01.04.2010) und in 5 der insgesamt 9 europäischen Ländern und damit in 5 der insgesamt 15 an der Studie beteiligten Ländern erhoben. Der EQ-5D ist ein standardisiertes Dokument zur Messung der gesundheitsbezogenen Lebensqualität und besteht aus zwei Teilen. [Rabin 2011] Im ersten Teil erfolgt eine deskriptive Beschreibung des Gesundheitszustands. Dazu beantworten die Patienten Fragen zu den fünf Domänen Beweglichkeit/Mobilität, Fähigkeit für sich selbst zu sorgen, alltägliche Tätigkeiten, Schmerzen/körperliche Beschwerden und Angst/Niedergeschlagenheit, welche jeweils in drei Kategorien, die von „kein Problem“ bis „ernsthaftes Problem“ reichen, aufgeteilt sind. Im zweiten Teil, der „*Visual Analogue Scale*“ (VAS), machen die Patienten eine generelle Angabe zu ihrem Gesundheitszustand, welchen sie auf einer Skala von 0 bis 100 (0 = *worst health you can imagine*, 100 = *best health you can imagine*) bewerten. Der EQ-5D-Fragebogen wurde vom Untersucher oder Studienkoordinator direkt an die Patienten gegeben. Die Erhebung des Fragebogens erfolgte bei der Screening Visite, bei Woche 13 und 25 und anschliessend alle weiteren 12 Wochen.

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 4-53: Bewertung des Verzerrungspotenzials für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ 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
AFFIRM	niedrig	ja	nein	nein	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Die Auswertungspopulation für den EQ-5D umfasste 214 Patienten, somit befinden sich, unabhängig von der Art und dem Zeitpunkt der der Auswertung, maximal 18 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt und das Verzerrungspotenzial für den Endpunkt „gesundheitsbezogene Lebensqualität nach EQ-5D“ wurde mit „hoch“ bewertet. Gründe für die niedrige Patientenanzahl in der Auswertung sind, dass die Erhebung der Lebensqualität mittels EQ-5D erst mit einer Protokolländerung nach Studienbeginn eingeführt wurde (*protocol amendment 2* am 01.04.2010) und es sich beim EQ-5D um ein Europa-spezifisches Instrument zur Beurteilung der Lebensqualität handelt, der Fragebogen somit nur in Studienzentren in europäischen Ländern zum Einsatz kam. Sonstige das Verzerrungspotenzial beeinflussende Faktoren wurden nicht identifiziert.

Stellen Sie die Ergebnisse für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ 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 Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Gesamtpunktzahl basierend auf der VAS)

Studie	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	Baseline	Ende	Delta	N[a]	Baseline	Ende	Delta		
Endpunkt Gesundheitsbezogene Lebensqualität (EQ-5D) Visual Analog Scale Gesamtpunktzahl Woche 49										
AFFIRM	23	74,7 (17,57)	72,4 (19,06)	-0,5 (15,38)	2	80,0 (0,00)	75,7 (22,28)	-11,5 (26,16)	0,85 [-0,69; 2,40]	p = 0,26
Endpunkt Gesundheitsbezogene Lebensqualität (EQ-5D) Visual Analog Scale Gesamtpunktzahl – letzte verfügbare Erhebung										
AFFIRM	124	68,7 (18,28)	64,7 (19,90)	-4,1 (17,32)	50	63,0 (19,73)	59,8 (20,36)	-4,1 (14,67)	-0,03 [-0,36; 0,30]	p = 0,86

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Die Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für EQ-5D zum jeweiligen Zeitpunkt auswertbaren Patienten.

Bei der Betrachtung der „gesundheitsbezogenen Lebensqualität nach der EQ-5D“ VAS Gesamtpunktzahl ergab sich in Woche 49 gegenüber Studienbeginn mit einer Abnahme der mittleren Punktzahl um 0,5 Punkte in der Enzalutamid-Gruppe bzw. einer Abnahme um 11,5 Punkte in der Placebogruppe eine SMD von 0,85 (95 %-KI [-0,69; 2,40]) und damit ein Vorteil zugunsten der Enzalutamid-Therapie. Dieser war aber mit einem p-Wert von p = 0,26 nicht statistisch signifikant. Wurde die letzte verfügbare Erhebung gegenüber Studienbeginn betrachtet, fand sich kein Unterschied zwischen den Therapiearmen.

Tabelle 4-55: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Shift-Tabelle für die Veränderung von Studienbeginn zu Woche 49)

		Enzalutamid			**Placebo**		
Domäne	**Studienbeginn**	***Level 1***	***Level 2***	***Level 3***	***Level 1***	***Level 2***	***Level 3***
Beweglichkeit, Mobilität	*Level 1*	16 (64,0 %)	3 (12,0 %)	0	1 (50 %)	1 (50 %)	0
	Level 2	2 (8,0 %)	4 (16,0 %)	0	0	0	0
	Level 3	0	0	0	0	0	0
	Fehlend	5	3	0	1	0	0
Fähigkeit für sich selbst zu sorgen	*Level 1*	21 (84,0 %)	2 (8,0 %)	0	2 (100,0 %)	0	0
	Level 2	0	2 (8,0 %)	0	0	0	0
	Level 3	0	0	0	0	0	0
	Fehlend	8	0	0	1	0	0
Alltägliche Tätigkeiten	*Level 1*	15 (60,0 %)	2 (8,0 %)	0	1 (50,0 %)	1 (50,0 %)	0
	Level 2	3 (12,0 %)	3 (12,0 %)	2 (8,0 %)	0	0	0
	Level 3	0	0	0	0	0	0
	Fehlend	6	2	0	1	0	0
Schmerzen, körperliche Beschwerden	*Level 1*	12 (48,0 %)	1 (4,0 %)	0	0	0	0
	Level 2	4 (16,0 %)	8 (32,0 %)	0	1 (50,0 %)	1 (50,0 %)	0
	Level 3	0	0	0	0	0	0
	Fehlend	2	5	1	1	0	0
Angst, Niedergeschlagenheit	*Level 1*	16 (64,0 %)	2 (8,0 %)	0	1 (50,0 %)	1 (50,0 %)	0
	Level 2	1 (4,0 %)	6 (24,0 %)	0	0	0	0
	Level 3	0	0	0	0	0	0
	Fehlend	6	2	0	1	0	0

Bei Betrachtung der einzelnen Domänen des EQ-5D (Beweglichkeit und Mobilität, Fähigkeit für sich selbst zu sorgen, alltägliche Tätigkeiten, Schmerzen und körperliche Beschwerden, sowie Angst und Niedergeschlagenheit) zeigte sich, dass sich im Vergleich von Woche 49 gegenüber Studienbeginn in der Enzalutamid-Gruppe für den Großteil der Patienten (72 % bis 92 % der Patienten) in den einzelnen Domänen keine Veränderungen ergaben. Für die Patienten mit Veränderungen traten Verbesserungen um ein *Level* am häufigsten im Bereich Schmerzen und körperliche Beschwerden auf (4 Patienten), Verschlechterungen um einen *Level* fanden am häufigsten im Bereich alltägliche Tätigkeiten (4 Patienten) statt. In der Placebogruppe befanden sich insgesamt lediglich zwei Patienten.

Tabelle 4-56: Ergebnisse für Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ in RCT mit dem zu bewertenden Arzneimittel (Shift-Tabelle für die Veränderung von Studienbeginn zu letzte verfügbare Erhebung)

		Enzalutamid			**Placebo**		
Domäne	**Studienbeginn**	***Level 1***	***Level 2***	***Level 3***	***Level 1***	***Level 2***	***Level 3***
Beweglichkeit, Mobilität	*Level 1*	59 (45,0 %)	21 (16,0 %)	1 (0,8 %)	25 (47,2 %)	6 (11,3 %)	0
	Level 2	12 (9,2 %)	36 (27,5 %)	2 (1,5 %)	5 (9,4 %)	16 (30,2 %)	0
	Level 3	0	0	0	0	0	1 (1,9 %)
	Fehlend	9	8	1	2	2	1
Fähigkeit für sich selbst zu sorgen	*Level 1*	96 (74,4 %)	18 (14,0 %)	2 (1,6 %)	39 (73,6 %)	9 (17,0 %)	0
	Level 2	3 (2,3 %)	9 (7,0 %)	1 (0,8 %)	1 (1,9 %)	3 (5,7 %)	0
	Level 3	0	0	0	0	1 (1,9 %)	0
	Fehlend	16	2	2	3	2	0
Alltägliche Tätigkeiten	*Level 1*	49 (37,7 %)	21 (16,2 %)	4 (3,1 %)	26 (49,1 %)	7 (13,2 %)	0
	Level 2	12 (9,2 %)	31 (23,8 %)	7 (5,4 %)	2 (3,8 %)	10 (18,9 %)	4 (7,5 %)
	Level 3	2 (1,5 %)	2 (1,5 %)	2 (1,5 %)	0	3 (5,7 %)	1 (1,9 %)
	Fehlend	10	8	1	2	2	1
Schmerzen, körperliche Beschwerden	*Level 1*	20 (15,4 %)	18 (13,8 %)	1 (0,8 %)	6 (11,3 %)	9 (17,0 %)	1 (1,9 %)
	Level 2	17 (13,1 %)	58 (44,6 %)	9 (6,9 %)	4 (7,5 %)	26 (49,1 %)	3 (5,7 %)
	Level 3	0	5 (3,8 %)	2 (1,5 %)	1 (1,9 %)	1 (1,9 %)	2 (3,8 %)
	Fehlend	3	15	1	2	2	1
Angst, Niedergeschlagenheit	Level 1	65 (49,6 %)	20 (15,3 %)	0	20 (37,7 %)	13 (24,5 %)	0
	Level 2	10 (7,6 %)	26 (19,8 %)	4 (3,1 %)	3 (5,7 %)	13 (24,5 %)	0
	Level 3	0	5 (3,8 %)	1 (0,8 %)	0	3 (5,7 %)	1 (1,9 %)
	Fehlend	9	9	0	3	1	1

Hinsichtlich des Vergleichs der einzelnen Domänen des EQ-5D (Beweglichkeit und Mobilität, Fähigkeit für sich selbst zu sorgen, alltägliche Tätigkeiten, Schmerzen und körperliche Beschwerden sowie Angst und Niedergeschlagenheit) zwischen der letzten verfügbaren Erhebung und Studienbeginn fanden sich in der Enzalutamid-Gruppe für 62 % bis 82 % der Patienten und in der Placebogruppe für 64 % bis 79 % der Patienten keine Veränderungen in den einzelnen Domänen. Verbesserungen in den einzelnen Domänen ergaben sich für 2 % bis 17 % der Patienten in der Enzalutamid-Gruppe und für 4 % bis 11 % in der Placebogruppe. Verschlechterungen in den einzelnen Domänen traten bei 16 % bis 34 % der Patienten unter Enzalutamid-Therapie und bei 11 % bis 25 % der Patienten unter Placebo-Therapie auf.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7 Endpunkt „Sicherheit“ - RCT

4.3.1.3.1.7.1 Endpunkt „Unerwünschte Ereignisse“ - RCT

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

Tabelle 4-57: Operationalisierung von Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Ein „unerwünschtes Ereignis“ (UE) war definiert als jedes ungünstige und unbeabsichtigte Zeichen, Symptom oder jede Erkrankung (einschließlich klinisch signifikanter Laborbefunde welche anhand der *National Cancer Institute's Common Toxicity Criteria for Adverse Events* (NCI-CTCAE), Version 4.0 als Grad ≥ 3 klassifiziert waren), die im Rahmen der klinischen Studie bei einem Patienten oder einer Versuchsperson auftraten oder sich verschlechterten, unabhängig davon, ob das Auftreten des unerwünschten Ereignisses mit dem medizinischen Produkt verknüpft war oder nicht. „Unerwünschte Ereignisse“ wurden nach dem MedDRA Kodiersystem (Version 12) kodiert. Sie wurden bei der Visite auf unspezifisches Nachfragen des Untersuchers hin festgestellt oder wenn der Patient Symptome schilderte oder sie bei der Untersuchung vorlagen. In zusätzlichen Tests wurden klinisch signifikante Anomalitäten bezüglich Vitalzeichen, EKG, Labortests und anderen diagnostischen Prozeduren untersucht. Nicht schwerwiegende UEs wurden im Zeitraum von der Einnahme der ersten Studienmedikation bis zum 30sten Tag nach der Einnahme der letzten Studienmedikation oder beim Absetzen der Studienmedikation bis zum Tag vor der Einnahme einer anderen Medikation, je nachdem welches Ereignis zuerst eintrat, registriert. Die Registrierung schwerwiegender UEs hingegen begann zum Zeitpunkt der Unterzeichnung der Einverständniserklärung.

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 4-58: Bewertung des Verzerrungspotenzials für Endpunkt „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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Unerwünschte Ereignisse“ wurde daher mit „niedrig“ bewertet.

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 4-59: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Unerwünschte Ereignisse gesamt					
AFFIRM	800	785 (98,1)	399	390 (97,7)	1,00 [0,99; 1,02] p = 0,67

a: Anzahl Patienten in der Auswertung
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

In der AFFIRM-Studie traten bei 98,1 % (785/800) der Patienten unter Enzalutamid-Therapie „unerwünschte Ereignisse“ auf im Vergleich zu 97,7 % (390/399) der Patienten in der Placebogruppe. Das Relative Risiko betrug 1,00 (95 %-KI [0,99; 1,02], p = 0,67). Für den Endpunkt „Unerwünschte Ereignisse“ ergaben sich somit keine Unterschiede zwischen den Behandlungsarmen.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7.2 Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ - RCT

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

Tabelle 4-60: Operationalisierung von Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Unerwünschte Ereignisse wurden wie unter dem Endpunkt ”Unerwünschte Ereignisse” (Abschnitt 4.3.1.3.1.7.1) beschrieben, definiert und registriert. Anhand der Kriterien des *NCI*-CTCAE, Version 4.0, wurden sie in die fünf Schweregrade 1 bis 5 eingeteilt. Änderte sich der Schweregrad eines UEs über die Zeit, wurde der Schwerste der Schweregrade für dieses UE registriert. Innerhalb des Endpunkts „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ werden alle UEs nach CTCAE Grad ≥ 3, welche mit einer relativen Häufigkeit von ≥ 3 % in mindestens einer Behandlungsgruppe auftraten, dargestellt.

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 4-61: Bewertung des Verzerrungspotenzials für Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ wurde daher mit „niedrig“ bewertet.

Stellen Sie die Ergebnisse für den Endpunkt „Unerwünschte Ereignisse nach CTCAE *Grad ≥ 3“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-62: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“, Enzalutamid vs. Placebo in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Unerwünschte Ereignisse mit CTCAE Grad ≥ 3					
AFFIRM	800	362 (45,3)	399	212 (53,1)	0,85 [0,76; 0,96] p = 0,009

a: Anzahl Patienten in der Auswertung

CTCAE: *Common Toxicity Criteria for Adverse Events*

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation. Dargestellt sind alle UEs nach CTCAE Grad ≥ 3, welche mit einer relativen Häufigkeit von ≥ 3 % in mindestens einer Behandlungsgruppe auftraten.

In der AFFIRM-Studie traten bei 45,3 % (362/800) der Patienten unter Enzalutamid-Therapie unerwünschte Ereignisse nach CTCAE Grad ≥ 3 auf, in der Placebogruppe traten diese Ereignisse hingegen bei 53,1 % (212/399) der Patienten auf. Das Relative Risiko lag mit einem Wert von 0,85 (95 %-KI [0,76; 0,96], p = 0,009) statistisch signifikant zugunsten der Enzalutamid-Gruppe.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7.3 Endpunkt „Schwerwiegende Unerwünschte Ereignisse“ - RCT

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

Tabelle 4-63: Operationalisierung von Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Ein „schwerwiegendes unerwünschtes Ereignis“ war definiert als jedes unerwünschte medizinische Vorkommnis, welches mindestens eines der folgenden Kriterien erfüllte: • tödlich verlief; • lebensbedrohlich war, wobei der Patient mit Auftreten des unerwünschten Ereignisses unmittelbar vom Tod bedroht war, mit Ausnahme von Ereignissen, welche, wären sie in einer schwereren Form aufgetreten oder unbehandelt geblieben, möglicherweise zum Tod geführt hätten; • die Hospitalisierung oder eine Verlängerung der Hospitalisierung des Patienten nach sich zog; • eine bleibende oder substantielle Behinderung nach sich zog (beispielsweise eine erhebliche Beeinträchtigung der Fähigkeit, die üblichen Aktivitäten des alltäglichen Lebens zu verrichten); • bei Nachkommen von Personen, welche die Prüfsubstanz erhielten, zu einer angeborenen Anomalie/einem angeborenen Geburtsgebrechen führte; Die Registrierung schwerwiegender UEs begann zum Zeitpunkt der Unterzeichnung der Einverständniserklärung bis zum 30sten Tag nach der Einnahme der letzten Studienmedikation, oder der Zeit, an der eine andere systemische, antineoplastische Therapie begann.

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 4-64: Bewertung des Verzerrungspotenzials für Endpunkt „Schwerwiegende 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ wurde daher mit „niedrig“ bewertet.

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

Tabelle 4-65: Ergebnisse für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“, Enzalutamid vs Placebo in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Schwerwiegende unerwünschte Ereignisse					
AFFIRM	800	268 (33,5)	399	154 (38,6)	0,87 [0,74; 1,02] p = 0,08

a: Anzahl Patienten in der Auswertung

RR: Relatives Risoko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation. Dargestellt sind alle Ereignisse, welche mit einem absoluten Unterschied von ≥ 0,5 % in der Ereignisrate zwischen den Behandlungsarmen auftraten.

„Schwerwiegende unerwünschte Ereignisse“ traten in der AFFIRM-Studie bei 33,5 % (268/800) der Patienten unter Enzalutamid-Therapie auf, im Gegensatz zu 38,6 % (154/399) der Patienten in der Placebogruppe. Das Relative Risiko betrug 0,87 (95 %-KI [0,74; 1,02], p = 0,08) zugunsten der Enzalutamid-Gruppe, der Unterschied zwischen den Gruppen war aber nicht statistisch signifikant.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7.4 Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ - RCT

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

Tabelle 4-66: Operationalisierung von Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Unerwünschte Ereignisse wurden wie unter dem Endpunkt ”Unerwünschte Ereignisse” (Abschnitt 4.3.1.3.1.7.1) beschrieben, definiert und registriert. Als „Studienabbruch aufgrund von unerwünschten Ereignissen“ wurde jeder Therapieabbruch gewertet, der aufgrund von unerwünschten Ereignissen und nicht aufgrund der Krankheitsprogression erfolgte. Diese Unterscheidung wurde vom Untersucher vorgenommen.

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 Endpunkt „Studienabbruch aufgrund unerwünschter 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ wurde daher mit „niedrig“ bewertet.

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

Tabelle 4-68: Ergebnisse für den Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Studienabbruch aufgrund unerwünschter Ereignisse					
AFFIRM	800	61 (7,6)	399	39 (9,8)	0,78 [0,53; 1,15] p = 0,20

a: Anzahl Patienten in der Auswertung

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation. Dargestellt sind alle Ereignisse, welche mit einem absoluten Unterschied von ≥ 0,5 % in der Ereignisrate zwischen den Behandlungsarmen auftraten.

In der AFFIRM-Studie traten bei 7,6 % (61/800) der Patienten unter Enzalutamid-Therapie „Studienabbrüche aufgrund unerwünschter Ereignisse“ auf. In der Placebogruppe war dies bei 9,8 % (39/399) der Patienten der Fall. Das Relative Risiko betrug 0,78 (95 %-KI [0,53; 1,15], p = 0,20) zugunsten der Enzalutamid-Gruppe. Ein statistisch signifikanter Unterschied zwischen den Gruppen war allerdings nicht vorhanden.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7.5 Endpunkt „Tod aufgrund unerwünschter Ereignisse“ - RCT

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

Tabelle 4-69: Operationalisierung von Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Unerwünschte Ereignisse wurden wie unter dem Endpunkt ”Unerwünschte Ereignisse” (Abschnitt 4.3.1.3.1.7.1) beschrieben, definiert und registriert. Alle Todesfälle, die nicht aufgrund einer Krankheitsprogression einhergingen oder nicht eine unbekannte Ursache hatten, wurden als „Todesfälle aufgrund unerwünschter Ereignisse“ erfasst.

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 4-70: Bewertung des Verzerrungspotenzials für Endpunkt „Tod aufgrund unerwünschter 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
AFFIRM	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet. Aus dem Studienbericht ergaben sich keine Hinweise auf ergebnisgesteuerte Berichterstattung. Das ITT-Prinzip wurde adäquat umgesetzt. Es wurden keine sonstigen das Verzerrungspotenzial beeinflussende Faktoren identifiziert. Das Verzerrungspotenzial für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ wurde daher mit „niedrig“ bewertet.

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

Tabelle 4-71: Ergebnisse für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI] p-Wert**
Tod aufgrund unerwünschter Ereignisse					
AFFIRM	800	23 (2,9)	399	14 (3,5)	0,82 [0,43; 1,58] p = 0,55

a: Anzahl Patienten in der Auswertung
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

In der AFFIRM-Studie starben 2,9 % (23/800) der Patienten unter Enzalutamid-Therapie und 3,5 % (14/399) der Patienten der Placebogruppe aufgrund unerwünschter Ereignisse. Das Relative Risiko betrug 0,82 (95 %-KI [0,43; 1,58], p = 0,55) zugunsten der Enzalutamid-Gruppe, war aber nicht statistisch signifikant.

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.

Meta-Analysen konnten nicht durchgeführt werden, da nur eine Studie vorlag.

4.3.1.3.1.7.6 Endpunkt „Krampfanfälle“ - RCT

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

Tabelle 4-72: Operationalisierung von Endpunkt „Krampfanfälle“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Operationalisierung
AFFIRM	Die separate Erhebung des Endpunkts „Krampfanfälle“ war im Rahmen der AFFIRM-Studie nicht vorgesehen, weshalb auch keine allgemeine Operationalisierung vorgenommen werden kann. Die Einzelereignisse sind im folgenden Abschnitt separat gelistet. Sie wurden auf der Basis der Einträge für schwere unerwünschte Ereignisse (mussten durch den Arzt vorgenommen werden) und deren Kodierung mittels MedrA identifiziert.

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

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

Tabelle 4-73: Bewertung des Verzerrungspotenzials für Endpunkt „Krampfanfälle“ 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
AFFIRM	niedrig	ja	a)	nein	ja	„hoch“

a) Es erfolgte keine statistische Auswertung, bei der deskriptiven Darstellung wurden alle randomisierten Patienten berücksichtigt.

Begründen Sie für jede Studie die abschließende Einschätzung.

Für das Ereignis „Krampfanfälle“ konnte keine statistische Auswertung durchgeführt werden, da es sich um seltene, heterogene Einzelereignisse handelt. Eine Auswertung wäre mit einem hohen Verzerrungspotential zu bewerten. Es wurden aus diesem Endpunkt keine Aussagen zum Zusatznutzen abgeleitet.

Krampfanfälle stellen ein ernstzunehmendes unerwünschtes Ereignis dar. In der Phase I/II Studie zur Untersuchung der Antitumor-Aktivität von Enzalutamid bei kastrationsresistentem Prostatakarzinom [Scher 2010] trat bei zwei Patienten ein Krampfanfall und bei einem Patienten ein möglicher Krampfanfall als unerwünschtes Ereignis auf. Zwei der Krampfanfälle wurden beobachtet und traten unter Einnahme von Enzalutamid in einer Dosis von 600 mg/Tag bzw. 360 mg/Tag auf. Unklar ist, ob Enzalutamid für das Auftreten der Krampfanfälle verantwortlich war, da beide Patienten eine Komedikation mit Medikamenten, welche die Krampfschwelle herabsetzen, erhielten und zudem komplexe medizinische Probleme, welche ebenfalls zum Auftreten des Anfalls geführt haben könnten, aufwiesen. [Scher 2010] Der dritte Patient erhielt eine Enzalutamid-Dosis von 480 mg/Tag, der Krampfanfall wurde nicht bezeugt (möglicher Krampfanfall), aber als solcher gewertet. In allen drei Fällen lag die verabreichte Enzalutamid-Dosis über der zugelassenen Dosis von 160 mg/Tag.

Aufgrund dieser Beobachtungen wurde in der AFFIRM-Studie ein besonderes Augenmerk darauf gelegt, um zum einen das Risko für das Auftreten eines Krampfanfalls zu reduzieren und zum anderen die Aufmerksamkeit der Untersucher hinsichtlich dieses Ereignisses zu schärfen. Patienten, welche unter Krampfanfällen litten oder eine Prädisposition dafür aufwiesen, wurden aus der Studie ausgeschlossen, ebenso wie Patienten, welche die Krampfschwelle senkende Medikamente einnahmen. Die Einnahme solcher Medikamente als Begleitmedikation war ebenfalls nicht erlaubt. Weiterhin mussten im Rahmen der Studie auftretende Krampfanfälle als schwerwiegendes unerwünschtes Ereignis dokumentiert werden. [Astellas Pharma GmbH 2013b] In der AFFIRM-Studie traten bis zum *data cutoff date* (25. September 2011) in der Enzalutamid-Gruppe bei fünf Patienten (0,6 %) Krampfanfälle auf, wohingegen es in der Placebogruppe keinen Fall gab. [Astellas Pharma GmbH 2013b; Scher 2012] Vier der Krampfanfälle waren selbstlimitierend und traten nach Absetzen der Studienmedikation nicht wieder auf. In einem Fall kam es zu einem *Status epilepticus*, welcher eine medizinische Behandlung erforderlich machte. Potentiell prädisponierende Faktoren waren Hirnmetastasen (zwei Patienten), eine versehentliche intravenöse Gabe von Lidocain (ein Patient), sowie übermäßiger Alkoholkonsum zusammen mit einer Haloperidol-Therapie und einer Gehirnatropie (ein Patient). [Medivation Inc. 2012a; Scher 2012] Nach dem *data cutoff date* (25. September 2011) der AFFIRM-Studie trat bei zwei weiteren Patienten in der Enzalutamid-Gruppe ein Krampfanfall auf, so dass insgesamt 7 von 800 (0,9 %) Patienten Krampfanfälle erlitten (siehe Tabelle 4-74). [Astellas Pharma GmbH 2013b; Medivation Inc. 2012a]

Tabelle 4-74: Übersicht über die im Enzalutamid-Arm auftretenden Krampfanfälle. [Astellas Pharma GmbH 2013b]

Enzalutamid Tägliche Dosis	**Art des Krampfanfalls**	**Beobachtet (ja/nein)**	**Mögliche Faktoren**
160 mg	Fokaler Beginn	Ja	Hirnmetastasen mit Blutung, gefäßbedingtem Ödem und Massenverschiebung
160 mg	Partieller Anfall	Ja	Irrtümliche Gabe von 40 mg Lidocain i.v. direkt vor Anfall; gleichzeitig Propafenone
160 mg	Status epilepticus (Verwirrung Verbindung mit partiell komplexen Status)	Ja	Hirnatrophie im Kopf CT
160 mg	Fokaler Beginn	Ja	ZNS Metastasen:Auge, Meningen, Zerebellum
160 mg	unbekannt	Nein	Schwerer Alkoholismus, 7 Tage vor Anfall, Beginn mit Haloperidol aufgrund von Übelkeit
160 mg	Synkope mit Merkmalen eines Anfalls	Ja	Keine bekannt
160 mg	Krampfanfall und komplexe partielle Anfälle	ja	Keine bekannt

4.3.1.3.2 Subgruppenanalysen – RCT

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

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten.

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

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

Alle Wirksamkeits– und Sicherheitsendpunkte der Studie AFFIRM wurden *a posteriori* anhand verschiedener Subgruppen beurteilt. Diese umfassten die Gruppen „Gabe von Bisphosphonaten während der Studie“, „Gabe von Steroiden während der Studie“, „ECOG-*Performance*-Status zu Studienbeginn“, „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“, „Alter“, „Anzahl vorheriger Chemotherapien“ und „viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des *Screenings*“.

A priori waren weitere Effektmodifikatoren (Hämoglobinwert zu Studienbeginn, *Gleason Score* zu Studienbeginn, LDH-Wert zu Studienbeginn, PSA-Wert zu Studienbeginn, Progressionsart zu Studienbeginn, sowie die geographische Region) hinsichtlich ihres Einflusses auf den Endpunkt „Gesamtüberleben“ untersucht worden. Sicherheitsbezogene Ereignisse (UE, UE nach CTCAE Grad ≥ 3, SUE, Studienabbruch aufgrund UE, Tod aufgrund UE) wurden zudem *a priori* in folgenden Subgruppen ausgewertet: „Alter (< 65, ≥ 65, ≥ 75)“; „Anamnese einer Hypertonie (ja vs. nein)“ und „Anamnese kardiovaskulärer Erkrankungen (ja vs. nein)“.

Im Folgenden werden die Ergebnisse der verschiedenen Subgruppenanalysen tabellarisch dargestellt. Eine Zusammenfassung findet sich in Abschnitt 4.3.1.3.2.2.

4.3.1.3.2.1 Subgruppenanalysen zur Studienpopulation-RCT

4.3.1.3.2.1.1 „Interaktionstests“

Wie im Abschnitt 4.2.5.5 beschrieben, wurden in einem ersten Schritt für alle Subgruppen Tests auf Interaktionen durchgeführt. Angaben zur Methodik finden sich im *Statistical*

Analysis Plan des CSR bzw. im *Statistical Analysis Plan* zu den Zusatzauswertungen (beides bei den Unterlagen zu den Studienberichten im Modul 5 abgespeichert). Bei Endpunkten, die mittels *Responder*-Analysen oder als dichotome Ereignisse ausgewertet wurden, wurden die Interaktionsterme der Subgruppen zusätzlich mittels der *Software Review Manager, Version 5.1*, Methode nach Deeks berechnet. [Deeks 2010] Bezüglich der Relativen Risiken sowie der dazugehörigen p-Werte und Konfidenzintervalle ergaben sich dadurch keine Unterschiede in den Ergebnissen. Für die Interaktionsterme wurden jedoch unterschiedliche Werte erhalten. Im Falle einer Abweichung der Interaktionsterme bezüglich der Kategorien „Beleg“ und „Hinweis“ wurden jeweils beide Interaktionsterme berichtet. Das Ergebnis für den jeweiligen Interaktionstest wurde als „Beleg“ auf eine Interaktion zwischen den Subgruppen gewertet, wenn ein p-Wert von $\leq 0{,}05$ vorlag, und als „Hinweis“ auf eine Interaktion zwischen den Subgruppen, wenn ein p-Wert $\leq 0{,}2$ vorlag. Lag mindestens ein Hinweis oder aber ein Beleg auf unterschiedliche Effekte in den Subgruppen vor, so wurden neben dem Gesamteffekt auch die einzelnen Subgruppenergebnisse in die Interpretation der Ergebnisse einbezogen. [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011]

Die Angaben, welche Subgruppen a priori und welche Subgruppen post hoc definiert wurden, befinden sich in Abschnitt 4.2.5.5. Für das Gesamtüberleben waren alle Subgruppen prädefiniert, bis auf die beiden Subgruppen nach Steroid – bzw. Bisphosphonatgabe.

Auf die Subgruppenergebnisse soll im Folgenden näher eingegangen werden. In Tabelle 4-75 findet sich eine Übersicht über alle durchgeführten Interaktionstests. Die Ergebnisse der Subgruppenanalysen, für die Interaktionen vorlagen, werden in den nachfolgenden Tabellen dargestellt.

Sämtliche durchgeführten Subgruppenanalysen sind vollständig im Modul 5 im Ordner Studienberichte inklusive eines SAP zu den ausführlichen Methoden der Analysen hinterlegt.

Tabelle 4-75: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests (p-Werte) der randomisiert kontrollierten Studie AFFIRM

Subgruppen **Endpunkte**	**Gabe von Bisphosphonaten während der Studie (ja vs. nein)**	**Gabe von Steroiden während der Studie (ja vs. nein)**	**ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)**	**Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4)**	**Alter (< 65 vs. ≥ 65)**	**Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)**	**Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein)**
Gesamtüberleben	0,92	0,06[a]	0,92	0,34	0,98	0,23	0,15[a]
Radiographisches progressionsfreies Überleben	0,61	< 0,0001[b]	0,33	0,0015[b]	0,18[a]	0,10[a]	0,21
Zeit bis zur ersten skelettbezogenen Komplikation	0,57	0,12[a]	0,93	0,85	0,20[a]	0,85	0,21
Schmerzlinderung	0,64[c]	0,43[c]	0,64[c]	NA[c]	0,89[c]	0,77[c]	NA[c]
Schmerzprogressionsrate (*Responder*-Analyse)	0,94	0,15[a]	0,67	0,003[b]	0,16[a]	0,41	0,88
Schmerzprogressionsrate (Veränderung zu Studienbeginn)	0,71	0,60	0,14[a]	0,10[a]	0,21	0,76	0,95
Zeit bis zur Schmerzprogression	0,42	0,66	0,68	0,66	0,56	0,60	0,36
Gesundheitsbezogene Lebensqualität nach FACT-P (*Responder*-Analyse)	0,84	0,09[a]	0,61	0,86	0,60	0,63	0,35
Gesundheitsbezogene Lebensqualität nach FACT-P							
Veränderung Woche 49 gegenüber Studienbeginn	0,86	0,16[a]	0,93	0,31	0,02[b]	0,99	0,27
Veränderung letzte verfügbare Erhebung gegenüber	0,53	0,03[b]	0,10[a]	0,91	0,74	0,47	0,93

Subgruppen **Endpunkte**	**Gabe von Bisphosphonaten während der Studie (ja vs. nein)**	**Gabe von Steroiden während der Studie (ja vs. nein)**	**ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)**	**Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4)**	**Alter (< 65 vs. ≥ 65)**	**Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)**	**Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein)**
Studienbeginn							
Gesundheitsbezogene Lebensqualität nach FACT-G							
Veränderung Woche 49 gegenüber Studienbeginn	0,63	0,17[a]	0,85	0,77	0,04[b]	0,78	0,31
Veränderung letzte verfügbare Erhebung gegenüber Studienbeginn	0,37	0,05[b]	0,29	0,45	0,80	0,41	0,92
Gesundheitsbezogene Lebensqualität nach EQ-5D							
Veränderung Woche 49 gegenüber Studienbeginn	NA	0,06[a]	NA	0,02[b]	NA	NA	NA
Veränderung letzte verfügbare Erhebung gegenüber Studienbeginn	0,83	0,85	0,28	0,20[a]	0,27	0,03[b]	0,41
Unerwünschte Ereignisse	0,33	0,62	NA (0,34)[c]	0,59	0,40	0,52	NA (0,23)[c]
Unerwünschte Ereignisse nach CTCAE Grad ≥ 3	0,10[a]	0,10[a] (0,02[b])[c]	0,31 (0,14[a])[c]	0,37 (0,18[a])[c]	0,65	0,98	0,90
Schwerwiegende unerwünschte Ereignisse	0,65	0,25 (0,14[a])[c]	0,04[b]	0,28 (0,19[a])[c]	0,68	0,88	0,88

Subgruppen **Endpunkte**	**Gabe von Bisphosphonaten während der Studie (ja vs. nein)**	**Gabe von Steroiden während der Studie (ja vs. nein)**	**ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)**	**Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4)**	**Alter (< 65 vs. ≥ 65)**	**Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)**	**Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein)**
Studienabbruch aufgrund unerwünschter Ereignisse	0,91	0,81	0,60	0,51	0,79	0,86	0,75
Tod aufgrund unerwünschter Ereignisse	0,03[b]	0,43	0,85	0,80	0,69	1,00	0,82

a: Hinweis auf Interaktion

b: Beleg für Interaktion

c: Berechnet mittels *Review Manager, Version 5.1* [Deeks 2010]

ECOG: *Eastern Cooperative Oncology Group*

CTCAE: *Common Toxicity Criteria for Adverse Events*

BPI: *Brief Pain Inventory*

NA: nicht anwendbar

4.3.1.3.2.1.2 Endpunkt „Gesamtüberleben“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-76: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI] p-Wert**
Gesamtüberleben					
AFFIRM					
Ja[b]	382	213 (55,8)	182	124 (68,1)	0,70 [0,56; 0,88] p = 0,002
Nein[c]	418	95 (22,7)	217	88 (40,6)	0,49 [0,37; 0,65] p < 0,0001
Interaktionstest					p = 0,06

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Steroide bekommen haben
c: Patienten, die während der Studie keine Steroide bekommen haben
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

Tabelle 4-77: Subgruppenanalyse zur Studienpopulation „Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des *Screenings* (ja vs. nein)“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 % KI] p-Wert**
Gesamtüberleben					
AFFIRM					
Ja[b]	196	105 (53,6)	82	50 (61,0)	0,78 [0,56; 1,09] p = 0,15
Nein[c]	604	203 (33,6)	317	162 (51,1)	0,57 [0,46; 0,70] p < 0,0001
Interaktionstest					p = 0,15

a: Anzahl Patienten in der Auswertung
b: Patienten, die zum Zeitpunkt des Screenings eine viszerale Erkrankung hatten
c: Patienten, die zum Zeitpunkt des Screenings keine viszerale Erkrankung hatten
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

4.3.1.3.2.1.3 Endpunkt „Radiographisches progressionsfreies Überleben“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-78: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI]** **p-Wert**
Radiographisches progressionsfreies Überleben					
AFFIRM					
Ja[b]	382	295 (77,2)	182	160 (87,9)	0,60 [0,49; 0,73] p < 0,0001
Nein[c]	418	229 (54,8)	217	177 (81,6)	0,26 [0,21; 0,33] p < 0,0001
Interaktionstest					p < 0,0001

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Steroide bekommen haben
c: Patienten, die während der Studie keine Steroide bekommen haben
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

Tabelle 4-79: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI]** **p-Wert**
Radiographisches progressionsfreies Überleben					
AFFIRM					
Schmerzpunktzahl < 4	574	356 (62,0)	284	235 (82,7)	0,34 [0,28; 0,40] p < 0,0001
Schmerzpunktzahl ≥ 4	226	168 (74,3)	115	102 (88,7)	0,58 [0,45; 0,75] p < 0,0001
Interaktionstest					p < 0,0015

a: Anzahl Patienten in der Auswertung

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der ITT-Population.

Tabelle 4-80: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI] p-Wert**
Radiographisches progressionsfreies Überleben					
AFFIRM					
Alter < 65	232	164 (70,7)	130	110 (84,6)	0,46 [0,36; 0,59] p < 0,0001
Alter ≥ 65	568	360 (63,4)	269	227 (84,4)	0,38 [0,32; 0,45] p < 0,0001
Interaktionstest					p = 0,18

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

Tabelle 4-81: Subgruppenanalyse zur Studienpopulation „Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)“ für den Endpunkt „Radiographisches progressionsfreies Überleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI] p-Wert**
Radiographisches progressionsfreies Überleben					
AFFIRM					
Vorherige Chemotherapie = 1	579	361 (62,3)	296	249 (84,1)	0,38 [0,32; 0,44] p < 0,0001
Vorherige Chemotherapie ≥ 2	221	163 (73,8)	103	88 (85,4)	0,48 [0,36; 0,63] p < 0,0001
Interaktionstest					p = 0,10

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

4.3.1.3.2.1.4 Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-82: Subgruppenanalyse zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI]** **p-Wert**
Zeit bis zur ersten skelettbezogenen Komplikation					
AFFIRM					
Ja[b]	382	187 (49,0)	182	91 (50,0)	0,78 [0,61; 1,01] p = 0,06
Nein[c]	418	100 (23,9)	217	70 (32,3)	0,56 [0,41; 0,75] p = 0,0001
Interaktionstest					p = 0,12

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Steroide bekommen haben
c: Patienten, die während der Studie keine Steroide bekommen haben
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

Tabelle 4-83: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 %-KI]** **p-Wert**
Zeit bis zur ersten skelettbezogenen Komplikation					
AFFIRM					
Alter < 65	232	93 (40,1)	130	65 (50,0)	0,60 [0,44; 0,83] p = 0,002
Alter ≥ 65	568	194 (34,2)	269	96 (35,7)	0,77 [0,60; 0,99] p = 0,04
Interaktionstest					p = 0,20

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

4.3.1.3.2.1.5 Endpunkt „Schmerzprogressionsrate“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-84: Subgruppenanalysen zur Studienpopulation nach „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schmerzprogressionsrate (*Responder*-Analyse)					
AFFIRM					
Patienten, die während der Studie Steroide bekommen haben	279	101 (36,2)	109	47 (43,1)	0,84 [0,64; 1,10] p = 0,20
Patienten, die während der Studie keine Steroide bekommen haben	346	73 (21,1)	150	54 (36,0)	0,59 [0,44; 0,79] p = 0,0004
Interaktionstest					p = 0,15

a: Anzahl Patienten in der Auswertung

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzprogressionsrate.

Tabelle 4-85: Subgruppenanalysen zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schmerzprogressionsrate (*Responder*-Analyse)					
AFFIRM					
Schmerzpunktzahl < 4	471	145 (30,8)	203	74 (36,5)	0,85 [0,67; 1,06] p = 0,14
Schmerzpunktzahl ≥ 4	154	29 (18,8)	56	27 (48,2)	0,39 [0,26; 0,60] p < <0,0001
Interaktionstest					p = 0,003

a: Anzahl Patienten in der Auswertung
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzprogressionsrate.

Tabelle 4-86: Subgruppenanalysen zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt „Schmerzprogressionsrate (*Responder*-Analyse)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schmerzprogressionsrate (*Responder*-Analyse)					
AFFIRM					
Alter < 65	181	49 (27,1)	83	38 (45,8)	0,59 [0,42; 0,83] p = 0,002
Alter ≥ 65	444	125 (28,2)	176	63 (35,8)	0,79 [0,61;1,01] p = 0,06
Interaktionstest					p = 0,16

a: Anzahl Patienten in der Auswertung
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzprogressionsrate.

Tabelle 4-87: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende[b]**	**Delta**	**N[a]**	**Studienbeginn**	**Ende[b]**	**Delta**		
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)										
AFFIRM										
ECOG = 0 oder 1	576	2,25 (2,04)	1,96 (2,12)	-0,29 (1,94)	242	2,10 (2,04)	2,74 (2,46)	0,64 (2,16)	-0,46 [-0,61; -0,31]	p < 0,0001
ECOG = 2	44	4,46 (2,47)	3,28 (2,36)	-1,18 (2,79)	17	3,96 (2,90)	4,60 (2,66)	0,64 (2,31)	-0,68 [-1,22; -0,14]	p = 0,02
Interaktionstest										p = 0,14

a: Anzahl Patienten in der Auswertung und Relevanz für Delta
b: Woche 13
SD: *Standard Deviation*
SMD: *Standardized Mean Difference*
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzprogressionsrate.

Tabelle 4-88: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende[b]**	**Delta**	**N[a]**	**Studienbeginn**	**Ende[b]**	**Delta**		
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)										
AFFIRM										
Schmerzpunktzahl < 4	467	1,40 (1,23)	1,53 (1,85)	0,13 (1,64)	203	1,33 (1,32)	2,28 (2,28)	0,95 (1,99)	-0,46 [-0,62; -0,30]	p < 0,0001
Schmerzpunktzahl ≥ 4	153	5,50 (1,20)	3,64 (2,26)	-1,83 (2,34)	56	5,46 (1,31)	4,99 (2,17)	-0,47 (2,40)	-0,58 [-0,88; -0,28]	p < 0,001
Interaktionstest										p = 0,10

a: Anzahl Patienten in der Auswertung und Relevanz für Delta
b: Woche 13
SD: *Standard Deviation*
SMD: *Standardized Mean Difference*
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für die Schmerzprogressionsrate.

4.3.1.3.2.1.6 Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-89: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ (*Responder*-Analyse) in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Relatives Risiko [95 %-KI]** **p-Wert**
Gesundheitsbezogene Lebensqualität nach FACT-P, *Responder*-Analyse					
AFFIRM					
Ja	284	117 (41,2)	102	20 (19,6)	2,10 [1,38; 3,19] p = 0,0005
Nein	346	147 (42,5)	139	16 (11,5)	3,69 [2,29; 5,95] p < 0,0001
Interaktionstest					p = 0,09

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Auswertungspopulation für FACT-P.

Tabelle 4-90: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ (Veränderung gegenüber Studienbeginn) in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-P Gesamtpunktzahl - Woche 49										
AFFIRM										
Ja[b]	66	109,24 (19,95)	103,21 (23,29)	-4,98 (19,14)	11	109,26 (20,49)	110,04 (23,07)	-0,76 (21,25)	-0,09 [-0,77; 0,59]	p = 0,79
Nein[d]	155	114,95 (18,66)	114,77 (21,41)	-0,11 (15,51)	10	126,90 (18,67)	117,91 (20,38)	-9,01 (8,09)	0,57 [-0,08; 1,22]	p = 0,08
Interaktionstest										p = 0,16
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl - Woche 49										
AFFIRM										
Ja[b]	66	78,74 (13,99)	74,14 (16,44)	-4,02 (13,20)	12	80,71 (13,11)	79,14 (14,61)	-1,57 (13,28)	-0,13 [-0,77; 0,52]	p = 0,70
Nein[c]	158	82,44 (13,82)	81,66 (16,11)	-0,49 (12,32)	10	90,24 (13,77)	83,42 (15,06)	-6,47 (4,53)	0,49 [-0,16; 1,13]	p = 0,14
Interaktionstest										p = 0,17

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-P Gesamtpunktzahl – letzte verfügbare Erhebung										
AFFIRM										
Ja[b]	276	105,30 (22,38)	96,49 (24,21)	-8,87 (18,57)	102	108,00 (21,53)	95,34 (23,23)	-13,19 (18,52)	0,22 [-0,01; 0,45]	p = 0,06
Nein[c]	337	111,81 (19,68)	109,69 (21,83)	-1,95 (16,21)	139	111,82 (20,44)	99,76 (23,52)	-11,64 (16,72)	0,61 [0,41; 0,81]	p < 0,0001
Interaktionstest										p = 0,03
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl – letzte verfügbare Erhebung										
AFFIRM										
Ja[b]	279	76,29 (16,00)	69,68 (17,48)	-6,64 (13,53)	107	78,27 (14,83)	68,99 (16,54)	-9,36 (13,32)	0,19 [-0,04; 0,41]	p = 0,10
Nein[c]	344	80,11 (14,59)	78,28 (16,10)	-1,86 (11,80)	142	79,78 (14,68)	71,65 (17,18)	-7,96 (12,74)	0,53 [0,33; 0,72]	p < 0,0001
Interaktionstest										p = 0,05

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

b: Patienten, die während der Studie Steroide bekommen haben

c: Patienten, die während der Studie keine Steroide bekommen haben

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für FACT-P zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

Tabelle 4-91: Subgruppenanalyse zur Studienpopulation „Alter (< 65 vs. ≥ 65)“ für den Endpunkt Gesundheitsbezogene Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-P Gesamtpunktzahl - Woche 49										
AFFIRM										
Alter < 65	59	112,75 (18,96)	109,87 (22,20)	-2,22 (19,18)	3	98,28 (21,43)	118,83 (11,28)	17,83 (21,10)	-0,86 [-2,08; 0,37]	p = 0,17
Alter ≥ 65	162	113,48 (19,31)	111,88 (22,73)	-1,33 (15,88)	18	120,89 (19,90)	112,74 (23,40)	-8,44 (12,82)	0,47 [-0,03; 0,96]	p = 0,06
Interaktionstest										p = 0,02
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl - Woche 49										
AFFIRM										
Alter < 65	60	80,70 (13,82)	78,91 (16,00)	-1,34 (13,42)	4	73,21 (11,96)	81,83 (7,77)	8,63 (9,41)	-0,67 [-1,72; 0,38]	p = 0,21
Alter ≥ 65	164	81,61 (14,02)	79,64 (16,76)	-1,60 (12,41)	18	87,67 (13,24)	81,05 (15,88)	-6,56 (8,52)	0,41 [-0,08; 0,90]	p = 0,10
Interaktionstest										p = 0,04

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für FACT-P zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

Tabelle 4-92: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-P-Gesamtpunktzahl - letzte verfügbare Erhebung										
AFFIRM										
ECOG = 0 oder 1	575	110,49 (20,23)	104,74 (23,73)	-5,73 (17,41)	225	111,08 (20,88)	98,81 (23,63)	-12,37 (17,57)	0,38 [0,23; 0,54]	p < 0,0001
ECOG = 2	38	86,28 (21,46)	90,35 (21,43)	5,04 (18,18)	16	97,90 (18,41)	85,34 (16,81)	-11,20 (16,73)	0,86 [0,25; 1,47]	p = 0,006
Interaktionstest										p = 0,10

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für FACT-P zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

4.3.1.3.2.1.7 Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-93: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach EQ-5D Gesamtpunktzahl - Woche 49										
AFFIRM										
Ja[b]	3	81,67 (2,89)	64,29 (27,15)	-8,33 (14,43)	1	80,00 (NA)	70,00 (28,28)	-30,00 (NA)	NA [NA; NA]	p = NA
Nein[c]	20	73,60 (18,64)	74,72 (16,14)	0,65 (15,52)	1	80,00 (NA)	87,00 (NA)	7,00 (NA)	-0,75 [-2,92; 1,43]	p = 0,48
Interaktionstest										p = 0,06

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

b: Patienten, die während der Studie Steroide bekommen haben

c: Patienten, die während der Studie keine Steroide bekommen haben

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

NA: nicht anwendbar

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für EQ-5D zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

Tabelle 4-94: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach EQ-5D Gesamtpunktzahl - Woche 49										
AFFIRM										
Schmerzpunktzahl < 4	17	81,00 (14,21)	72,92 (18,05)	-4,53 (13,55)	1	80,00 (NA)	88,50 (2,12)	7,00 (NA)	-0,85 [-3,03; 1,33]	p = 0,42
Schmerzpunktzahl ≥ 4	6	56,67 (13,66)	70,71 (23,88)	10,83 (15,63)	1	80,00 (NA)	50,00 (NA)	-30,00 (NA)	2,61 [-0,16; 5,39]	p = 0,06
Interaktionstest										p = 0,02
Gesundheitsbezogene Lebensqualität nach EQ-5D Gesamtpunktzahl – letzte verfügbare Erhebung										
AFFIRM										
Schmerzpunktzahl < 4	92	74,08 (15,95)	67,69 (18,57)	-6,21 (16,00)	41	65,61 (18,46)	62,29 (19,56)	-4,02 (13,90)	-0,14 [-0,51; 0,23]	p = 0,46
Schmerzpunktzahl ≥ 4	32	53,19 (15,63)	55,56 (21,20)	1,81 (19,73)	9	51,11 (22,05)	46,67 (20,62)	-4,44 (18,78)	0,35 [-0,42; 1,11]	p = 0,36
Interaktionstest										p = 0,20

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für EQ-5D zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

Tabelle 4-95: Subgruppenanalyse zur Studienpopulation „Anzahl vorheriger Chemotherapien (1 vs. ≥ 2)“ für den Endpunkt „Gesundheitsbezogene Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn)“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid Mittelwert (SD)**				**Placebo Mittelwert (SD)**				**SMD [95 %-KI]**	**p-Wert**
	N[a]	**Studienbeginn**	**Ende**	**Delta**	**N[a]**	**Studienbeginn**	**Ende**	**Delta**		
Gesundheitsbezogene Lebensqualität nach FACT-G Gesamtpunktzahl – letzte verfügbare Erhebung										
AFFIRM										
1	95	70,40 (18,29)	63,66 (20,87)	-6,73 (17,69)	37	64,14 (19,36)	62,26 (21,49)	-3,19 (13,91)	-0,22 [-0,61; 0,16]	p = 0,25
≥ 2	29	63,07 (17,39)	68,26 (15,60)	4,34 (13,04)	13	59,77 (21,19)	53,00 (15,45)	-6,69 (16,99)	0,67 [-0,01; 1,36]	p = 0,05
Interaktionstest										p = 0,03

a: Anzahl Patienten in der Auswertung und Relevanz für Delta

SD: *Standard Deviation*

SMD: *Standardized Mean Difference*

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der innerhalb der Auswertungspopulation für EQ-5D zum jeweiligen Zeitpunkt und für die jeweilige Fragestellung auswertbaren Patienten.

4.3.1.3.2.1.8 Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-96: Subgruppenanalyse zur Studienpopulation „Gabe von Bisphosphonaten während der Studie (ja vs. nein)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Unerwünschte Ereignisse nach CTCAE Grad ≥ 3					
AFFIRM					
Ja[b]	383	159 (41,5)	183	100 (54,6)	0,76 [0,64; 0,91] p = 0,002
Nein[c]	417	203 (48,7)	216	112 (51,9)	0,94 [0,80; 1,10] p = 0,45
Interaktionstest					p = 0,10

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Biphosphonate bekommen haben
c: Patienten, die während der Studie keine Biphosphonate bekommen haben
RR: Relatives Risiko
KI: Konfidenzintervall
CTCAE: *Common Toxicity Criteria for Adverse Events*
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

Tabelle 4-97: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Unerwünschte Ereignisse nach CTCAE Grad ≥ 3					
AFFIRM					
Ja[b]	382	237 (62,0)	182	119 (65,4)	0,95 [0,83; 1,08] p = 0,43
Nein[c]	418	125 (29,9)	217	93 (42,9)	0,70 [0,56; 0,86] p = 0,0009
Interaktionstest					p = 0,10 p = 0,02[d]

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Steroide bekommen haben
c: Patienten, die während der Studie keine Steroide bekommen haben
d: Berechnet nach Deeks 2010 [Deeks 2010]
RR: Relatives Risiko
KI: Konfidenzintervall
CTCAE: *Common Toxicity Criteria for Adverse Events*
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

Tabelle 4-98: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Unerwünschte Ereignisse nach CTCAE Grad ≥ 3					
AFFIRM					
ECOG = 0 oder 1	730	312 (42,7)	367	190 (51,8)	0,83 [0,73; 0,94] p = 0,004
ECOG = 2	70	50 (71,4)	32	22 (68,8)	1,04 [0,79; 1,37] p = 0,79
Interaktionstest					p = 0,31 p = 0,14[b]

a: Anzahl Patienten in der Auswertung
b: Berechnet nach Deeks 2010 [Deeks 2010]
RR: Relatives Risiko
KI: Konfidenzintervall
CTCAE: *Common Toxicity Criteria for Adverse Events*
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

Tabelle 4-99: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Unerwünschte Ereignisse nach CTCAE Grad ≥ 3					
AFFIRM					
Schmerzpunktzahl < 4	574	228 (39,7)	284	140 (49,3)	0,81 [0,69; 0,94] p = 0,006
Schmerzpunktzahl ≥ 4	226	134 (59,3)	115	72 (62,6)	0,95 [0,79; 1,13] p = 0,55
Interaktionstest					p = 0,37 p = 0,18[b]

a: Anzahl Patienten in der Auswertung
b: Berechnet nach Deeks 2010 [Deeks 2010]
RR: Relatives Risiko
KI: Konfidenzintervall
CTCAE: *Common Toxicity Criteria for Adverse Events*
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

4.3.1.3.2.1.9 Endpunkt „Schwerwiegende unerwünschte Ereignisse“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-100: Subgruppenanalyse zur Studienpopulation „Gabe von Steroiden während der Studie (ja vs. nein)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schwerwiegende unerwünschte Ereignisse					
AFFIRM					
Ja[b]	382	176 (46,1)	182	89 (48,9)	0,94 [0,78; 1,13] p = 0,53
Nein[c]	418	92 (22,0)	217	65 (30,0)	0,74 [0,56; 0,96] p = 0,03
Interaktionstest					p = 0,25 p = 0,14[d]

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Steroide bekommen haben
c: Patienten, die während der Studie keine Steroide bekommen haben
d: Berechnet nach Deeks 2010 [Deeks 2010]
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

Tabelle 4-101: Subgruppenanalyse zur Studienpopulation „ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schwerwiegende unerwünschte Ereignisse					
AFFIRM					
ECOG = 0 oder 1	730	227 (31,1)	367	140 (38,1)	0,82 [0,69; 0,97] p = 0,02
ECOG = 2	70	41 (58,6)	32	14 (43,8)	1,34 [0,86; 2,08] p = 0,19
Interaktionstest					p = 0,04

a: Anzahl Patienten in der Auswertung
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

Tabelle 4-102: Subgruppenanalyse zur Studienpopulation „Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI – *short form* bei Studienbeginn (< 4 vs. ≥ 4)“ für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Schwerwiegende unerwünschte Ereignisse					
AFFIRM					
Schmerzpunktzahl < 4	574	161 (28,0)	284	99 (34,9)	0,81 [0,66; 0,99] p = 0,04
Schmerzpunktzahl ≥ 4	226	107 (47,3)	115	55 (47,8)	0,99 [0,78; 1,25] p = 0,93
Interaktionstest					p = 0,28 p = 0,19[b]

a: Anzahl Patienten in der Auswertung

b: Berechnet nach Deeks 2010 [Deeks 2010]

RR: Relatives Risiko

KI: Konfidenzintervall

Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

4.3.1.3.2.1.10 Endpunkt „Tod aufgrund unerwünschter Ereignisse“ – Subgruppenanalysen zur Studienpopulation – RCT

Tabelle 4-103: Subgruppenanalyse zur Studienpopulation „Gabe von Bisphosphonaten während der Studie (ja vs. nein)“ für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

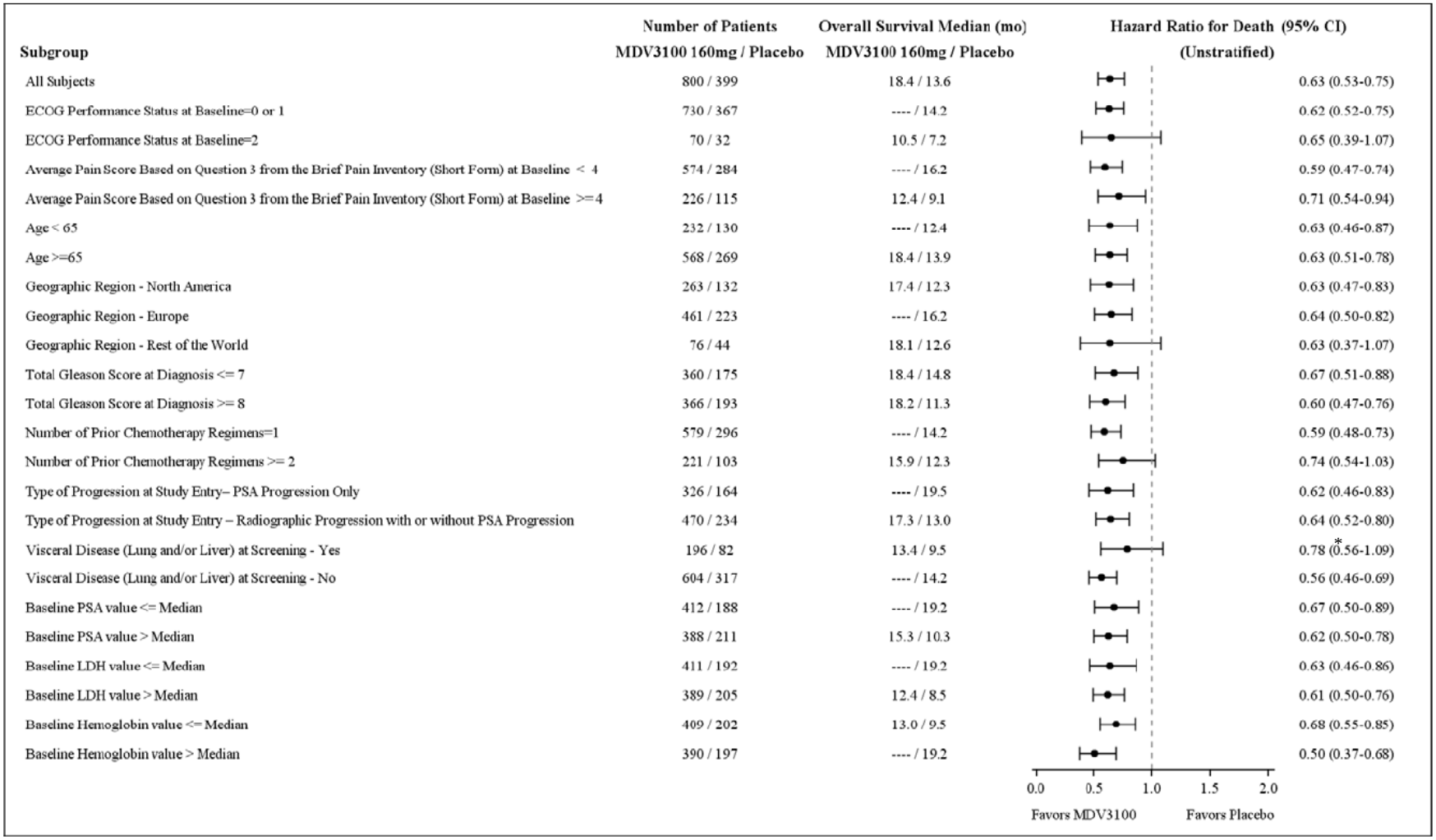
Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**RR [95 %-KI]** **p-Wert**
Tod aufgrund unerwünschter Ereignisse					
AFFIRM					
Ja[b]	383	4 (1,0)	183	7 (3,8)	0,27 [0,08; 0,92] p = 0,04
Nein[c]	417	19 (4,6)	216	7 (3,2)	1,41 [0,60; 3,29] p = 0,43
Interaktionstest					p = 0,03

a: Anzahl Patienten in der Auswertung
b: Patienten, die während der Studie Biphosphonate bekommen haben
c: Patienten, die während der Studie keine Biphosphonate bekommen haben
RR: Relatives Risiko
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Sicherheitspopulation.

4.3.1.3.2.2 *A priori* definierte Subgruppenanalysen zur Studienpopulation

4.3.1.3.2.2.1 Endpunkt „Gesamtüberleben“ – *a priori* definierte Subgruppenanalysen zur Studienpopulation – RCT

Für den Endpunkt „Gesamtüberleben“ wurden bei der Auswertung der Studie zusätzlich zu den im vorherigen Abschnitt bereits berichteten Subgruppenmerkmalen (Abschnitt 4.3.1.3.2.1) die prädefinierten Effektmodifikatoren „Hämoglobinwert zu Studienbeginn“, „*Gleason Score* zu Studienbeginn“, „LDH-Wert zu Studienbeginn“, „PSA-Wert zu Studienbeginn“, „Progressionsart zu Studienbeginn“ sowie die „geographische Region“ auf ihren Einfluss auf das Gesamtergebnis hin untersucht. In Abbildung 2 findet sich eine Übersicht über alle im Rahmen der Studie für den Endpunkt „Gesamtüberleben“ durchgeführten Subgruppenanalysen mit Angaben zur Anzahl der Patienten, dem medianen Gesamtüberleben und der Hazard Ratio für die jeweiligen Kategorien der Subgruppen. Für die prädefinierten Subgruppenmerkmale wurden zusätzlich Interaktionsterme berechnet (siehe Tabelle 4-104) und bei Hinweisen auf oder Belegen für Interaktionen wurden die Ergebnisse der Analysen im Detail dargestellt.



Subgroup	Number of Patients MDV3100 160mg / Placebo	Overall Survival Median (mo) MDV3100 160mg / Placebo	Hazard Ratio for Death (95% CI) (Unstratified)
All Subjects	800 / 399	18.4 / 13.6	0.63 (0.53-0.75)
ECOG Performance Status at Baseline=0 or 1	730 / 367	---- / 14.2	0.62 (0.52-0.75)
ECOG Performance Status at Baseline=2	70 / 32	10.5 / 7.2	0.65 (0.39-1.07)
Average Pain Score Based on Question 3 from the Brief Pain Inventory (Short Form) at Baseline < 4	574 / 284	---- / 16.2	0.59 (0.47-0.74)
Average Pain Score Based on Question 3 from the Brief Pain Inventory (Short Form) at Baseline >=4	226 / 115	12.4 / 9.1	0.71 (0.54-0.94)
Age < 65	232 / 130	---- / 12.4	0.63 (0.46-0.87)
Age >=65	568 / 269	18.4 / 13.9	0.63 (0.51-0.78)
Geographic Region - North America	263 / 132	17.4 / 12.3	0.63 (0.47-0.83)
Geographic Region - Europe	461 / 223	---- / 16.2	0.64 (0.50-0.82)
Geographic Region - Rest of the World	76 / 44	18.1 / 12.6	0.63 (0.37-1.07)
Total Gleason Score at Diagnosis <= 7	360 / 175	18.4 / 14.8	0.67 (0.51-0.88)
Total Gleason Score at Diagnosis >= 8	366 / 193	18.2 / 11.3	0.60 (0.47-0.76)
Number of Prior Chemotherapy Regimens=1	579 / 296	---- / 14.2	0.59 (0.48-0.73)
Number of Prior Chemotherapy Regimens >= 2	221 / 103	15.9 / 12.3	0.74 (0.54-1.03)
Type of Progression at Study Entry– PSA Progression Only	326 / 164	---- / 19.5	0.62 (0.46-0.83)
Type of Progression at Study Entry – Radiographic Progression with or without PSA Progression	470 / 234	17.3 / 13.0	0.64 (0.52-0.80)
Visceral Disease (Lung and/or Liver) at Screening - Yes	196 / 82	13.4 / 9.5	0.78 (0.56-1.09)*
Visceral Disease (Lung and/or Liver) at Screening - No	604 / 317	---- / 14.2	0.56 (0.46-0.69)
Baseline PSA value <= Median	412 / 188	---- / 19.2	0.67 (0.50-0.89)
Baseline PSA value > Median	388 / 211	15.3 / 10.3	0.62 (0.50-0.78)
Baseline LDH value <= Median	411 / 192	---- / 19.2	0.63 (0.46-0.86)
Baseline LDH value > Median	389 / 205	12.4 / 8.5	0.61 (0.50-0.76)
Baseline Hemoglobin value <= Median	409 / 202	13.0 / 9.5	0.68 (0.55-0.85)
Baseline Hemoglobin value > Median	390 / 197	---- / 19.2	0.50 (0.37-0.68)

Abbildung 2: Ergebnisse der im CSR dargestellten Subgruppenanalysen für den Endpunkt „Gesamtüberleben“.

*Der numerische Unterschied zu dem in Tabelle 4-77 angegebenen HR (95 %-KI) für die Subgruppe „Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des *Screenings* (ja vs. nein)“ basiert auf unterschiedlicher Rundung.

Tabelle 4-104: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests für prädefinierte Subgruppen-Analysen zum Endpunkt „Gesamtüberleben“ der randomisiert kontrollierten Studie AFFIRM

Subgruppen **Endpunkte**	**Hämoglobin zu Studienbeginn**	**Gesamt Gleason Score zu Studienbeginn**	**LDH zu Studienbeginn**	**PSA zu Studienbeginn**	**Progressionstyp zu Studienbeginn**	**Geographische Region [a]**	**Geographische Region [b]**
Gesamtüberleben	0,14[c]	0,54	0,77	0,69	0,89	0,88	0,88

a: Europa / Nordamerika
b: Europa / übrige Welt
c: Hinweis auf Interaktion
LDH: Laktat-Dehydrogenase
PSA: Prostata-spezifisches Antigen

Weitere Interaktionsterme, die sich auf den Endpunkt “Gesamtüberleben“ beziehen, sind in der Tabelle 4-75 berichtet.

Tabelle 4-105: Subgruppenanalyse zur Studienpopulation „Hämoglobin zu Studienbeginn“ für den Endpunkt „Gesamtüberleben“ in RCT mit dem zu bewertenden Arzneimittel

Endpunkt **Studie** **Subgruppe**	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Hazard Ratio [95 % KI]** **p-Wert**
Gesamtüberleben					
AFFIRM					
Hämoglobinwert ≤ Median	409	220 (53,8)	202	135 (66,8)	0,68 [0,55; 0,85] p = 0,0005
Hämoglobinwert > Median	390	87 (22,3)	197	77 (39,1)	0,50 [0,37; 0,68] p < 0,0001
Interaktionstest					p = 0,14

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der ITT-Population.

4.3.1.3.2.2.2 Endpunkt „Sicherheitsbezogene Ereignisse“ – *a priori* definierte Subgruppenanalysen zur Studienpopulation – RCT

Die sicherheitsbezogenen Ereignisse UE (unter der Behandlung auftretend), SUE (unter der Behandlung auftretend), Studienabbruch aufgrund UE (basierend auf dem Grund für den Behandlungsabbruch laut dem Dokumentationsbogen für Behandlungsabbruch), Studienabbruch aufgrund UE (basierend auf der bezüglich der Studienmedikation ergriffenen Massnahme laut dem Dokumentationsbogen für UEs) und Tod aufgrund UE (unter der Behandlung auftretend) wurden zudem *a priori* in folgenden Subgruppen ausgewertet: „Alter (< 65, ≥ 65, ≥ 75)“; „Anamnese einer Hypertonie (ja vs. nein)“ und „Anamnese kardiovaskulärer Erkrankungen (ja vs. nein)“.

Für die prädefinierten Subgruppenmerkmale wurden zusätzlich Interaktionsterme berechnet (siehe Tabelle 4-106) und bei Hinweisen auf oder Belegen für Interaktionen wurden die Ergebnisse der Analysen im Detail dargestellt.

Tabelle 4-106: Enzalutamid vs. Placebo Zusammenfassung der Interaktionstests für prädefinierte Subgruppen-Analysen zu sicherheitsbezogenen Ereignissen, die unter der Behandlung auftraten, aus der randomisiert kontrollierten Studie AFFIRM

Subgruppen **Endpunkte**	**Alter (< 75 vs. ≥ 75)**	**Anamnese einer Hypertonie (ja vs. nein)**	**Anamnese kardiovaskulärer Erkrankungen (ja vs. nein)**
UE (unter der Behandlung auftretend)	p = 0,03[b]	p = 0,18[a]	p = 0,24
SUE (unter der Behandlung auftretend)	p = 0,51	p = 0,37	p = 0,30
Studienabbruch aufgrund UE (basierend auf dem Grund für den Behandlungsabbruch laut dem Dokumentationsbogen für Behandlungsabbruch)	p = 0,91	p = 0,52	p = 0,84
Studienabbruch aufgrund UE (basierend auf der bezüglich der Studienmedikation ergriffenen Massnahme laut dem Dokumentationsbogen für UEs)	p = 0,63	p = 0,92	p = 0,62
Tod aufgrund UE (unter der Behandlung auftretend)	p = 0,98	p = 0,85	p = 0,21
a: Hinweis auf Interaktion b: Beleg für Interaktion			

Tabelle 4-107 Subgruppenanalyse zur Studienpopulation nach „Alter (< 75 vs. ≥ 75)“ für sicherheitsbezogene Ereignisse in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Relatives Risiko [95 %-KI] p-Wert**
UE (unter der Behandlung auftretend)					
Alter < 75	601	587 (97,7)	295	290 (98,3)	0,99 (0,97; 1,01) p = 0,51
Alter ≥ 75	199	198 (99,5)	104	100 (96,2)	1,03 (0,99; 1,08) p = 0,09
Interaktionstest					p = 0,03

a: Anzahl Patienten in der Auswertung
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Safety-Population.

Tabelle 4-108: Subgruppenanalyse zur Studienpopulation nach „Anamnese einer Hypertonie (ja vs. nein)“ für sicherheitsbezogene Ereignisse in RCT mit dem zu bewertenden Arzneimittel

Endpunkt Studie Subgruppe	**Enzalutamid**		**Placebo**		**Enzalutamid vs. Placebo**
	N[a]	**Patienten mit Ereignis n (%)**	**N[a]**	**Patienten mit Ereignis n (%)**	**Relatives Risiko [95 %-KI] p-Wert**
UE (unter der Behandlung auftretend)					
Ja[b]	425	421 (99,1)	219	214 (97,7)	1,01 (0,99; 1,04) p = 0,23
Nein[c]	375	364 (97,1)	180	176 (97,8)	0,99 (0,97; 1,02) p = 0,61
Interaktionstest					p = 0,18

a: Anzahl Patienten in der Auswertung
b: Patienten, die zu Studienbeginn unter einer Hypertonie litten
c: Patienten, die zu Studienbeginn nicht unter einer Hypertonie litten
KI: Konfidenzintervall
Alle Angaben beziehen sich auf die Analyse der Safety-Population.

4.3.1.3.2.3 Zusammenfassung der Ergebnisse der Subgruppenanalysen aus RCT

Wie bereits oben im Abschnitt 4.3.1.3.2.1.1 beschrieben, wurden bei Hinweisen auf oder Belegen für unterschiedliche Effekte in den Subgruppen neben dem Gesamteffekt auch die einzelnen Subgruppenergebnisse in die Interpretation der Ergebnisse einbezogen. [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011]

Tabelle 4-109 zeigt die Ergebnisse für alle Endpunkte, bei denen ein Hinweis oder ein Beleg auf eine Interaktion vorlag. (Sämtliche Analysen zur Untersuchung der Effektmodifikatoren sind vollständig im Modul 5 im Ordner Studienberichte gespeichert.) Diese werden im Folgenden zusammengefasst dargestellt:

Ergebnisse der *a posteriori* definierten Subgruppenanalysen

Schlussfolgerungen aus beobachteten Hinweisen/Belegen unterliegen den bekannten Einschränkungen für *post hoc*-Analysen und multiplem Testen. Zudem war die Fallzahl der Studie nicht dafür berechnet, Effekte in Subgruppen zu zeigen.

Weiterhin ist zu beachten, dass die Subgruppenmerkmale viszerale Erkrankung (ja/nein) und Steroidgabe (ja/nein) bereits mit einer unterschiedlichen Prognose für die Patienten verbunden sind. Patienten mit viszeraler Erkrankung weisen *per se* eine schlechtere Prognose als Patienten ohne diese Erkrankung auf. Die Steroidgabe durch den Arzt stellt gleichzeitig einen Hinweis auf einen schlechteren Allgemeinzustand des Patienten dar.

Hinsichtlich des Endpunkts „Gesamtüberleben“ fand sich für die Subgruppen „Gabe von Steroiden während der Studie“ und „viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings“ jeweils ein Hinweis auf eine Interaktion (Tabelle 4-109). Während in der Subgruppe „Gabe von Steroiden“ beide Arme der Subgruppe einen signifikanten Vorteil einer Enzalutamid-Gabe gegenüber einer Placebo-Gabe aufwiesen (Gabe von Steroiden - Ja: HR = 0,70; 95 %-KI [0,56; 0,88], p = 0,002; Gabe von Steroiden – Nein: HR = 0,49; 95 %-KI [0,37; 0,65], p < 0,0001), profitierten in der Subgruppe „viszerale Erkrankung zum Zeitpunkt des Screenings“ vor allem die Patienten ohne vorherige Erkrankung von einer Enzalutamid-Therapie (Viszerale Erkrankung – Ja: HR = 0,78; 95 %-KI [0,56; 1,09], p = 0,15; viszerale Erkrankung – Nein: HR = 0,57; 95 %-KI [0,46; 0,70], p < 0,0001).

Für den Endpunkt „radiographisch progressionsfreies Überleben“ konnten für die Subgruppen „Gabe von Steroiden während der Studie“ und „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ jeweils Belege für eine Interaktion und für die Subgruppen „Alter“ und „Anzahl vorheriger Chemotherapien“ jeweils Hinweise auf eine Interaktion erbracht werden (Tabelle 4-109). Allerdings erwies sich innerhalb dieser Subgruppen jeweils für beide Kategorien eine Enzalutamid-Therapie als signifikant überlegen gegenüber einer Placebo-Therapie. (Gabe von Steroiden – Ja: HR = 0,60; 95 %-KI [0,49; 0,73], p < 0,0001; Gabe von Steroiden – Nein: HR = 0,26; 95 %-KI [0,21; 0,33], p < 0,0001; Schmerzpunktzahl < 4: HR = 0,34; 95 %-KI [0,28; 0,40], p < 0,0001; Schmerzpunktzahl ≥ 4 HR = 0,58; 95 %-KI [0,45; 0,75], p < 0,0001; Alter < 65:

HR = 0,46; 95 %-KI [0,36; 0,59], p < 0,0001; Alter ≥ 65: HR = 0,38; 95 %-KI [0,32; 0,45], p < 0,0001; Anzahl vorheriger Chemotherapien = 1: HR = 0,38; 95 %-KI [0,32; 0,44], p < 0,0001; Anzahl vorheriger Chemotherapien ≥ 2: HR = 0,48; 95 %-KI [0,36; 0,63], p < 0,0001).

Bezüglich des Endpunkts „Zeit bis zur ersten skelettbezogenen Komplikation" konnte für die Subgruppen „Gabe von Steroiden während der Studie" und „Alter" jeweils ein Hinweis auf eine Interaktion festgestellt werden (Tabelle 4-109). Innerhalb der Subruppe „Gabe von Steroiden" war die Therapie mit Enzalutamid vor allem für Patienten ohne begleitende Steroidgabe von Vorteil gegenüber einer Placebo-Gabe (Gabe von Steroiden – ja: HR = 0,78; 95 %-KI [0,61; 1,01], p = 0,06; Gabe von Steroiden – nein: HR = 0,56; 95 %-KI [0,41; 0,75], p = 0,0001). Bezüglich der Subgruppe „Alter" fand sich für beide Alterskategorien (< 65 und ≥ 65 Jahre) ein statistisch signifikanter Vorteil für Enzalutamid gegenüber Placebo (Alter < 65: HR = 0,60; 95 %-KI [0,44; 0,83], p = 0,002; Alter ≥ 65: HR = 0,77; 95 %-KI [0,60; 0,99], p = 0,04).

Für den Endpunkt „Schmerzprogressionsrate" fand sich in der *Responder*-Analyse für die Subgruppe „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn" ein Beleg für eine Interaktion und für die Subgruppen „Gabe von Steroiden während der Studie" und „Alter" jeweils ein Hinweis auf eine Interaktion (Tabelle 4-109). Im Einzelnen erwies sich Enzalutamid im Vergleich zu Placebo dabei für Patienten mit einer duchschnittlichen Schmerzpunktzahl ≥ 4 (Schmerzpunktzahl < 4: RR = 0,85 [0,67; 1,06], p = 0,14; Schmerzpunktzahl ≥ 4: RR = 0,39; 95 %-KI [0,26; 0,60], p < < 0,0001), für Patienten ohne begleitende Steroidgabe (Gabe von Steroiden – ja: RR = 0,84; 95 %-KI [0,64; 1,10], p = 0,20; Gabe von Steroiden – nein: RR = 0,59; 95 %-KI [0,44; 0,79], p = 0,0004) und für Patienten < 65 Jahre (Alter < 65: RR = 0,59; 95 %-KI [0,42; 0,83], p = 0,002; Alter ≥ 65: RR = 0,79; 95 %-KI [0,61;1,01], p = 0,06) als statistisch signifikant überlegen.

Bezüglich der Auswertung des Endpunkts „Schmerzprogressionsrate" als Veränderung gegenüber Studienbeginn fand sich für die Subgruppen „ECOG-*Performance*-Status zu Studienbeginn" und „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn" jeweils ein Hinweis auf eine Interaktion (Tabelle 4-109). Innerhalb dieser beiden Subgruppen ergab sich jedoch jeweils für beide Kategorien ein statistisch signifikanter Vorteil für Enzalutamid gegenüber Placebo (ECOG-*Performance*-Status 0 oder 1: SMD = -0,46; 95 %-KI [-0,61; -0,31], p < 0,0001; ECOG-*Performance*-Status 2: SMD = -0,68; 95 %-KI [-1,22; -0,14], p = 0,02; Schmerzpunktzahl < 4: SMD = -0,46; 95 %-KI [-0,62; -0,30], p < 0,0001; Schmerzpunktzahl ≥ 4: SMD = -0,58; 95 %-KI [-0,88; -0,28], p < 0,001).

Bezüglich des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P" (*Responder*-Analyse) konnte für die Subgruppe „Gabe von Steroiden während der Studie" ein Hinweis auf eine Interaktion festgestellt werden (Tabelle 4-109). Allerdings erwies sich sowohl für Patienten, welche studienbegleitend Steroide bekamen, als auch für Patienten, welche keine Steroide bekamen, eine Enzalutamid-Therapie hinsichtlich der Auswertung des Endpunkts als *Responder*-Analyse signifikant überlegen gegenüber einer Placebo-Gabe (Gabe von Steroiden

– ja: RR = 2,10; 95 %-KI [1,38; 3,19], p < 0,0005; Gabe von Steroiden – nein: RR = 3,69; 95 %-KI [2,29; 5,95], p < 0,0001).

Hinsichtlich des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ (Veränderung der Gesamtpunktzahl FACT-P bzw. FACT-G der letzten verfügbaren Erhebung gegenüber Studienbeginn) ergab sich für die Subgruppe „Gabe von Steroiden während der Studie“ sowohl für die Betrachtung des FACT-P als auch für die Betrachtung des FACT-G ein Beleg für eine Interaktion. Weiterhin fand sich für die Subgruppe „ECOG-*Performance*-Status zu Studienbeginn“ bei Betrachtung des FACT-P ein Hinweis auf Interaktion (Tabelle 4-109). Im Rahmen der Subgruppe „Gabe von Steroiden während der Studie“ erwies sich Enzalutamid sowohl für die Betrachtung des FACT-P als auch des FACT-G im Vergleich zu Placebo für Patienten ohne begleitende Steroidgabe als signifikant überlegen (FACT-P: Gabe von Steroiden – ja: SMD = 0,22; 95 %-KI [-0,01; 0,45], p = 0,06; Gabe von Steroiden – nein: SMD = 0,61; 95 %-KI [0,41; 0,81], p < 0,0001; FACT-G: Gabe von Steroiden – ja: SMD = 0,19; 95 %-KI [-0,04; 0,41], p = 0,10; Gabe von Steroiden – nein: SMD = 0,53; 95 %-KI [0,33; 0,72], p < 0,0001). Für die Subgruppe „ECOG-*Performance*-Status zu Studienbeginn“ (Betrachtung FACT-P) war Enzalutamid in beiden Armen der Subgruppe von signifikantem Vorteil gegenüber einer Placebo-Therapie (ECOG-*Performance*-Status 0 oder 1: SMD = 0,38; 95 %-KI [0,23; 0,54], p < 0,0001; ECOG-*Performance*-Status 2: SMD = 0,86; 95 %-KI [0,25; 1,47], p = 0,006).

Hinsichtlich der Auswertung des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ (Veränderung der Gesamtpunktzahl des FACT-P bzw. FACT-G in Woche 49 gegenüber Studienbeginn) fand sich sowohl bei der Betrachtung der FACT-P- als auch der FACT-G-Gesamtpunktzahl für die Subgruppe „Gabe von Steroiden während der Studie“ ein Hinweis auf eine Interaktion und für die Subgruppe „Alter“ ein Beleg für eine Interaktion (Tabelle 4-109). In beiden Subgruppen zeigten sich für jeweils beide Betrachtungen in den jeweiligen Kategorien allerdings keine signifikanten Unterschiede zwischen den beiden Behandlungsarmen (FACT-P: Gabe von Steroiden – ja: SMD = -0,09; 95 %-KI [-0,77; 0,59], p = 0,79; Gabe von Steroiden – nein: SMD = 0,57; 95 %-KI [-0,08; 1,22], p = 0,08; Alter < 65: SMD = -0,86; 95 %-KI [-2,08; 0,37], p = 0,17; Alter ≥ 65: SMD = 0,47; 95 %-KI [-0,03; 0,96], p = 0,06; FACT-G: Gabe von Steroiden – ja: SMD = -0,13; 95 %-KI [-0,77; 0,52], p = 0,70; Gabe von Steroiden – nein: SMD = 0,49; 95 %-KI [-0,16; 1,13], p = 0,14; Alter < 65: SMD = -0,67; 95 %-KI [-1,72; 0,38], p = 0,21; Alter ≥ 65: SMD = 0,41; 95 %-KI [-0,08; 0,90], p = 0,10).

Für den Endpunkt „Gesundheitsgezogene Lebensqualität nach EQ-5D“ ergab sich bei der Betrachtung der letzten verfügbaren Erhebung gegenüber Studienbeginn für die Subgruppe „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ ein Hinweis auf eine Interaktion und für die Subgruppe „Anzahl vorheriger Chemotherapien“ ein Beleg für eine Interaktion (Tabelle 4-109). Allerdings ergaben sich für die jeweiligen Kategorien der beiden Subgruppen keine signifikanten Unterschiede zwischen den Therapiearmen (Schmerzpunktzahl < 4: SMD = -0,14; 95 %-KI [-0,51; 0,23], p = 0,46; Schmerzpunktzahl ≥ 4: SMD = 0,35; 95 %-KI [-0,42; 1,11], p = 0,36; Anzahl vorheriger

Chemotherapien = 1: SMD = -0,22; 95 %-KI [-0,61; 0,16], p = 0,25; Anzahl vorheriger Chemotherapien ≥ 2: SMD = 0,67; 95 %-KI [-0,01; 1,36], p = 0,05).

Bei der Betrachtung der Veränderung der Gesamtpunktzahl des EQ-5D in Woche 49 gegenüber Studienbeginn fand sich für die Subgruppe „Gabe von Steroiden während der Studie“ ein Hinweis auf eine Interaktion und für die Subgruppe „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ ein Beleg für eine Interaktion (Tabelle 4-109). Allerdings zeigten sich innerhalb der jeweiligen Kategorien dieser Subgruppen keine signifikanten Unterschiede zwischen den Therapiearmen (Gabe von Steroiden – ja: SMD = NA [NA; NA], p = NA; Gabe von Steroiden – nein: SMD = -0,75; 95 %-KI [-2,92; 1,43], p = 0,48; Schmerzpunktzahl < 4: SMD = -0,85; 95 %-KI [-3,03; 1,33], p = 0,42; Schmerzpunktzahl ≥ 4: SMD = 2,61; 95 %-KI [-0,16; 5,39], p = 0,06).

Hinsichtlich des Endpunkts „unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ ergab sich für die Subgruppen „Gabe von Bisphosphonaten während der Studie“ sowie „Gabe von Steroiden während der Studie“ jeweils ein Hinweis für eine Interaktion. Bei der Berechnung der Interaktionsterme mittels *Review Manager, Version 5.1* [Deeks 2010] wurde der Hinweis auf Interaktion in der Subgruppe „Gabe von Steroiden während der Studie“ zu einem Beleg und zusätzlich ergaben sich Hinweise auf Interaktionen für die Subgruppen „ECOG-*Performance*-Status zu Studienbeginn“ und „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ (Tabelle 4-109). Innerhalb der Subgruppe „Gabe von Bisphosphonaten“ war Enzalutamid bei Patienten mit begleitender Bisphosphonatgabe von signifikantem Vorteil (Bisphosphonate – ja: RR = 0,76; 95 %-KI [0,64; 0,91], p = 0,002 ; Bisphosphonate – nein: RR = 0,94; 95 %-KI [0,80; 1,10], p = 0,45). Im Rahmen der Subgruppe „Gabe von Steroiden“ erwies sich Enzalutamid bei Patienten ohne Steroide als Begleitmedikation als signifikant überlegen (Gabe von Steroiden – ja: RR = 0,95; 95 %-KI [0,83; 1,08], p = 0,43 ; Gabe von Steroiden – nein: RR = 0,70; 95 %-KI [0,56; 0,86], p = 0,0009). Unter Einbezug der Berechnung nach Deeks ergab sich unter Enzalutamid-Therapie innerhalb der Subgruppe „ECOG-*Performance*-Status“ ein Vorteil für Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 (ECOG-*Performance*-Status 0 oder 1: RR = 0,83; 95 %-KI [0,73; 0,94], p = 0,004 ; ECOG-*Performance*-Status 2: RR = 1,04; 95 %-KI [0,79; 1,37], p = 0,79) und innerhalb der Subgruppe „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ ein Vorteil für Patienten mit einer Schmerzpunktzahl < 4 (Schmerzpunktzahl < 4: RR = 0,81; 95 %-KI [0,69; 0,94], p = 0,006 ; Schmerzpunktzahl ≥ 4: RR = 0,95; 95 %-KI [0,79; 1,13], p = 0,55).

In Bezug auf den Endpunkt „schwerwiegende unerwünschte Ereignisse“ konnte für die Subgruppe „ECOG-*Performance*-Status zu Studienbeginn“ ein Beleg auf Interaktion festgestellt werden. Bei Betrachtung dieser Subgruppe erwies sich die Behandlung mit Enzalutamid für Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 einer Behandlung mit Placebo als signifikant überlegen (ECOG-*Performance*-Status von 0 oder 1: RR = 0,82; 95 %-KI [0,69; 0,97], p = 0,02; ECOG-*Performance*-Status 2: RR = 1,34; 95 %-KI [0,86; 2,08], p = 0,19). Bei der Berechnung der Interaktionsterme mittels *Review Manager, Version 5.1* [Deeks 2010] ergab sich zusätzlich für die Subgruppen „Gabe von Steroiden

während der Studie“ und „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ jeweils ein Hinweis auf eine Interaktion (Tabelle 4-109). Hierbei zeigte sich innerhalb der Subgruppe „Gabe von Steroiden“ für Patienten ohne begleitende Steroidgabe (Gabe von Steroiden – ja: RR = 0,94; 95 %-KI [0,78; 1,13], p = 0,53; Gabe von Steroiden – nein: RR = 0,74; 95 %-KI [0,56; 0,96], p = 0,03) und innerhalb der Subgruppe „durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn“ für Patienten mit einer Schmerzpunktzahl < 4 (Schmerzpunktzahl < 4: RR = 0,81; 95 %-KI [0,66; 0,99], p = 0,04; Schmerzpunktzahl ≥ 4: RR = 0,99; 95 %-KI [0,78; 1,25], p = 0,93) ein signifikanter positiver Effekt von Enzalutamid gegenüber Placebo.

Für den Endpunkt „Tod aufgrund unterwünschter Ereignisse“ fand sich für die Subgruppe „Gabe von Bisphosphonaten während der Studie“ ein Beleg für eine Interaktion (Tabelle 4-109). Bei Patienten, welche während der Studie Bisphosphonate bekamen, war das relative Risiko aufgrund eines unerwünschten Ereignisses zu sterben unter Enzalutamid-Therapie signifikant erniedrigt (Gabe von Bisphosphonaten – ja: RR = 0,27; 95 %-KI [0,08; 0,92], p = 0,04; Gabe von Bisphosphonaten – nein: RR = 1,41; 95 %-KI [0,60; 3,29], p = 0,43).

Ergebnisse der zusätzlichen, *a priori* definierten Subgruppenanalysen

Für die *a priori* definierten potentiellen Effektmodifikatoren zeigte sich für das „Gesamtüberleben“ ein Hinweis auf Interaktion für die Subgruppe „Hämoglobinwert zu Studienbeginn“. Dieser hatte allerdings keinen Einfluss auf das Ergebnis der Hauptanalyse, da beide Kategorien der Subgruppe einen signifikanten Vorteil für Patienten im Enzalutamid-Arm zeigten. Im Rahmen der Auswertung der „sicherheitsbezogenen Ereignisse“ ergab sich für die „Unerwünschten Ereignisse (unter der Behandlung auftretend)“ für die Subgruppe „Alter (< 75 vs. ≥ 75)“ ein Beleg für eine Interaktion und für die Subgruppe „Anamnese einer Hypertonie (ja vs. nein)“ ein Hinweis auf eine Interaktion. Allerdings zeigten sich jeweils in beiden Kategorien der Subgruppen keine signifikanten Unterschiede zwischen den Behandlungsarmen.

Tabelle 4-109: Zusammenfassung der Effektschätzer für Subgruppen, bei denen ein Beleg oder ein Hinweis auf eine Interaktion vorlagen, aus der randomisiert kontrollierten Studie AFFIRM

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, [95 %-KI]** **p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Gesamtüberleben			
Gabe von Steroiden während der Studie	Ja	HR = 0,70 [0,56; 0,88], p = 0,002	p = 0,06
	Nein	HR = 0,49 [0,37; 0,65], p < 0,0001	Hinweis
Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings	Ja	HR = 0,78 [0,56; 1,09], p = 0,15	p = 0,15
	Nein	HR = 0,57 [0,46; 0,70], p < 0,0001	Hinweis
Hämoglobin zu Studienbeginn[b]	≤ Median	HR = 0,68 [0,55; 0,85], p = 0,0005	p = 0,14
	> Median	HR = 0,50 [0,37; 0,68], p < 0,0001	Hinweis
Radiographisches progressionsfreies Überleben			
Gabe von Steroiden während der Studie	Ja	HR = 0,60 [0,49; 0,73], p < 0,0001	p < 0,0001
	Nein	HR = 0,26 [0,21; 0,33], p < 0,0001	Beleg
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	HR = 0,34 [0,28; 0,40], p < 0,0001	p < 0,0015
	≥ 4	HR = 0,58 [0,45; 0,75], p < 0,0001	Beleg
Alter	< 65	HR = 0,46 [0,36; 0,59], p < 0,0001	p = 0,18
	≥ 65	HR = 0,38 [0,32; 0,45], p < 0,0001	Hinweis
Anzahl vorheriger Chemotherapien	1	HR = 0,38 [0,32; 0,44], p < 0,0001	p = 0,10
	≥ 2	HR = 0,48 [0,36; 0,63], p < 0,0001	Hinweis
Zeit bis zur ersten skelettbezogenen Komplikation			
Gabe von Steroiden während der Studie	Ja	HR = 0,78 [0,61; 1,01], p = 0,06	p = 0,12
	Nein	HR = 0,56 [0,41; 0,75], p = 0,0001	Hinweis
Alter	< 65	HR = 0,60 [0,44; 0,83], p = 0,002	p = 0,20

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, [95 %-KI]** **p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
	≥ 65	HR = 0,77 [0,60; 0,99], p = 0,04	Hinweis
Schmerzprogressionsrate (Responder-Analyse)			
Gabe von Steroiden während der Studie	Ja	RR = 0,84 [0,64; 1,10], p = 0,20	p = 0,15
	Nein	RR = 0,59 [0,44; 0,79], p = 0,0004	Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,85 [0,67; 1,06], p = 0,14	p = 0,003
	≥ 4	RR = 0,39 [0,26; 0,60], p < 0,0001	Beleg
Alter	< 65	RR = 0,59 [0,42; 0,83], p = 0,002	p = 0,16
	≥ 65	RR = 0,79 [0,61;1,01], p = 0,06	Hinweis
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)			
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	SMD = -0,46 [-0,61; -0,31], p < 0,0001	p = 0,14
	2	SMD = -0,68 [-1,22; -0,14], p = 0,02	Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	SMD = -0,46 [-0,62; -0,30], p < 0,0001	p = 0,10
	≥ 4	SMD = -0,58 [-0,88; -0,28], p < 0,001	Hinweis
Lebensqualität nach FACT-P (Responder-Analyse)			
Gabe von Steroiden während der Studie	Ja	RR = 2,10 [1,38; 3,19], p < 0,0005	p = 0,09
	Nein	RR = 3,69 [2,29; 5,95], p < 0,0001	Hinweis
Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – Woche 49			
Gabe von Steroiden während der Studie	Ja	SMD = -0,09 [-0,77; 0,59], p = 0,79	p = 0,16
	Nein	SMD = 0,57[-0,08; 1,22], p = 0,08	Hinweis
Alter	< 65	SMD = -0,86 [-2,08; 0,37], p = 0,17	p = 0,02
	≥ 65	SMD = 0,47 [-0,03; 0,96], p = 0,06	Beleg
Lebensqualität nach FACT-P (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, [95 %-KI]** **p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Gabe von Steroiden während der Studie	Ja	SMD = 0,22 [-0,01; 0,45], p = 0,06	p = 0,03
	Nein	SMD = 0,61 [0,41; 0,81], p < 0,0001	Beleg
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	SMD = 0,38 [0,23; 0,54], p < 0,0001	p = 0,10
	2	SMD = 0,86 [0,25; 1,47], p = 0,006	Hinweis
Lebensqualität nach FACT-G (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – Woche 49			
Gabe von Steroiden während der Studie	Ja	SMD = -0,13 [-0,77; 0,52], p = 0,70	p = 0,17
	Nein	SMD = 0,49 [-0,16; 1,13], p = 0,14	Hinweis
Alter	< 65	SMD = -0,67 [-1,72; 0,38], p = 0,21	p = 0,04
	≥ 65	SMD = 0,41 [-0,08; 0,90], p = 0,10	Beleg
Lebensqualität nach FACT-G (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			
Gabe von Steroiden während der Studie	Ja	SMD = 0,19 [-0,04; 0,41], p = 0,10	p = 0,05
	Nein	SMD = 0,53 [0,33; 0,72], p < 0,0001	Beleg
Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – Woche 49			
Gabe von Steroiden während der Studie	Ja	SMD = NA [NA; NA], p = NA	p = 0,06
	Nein	SMD = -0,75 [-2,92; 1,43], p = 0,48	Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	SMD = -0,85 [-3,03; 1,33], p = 0,42	p = 0,02
	≥ 4	SMD = 2,61 [-0,16; 5,39], p = 0,06	Beleg
Lebensqualität nach EQ-5D (Veränderung gegenüber Studienbeginn) Gesamtpunktzahl – letzte verfügbare Erhebung			
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	SMD = -0,14 [-0,51; 0,23], p = 0,46	p = 0,20
	≥ 4	SMD = 0,35 [-0,42; 1,11], p = 0,36	Hinweis
Anzahl vorheriger Chemotherapien	1	SMD = -0,22 [-0,61; 0,16], p = 0,25	p = 0,03
	≥ 2	SMD = 0,67 [-0,01; 1,36], p = 0,05	Beleg

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, [95 %-KI]** **p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
UE (unter der Behandlung auftretend)[b]			
Alter	< 75	RR = 0,99 [0,97; 1,01], p = 0,51	p = 0,03
	≥ 75	RR = 1,03 [0,99; 1,08], p = 0,09	Beleg
Anamnese einer Hypertonie	Ja	RR = 1,01 [0,99; 1,04], p = 0,23	p = 0,18
	Nein	RR = 0,99 [0,97; 1,02], p = 0,61	Hinweis
Unerwünschte Ereignisse nach* CTCAE *Grad ≥ 3			
Gabe von Bisphosphonaten während der Studie	Ja	RR = 0,76 [0,64; 0,91], p = 0,002	p = 0,10
	Nein	RR = 0,94 [0,80; 1,10], p = 0,45	Hinweis
Gabe von Steroiden während der Studie	Ja	RR = 0,95 [0,83; 1,08], p = 0,43	p = 0,10 bzw. p = 0,02[a]
	Nein	RR = 0,70 [0,56; 0,86], p = 0,0009	Hinweis bzw. Beleg
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	RR = 0,83 [0,73; 0,94], p = 0,004	p = 0,31 bzw. p = 0,14[a]
	2	RR = 1,04 [0,79; 1,37], p = 0,79	Keine Interaktion bzw. Hinweis
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,81 [0,69; 0,94], p = 0,006	p = 0,37 bzw. p = 0,18[a]
	≥ 4	RR = 0,95 [0,79; 1,13], p = 0,55	Keine Interaktion bzw. Hinweis
Schwerwiegende Unerwünschte Ereignisse			
Gabe von Steroiden während der Studie	Ja	RR = 0,94 [0,78; 1,13], p = 0,53	p = 0,25 bzw. p = 0,14[a]
	Nein	RR = 0,74 [0,56; 0,96], p = 0,03	Keine Interaktion bzw. Hinweis
ECOG-*Performance*-Status bei Studienbeginn	0 oder 1	RR = 0,82 [0,69; 0,97], p = 0,02	p = 0,04
	2	RR = 1,34 [0,86; 2,08], p = 0,19	Beleg
Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn	< 4	RR = 0,81 [0,66; 0,99], p = 0,04	p = 0,28 bzw. p = 0,19[a]
	≥ 4	RR = 0,99 [0,78; 1,25], p = 0,93	Keine Interaktion bzw. Hinweis
Tod aufgrund unerwünschter Ereignisse			

Endpunkt **Subgruppe**	**Kategorie**	**Effektschätzer, [95 %-KI]** **p-Wert** **Enzalutamid vs. Placebo**	**p-Wert der Interaktion**
Gabe von Bisphosphonaten während der Studie	Ja	RR = 0,27 [0,08; 0,92], p = 0,04	p = 0,03
	Nein	RR = 1,41 [0,60; 3,29], p = 0,43	Beleg

a: Berechnet nach Deeks 2010 [Deeks 2010]
b: *a priori* definierte Subgruppenanalysen
HR: Hazard Ratio
RR: Relatives Risiko
SMD: *Standardized Mean Difference*
ECOG: *Eastern Cooperative Oncology Group*
BPI: *Brief Pain Inventory*
FACT-P: *Functional Assessment of Cancer Therapy- Prostate*
FACT-G: *Functionel Assessment of Cancer Therapy- General*
EQ-5D: *Eastern Cooperative Oncology Group*
CTCAE*: Common Toxicity Criteria for Adverse Events*
NA: nicht anwendbar

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 Bewertung des medizinischen Nutzens und Zusatznutzens von Enzalutamid erfolgte anhand einer randomisierten kontrollierten, doppelblinden Phase-III-Studie mit insgesamt 1.199 erwachsenen Patienten.

Studienpopulationsebene (= Zulassungspopulationsebene)

Im Vergleich zu einer Placebo-Behandlung erwies sich die Behandlung mit Enzalutamid auf Studienpopulationsebene hinsichtlich der Endpunkte „Gesamtüberleben“, „Radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“, „Schmerzprogressionsrate“, „Zeit bis zur Schmerzprogression“, „Gesundheitsbezogene Lebensqualität nach FACT-P“ und hinsichtlich des Auftretens „Unerwünschter Ereignisse nach CTCAE Grad ≥ 3“ als statistisch signifikant überlegen. Bei den sonstigen Sicherheitsendpunkten gab es keine signifikanten Unterschiede zwischen Enzalutamid und Placebo.

- „Gesamtüberleben“

Die Patienten in der Enzalutamid-Gruppe hatten mit einem medianen Gesamtüberleben von 18,4 Monaten (95 %-KI [17,3; n.e.]) einen Überlebensvorteil von 4,8 Monaten gegenüber den Patienten in der Placebogruppe, deren medianes Gesamtüberleben nur 13,6 Monate (95 %-KI [11,3; 15,8]) betrug. Dieser Vorteil fiel mit einem Hazard Ratio (stratifizierte Analyse) von 0,63 (95 %-KI [0,53; 0,75]) und einem p-Wert von $p < 0,0001$ statistisch signifikant aus. Das Ergebnis wurde durch die Sensitivitätsanalyse bestätigt, in der sich für das unstratifizierte Hazard Ratio ein Wert von 0,63 (95 %-KI [0,53; 0,75]) ergab, mit einem ebenfalls statistisch signifikantem Vorteil von $p < 0,0001$.

- „Radiographisches progressionsfreies Überleben“

Für den Endpunkt „radiographisches progressionsfreies Überleben“ zeigte sich mit einem Hazard Ratio von 0,40 (95 %-KI [0,35; 0,47]) und einem p-Wert von $p < 0,0001$ ein statistisch signifikanter Vorteil zugunsten der Enzalutamid-Therapie. Patienten in der Enzalutamid-Gruppe hatten dabei im Vergleich zu Patienten in der Placebogruppe ein um 5,4 Monate länger andauerndes medianes radiographisches progressionsfreies Überleben. Die Sensitivitätsanalysen für diesen Endpunkt zeigten bei Beurteilung der Progression durch den Investigator und modifizierte Annahmen zur Zensierung ein Hazard Ratio von 0,41 (95 %-KI

[0,35; 0,47]), bei Beurteilung als *derived* Progression und primäre Annahmen für die Zensierung ein Hazard Ratio von 0,40 (95 %-KI [0,35; 0,47]) und bei Beurteilung als *derived* Progression und modifizierte Annahmen zur Zensierung ein Hazard Ratio von 0,41 (95 % KI [0,35; 0,47]) und bestätigten somit die Ergebnisse der Hauptanalyse.

- „Zeit bis zur ersten skelettbezogenen Komplikation“

Die mediane „Zeit bis zur ersten skelettbezogenen Komplikation“ betrug bei Patienten der Enzalutamid-Gruppe 16,7 Monate (95 %-KI [14,6; 19,1]) im Vergleich zu 13,3 Monaten (95 %-KI [9,9; n.e.]) in der Placebogruppe. Bei Patienten unter Enzalutamid-Therapie trat die erste skelettbezogene Komplikation demnach 3,4 Monate später auf als bei Patienten, die Placebo erhielten. Das Hazard Ratio lag mit einem Wert von 0,69 (95 %-KI [0,57; 0,84], p = 0,0001) statistisch signifikant zugunsten der Enzalutamid-Gruppe. Innerhalb der Unterteilung des Endpunkts „Zeit bis zur ersten skelettbezogenen Komplikation“ in die Bereiche „Zeit bis zur Bestrahlung des Knochens“, „Zeit bis zur Knochenoperation“, „Zeit bis zur pathologischen Knochenfraktur“, „Zeit bis zur Rückenmarkkompression“ und „Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen“ ergab sich mit einem Hazard Ratio von 0,71 (95 %-KI [0,55; 0,91]) und einem p-Wert von p = 0,0058 ein signifikanter Vorteil zugunsten von Enzalutamid die Zeit bis zur Bestrahlung des Knochens betreffend. Das Ergebnis der Sensitivitätsanalyse, in welcher Patienten, deren Therapie sich aufgrund von Knochenschmerzen bei Knochenmetastasen änderte, nicht mitgezählt wurden, bestätigte das Ergebnis der Hauptanalyse. Auch hier lag das Hazard Ratio mit einem Wert von 0,74 (95 %-KI [0,61; 0,91]) und einem p-Wert von p = 0,004 statistisch signifikant zugunsten der Enzalutamid-Gruppe.

- „Schmerzprogressionsrate“

Innerhalb des Endpunkts „Schmerzprogressionsrate“ zeigte sich sowohl für die *Responder*-Analyse als auch für die Veränderung der Schmerzpunktzahl gegenüber Studienbeginn ein signifikanter Unterschied zugunsten der Enzalutamid-Therapie. Die *Responder*-Analyse ergab für das Relative Risiko, eine Schmerzprogression zu erleiden, einen Wert von 0,71 (95 %-KI [0,59; 0,87], p = 0,0008), welcher statistisch signifikant zugunsten der Enzalutamid-Gruppe lag. Die SMD zwischen den beiden Therapiegruppen betrug für den Unterschied der Schmerzpunktzahl zwischen Studienbeginn und Woche 13 -0,48; 95 %-KI [-0,62; -0,34] und fiel mit einem p-Wert von p < 0,0001 ebenfalls statistisch signifikant zugunsten der Enzalutamid-Gruppe aus.

- „Zeit bis zur Schmerzprogression“

Die mediane „Zeit bis zur Schmerzprogression“ betrug in der Placebogruppe 13,8 Monate (95 %-KI [13,8; n.e.]). In der Enzalutamid-Gruppe war der Median zum Zeitpunkt der Datenerhebung noch nicht erreicht. Das Hazard Ratio von 0,56 (95 %-KI [0,41; 0,78]) mit einem p-Wert von p = 0,0004 ergab dabei einen statistisch signifikanten Vorteil der Enzalutamid-Therapie gegenüber einer Placebo-Behandlung.

- „Gesundheitsbezogene Lebensqualität nach FACT-P“

Bei der Untersuchung des Endpunkts „Gesundheitsbezogenen Lebensqualität nach FACT-P“ ergaben sich, je nach Art der durchgeführten Analyse, unterschiedliche Ergebnisse. Im Rahmen der Auswertung des Fragebogens als *Responder*-Analyse zeigte sich unter Enzalutamid-Gabe bei einer signifikant größeren Anzahl von Patienten eine Verbesserung der Lebensqualität gegenüber Placebo (Relatives Risiko von 2,81 (95 %-KI [2,05; 3,84], p < 0,0001)). Bei Betrachtung der Veränderung der FACT-P-Gesamtpunktzahl der letzten verfügbaren Erhebung gegenüber Studienbeginn zeigte sich ebenfalls ein signifikanter Vorteil von Enzalutamid (SMD = 0,42; 95 %-KI [0,27; 0,56], p < 0,0001). Dieser Vorteil konnte weiterhin auf Ebene der Gesamtpunktzahl FACT-G (SMD = 0,36; 95 %-KI [0,21; 0,50], p < 0,0001) und, mit Ausnahme der Subskala „Verhältnis zu Freunden, Bekannten und Familie“ (SMD = 0,06; 95 %-KI [-0,08; 0,21], p = 0,41), auch auf Ebene aller einzelnen Subskalen nachgewiesen werden (Körperliches Wohlbefinden (SMD = 0,26; 95 %-KI [0,11; 0,40], p = 0,0006), Seelisches Wohlbefinden (SMD = 0,33; 95 %-KI [0,18; 0,47], p < 0,0001), Funktionelles Wohlbefinden (SMD = 0,30; 95 %-KI [0,15; 0,45], p < 0,0001), Prostatakarzinom-Subskala (SMD = 0,34; 95 %-KI [0,20; 0,49], p < 0,0001). Bei Betrachtung der Veränderung der Punktzahlen für FACT-P, FACT-G sowie der einzelnen Subskalen in Woche 49 gegenüber Studienbeginn ergaben sich jedoch auf keiner Ebene signifikante Unterschiede zwischen den beiden Behandlungsgruppen.

- „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“

Im Rahmen der Studie traten bei 45,3 % (362/800) der Patienten unter Enzalutamid-Therapie unerwünschte Ereignisse nach CTCAE Grad ≥ 3 auf, wohingegen Ereignisse dieser Art in der Placebogruppe bei 53,1 % (212/399) der Patienten auftraten. Das Relative Risiko, ein unerwünschtes Ereignis nach CTCAE Grad ≥ 3 zu erleiden, zeigte mit einem Wert von 0,85 (95 %-KI [0,76; 0,96], p = 0,009) einen signifikanten Vorteil für Patienten in der Enzalutamid-Gruppe.

Für alle anderen Endpunkte („Schmerzlinderung“ „Gesundheitsbezogene Lebensqualität nach EQ-5D“, „Unerwünschte Ereignisse“, „Schwerwiegende unerwünschte Ereignisse“ „Studienabbruch aufgrund unerwünschter Ereignisse“ und „Tod aufgrund unerwünschter Ereignisse“) gab es keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen.

Subgruppenebene

Für alle Endpunkte wurden Analysen auf Subgruppenebene durchgeführt. Dabei waren lediglich Subgruppenanalysen für den Endpunkt „Gesamtüberleben“ und sicherheitsbezogene Ereignisse präspezifiziert, alle anderen Subgruppenanalysen wurden *post hoc* durchgeführt (Details siehe Abschnitt 4.2.5.5). Die Ergebnisse zu allen Subgruppenanalysen wurden sowohl tabellarisch als auch deskriptiv im Abschnitt 4.3.1.3.2.3 dargestellt. Im Folgenden werden alle Endpunkte im Hinblick auf die Ergebnisse der Subgruppenanalysen beleuchtet. Dabei werden zunächst die Endpunkte mit statistisch signifikantem Vorteil von Enzalutamid

gegenüber Placebo auf Studienpopulationsebene betrachtet, anschließend folgen die Endpunkte ohne signifikante Unterschiede auf Studienpopulationsebene.

Schlussfolgerungen aus beobachteten Hinweisen/Belegen unterliegen den bekannten Einschränkungen für *post hoc*-Analysen und multiplem Testen. Zudem war die Fallzahl der Studie nicht dafür berechnet, Effekte in Subgruppen zu zeigen.

Weiterhin ist zu beachten, dass die Subgruppenmerkmale viszerale Erkrankung (ja/nein) und Steroidgabe (ja/nein) bereits mit einer unterschiedlichen Prognose für die Patienten verbunden sind. Patienten mit viszeraler Erkrankung weisen *per se* eine schlechtere Prognose als Patienten ohne diese Erkrankung auf. Die Steroidgabe durch den Arzt stellt gleichzeitig einen Hinweis auf einen schlechteren Allgemeinzustand des Patienten dar.

Endpunkte mit statistisch signifikantem Vorteil von Enzalutamid gegenüber Placebo auf Studienpopulationsebene

- „Gesamtüberleben“

Für den Endpunkt „Gesamtüberleben“ zeigte die Behandlung mit Enzalutamid für die Patienten der Studienpopulation einen statistisch signifikanten Vorteil gegenüber der Placebo-Behandlung.

- „Radiographisches progressionsfreies Überleben“

Bezüglich des Endpunkts „radiographisches progressionsfreies Überleben“ konnte für die gesamte Studienpopulation ein signifikanter Vorteil von Enzalutamid gegenüber Placebo nachgewiesen werden.

- „Zeit bis zur ersten skelettbezogenen Komplikation“

Hinsichtlich der Zeit bis zum Auftreten der ersten skelettbezogenen Komplikation hatten Patienten in der Enzalutamid-Gruppe einen signifikanten Vorteil gegenüber Patienten in der Placebogruppe.

- „Schmerzprogressionsrate“

Im Hinblick auf die Auswertung des Endpunkts „Schmerzprogressionsrate“ als *Responder*-Analyse ist Enzalutamid für Patienten mit einer Schmerzpunktzahl ≥ 4, für Patienten ohne begleitende Steroidmedikation und für Patienten < 65 Jahre von signifikantem Vorteil. Im Rahmen der Auswertung des Endpunkts als Veränderung gegenüber Studienbeginn zeigte sich jedoch für alle Patienten in der Studienpopulation Enzalutamid gegenüber Placebo als signifikant überlegen.

- „Zeit bis zur Schmerzprogression“

Für die gesamte Studienpopulation ist bezüglich dieses Endpunkts eine Enzalutamid-Therapie von signifikantem Vorteil im Vergleich zu einer Placebo-Behandlung.

- „Gesundheitsbezogene Lebensqualität nach FACT-P“

Hinsichtlich der Auswertung des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ als *Responder*-Analyse zeigte sich für die gesamte Studienpopulation ein signifikanter Vorteil für Patienten der Enzalutamid-Gruppe im Vergleich zu Patienten in der Placebogruppe.

Bezüglich der Auswertung des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ als Veränderung der letzten verfügbaren Erhebung gegenüber Studienbeginn erwies sich, sowohl bei Betrachtung der FACT-P- als auch der FACT-G-Gesamtpunktzahl, Enzalutamid bei Patienten ohne begleitende Steroidgabe als signifikant überlegen gegenüber einer Placebo-Therapie. Bei Betrachtung der Veränderung der beiden Gesamtpunktzahlen (FACT-P bzw. des FACT-G) in Woche 49 gegenüber Studienbeginn ergab sich innerhalb der Studienpopulation für Enzalutamid jedoch kein statistisch signifikanter Vorteil im Vergleich zu Placebo.

- „Unerwünschter Ereignisse nach CTCAE Grad ≥ 3“

Hinsichtlich des Endpunkts „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ haben Patienten mit begleitender Bisphosphonatgabe und Patienten ohne begleitende Steroidmedikation unter Enzalutamid-Therapie ein signifikant geringeres relatives Risiko für das Auftreten unerwünschter Ereignisse nach CTCAE Grad ≥ 3 als unter Placebo-Therapie. Unter Berücksichtigung der zusätzlichen Interaktionsterme nach Deeks [Deeks 2010] (siehe Tabelle 4-75) bezieht sich dieses Resultat ebenfalls auf Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 und auf Patienten mit einer Schmerzpunktzahl < 4 zu Studienbeginn.

Endpunkte ohne statistisch signifikantem Vorteil von Enzalutamid gegenüber Placebo auf Studienpopulationsebene

- „Schmerzlinderung“

Für den Endpunkt „Schmerzlinderung“ ergaben sich im Rahmen der Subgruppenanalysen keine Hinweise auf oder Belege für Interaktionen (Tabelle 4-109).

- „Gesundheitsgezogene Lebensqualität nach EQ-5D“

Auf Studienpopulationsebene bestand bezüglich der „gesundheitsgezogenen Lebensqualität nach EQ-5D“ kein Unterschied zwischen den beiden Therapiegruppen. Im Rahmen der Subgruppenanalysen ergaben sich zwar Hinweise und Belege auf Interaktionen, jedoch traten in keiner der betrachteten Subgruppen signifikante Unterschiede zwischen den Therapiearmen auf.

- „Unerwünschte Ereignisse“

Im Rahmen der Subgruppenanalysen für den Endpunkt „unerwünschte Ereignisse“ ergaben sich keine Hinweise oder Belege für Interaktionen (Tabelle 4-109). Bezüglich des Auftretens unerwünschter Ereignisse ergaben sich somit keine Unterschiede zwischen den Behandlungsarmen.

- „Schwerwiegende unerwünschte Ereignisse“

Obwohl auf Studienpopulationsebene für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ kein Unterschied zwischen den Behandlungsgruppen bestand, zeigte sich im Rahmen der Subgruppenanalysen, dass das relative Risiko für das Auftreten eines schwerwiegenden unerwünschten Ereignisses unter Enzalutamid-Therapie bei Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 signifikant geringer ausfiel als unter Placebo-Therapie. Unter Berücksichtigung der zusätzlichen Interaktionsterme nach Deeks [Deeks 2010] (siehe Tabelle 4-75) fiel dieses Risiko außerdem für Patienten ohne begleitende Steroidgabe und für Patienten mit einer Schmerzpunktzahl < 4 zu Studienbeginn unter Enzalutamid-Therapie statistisch signifikant geringer aus als unter Placebo-Therapie.

- „Studienabbruch aufgrund unerwünschter Ereignisse“

Im Rahmen der Subgruppenanalysen für den Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ ergaben sich keine Hinweise oder Belege für Interaktionen (Tabelle 4-109). Bezüglich des Abbrechens der Studie aufgrund unerwünschter Ereignisse ergaben sich somit keine Unterschiede zwischen den Behandlungsarmen.

- „Tod aufgrund unerwünschter Ereignisse“

Obwohl auf Studienpopulationsebene für den Endpunkt „Tod aufgrund unerwünschter Ereignisse“ kein Unterschied zwischen den Behandlungsgruppen bestand, zeigte sich im Rahmen der Subgruppenanalysen, dass für Patienten, welche studienbegleitend Bisphosphonate bekamen, das relative Risiko aufgrund eines unerwünschten Ereignisses zu sterben im Enzalutamid-Arm signifikant geringer ausfiel als im Placebo-Arm.

Deskriptiv berichtete Endpunkte

- „Lähmungen und lähmungsbedingte Harninkontinenz“

Im Rahmen der Studie traten in beiden Behandlungsarmen Lähmungserscheinungen gleich häufig auf (bei jeweils 1,75 % der Patienten, siehe Tabelle 4-47), wobei nicht beurteilt werden kann, inwieweit diese Lähmungserscheinungen auf Metastasierungen zurückzuführen sind oder andere Gründe haben. Insofern konnte keine Aussage darüber getroffen werden, ob und inwieweit durch die Therapie metastasenbedinge Lähmungserscheinungen gelindert wurden. Es ist anzunehmen, dass es diesbezüglich keine Unterschiede zwischen den Therapiegruppen gab.

Unter den in der Studie auftetenden 40 Fällen von Harninkontinenz wurde bei keinem Patienten explizit eine durch Metastasierung hervorgerufene Lähmungserscheinung als Ursache registriert. Im Enzalutamid-Arm traten zwei Fälle von Harninkontinenz auf, welche mit einer Rückenmarkkompression einhergingen, und von daher im Zusammenhang mit einer fortschreitende Metastasierung stehen könnten. Bei zwei weiteren Patienten im Enzalutamid-Arm trat eine Harninkontinenz ein bzw. zwei Monate nach dem Ereignis Rückenmarkkompression auf, stand also nicht in unmittelbarem Zusammenhang mit diesem. Aufgrund dieser Erhebungen lässt sich keine Aussage darüber treffen, ob in den Therapiearmen Unterschiede im Auftreten lähmungsbedingter Harnikontinenz bestanden.

- „Krampfanfälle“

In der AFFIRM-Studie traten im Enzalutamid-Arm bei sieben Patienten (0,9 %) Krampfanfälle als schwerwiegendes unerwünschtes Ereignis auf. Bei vier der Patienten konnten potentiell prädisponierende Faktoren festgestellt werden, insofern kann die Ursache des Auftretens dieser Fälle nicht ausschließlich auf die Studienmedikation zurückgeführt werden. Die Seltenheit der Ereignisse lässt allerdings keine sichere Aussage darüber zu, ob und inwieweit sich die Therapiearme hinsichtlich des Auftreten von Krampfanfällen tatsächlich unterscheiden.

Tabelle 4-110: Enzalutamid vs. Placebo: Zusammenfassung der Effekte auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM

	Enzalutamid vs. Placebo	
Zielgröße	**Effektschätzer [95 %-KI]**	**p-Wert**
Mortalität		
Gesamtüberleben	HR = 0,63 [0,53; 0,75] absolute Differenz: 4,8 Monate	p < 0,0001
Morbidität		
Radiographisches progressionsfreies Überleben (rPFS)	HR = 0,40 [0,35; 0,47] absolute Differenz: 5,4 Monate	p < 0,0001
Zeit bis zur ersten skelettbezogenen Komplikation	HR = 0,69 [0,57; 0,84] absolute Differenz: 3,4 Monate	p = 0,0001
- Zeit bis zur Bestrahlung des Knochens	HR = 0,71 [0,55; 0,91] absolute Differenz: n.e. p = 0,0058	p = 0,007
- Zeit bis zur Knochenoperation	HR = 2,49 [0,30; 20,79] absolute Differenz: n.e. p = 0,38	p = 0,28
- Zeit bis zur pathologischen Knochenfraktur	HR = 0,85 [0,47; 1,56] absolute Differenz: n.e. p = 0,60	p = 0,99
- Zeit bis zur Rückenmarkkom-pression	HR = 1,00 [0,64; 1,55] absolute Differenz: n.e. p = 0,99	p = 0,49
- Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen	HR = 0,54 [0,28; 1,04] absolute Differenz: n.e. p = 0,06	p = 0,15
Schmerzlinderung	RR = 6,73 [0,99; 45,88]	p > 0,05
Schmerzprogressionsrate (*Responder*-Analyse)	RR = 0,71 [0,59; 0,87]	p = 0,0008
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)	SMD = -0,48 [-0,62; -0,34]	p < 0,0001
Zeit bis zur Schmerzprogression	HR = 0,56 [0,41; 0,78] absolute Differenz: n.e.	p = 0,0004
Gesundheitsbezogene Lebensqualität		
FACT-P (*Responder*-Analyse)	RR = 2,81 [2,05; 3,84]	p < 0,0001

	Enzalutamid vs. Placebo	
Zielgröße	**Effektschätzer [95 %-KI]**	**p-Wert**
FACT-P (Veränderung Woche 49 gegenüber Studienbeginn)		
- FACT-P Gesamtpunktzahl	SMD = 0,22 [-0,23, 0,67]	p = 0,33
- FACT-G Gesamtpunktzahl	SMD = 0,19 [-0,25; 0,63]	p = 0,40
- FACT-P Körperliches Wohlbefinden	SMD = -0,07 [-0,50; 0,36]	p = 0,76
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,23 [-0,20; 0,66]	p = 0,29
- FACT-P Seelisches Wohlbefinden	SMD = 0,19 [-0,25; 0,63]	p = 0,40
- FACT-P Funktionelles Wohlbefinden	SMD = 0,09 [-0,35; 0,53]	p = 0,70
- FACT-P Prostatakarzinom-Subskala	SMD = 0,05 [-0,39; 0,49]	p = 0,83
FACT-P (Veränderung letzte verfügbare Erhebung gegenüber Studienbeginn)		
- FACT-P Gesamtpunktzahl	SMD = 0,42 [0,27; 0,56]	p < 0,0001
- FACT-G Gesamtpunktzahl	SMD = 0,36 [0,21; 0,50]	p < 0,0001
- FACT-P Körperliches Wohlbefinden	SMD = 0,26 [0,11; 0,40]	p = 0,0006
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,06 [-0,08; 0,21]	p = 0,41
- FACT-P Seelisches Wohlbefinden	SMD = 0,33 [0,18; 0,47]	p < 0,0001
- FACT-P Funktionelles Wohlbefinden	SMD = 0,30 [0,15; 0,45]	p < 0,0001
- FACT-P Prostatakarzinom-Subskala	SMD = 0,34 [0,20; 0,49]	p < 0,0001
EQ-5D VAS Gesamtpunktzahl (Veränderung Woche 49 gegenüber Studienbeginn)	SMD = 0,85 [-0,69; 2,40]	p = 0,26
EQ-5D VAS	SMD = -0,03 [-0,36; 0,30]	p = 0,86

	Enzalutamid vs. Placebo	
Zielgröße	**Effektschätzer [95 %-KI]**	**p-Wert**
Gesamtpunktzahl (Veränderung letzte verfügbare Erhebung gegenüber Studienbeginn)		
Sicherheit		
Unerwünschte Ereignisse	RR = 1,00 [0,99; 1,02]	p = 0,67
Unerwünschte Ereignisse CTCAE Grad ≥ 3	RR = 0,85 [0,76; 0,96]	p = 0,009
Schwerwiegende unerwünschte Ereignisse	RR = 0,87 [0,74; 1,02]	p = 0,08
Studienabbruch aufgrund unerwünschter Ereignisse	RR = 0,78 [0,53; 1,15]	p = 0,20
Tod aufgrund unerwünschter Ereignisse	RR = 0,82 [0,43; 1,58]	p = 0,55

n.e.: nicht erreicht
KI: Konfidenzintervall
HR: *Hazard Ratio*
SMD: *Standardized Mean Difference*
RR: Relatives Risiko
FACT-P: *Functional Assessment of Cancer Therapy- Prostate*
FACT-G: *Functional Assessment of Cancer Therapy- General*
EQ-5D VAS: *European Quality of Live Five Domain Scale Visual Analogue Scale*
CTCAE: *Common Toxicity Criteria for adverse Events*

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*

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.

Indirekte Vergleiche wurden nicht 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 4-111: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

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

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

Tabelle 4-112: 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.

Nicht zutreffend

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

Tabelle 4-113: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-114: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend

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

Tabelle 4-115: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

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

Nicht zutreffend

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

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

Indirekte Vergleiche wurden nicht 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*

Nicht randomisierte vergleichende Studien wurden nicht verwendet.

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.

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 4-116: 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 4-117: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

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

Tabelle 4-118: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

1) Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar.

Nicht zutreffend

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

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

Nicht randomisierte vergleichende Studien wurden nicht verwendet.

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*

Es wurden keine weiteren Untersuchungen durchgeführt.

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.

Nicht zutreffend

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-119: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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.

Nicht zutreffend

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung.

Nicht zutreffend

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

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

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

Es wurden keine weiteren Untersuchungen durchgeführt.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

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

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Evidenzklassifizierung

Zur Beurteilung des Nutzens und Zusatznutzens von Enzalutamid gegenüber Placebo in der Indikation metastasiertes kastrationsresistentes Prostatakarzinom bei erwachsenen Männern, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet, lag eine randomisierte, kontrollierte klinische Studie vor. Es handelte sich um eine placebo-

kontrollierte Phase-III-Studie, in welcher insgesamt 1.199 Patienten, aufgeteilt im Verhältnis 2:1 auf die beiden Behandlungsarme Enzalutamid und Placebo, untersucht wurden. Die Studie wurde doppelblind durchgeführt, d.h. weder Patient noch behandelnder Arzt waren über die zugewiesene Behandlungsgruppe informiert. Die Studie wurde durch das Erreichen eines signifikanten Ergebnisses zur Interimanalyse, wie für diesen Fall geplant, beendet. Die Studie ist somit hinsichtlich ihres Designs der Evidenzstufe Ib zuzuordnen. [Gemeinsamer Bundesausschuss (G-BA) 2013]

Studienqualität

Verzerrungspotential auf Studienebene

In der Studie wurden keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder andere Aspekte, die das Verzerrungspotenzial beeinflussen könnten, identifiziert. Das Studiendesign (Randomisierung, Placebokontrolle, verdeckte Gruppenzuteilung, Verblindung von Patienten und Untersuchern) wurde konsequent umgesetzt, und die Auswertung erfolgte umfassend und transparent. Die Studie und ihre Ergebnisse wurden in einer *peer-reviewed* Fachzeitschrift, dem *New England Journal of Medicine*, veröffentlicht. Die Zahlen in der Publikation sind deckungsgleich mit den Zahlen im CSR. Das Verzerrungspotenzial auf Studienebene wurde daher für die Studie mit „niedrig“ beurteilt, und es kann davon ausgegangen werden, dass das Gesamtergebnis der Studie nicht relevant verzerrt ist.

Validität der Endpunkte und Verzerrungspotential auf Endpunktebene

Neben den Sicherheitsparametern „Unerwünschte Ereignisse“, „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“, „Schwerwiegende unerwünschte Ereignisse“, „Studienabbruch aufgrund unerwünschter Ereignisse“ und „Tod aufgrund unerwünschter Ereignisse“ wurden zusätzlich die Endpunkte „Gesamtüberleben“, „radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“, „Schmerzlinderung“, „Schmerzprogressionsrate“, „Zeit bis zur Schmerzprogression“, „Gesundheitsbezogene Lebensqualität nach FACT-P“ und „Gesundheitsbezogene Lebensqualität nach EQ-5D“ berücksichtigt. Diese Endpunkte sind etabliert, wurden als patientenrelevant eingestuft (siehe Abschnitt 4.2.5.2) und sind im Einklang mit den Empfehlungen der Europäischen Arzneimittelagentur EMA [European Medicines Agency (EMA) 2012] und der Prostate Cancer Working Group [Bubley 1999; Scher 2008] zur Planung klinischer Studien in der Onkologie bzw. beim metastasierten kastrationsresistenten Prostatakarzinom. Die Endpunkte „Lähmungen“, „lähmungsbedingte Harninkontinenz“ und „Krampfanfälle“ wurden aufgrund der Seltenheit der auftretenden Ereignisse nicht statistisch ausgewertet sondern lediglich deskriptiv berichtet.

Das Verzerrungspotenzial auf Endpunktebene wurde für alle Endpunkte mit Ausnahme des Endpunkts „Schmerzlinderung“, des Endpunkts „Schmerzprogressionsrate“, des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ und des Endpunkts „Gesundheitsbezogene Lebensqualität nach EQ-5D“ als niedrig eingestuft.

Für den Endpunkt „Schmerzlinderung“ konnte lediglich eine stark reduzierte Patientenzahl analysiert werden (64 (5 %) von 1.199 Patienten), da nur Patienten in die Auswertung aufgenommen wurden, die bereits zu Studienbeginn eine ausreichend hohe Schmerzbelastung aufwiesen und damit ein entsprechendes Potential zur Schmerzlinderung besaßen (die durchschnittliche Schmerzpunktzahl zu Studienbeginn musste mindestens 4 Punkte betragen).

In die Auswertung des Endpunkts „Schmerzprogressionsrate“ gingen nur Patienten ein, welche an mindestens 4 von 7 Tagen während der *run-in phase* zu Studienbeginn und an den 7 Tagen vor der Woche-13-Visite gültige Eintragungen im Schmerztagebuch hatten. Aufgrund dieser Kriterien reduzierte sich die Anzahl der Patienten in der Auswertungspopulation auf 884 (74 %) von 1.199 Patienten.

Für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ konnte ebenfalls nur eine reduzierte Anzahl an Patienten in die Auswertung aufgenommen werden, da nur Patienten für die Auswertung in Frage kamen, für welche zu Studienbeginn und zum jeweiligen Auswertezeitpunkt gültige Fragebögen zur Verfügung standen. Für die *Responder*-Analyse standen 871 (73 %) von 1.199 Patienten zur Verfügung, für die Auswertung der Veränderung der Gesamtpunktzahl FACT-P in Woche 49 bzw. zum Zeitpunkt der letzten verfügbaren Erhebung gegenüber Studienbeginn waren dies 242 (20 %) bzw. 854 (71 %) von 1.199 Patienten.

Der Fragebogen EQ-5D ist ein Europa-spezifisches Instrument zur Beurteilung der Lebensqualität, insofern kam er nur in Studienzentren in europäischen Ländern zum Einsatz. Die Erfassung der Lebensqualität wurde mittels dieses Fragebogens erst im Rahmen einer Protokolländerung nach Studienbeginn in die Studie integriert. Die Anzahl der Patienten, welche den Fragebogen ausfüllen konnten, nahm über die Zeit ab. Daher stand für die Auswertung der Veränderung der Gesamtpunktzahl in Woche 49 bzw. zum Zeitpunkt des *last available assessment* gegenüber Studienbeginn mit 25 (2 %) bzw. 174 (15 %) von 1.199 Patienten nur noch eine geringe Anzahl an Patienten zur Verfügung.

Für randomisierte klinische Studien gilt nach VerfO die Evidenzstufe Ib [Gemeinsamer Bundesausschuss (G-BA) 2013]. Die Qualität der Studie und damit ihre Aussagekraft kann aufgrund des Studiendesigns und des niedrigen Verzerrungspotenzials auf Studien- und größtenteils auch auf Endpunktebene als hoch eingeschätzt werden, d.h. die Studie besitzt eine hohe qualitative Ergebnissicherheit.

Beleglage

Bei der Beurteilung der Beleglage stellt das IQWiG an die Ableitung des Belegs für einen Zusatznutzen besondere Anforderungen. „[…] *eine einzelne Studie mit hoher qualitativer Ergebnissicherheit kann trotz statistisch signifikantem Effekt […] in der Regel nur einen Hinweis liefern. Soll aus lediglich einer Studie im Ausnahmefall ein Nutzenbeleg abgeleitet werden, so sind an eine solche Studie und ihre Ergebnisse besondere Anforderungen zu stellen.*“ [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2013] Als besondere Anforderungen zitiert das IQWiG die „*Points to consider on application with:*

1. meta-analyses; 2. one pivotal study" der EMA (Abbildung 3) [European Medicines Agency (EMA) 2001]:

III.2 Prerequisites for one pivotal study applications

In cases where the confirmatory evidence is provided by one pivotal study only, this study will have to be exceptionally compelling, and in the regulatory evaluation special attention will be paid to:

- The internal validity. There should be no indications of a potential bias.
- The external validity. The study population should be suitable for extrapolation to the population to be treated.
- Clinical relevance. The estimated size of treatment benefit must be large enough to be clinically valuable.
- The degree of statistical significance. Statistical evidence considerably stronger than $p<0.05$ is usually required, accompanied by precise estimates of treatment effects, i.e. narrow confidence intervals. The required degree of significance will depend on factors such as the therapeutic indication, the primary endpoint, the amount of supportive data and whether the alternative analyses demonstrating consistency are pre-specified. When the aim is to demonstrate non-inferiority, one study is more likely to be accepted if the lower 95% confidence bound is well away from the non-inferiority margin.
- Data quality.
- Internal consistency. Similar effects demonstrated in different pre-specified sub-populations. All-important endpoints showing similar findings.
- Centre effects. None of the study centres should dominate the overall result, neither in terms of number of subjects nor in terms of magnitude of effect.
- The plausibility of the hypothesis tested.

Abbildung 3: "*Points to consider on application with: 1. meta-analyses; 2. one pivotal study*" [European Medicines Agency (EMA) 2001]

Diese Anforderungen wurden von der AFFIRM-Studie erfüllt. Es handelte sich um eine randomisierte, kontrollierte, doppelblinde Studie mit 1.119 Patienten, die in 15 Ländern in Nordamerika, Südamerika, Europa, Australien und Südafrika in 156 Zentren durchgeführt wurde. Wie im Abschnitt 4.3.1.2.2 dargelegt, wies die Studie die geforderte interne Validität auf, da es keine Hinweise auf ein Verzerrungspotential auf Studienebene gab.

Hinsichtlich der Auswahl der Patienten und somit der externen Validität wurde angestrebt, die neue Therapie unter möglichst realistischen klinischen Bedingungen zu testen. Begründet durch das Studiendesign sind dabei allerdings gewisse generelle Einschränkungen unvermeidlich. So ist es beispielsweise erforderlich, dass die Patienten ein Mindestmaß an Mobilität besitzen, um an den Untersuchungen teilzunehmen. Hinsichtlich der Heterogenität der Studienpopulation ist zu berücksichtigen, dass das Vermögen der Studie, Unterschiede zwischen den Behandlungsarmen aufzuzeigen, gewährleistet bleibt, wobei gleichzeitig eine klinisch relevante und dem Versorgungskontext entsprechende Patientenpopulation abgebildet werden soll. [European Medicines Agency (EMA) 2012] In der Studie wurde versucht, die Patientenpopulation möglichst repräsentativ zu gestalten, d.h. die Übertragbarkeit der Studienergebnisse auf den realen Versorgungsalltag zu gewährleisten.

Die im Rahmen der Studie AFFIRM erzielten Ergebnisse für die Behandlung von Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet, zeigten sich insgesamt betrachtet als sehr konsistent. Für alle statistisch betrachteten Wirksamkeitsendpunkte und Sicherheitsparameter wurden im Enzalutamid-Arm entweder signifikante Vorteile gegenüber dem Placebo-Arm erzielt oder aber eine Vergleichbarkeit mit Placebo erreicht. Insbesondere für den primären Endpunkt „Gesamtüberleben“ und die sekundären Endpunke „radiographisches progressionsfreies Überleben“, „Zeit bis zur ersten skelettbezogenen Komplikation“ und „Zeit bis zur Schmerzprogression“ konnten die Vorteile einer Enzalutamid-Therapie deutlich aufgezeigt werden. Auch für die anderen Aspekte des Endpunkts „Schmerzen“ sowie bei der Betrachtung der „Lebensqualität nach FACT-P“ ergaben sich Vorteile unter Enzalutamid-Gabe, auch wenn diese aufgrund der in die Analysen eingegangenen Patientenzahlen mit einer gewissen Unsicherheit behaftet sind. Im Hinblick auf die Sicherheit zeigte sich Enzalutamid besonders bei den Parametern welche gravierendere Ereignisse betrachten (UEs nach CTCAE Grad ≥ 3, SUEs und Tod aufgrund UEs) innerhalb einiger Patientengruppen als deutlich nebenwirkungsärmer und für alle anderen Patienten als vergleichbar mit Placebo. Einzig die Krampfanfälle stellen eine relevante Nebenwirkung im Vergleich zu Placebo dar. Insgesamt betrachtet sind die Ergebnisse der Studie konsistent und es ergeben sich für eine Vielzahl der patientenrelevanten Endpunkte deutlich statistisch signifikante und klinisch relevante Verbesserungen, welche durch die Ergebnisse der Sensitivitätsanalysen und der prä-spezifizierten Subgruppenanalysen untermauert werden. Eine vollkommene Kongruenz der Ergebnisse konnte angesichts der Vielzahl der ausgewerteten Endpunkte und der *post hoc* definierten Subgruppenauswertungen nicht erwartet werden.

Zusammenfassend ist festzuhalten, dass aufgrund der hohen Patientenzahl, des randomisierten, kontrollierten, doppelblinden Designs, der vorliegenden internen wie externen Validität sowie der klinischen Relevanz, statistischen Signifikanz, Datenqualität und Konsistenz der Ergebnisse die Studie den Kriterien der EMA für eine pivotale Studie entspricht, was durch die Erteilung der Zulassung durch die EMA bestätigt wurde, und somit als Einzelstudie zur Ableitung eines Belegs für einen Zusatznutzen herangezogen werden kann.

Interpretation der Ergebnisse

In der Studie AFFIRM wurde die Wirksamkeit und Sicherheit von Enzalutamid + *Best Supportive Care* gegenüber *Best Supportive Care* untersucht, wobei die in der Studie eingesetzte Vergleichstherapie der vom G-BA festgelegten zweckmäßigen Vergleichstherapie entsprach. In der Studie konnte gezeigt werden, dass Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet und die daraufhin mit Enzalutamid behandelt wurden, eine klinisch relevante und statistisch signifikante Verlängerung der Überlebenszeit erfuhren. Dieser Gewinn an Überlebenszeit konnte auf Ebene der gesamten Studienpopulation

gezeigt werden und auch über fast alle untersuchten Subgruppen hinweg nachgewiesen werden. Lediglich für die Subgruppe der Patienten, welche zu Studienbeginn unter einer Erkrankung der Lunge und/oder Leber litten, konnte der Überlebensvorteil statistisch nicht belegt werden. Gleichwohl zeigte sich aber auch in dieser Gruppe mit einem Hazard Ratio von 0,78 (95 %-KI [0,56; 1,09], p = 0,15) die Tendenz eines Überlebensvorteils. Ein Grund für den fehlenden statistischen Nachweis könnte die relativ kleine Gruppengröße von weniger als 25 % der Gesamtpatientenzahl innerhalb dieser Subgruppe sein, welche sich auf die Größe des Konfidenzintervalls auswirkt.

Weiterhin konnte über die gesamte Studienpopulation hinweg ein verlängertes radiographisches progressionsfreies Überleben für Patienten unter Enzalutamid-Behandlung gezeigt werden. Zwar ergaben sich im Rahmen der Subgruppenanalysen Hinweise und Belege für den Einfluss verschiedener Faktoren auf das radiographische progressionsfreie Überleben, jedoch wurde der Vorteil der Enzalutamid-Behandlung dadurch nicht beeinträchtigt, so dass dieser über die gesamte Studienpopulation hinweg statistisch signifikant ausfiel.

Für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ erwies sich bei Betrachtung der Studienpopulationsebene Enzalutamid für alle Patienten als signifikant überlegen. Die Analyse von Einzelereignissen in dieser Gruppe („Zeit bis zur Bestrahlung des Knochens“, „Zeit bis zur Knochenoperation“, „Zeit bis zur pathologischen Knochenfraktur“, „Zeit bis zur Rückenmarkkompression“ und „Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen“) ergab, dass dieser Unterschied wesentlich durch die „Zeit bis zur Bestrahlung des Knochens“ bedingt war. Im Rahmen der *post hoc* durchgeführten Subgruppenanalysen zeigte sich, dass der Vorteil für Patienten, welche keine begleitende Steroidmedikation bekamen, besonders ausgeprägt war. Patienten mit begleitender Steroidgabe profitieren mit einem Hazard Ratio von 0,78 (95 %-KI [0,61; 1,01], p = 0,06) und dem nur knapp die Eins überschreitendem Konfidenzintervall sehr wahrscheinlich ebenfalls von der Enzalutamid-Behandlung, allerdings fiel für diese Subgruppe der Effekt statistisch nicht signifikant aus.

Da die Gabe von Bisphosphonaten das Auftreten skelettaler Komplikationen beeinflussen kann, wird dieser Punkt im Folgenden nochmals gesondert aufgegriffen: Ein Einschlusskriterium der Studie war, dass Patienten unter Bisphosphonat-Gabe für mindestens vier Wochen eine stabile Dosis erhalten haben mussten. Sowohl zu Studienbeginn als auch während der Studie (siehe Tabelle 4-12) erhielt eine in beiden Armen vergleichbare Anzahl an Patienten Bisphosphonate. Deshalb sollte der Einfluss von Bisphosphonaten als Begleitmedikation auf das Auftreten skelettbezogener Komplikationen in den beiden Therapiearmen vergleichbar sein und die Aussagekraft des Nachweises nicht beeinträchtigen. Bei Betrachtung der Subgruppe „Gabe von Bisphosphonaten während der Studie“ ergab sich diesbezüglich auch keine Interaktion (siehe Tabelle 4-75).

Um die Wirkung von Enzalutamid im Vergleich zu Placebo auf die Schmerzbelastung der Patienten zu untersuchen, wurden verschiedene Parameter wie „Schmerzlinderung“, „Schmerzprogressionsrate“ und die „Zeit bis zur Schmerzprogression“ analysiert. Im Bezug auf die „Schmerzlinderung“ zeigten die Patienten einen deutlichen Trend zur Reduktion der

Schmerzbelastung unter Enzalutamid-Therapie. Im Enzalutamid-Arm erfuhren 44,9 % (22/49) der Patienten eine Schmerzlinderung, während dies bei nur 6,7 % (1/15) der Patienten in der Placebogruppe der Fall war. Das Relative Risiko von 6,73 (95 %-KI [0,99; 45,88]) spiegelt den deutlichen nummerischen Unterschied zwischen den Behandlungsgruppen wider, erreichte allerdings mit einem p-Wert von p = 0,05 keine statistische Signifikanz. Ein möglicher Grund hierfür könnte die geringe Anzahl an Patienten in der Auswertung sein. Da die Patienten zunächst eine gewisse Schmerzbelastung aufweisen mussten, um eine Linderung messbar zu machen, standen für diese Analyse lediglich 5 % der Patienten zur Verfügung. Um dem Hinweis auf eine verbesserte Schmerzlinderung unter Enzalutamid-Therapie nachzugehen, wurden des Weiteren die Endpunkte „Schmerzprogressionsrate“ und „Zeit bis zur Schmerzprogression“ ausgewertet. Die Betrachtung des Endpunkts „Schmerzprogressionsrate“ zeigte einen signifikanten Vorteil von Enzalutamid auf Studienpopulationsebene, welcher sowohl im Rahmen der *Responder*-Analyse als auch bei Betrachtung der Veränderung der Schmerzpunktzahl gegenüber Studienbeginn nachgewiesen werden konnte. Für die Auswertung standen hier ca. 75% der Patienten der Studienpopulation zur Verfügung (es gingen nur Patienten mit zu den entsprechenden Zeitpunkten gültigen Eintragungen im Schmerztagebuch in die Auswertung ein), weshalb diese Ergebnisse auch hier als möglicherweise hoch verzerrt eingestuft wurden. Eine differenziertere Betrachtung des Endpunkts auf Subgruppenebene zeigte einen Einfluss diverser Faktoren auf die Ergebnisse der beiden Auswertungen. Die Auswertung als *Responder*-Analyse zeigte, dass insbesondere Patienten, welche keine begleitenden Steroide erhielten, Patienten, die bereits zu Studienbeginn einer hohen Schmerzbelastung ausgesetzt waren und Patienten, die jünger als 65 Jahre alt waren, statistisch signifikant von einer Enzalutamid-Therapie profitierten. Auch alle anderen Patienten zeigten unter Enzalutamid-Therapie ein geringeres relatives Risiko, einen Schmerzanstieg zu erleiden, jedoch wurde hier keine statistische Signifikanz erreicht. Im Rahmen der Auswertung des Endpunkts als Veränderung der Schmerzpunktzahl gegenüber Studienbeginn zeigten die Faktoren ECOG-*Performance*-Status zu Studienbeginn und durchschnittliche Schmerzpunktzahl zu Studienbeginn Hinweise auf Beeinflussung des Ergebnisses, jedoch wurde die Signifikanz des Vorteils der Enzalutamid-Behandlung dadurch nicht beeinträchtigt, so dass dieser über die gesamte Studienpopulation hinweg als gegeben angesehen werden kann. Für den Endpunkt „Zeit bis zur Schmerzprogression“ wurde die gesamte ITT-Population betrachtet, wobei sich Enzalutamid mit einer signifikant verlängerten Zeit bis zur Schmerzprogression für die gesamte Studienpopulation als die überlegene Therapie erwies. Bei Betrachtung der Subgruppen ergaben sich keine Hinweise auf oder Belege für Interaktionen. Bei der Betrachtung des Analgetika-Verbrauchs, wie er im Rahmen der Studie erfasst wurde (siehe Tabelle 4-12), zeigte sich hinsichtlich der Anzahl der Patienten, welche mindestens ein begleitendes Schmerzmittel verabreicht bekamen, nur ein geringer Unterschied. Im Placebo-Arm bekamen um 1,8 Prozentpunkte mehr Patienten ein Analgetikum als im Enzalutamid-Arm. Bei einer Betrachtung auf Ebene der einzelnen Medikamente fiel der Unterschied zwischen den Therapiearmen jedoch deutlicher aus. Während mehr Patienten im Enzalutamid-Arm das relativ leichte Schmerzmittel Paracetamol erhielten (Unterschied zwischen den Armen: 1,9 Prozentpunkte), wurden im Placebo-Arm mehr Patienten mit wesentlich stärker schmerzlindernd wirksamen Opiaten behandelt. Dies war für alle verabreichten Opiate der Fall, wobei die Unterschiede zwischen den Armen bis zu

4,3 Prozentpunkte (Oxycodonhydrochlorid) betrugen. Da die Studie doppelt verblindet durchgeführt wurde, ist eine Beeinflussung des behandelnden Arztes hinsichtlich der Schmerzmedikation auszuschließen.

Um Informationen über die Veränderungen der gesundheitsbezogenen Lebensqualität unter Enzalutamid- bzw. Placebo-Therapie zu erhalten, kamen in der Studie die beiden Fragebögen FACT-P und EQ-5D zum Einsatz. Für die Auswertung des FACT-P wurde sowohl eine *Responder*-Analyse durchgeführt, als auch die Veränderung der Punktzahl gegenüber Studienbeginn betrachtet, wobei hier einmal eine Betrachtung der Veränderung gegenüber der letzten zur Verfügung stehenden Erhebung (*last available assessment*) und einmal eine Betrachtung der Veränderung gegenüber Woche 49 erfolgte. Die Auswertung des EQ-5D erfolgte ebenfalls über eine Betrachtung der Veränderung der Gesamtpunktzahl, basierend auf der VAS, zu den selben Zeitpunkten. Die Auswertung des FACT-P zeigte in der *Responder*-Analyse bei Patienten in der Enzalutamid-Gruppe eine signifikant höhere Ansprechrate für die Verbesserung der Lebensqualität. Auch bei Betrachtung der Veränderung der letzten verfügbaren Erhebung gegenüber Studienbeginn ergab sich für diese Patientengruppe ein signifikanter Vorteil. Dieser erstreckte sich neben der Gesamtpunktzahl auch auf die Punktzahlen für FACT-G und alle Subskalen mit Ausnahme der Subskala „Verhältnis zu Freunden, Bekannten und Familie“. Insofern basiert der für den Enzalutamid-Arm gezeigte Vorteil in der Lebensqualität auf Vorteilen in fast allen untersuchten Bereichen. Für die Auswertungen standen ca. 73 % der Patientenpopulation zur Verfügung, weshalb diese Ergebnisse als möglicherweise hoch verzerrt eingestuft wurden. Bei Betrachtung der Veränderung der Punktzahlen in Woche 49 gegenüber Studienbeginn ergaben sich auf keiner Ebene signifikante Unterschiede zwischen den beiden Behandlungsgruppen. Zu diesem späten Zeitpunkt befanden sich nur noch ca. 20 % der Patienten in der Analyse, weshalb dieses Ergebnis aufgrund seiner geringen Aussagekraft nicht in die Bewertung des Zusatznutzens einging. Die differenzierte Betrachtung auf Subgruppenebene zeigte, dass der positive Effekt von Enzalutamid hinsichtlich der Verbesserung der Lebensqualität in der *Responder*-Analyse trotz eines Hinweises auf Interaktion durch die Gabe von Steroiden robust und über die gesamte Studienpopulation hinweg signifikant zu finden war. Bezüglich der Auswertung als Veränderung der letzten verfügbaren Erhebung gegenüber Studienbeginn ergab sich, sowohl bei Betrachtung der FACT-P als auch der FACT-G Gesamtpunktzahl, in dieser Subgruppe ein Beleg für eine Interaktion. Enzalutamid erwies sich vor allem bei Patienten ohne begleitende Steroidgabe als signifikant überlegen gegenüber einer Placebo-Therapie, allerdings zeigte sich auch für Patienten mit Steroiden als Begleitmedikation ein Trend zugunsten Enzalutamid. Die Auswertung des Fragebogens EQ-5D ergab für keinen der beiden untersuchten Zeitpunkte einen signifikanten Unterschied in der gesundheitsgezogenen Lebensqualität zwischen den beiden Therapiegruppen. Im Rahmen der Subgruppenanalysen zeigte sich mit p-Werten von $p = 0,05$ bzw. $p = 0,06$ jeweils ein starker, aber nicht signifikant ausfallender Trend zugunsten von Enzalutamid für Patienten mit mindestens zwei vorangegangenen Chemotherapien (Auswertung letzte verfügbare Erhebung) bzw. mit hoher Schmerzbelastung (Auswertung Woche 49). Bedingt durch die partielle und erst nach Studienbeginn erfolgte Einführung des Fragebogens standen für die Auswertung zu den beiden Zeitpunkten lediglich 15 % bzw. 2 % der Patienten für eine Auswertung zur

Verfügung. Aufgrund des daraus resultierenden hohen Verzerrungspotentials gingen die Ergebnisse aus dem EQ-5D nicht in die Bewertung des Zusatznutzens ein.

Neben den Wirksamkeitsendpunkten wurde in der Studie auch die Sicherheit von Enzalutamid untersucht und diesbezüglich verschiedene Parameter betrachtet. Auf Studienpopulationsebene zeigte sich, dass unter Enzalutamid-Gabe signifikant weniger unerwünschte Ereignisse nach CTCAE Grad ≥ 3 auftraten als unter Placebo-Gabe. Mittels Subgruppenanalysen ergab sich, dass das im Enzalutamid-Arm verminderte Relative Risiko ein unerwünschtes Ereignis nach CTCAE Grad ≥ 3 zu erleiden besonders auf Patienten, die während der Studie Bisphosphonate erhielten, Patienten, welche keine Steroide erhielten, Patienten mit einem ECOG-*Performance*-Status von 0 bzw. 1 zu Studienbeginn und auf Patienten mit einer geringen Schmerzbelastung zurückzuführen war, da für diese Gruppen jeweils signifikante Unterschiede zugunsten von Enzalutamid auftraten. Für alle anderen betrachteten Sicherheitsendpunkte („Unerwünschte Ereignisse“, „Schwerwiegende unerwünschte Ereignisse“ „Studienabbruch aufgrund unerwünschter Ereignisse“ und „Tod aufgrund unerwünschter Ereignisse“) ergaben sich auf Studienpopulationsebene keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen, jedoch zeigte sich bei Betrachtung der Werte der Relativen Risiken für die Endpunkte „Schwerwiegende unerwünschte Ereignisse“ „Studienabbruch aufgrund unerwünschter Ereignisse“ und „Tod aufgrund unerwünschter Ereignisse“ jeweils ein Trend zugunsten von Enzalutamid, welcher mit einem p-Wert von $p = 0,08$ für die „Schwerwiegenden unerwünschten Ereignisse“ besonders ausgeprägt war. Für diesen Endpunkt konnte der positive Trend unter Enzalutamid-Gabe auf diejenigen Patienten, welche keine Steroide erhielten, die einen ECOG-*Performance*-Status von 0 bzw. 1 zu Studienbeginn aufwiesen und auf Patienten mit einer geringen Schmerzbelastung zurückgeführt werden, da innerhalb der Subgruppenanalysen für diese Patienten ein signifikant geringeres relatives Risiko, ein schwerwiegendes Ereignis zu erleiden, nachgewiesen werden konnte. Bezüglich des Endpunkts „Tod aufgrund unerwünschter Ereignisse“ konnte auf Subgruppenebene für Patienten mit begleitender Bisphosphonatgabe ebenfalls ein signifikanter Vorteil für Patienten im Enzalutamid-Arm nachgewiesen werden.

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.

Da es sich bei der Datengrundlage mit der Studie AFFIRM mit insgesamt 1.199 teilnehmenden Patienten um eine der größten bisher durchgeführten randomisierten, kontrollierten Studien mit niedrigem Verzerrungspotenzial in dieser Indikation handelte (Enzalutamid = 800; Placebo = 399), die auch weitere Anforderungen der EMA an eine *pivotal study* erfüllen und vom IQWiG gefordert werden, wurde die Aussagekraft für den Zusatznutzen als „Beleg“ gewertet (siehe auch Abschnitt 4.4.1, Aussagekraft der Nachweise).

Für die Interpretation der Ergebnisse von Subgruppenanalysen wird im IQWiG-Methodenpapier [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011] angegeben: *„Ergebnisse von Subgruppenanalysen sollten […]. nicht das Ergebnis der primären Analyse dominieren; […].“* Die vorliegenden multiplen Subgruppen-Analysen wurden *a posteriori* definiert und durchgeführt, die Fallzahlberechnung war nicht auf diese Subgruppen abgestimmt und die Randomisierung war nur nach zwei Merkmalen der verwendeten Subgruppenmerkmale stratifiziert (ECOG-*Performance*-Status und Punktzahl der Frage 3 des BPI). Die Subgruppen „Gabe von Bisphosphonaten bzw. von Steroiden“ während der Studie weisen zudem eine eingeschränkte Aussagekraft auf, da sie Merkmale verwenden, die sich während der Studie verändern bzw. durch die Therapie selbst

(Enzalutamid bzw. *Best Supportive Care*) beeinflusst werden können. Unter Berücksichtigung dieser Aspekte wurden Aussagen bzgl. eines Zusatznutzens von Enzalutamid in Subgruppen in der folgenden Ableitung des Zusatznutzens nur als Hinweis bewertet. Sie sind demzufolge den Ergebnissen, die sich auf die Ebene der Studienpopulation beziehen, untergeordnet.

Der Zusatznutzen für Enzalutamid sowie dessen Begründung wurde sowohl auf Basis der Ergebnisse der Studie AFFIRM abgeleitet und ergänzt durch den Vergleich dieser Ergebnisse mit der Einordnung des Zusatznutzens für Abirateronacetat und Cabazitaxel. [Bundesministerium für Gesundheit (BMG) 2012a; Bundesministerium für Gesundheit (BMG) 2012b] Auf Studienpopulations- und Subgruppenebene ergab sich für den Vergleich einer Behandlung mit *Best Supportive Care* in der Placebogruppe bzw. der Behandlung mit Enzalutamid + *Best Supportive Care* folgender Zusatznutzen pro Endpunkt:

Gesamtmortaliät

„Gesamtüberleben“

Patienten in der Enzalutamid-Gruppe hatten mit einem medianen Gesamtüberleben von 18,4 Monaten (95 %-KI [17,3; n.e.]) einen Überlebensvorteil von 4,8 Monaten gegenüber Patienten in der Placebogruppe, deren medianes Gesamtüberleben nur 13,6 Monate (95 %-KI [11,3; 15,8]) betrug. Die Hazard Ratio (stratifiziert) betrug 0,63 (95 %-KI [0,53; 0,75], $p < 0,0001$) und der Vorteil war statistisch signifikant. Dies stellt eine bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens gemäß § 5 Absatz 7 der AM-NutzenV dar. Sowohl HR wie der mediane Überlebensvorteil zeigten bessere Werte als bei der Anwendung von Abirateronacetat (HR = 0,74; 95 %-KI [0,64; 0,86] absolute Differenz 4,6 Monate, $p < 0,001$) bzw. Cabazitaxel (HR = 0,70; 95 %-KI [0,59; 0,83] absolute Differenz 2,4 Monate, $p < 0,001$).

Für das Gesamtüberleben lag für drei Subgruppenmerkmale ein Hinweis auf eine Interaktion vor: „Gabe von Steroiden während der Studie“, „Hämoglobinwert zu Studienbeginn“ und „viszerale Erkrankung“. Ein signifikanter Vorteil für Enzalutamid zeigte sich jedoch auch in allen diesen Subgruppen, mit Ausnahme derjenigen Patienten, welche zum Zeitpunkt des Screenings unter einer viszeralen Erkrankung (Lunge/Leber) litten, die keinen statistisch signifikanten Vorteil gegenüber der Behandlung mit *Best Supportive Care* hatten. Das Ausmaß des Zusatznutzens war in allen diesen Subgruppen entweder „beträchtlich“ oder „erheblich“.

Insgesamt wird der **Zusatznutzen für den Endpunkt „Gesamtüberleben“** als ***erheblich (Beleg)*** auf der Ebene der Studienpopulation eingestuft.

Sicherheit: „Tod aufgrund unerwünschter Ereignisse“

Auf Studienpopulationsebene bestand für diesen Endpunkt kein signifikanter Unterschied zwischen den Behandlungsgruppen. Im Rahmen der Subgruppenanalysen wurde ein Beleg für eine Subgruppen-Interaktion in der Gruppe nach Gabe von Bisphosphonaten gesehen, so dass für Patienten, welche studienbegleitend Bisphosphonate bekamen, das relative Risiko

aufgrund eines unerwünschten Ereignisses zu sterben, im Enzalutamid-Arm signifikant geringer ausfiel als im Arm unter *Best Supportive Care* (HR = 0,27; 95 %-KI [0,08; 0,92]). Der Zusatznutzen bzw. verminderte Schaden weist für diese Subgruppe (studienbegleitende Gabe von Bisphosphonaten) für den „Tod aufgrund eines unerwünschten Ereignisses“ ein ***beträchtliches Ausmaß (Hinweis)*** auf**.**

Schwerwiegende Symptome und Nebenwirkungen sowie Lebensqualität

„Radiographisches progressionsfreies Überleben“

Für den Endpunkt „radiographisches progressionsfreies Überleben“ zeigte sich ein statistisch signifikanter Vorteil zugunsten der Enzalutamid-Therapie gegenüber *Best Supportive Care* (Hazard Ratio 0,40 (95 %-KI [0,35; 0,47]) $p < 0,0001$). Patienten in der Enzalutamid-Gruppe hatten dabei im Vergleich zu Patienten in der Placebogruppe mit *Best Supportive Care* ein um 5,4 Monate länger andauerndes medianes radiographisches progressionsfreies Überleben.

Die Subgruppen-Analysen ergaben auch für alle untersuchten Subgruppenmerkmale mit einem Hinweis oder Beleg auf eine Subgruppen-Interaktion einen signifikanten Unterschied mit Vorteil für die Therapiegruppe unter Enzalutamid. Das Ausmaß des Zusatznutzens entsprach für alle Subgruppen – bis auf eine (Subgruppenmerkmal Schmerzpunktzahl ≥ 4 mit KIs = 0,75 ein beträchtliches Ausmaß) – dem Ausmaß auf der Ebene der Studienpopulation.

Das PFS wird von der EMA als patientenrelevantes Ereignis angesehen: *„Prolonged PFS/DFS as such, however, is considered to be of benefit to the patient“* [European Medicines Agency (EMA) 2012] und von anderen Behörden ebenfalls als valider Endpunkt bewertet (siehe Abschnitt 4.5.4).“ Das radiographisch progressionsfreie Überleben wurde für die Ableitung des Zusatznutzens in der Kategorie „Morbidität“ ausgewertet und zeigt - analog dem Überlebensvorteil - eine bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens gemäß § 5 Absatz 7 der AM-NutzenV, was ebenfalls als ***erheblicher Zusatznutzen (Beleg)*** eingeordnet wird.

„Zeit bis zur ersten skelettbezogenen Komplikation“

Die mediane „Zeit bis zur ersten skelettbezogenen Komplikation“ betrug bei Patienten der Enzalutamid-Gruppe 16,7 Monate (95 %-KI [14,6; 19,1]) im Vergleich zu 13,3 Monaten (95 %-KI [9,9; n.e.]) in der Placebogruppe mit *Best Supportive Care* (absolute Differenz: 3,4 Monate; HR = 0,69 (95 %-KI [0,57; 0,84], $p = 0,0001$) und war statistisch signifikant zugunsten der Enzalutamid-Gruppe. Die Analyse von Einzelereignissen in dieser Gruppe (Zeit bis zur Bestrahlung des Knochens, Zeit bis zur Knochenoperation, Zeit bis zur pathologischen Knochenfraktur, Zeit bis zur Rückenmarkkompression und Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen) zeigte, dass der Unterschied wesentlich durch die Zeit bis zur Bestrahlung des Knochens bedingt war, die einen signifikanten Vorteil zugunsten von Enzalutamid (HR = 0,71 (95 %-KI [0,55; 0,91], $p = 0,0058$) zeigte. Da die Komedikation (siehe Tabelle 4-12) in beiden

Therapiegruppen bzgl. der Therapie mit Bisphosphonaten vergleichbar war, wurde eine Verzerrung durch die Komedikation auf diesen Endpunkt nicht angenommen.

Subgruppen-Analysen zeigten für Patienten, welche eine begleitende Steroidmedikation bekamen, keinen signifikanten Vorteil gegenüber Patienten in der Gruppe unter *Best Supportive Care.* Für alle anderen untersuchten Subgruppenmerkmale bestand ebenfalls ein signifikanter Unterschied zu Gunsten der Patienten unter Enzalutamid.

Der absolute Unterschied der Differenz war vergleichbar mit Abirateronacetat (HR = 0,62 (95 %-KI [0,48; 0,79]) p < 0,001), für Cabazitaxel wurde dieser Endpunkt nicht zur Nutzenbewertung herangezogen. Bei einem eindeutig signifikanten Unterschied sowie einer deutlichen medianen Verlängerung von über 4 Monaten wurde für den Endpunkt „Zeit bis zur ersten skelettbezogenen Komplikation“ ein ***beträchtlicher Zusatznutzen (Beleg)*** auf der Ebene der Studienpopulation angenommen.

„Gesundheitsbezogene Lebensqualität nach FACT-P und EQ-5D“

Auf Studienpopulationsebene bestand bezüglich der „gesundheitsgezogenen Lebensqualität nach EQ-5D“ kein Unterschied zwischen den beiden Therapiegruppen. Im Rahmen der Subgruppenanalysen ergaben sich zwar Hinweise und Belege auf Interaktionen, jedoch traten in keiner der betrachteten Subgruppen signifikante Unterschiede zwischen den Therapiearmen auf. Somit wird aus der Auswertung des EQ-5D kein Zusatznutzen oder -schaden für eine der beiden Therapien abgeleitet.

Bei der Untersuchung des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ ergaben sich signifikante Unterschiede mit Vorteil für die Patienten unter Enzalutamid: Sowohl in der *Responder*-Analyse (RR = 2,81 (95 %-KI [2,05; 3,84]; Gegenereignis: 0,68; 95 %-KI [0,63; 0,74], p < 0,0001)) wie auch bei der Betrachtung der Veränderung der FACT-P-Gesamtpunktzahl der letzten verfügbaren Erhebung gegenüber Studienbeginn zeigte sich ein signifikanter Vorteil von Enzalutamid gegenüber *Best Supportive Care* (SMD = 0,42; 95 %-KI; 95 %-KI [0,27; 0,56], p < 0,0001). Dieser Vorteil konnte ebenfalls für die Analyse der Gesamtpunktzahl FACT-G (SMD = 0,36; 95 %-KI [0,21; 0,50], p < 0,0001) und, mit Ausnahme der Subskala „Verhältnis zu Freunden, Bekannten und Familie“ (SMD = 0,06 95 %-KI [-0,08; 0,21], p = 0,41), auch auf Ebene aller einzelnen Subskalen nachgewiesen werden (körperliches Wohlbefinden (SMD = 0,26; 95 %-KI [0,11; 0,40,], p = 0,0006), seelisches Wohlbefinden (SMD = 0,33; 95 %-KI [0,18; 0,47], p < 0,0001), Funktionelles Wohlbefinden (SMD = 0,30; 95 %-KI [0,15; 0,45], p < 0,0001), Prostatakarzinom-Subskala (SMD = 0,34; 95 %-KI [0,20; 0,49], p < 0,0001).

Bezüglich der Auswertung des Endpunkts „Gesundheitsbezogene Lebensqualität nach FACT-P“ als Veränderung der letzten verfügbaren Erhebung gegenüber Studienbeginn erwies sich, sowohl bei Betrachtung der FACT-P- als auch der FACT-G-Gesamtpunktzahl, Enzalutamid bei Patienten ohne begleitende Steroidgabe als signifikant überlegen gegenüber einer Placebo-Therapie mit *Best Supportive Care*. Bei Betrachtung der Veränderung der beiden Gesamtpunktzahlen (FACT-P bzw. des FACT-G) in Woche 49 gegenüber Studienbeginn

ergab sich innerhalb der Studienpopulation für Enzalutamid jedoch kein statistisch signifikanter Vorteil im Vergleich zu Placebo, wobei diese Auswertung auch als hoch verzerrt beurteilt wurde, da über 30 % der randomisierten Patienten in Woche 49 nicht mehr in der Studie waren und den Fragebogen zu diesem Zeitpunkt nicht ausfüllen konnten.

Es leitet sich für den Endpunkt „Gesundheitsbezogene Lebensqualität nach FACT-P“ ein ***erheblicher Zusatznutzen (Hinweis)*** auf der Ebene der Studienpopulation ab.

Sicherheit: „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“

Bei 45,3 % (362/800) der Patienten unter Enzalutamid-Therapie traten unerwünschte Ereignisse nach CTCAE Grad ≥ 3 auf, wohingegen Ereignisse dieser Art bei 53,1 % (212/399) unter *Best Supportive Care* auftraten. Das Relative Risiko, ein unerwünschtes Ereignis nach CTCAE Grad ≥ 3 zu erleiden, zeigte mit einem Wert von 0,85 (95 %-KI [0,76; 0,96]) p = 0,009) einen signifikanten Vorteil für Patienten in der Enzalutamid-Gruppe. In den Subgruppen, die sich bzgl. der begleitendenden Gabe von Bisphosphonaten oder Steroiden unterschieden, sowie den Subgruppen nach ECOG-*Performance*-Status bzw. Schmerzen zu Studienbeginn ergaben sich Hinweise auf Interaktionen (Signifikanter Vorteil für folgende Subgruppen: Patienten mit begleitender Bisphosphonatgabe und Patienten ohne begleitende Steroidmedikation; Patienten mit einem ECOG-*Performance*-Status von 0 oder 1 und auf Patienten mit einer Schmerzpunktzahl < 4 zu Studienbeginn – Ausmaß des Zusatznutzens/geringeren Schadens jeweils *gering*). Insgesamt ergab sich ein geringerer Schaden für diesen Endpunkt „Unerwünschte Ereignisse nach CTCAE Grad ≥ 3“ unter Enzalutamid mit ***geringem Ausmaß (Beleg)*** auf der Ebene der Studienpopulation.

Sicherheit: „Schwerwiegende unerwünschte Ereignisse“

Auf Studienpopulationsebene bestand für diesen Endpunkt kein Unterschied zwischen den Behandlungsgruppen, im Rahmen der Subgruppenanalysen zeigte sich jedoch ein Hinweis auf eine Interaktion für die Subgruppen nach dem ECOG-*Performance*-Status, der begleitenden Steroidgabe sowie der Schmerzpunktzahl zu Studienbeginn: Das relative Risiko für das Auftreten eines schwerwiegenden unerwünschten Ereignisses unter Enzalutamid-Therapie war bei Patienten mit einem ECOG-*Performance*-Status von 0 oder 1, bei Patienten ohne begleitende Steroidgabe und für Patienten mit einer Schmerzpunktzahl < 4 zu Studienbeginn unter Enzalutamid-Therapie signifikant geringer als unter Therapie mit *Best Supportive Care*. Die obere Grenze des Konfidenzintervalls lag auch in den jeweiligen Subgruppen zwischen 0,90 und 1,00, so dass für diese Subgruppen ein geringerer Schaden mit dem Ausmaß ***gering (Beleg)*** für Enzalutamid gegenüber *Best Supportive Care* resultiert.

Studienabbruch aufgrund unerwünschter Ereignisse

Weder auf Studienpopulationsebene noch im Rahmen der Subgruppenanalysen ergaben sich für den Endpunkt „Studienabbruch aufgrund unerwünschter Ereignisse“ signifikante Unterschiede zwischen den Therapiegruppen bzw. Hinweise oder Belege für Interaktionen zwischen Subgruppen. Für diesen Endpunkt besteht daher kein Zusatznutzen oder –schaden ***(Beleg)***.

Nicht schwerwiegende Symptome und Nebenwirkungen:

„Schmerzen“

Das Relative Risiko für den Endpunkt „Patienten mit Schmerzlinderung“ betrug 6,73 (95 %-KI [0,99; 45,88], p = 0,05), wobei unter Enzalutamid 44,9 % der Patienten (22/49) und unter *Best Supportive Care* 6,7 % der Patienten (1/15) eine Schmerzlinderung erfuhren. Der deutliche Trend zu Gunsten der Enzalutamid-Gruppe erreichte jedoch keine statistische Signifikanz. Da nur ein geringer Teil der Studienpopulation in die Auswertung aufgenommen werden konnte, wurde aus diesem Endpunkt kein Zusatznutzen abgeleitet. Um dem Hinweis auf eine geringere Schmerzbelastung unter Enzalutamid-Therapie nachzugehen, wurden weitere Endpunkte ausgewertet.

Für den Endpunkt „Patienten mit Schmerzprogression“ zeigte die *Responder*-Analyse ein relatives Risiko von 0,71 (95 %-KI [0,59; 0,87], p = 0,0008). Die SMD zwischen den beiden Therapiegruppen betrug für den Unterschied der Schmerzpunktzahl zwischen Studienbeginn und Woche 13 -0,48 (95 %KI [-0,62; -0,34], p < 0,0001). Die Analyse der Subgruppen zeigte für die Subgruppenmerkmale „Schmerzpunktzahl ≥ 4“, „ohne begleitende Steroidmedikation“ und „Alter < 65“ Jahre ebenfalls einen signifikanten Vorteil für die Gruppe unter Enzalutamid, dessen Ausmaß für die ersten beiden **Subgruppen** jeweils in der Kategorie „***beträchtlich*“ *(Anhaltspunkt)*** eingeordnet werden kann.

Die mediane „Zeit bis zur Schmerzprogression“ betrug in der Gruppe unter *Best Supportive Care* 13,8 Monate (95 %-KI [13,8; n.e.]), in der Enzalutamid-Gruppe war der Median zum Zeitpunkt der Datenerhebung noch nicht erreicht (HR = 0,56 (95 %-KI [0,41; 0,78], p = 0,0004). Es ergaben sich keine Hinweise oder Belege für Interaktionen in Subgruppen. Somit ergibt sich für den Endpunkt „**Zeit bis zur Schmerzprogression**“ ein Zusatznutzen für Enzalutamid mit dem ***Ausmaß „beträchtlich“ (Beleg)*** auf der Ebene der Studienpopulation.

Insgesamt wurden die Vorteile für Enzalutamid gegenüber *Best Supportive Care* für den Endpunkt „Schmerzen“ unter Berücksichtigung der klinischen Bedeutung von Schmerzen, der der verschiedenen Ergebnisse für die unterschiedlichen Operationalisieungen sowie deren Aussagekraft (siehe auch Abschnitt 4.4.1) und Subgruppen als ***beträchtlicher Zusatznutzen (Beleg)*** im Vergleich zu *Best Supportive Care* bewertet.

„Unerwünschte Ereignisse“

In der AFFIRM-Studie betrug das Relative Risiko 1,00 (95 %-KI [0,99; 1,02], p = 0,67). Für diesen Endpunkt wurde daher ***kein Zusatznutzen oder -schaden (Beleg)*** abgeleitet.

Die Zusammenschau der signifikant positiven Effekte in den Bereichen Mortalität sowie Morbidität und Lebensqualität ergibt einen <u>Beleg</u> für einen <u>erheblichen</u> Zusatznutzen für Enzalutamid im Vergleich zu Best Supportive Care auf der Ebene der Zulassungspopulation, der nicht durch negative Effekte relativiert wird.

Tabelle 4-120: Enzalutamid vs. Placebo: Zusammenfassung des Zusatznutzens auf Endpunktebene für die Studienpopulation der randomisierten kontrollierten Studie AFFIRM

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
Mortalität		
Gesamtüberleben	HR = 0,63 [0,53; 0,75] absolute Differenz: 4,8 Monate absoluter Zeitraum: 18,4 vs. 13,6 Monate p < 0,0001	Endpunktkategorie: Überlebensdauer KIs < 0,85 Beleg für Zusatznutzen Ausmaß: erheblich
Morbidität		
Radiographisches progressionsfreies Überleben (rPFS)	HR = 0,40 [0,35; 0,47] absolute Differenz: 5,4 Monate absoluter Zeitraum: 8,3 vs. 2,9 Monate p < 0,0001	Endpunktkategorie: schwerwiegende /schwere Symptome / Folgekomplikationen KIs < 0,75 Beleg für einen Zusatznutzen Ausmaß: erheblich
Zeit bis zur ersten skelettbezogenen Komplikation	HR = 0,69 [0,57; 0,84] absolute Differenz: 3,4 Monate absoluter Zeitraum: 16,7 vs. 13,3 Monate p = 0,0001	Endpunktkategorie: schwerwiegende /schwere Symptome / Folgekomplikationen KIs < 0,90 Beleg für einen Zusatznutzen Ausmaß: beträchtlich
- Zeit bis zur Bestrahlung des Knochens	HR = 0,71 [0,55; 0,91] absolute Differenz: n.e. p = 0,0058	[a]
- Zeit bis zur Knochenoperation	HR = 2,49 [0,30; 20,79] absolute Differenz: n.e. p = 0,38	[a]
- Zeit bis zur pathologischen Knochenfraktur	HR = 0,85 [0,47; 1,56] absolute Differenz: n.e. p = 0,60	[a]
- Zeit bis zur Rückenmarkkompression	HR = 1,00 [0,64; 1,55] absolute Differenz: n.e. p = 0,99	[a]
- Zeit bis zur Änderung der antineoplastischen Therapie zur Behandlungvon Knochenschmerzen	HR = 0,54 [0,28; 1,04] absolute Differenz: n.e. p = 0,06	[a]
Schmerzlinderung	RR = 6,73 [0,99; 45,88] p > 0,05	Hinweis auf keinen Zusatznutzen oder –schaden

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
Schmerzprogressionsrate (*Responder*-Analyse)	RR = 0,71 [0,59; 0,87] p = 0,0008	Endpunktkategorie: Nicht schwerwiegende Symptome und Nebenwirkungen KIs < 0,90 Hinweis[c] auf einen Zusatznutzen Ausmaß: gering
Schmerzprogressionsrate (Veränderung gegenüber Studienbeginn)	SMD = -0,48 [-0,62; -0,34] p < 0,0001	Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
Zeit bis zur Schmerzprogression	HR = 0,56 [0,41; 0,78] absolute Differenz: n.e. p = 0,0004	Endpunktkategorie: Nicht schwerwiegende Symptome und Nebenwirkungen KIs < 0,80 Beleg für einen Zusatznutzen Ausmaß: beträchtlich
Gesundheitsbezogene Lebensqualität		
FACT-P (*Responder*-Analyse)	RR = 2,81 [2,05; 3,84] Gegenereignis: RR = 0,68 [0,63; 0,74][b] p < 0,0001	Endpunktkategorie: Lebensqualität KIs < 0,75 Hinweis[c] auf einen Zusatznutzen Ausmaß: erheblich
FACT-P (Veränderung Woche 49 gegenüber Studienbeginn)		Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-P Gesamtpunktzahl	SMD = 0,22 [-0,23, 0,67] p = 0,33	
- FACT-G Gesamtpunktzahl	SMD = 0,19 [-0,25; 0,63] p = 0,40	
- FACT-P Körperliches Wohlbefinden	SMD = -0,07 [-0,50; 0,36] p = 0,76	[a]
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,23 [-0,20; 0,66] p = 0,29	[a]
- FACT-P Seelisches Wohlbefinden	SMD = 0,19 [-0,25; 0,63] p = 0,40	[a]
- FACT-P Funktionelles Wohlbefinden	SMD = 0,09 [-0,35; 0,53] p = 0,70	[a]
- FACT-P Prostatakarzinom-Subskala	SMD = 0,05 [-0,39; 0,49] p = 0,83	[a]
FACT-P (Veränderung last assessment gegenüber Studienbeginn)		

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
- FACT-P Gesamtpunktzahl	SMD = 0,42 [0,27; 0,56] p < 0,0001	Endpunktkategorie: Lebensqualität Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-G Gesamtpunktzahl	SMD = 0,36 [0,21; 0,50] p < 0,0001	Endpunktkategorie: Lebensqualität Ableitung des Zusatznutzens auf Basis der *Responder*-Analyse[d]
- FACT-P Körperliches Wohlbefinden	SMD = 0,26 [0,11; 0,40] p = 0,0006	[a]
- FACT-P Verhältnis zu Freunden, Bekannten und Familie	SMD = 0,06 [-0,08; 0,21] p = 0,41	[a]
- FACT-P Seelisches Wohlbefinden	SMD = 0,33 [0,18; 0,47] p < 0,0001	[a]
- FACT-P Funktionelles Wohlbefinden	SMD = 0,30 [0,15; 0,45] p < 0,0001	[a]
- FACT-P Prostatakarzinom-Subskala	SMD = 0,34 [0,20; 0,49] p < 0,0001	[a]
EQ-5D VAS Gesamtpunktzahl (Veränderung Woche 49 gegenüber Studienbeginn)	SMD = 0,85 [-0,69; 2,40] p = 0,26	Endpunktkategorie: Lebensqualität Hinweis auf keinen Zusatznutzen oder –schaden
EQ-5D VAS Gesamtpunktzahl (Veränderung letzte verfügbare Erhebung gegenüber Studienbeginn)	SMD = -0,03 [-0,36; 0,30] p = 0,86	Endpunktkategorie: Lebensqualität Hinweis auf keinen Zusatznutzen oder –schaden
Sicherheit		
Unerwünschte Ereignisse	RR = 1,00 [0,99; 1,02] p = 0,67	Endpunktkategorie: Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden
Unerwünschte Ereignisse CTCAE Grad ≥ 3	RR = 0,85 [0,76; 0,96] p = 0,009	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen KIs < 1,00 Beleg für einen Zusatznutzen Ausmaß: gering
Schwerwiegende unerwünschte Ereignisse	RR = 0,87 [0,74; 1,02] p = 0,08	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden
Studienabbruch aufgrund unerwünschter Ereignisse	RR = 0,78 [0,53; 1,15] p = 0,20	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden

	Enzalutamid vs. Placebo	**Ausmaß des Zusatznutzens**
Zielgröße	**Effektschätzer [95 %-KI] p-Wert**	
Tod aufgrund unerwünschter Ereignisse	RR = 0,82 [0,43; 1,58] p = 0,55	Endpunktkategorie: schwerwiegende / schwere Nebenwirkungen Beleg für keinen Zusatznutzen oder –schaden
a: Die Subskalen werden nicht zur separaten Ableitung eines Zusatznutzens herangezogen, sondern dienen der Absicherung des Zusatznutzens auf Ebene der Gesamtskala. b: Berechnung des RR für das Gegenereignis c: Liegt ein hohes Verzerrungspotenzials vor, wurde die Aussagekraft um eine Kategorie gemindert, d.h. ein Beleg wurde in einem solchen Fall als Hinweis bewertet. d: Wie in den Methoden 4.0 [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011] vorgeschlagen, wurde zur Ableitung des Zusatznutens eine *Responder*-Analyse durchgeführt. KI: Konfidenzintervall KIs: obere Grenze Konfidenzintervall RR: Relatives Risiko HR: Hazard Ratio SMD: *Standardized Mean Difference* CTCAE: *Common Toxicity Criteria for Adverse Events* n.e.: nicht erreicht		

Zusammenfassende Tabellen zu den Ergebnissen der Subgruppen-Auswertungen siehe Tabelle 4-109.

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 4-121: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Männer mit metastasiertem kastrationsresistentem Prostatakarzinom, deren Erkrankung während oder nach einer Chemotherapie mit Docetaxel fortschreitet.	erheblich

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 zutreffend

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 4-122: 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.

Indirekte Vergleiche wurden nicht 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.

Nicht randomisierte vergleichende Studien wurden nicht verwendet.

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

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

Nicht zutreffend

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005[13], Molenberghs 2010[14]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006[15]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95 %-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[16]) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

[13] *Burzykowski T (Ed.: The evaluation of surrogate endpoints. New York: Springer; 2005.*
[14] *Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.*
[15] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*
[16] *Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.*

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.

Das progressionsfreie Überleben (PFS) ist definiert als das Zeitintervall von der Randomisierung bis zum ersten nachweisbaren Auftreten einer Progression (entweder radiographische Progression oder skelettbezogene Komplikation) oder bis zum Tod jedweder Ursache. [Medivation Inc. 2009]

Das progressionsfreie Überleben ist ein aussagekräftiger und relevanter Endpunkt, da ein Progress, je nach Tumorentität und Krankheitsstadium häufig Monate bis Jahre bevor das Überleben gemessen werden kann, auftritt. Das PFS ist somit ein früher Indikator für das Ansprechen des Patienten auf eine Therapie und wird zudem nicht durch den Effekt möglicher Folgetherapien beeinflusst. Aufgrund seiner Definition kommt es zu häufigerem Auftreten von Ereignissen im Vergleich zum Gesamtüberleben und lässt sich daher bereits bei kleineren Patientenpopulationen und kürzerer Studiendauer bestimmen.

Das progressionsfreie Überleben ist somit als patientenrelevanter Endpunkt anzusehen. Der aktuelle Entwurf der EMA *Guideline on the evaluation of anticancer medicinal products in man* [European Medicines Agency (EMA) 2012] sieht als akzeptable primäre Endpunkte die Heilungsrate, das Gesamtüberleben und das progressionsfreie bzw. krankheitsfreie Überleben vor. Auch in Fachkreisen [Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO) 2010] sowie vom Paul Ehrlich Institut wird die Erfassung des Endpunkts „progressionsfreies Überleben“ akzeptiert. [Paul Ehrlich Institut 2010] Der Gemeinsamen Bundesausschuss merkt dazu an: „*Von den verschiedenen […] Parametern eignet sich in erster Linie das progressionsfreie Überleben*.“ [Gemeinsamer Bundesausschuss (G-BA) 2010] Auch die im vorherigen Absatz angesprochenen ethischen Gesichtspunkte (kleinere Patientenpopulationen und kürzere Studiendauer) sprechen für eine Patientenrelevanz dieses Endpunktes.

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.

Studie AFFIRM:

Studienbericht:

AFFIRM: A Multinational Phase 3, Randomized,Double-Blind, Placebo-Controlled Efficacy and Safety Study of Oral MDV3100 in Patients with Progressive Castration-Resistant Prostate Cancer Previously Treated with Docetaxel-Based Chemotherapy.

Appendices zum Studienbericht:

Appendix 16.1.1, Appendix 16.1.2, Appendix 16.1.3, Appendix 16.1.4, Appendix 16.1.5, Appendix 16.1.6, Appendix 16.1.7, Appendix 16.1.8, Appendix 16.1.9, Appendix 16.1.10, Appendix 16.1.11, Appendix 16.1.12 und Appendix 16.2

Zusatzanalysen:

SAP Zusatzanalysen, Ergebnisse der Zusatzauswertungen

Publikation:

Scher, H.I.; Fizazi, K.; Saad, F.; Taplin, M.E.; Sternberg, C.N.; Miller, K. et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. N.Engl.J.Med. 27.9.2012. 367, (13), 1187-1197. [Scher 2012]

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

1. Aragon-Ching Jeanny B. Enzalutamide (formerly MDV3100) as a new therapeutic option for men with metastatic castration-resistant prostate cancer. Asian journal of andrology 2012; 14(6): 805-806.

2. Astellas Pharma Europe BV. Relative Bioavailability Study of Enzalutamide in Prostate Cancer Patients [online]. 2013 [Zugriff: 23.07.2013] URL: http://clinicaltrials.gov/ct2/show/study/NCT01902251?term=9785-CL-0003&rank=1

3. Astellas Pharma GmbH. Xtandi 40 mg Weichkapseln: Fachinformation [online]. 4.2013a [Zugriff: 22.07.2013] URL: www.fachinfo.de

4. Astellas Pharma GmbH. Studie AFFIRM - Tabellarische Darstellung von Krampfanfällen. 2013b

5. Atkinson, T.M.; Mendoza, T.R.; Sit, L.; Passik, S.; Scher, H.I.; Cleeland, C. et al. The Brief Pain Inventory and its "pain at its worst in the last 24 hours" item: clinical trial endpoint considerations. Pain Medicine 2010; 11(3): 337-346.

6. Berruti, A.; Generali, D.; Tampellini, M. Enzalutamide in prostate cancer after chemotherapy. New England Journal of Medicine 2012; 367(25): 2448.

7. Bubley, G.J.; Carducci, M.; Dahut, W.; Dawson, N.; Daliani, D.; Eisenberger, M. et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the Prostate-Specific Antigen Working Group. Journal of Clinical Oncology 1999; 17(11): 3461-3467.

8. Bundesministerium für Gesundheit (BMG). Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) Cabazitaxel. 29.3.2012a [Zugriff: 16.06.2013] URL: http://www.g-ba.de/informationen/nutzenbewertung/10/#tab/beschluesse

9. Bundesministerium für Gesundheit (BMG). Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) Abirateronacetat [online]. 29.3.2012b [Zugriff: 16.06.2013] URL: http://www.g-ba.de/informationen/nutzenbewertung/19/#tab/beschluesse

10. Cella, D.; Nichol, M.B.; Eton, D.; Nelson, J.B.; Mulani, P. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy--Prostate: results from a clinical trial of patients with metastatic hormone-refractory prostate cancer. Value in Health 2009; 12(1): 124-129.

11. Deeks, J. & Higgins, J. Statistical algorithms in Review Manager 5. 2010 August

12. Deutsche Gesellschaft für Hämatologie und Onkologie (DGHO). Frühe Nutzenbewertung onkologischer Arzneimittel. Diskussionspapier der DGHO zu den Fragen des Gemeinsamen Bundesausschusses anlässlich der Diskussionsrunde am 29. November 2010 [online]. 29.11.2010 [Zugriff: 26.06.2012] URL: http://www.dgho.de/informationen/grundlegende-dokumente/gesundheitspolitische-papiere/101126_Diskussionspapier_DGHO_29.11.2010_final.pdf

13. Dumas, L.; Payne, H.; Chowdhury, S. The evolution of antiandrogens: MDV3100 comes of age. Expert review of anticancer therapy 2012; 12(2): 131-133.

14. Esper, P.; Mo, F.; Chodak, G.; Sinner, M.; Cella, D.; Pienta, K.J. Measuring quality of life in men with prostate cancer using the functional assessment of cancer therapy-prostate instrument. Urology 1997; 50(6): 920-928.

15. European Medicines Agency (EMA). Points to consider on application with 1. meta-analyses; 2. one pivotal study [online]. 2001 [Zugriff: 23.04.2013] URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003657.pdf

16. European Medicines Agency (EMA). Guideline on the evaluation of anticancer medicinal products in man [online]. 13.12.2012 [Zugriff:16.04.2013] URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf

17. Fleming, M.T.; Scher, H.I.; Fizazi, K.; Taplin, M.E.; Forer, D.; Hirmand, M. et al. Long-term responders to enzalutamide (ENZA) during the phase III AFFIRM trial: Baseline characteristics and efficacy outcomes. Journal of Clinical Oncology 2013; 31(6).

18. Gemeinsamer Bundesausschuss (G-BA). Diskussionsveranstaltung zur Nutzenbewertung onkologischer Arzneimittel nach § 35a SGB V am 29. November 2010 [online]. 29.11.2010 [Zugriff am 26.06.2012] URL: http://www.g-ba.de/institution/service/veranstaltungen/veranstaltungen-amnog/onkologie

19. Gemeinsamer Bundesausschuss (G-BA). Bestimmung der zweckmäßigen Vergleichstherapie - Beratungsanforderung 2012-B-030 - Enzalutamide zur Behandlung des metastasierten Prostatakarzinoms. 12.9.2012

20. Gemeinsamer Bundesausschuss (G-BA). Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 21.2.2013 [Zugriff: 13.09.2012] URL: http://www.g-ba.de/downloads/62-492-634/VerfO_2012-06-21.pdf

21. Geynisman, D.M.; Szmulewitz, R.Z.; Stadler, W.M. A trial postmortem: Challenges in conducting a randomized, double-blind, phase 2 study in men with castration-resistant prostate cancer. European Urology 2012; 62(5): 864-866.

22. Hoffman-Censits, J.; Kelly, W.K. Enzalutamide: a novel anti-androgen for patients with castrate resistant prostate cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2013.

23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden Version 4.0 [online]. 23.9.2011 [Zugriff: 13.09.2012] URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf

24. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Aktualisierung einiger Abschnitte der Allgemeinen Methoden Version 4.0 sowie neue Abschnitte zur Erstellung Allgemeinen Methoden Version 4.1 - Entwurf vom 18.04.2013 [online]. 2013 [Zugriff: 22.07.2013] URL: https://www.iqwig.de/download/13-04-18_Entwurf_Aktualisierung_Allgemeine_Methoden_4-0.pdf

25. Kuznar, W. Enzalutamide extends survival, improves qol in advanced prostate cancer. American Health and Drug Benefits 2012; 5(SPL.ISS. 5).

26. Medivation Inc. AFFIRM (Amendment 1 Version 2.0): A Multinational Phase 3, Randomized, Double-Blind, Placebo-Controlled Efficacy and Safety Study of Oral MDV3100 in Patients with Progressive Castration-Resistant Prostate Cancer Previously Treated with Docetaxel-based Chemotherapy. 2009

27. Medivation Inc. AFFIRM (Clinical study report): A Multinational Phase 3, Randomized,Double-Blind, Placebo-Controlled Efficacy and Safety Study of Oral MDV3100 in Patients with Progressive Castration-Resistant Prostate Cancer Previously Treated with Docetaxel-Based Chemotherapy. 2012a

28. Medivation Inc. NCT00974311: Safety and Efficacy Study of MDV3100 in Patients With Castration-Resistant Prostate Cancer Who Have Been Previously Treated With Docetaxel-based Chemotherapy (AFFIRM) [online]. 2012b [Zugriff: 13.02.2013] http://www.clinicaltrials.gov/ct2/show/NCT00974311?term=Enzalutmide+OR+MDV3100+OR+Xtandi&phase=123&rank=6&submit_fld_opt=

29. Medivation Inc. EUCTR2009-013174-41-BE: A study of patients with prostate cancer who have previously been treated with docetaxel-based chemotherapy, where patients receive either study drug or placebo [online]. 2012c [Zugriff: 30.07.2013] http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2009-013174-41-BE

30. Miller, K.; Scher, H.I.; Fizazi, K.; Basch, E.; Sternberg, C.N.; Hirmand M et al. Effect of enzalutamide on health-related quality of life (Hrqol) in men with metastatic castration-resistant prostate cancer (mcrpc) following docetaxel-based therapy: results from the affirm study (Abstract 17, Poster). Journal of Clinical Oncolog 2013; 31(6).

31. Mulders, P.; Scher, H.; Fizazi, K.; Saad, F.; Taplin, M.; Sternberg, C. et al. MDV3100, an androgen receptor signaling inhibitor, improves overall survival in patients with prostate cancer post docetaxel: Results from the Phase 3 AFFIRM study. Urology 2012; 80(3): -S30.

32. Paul Ehrlich Institut. Schreiben vom 08. November 2010 an den Gemeinsamen Bundesausschuss zur Diskussionsrunde zur frühen Nutzenbewertung onkologischer Arzneimittel [online]. 2010 [Zugriff: 26.06.2012] URL: http://www.g-ba.de/downloads/17-98-3006/PEI.pdf

33. Payton, S. Prostate cancer: MDV3100 has antitumor activity in castration-resistant disease. Nature Reviews Urology 2010; 7(6): 300.

34. Rabin, R., Oemar, M., & Oppe, M. EQ-5D-3L User Guide - Basic information on how to use the EQ-5D-3L instrument (Version 4.0). 4.2011 [Zugriff: 24.07.2013] URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-3L.pdf

35. Saad, F.; de Bono, J.S.; Shore, N.D.; Fizazi, K.; Hirmand, M.; Forer, D. et al. Efficacy outcomes by baseline prostate-specific antigen (PSA): Results from the phase 3 affirm trial. Journal of Urology 2013; 189(4): -e324.

36. Scher, H.I.; Eisenberger, M.; D'Amico, A.V.; Halabi, S.; Small, E.J.; Morris, M. et al. Eligibility and outcomes reporting guidelines for clinical trials for patients in the state of a rising prostate-specific antigen: recommendations from the Prostate-Specific Antigen Working Group. Journal of Clinical Oncology 2004; 22(3): 537-556.

37. Scher, H.I.; Halabi, S.; Tannock, I.; Morris, M.; Sternberg, C.N.; Carducci, M.A. et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. Journal of Clinical Oncology 2008; 26(7): 1148-1159.

38. Scher, H.I.; Beer, T.M.; Higano, C.S.; Anand, A.; Taplin, M.E.; Efstathiou, E. et al. Antitumour activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. Lancet 2010; 375(9724): 1437-1446.

39. Scher, H.I.; Fizazi, K.; Saad, F.; Taplin, M.E.; Sternberg, C.N.; Miller, K. et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. New England Journal of Medicine 2012; 367(13): 1187-1197.

40. Scher, H.I.; Fizazi, K.; Saad, F.; Chi, K.N.; Taplin, M.E.; Sternberg, C.N. et al. Impact of on-study corticosteroid use on efficacy and safety in the phase III AFFIRM study of enzalutamide (ENZA), an androgen receptor inhibitor. Journal of Clinical Oncology 2013; 31(6).

41. Sternberg, C.N.; De Bono, J.S.; Chi, K.N.; Fizazi, K.; Mulders, P.; Hirmand, M. et al. Outcomes in elderly patients with metastatic castration-resistant prostate cancer (mCRPC) treated with the androgen receptor inhibitor enzalutamide: Results from the phase III AFFIRM trial. Journal of Clinical Oncology 2013; 31(6).

42. Webster, K.; Odom, L.; Peterman, A.; Lent, L.; Cella, D. The functional assessment of chronic illness therapy (FACIT) measurement system: Validation of version 4 of the core questionnaire. Quality of Life Research 1999; 8(7): 604.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50") und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

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

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

[17] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Datenbankname	The Cochrane Library
Suchoberfläche	http://onlinelibrary.wiley.com/cochranelibrary/search/ → Search → Search All Text
Datum der Suche	08.07.2013
Zeitsegment	Keine zeitliche Begrenzung
Suchfilter	Enzalutamide [SEARCH ALL TEXT] OR Xtandi [SEARCH ALL TEXT] OR MDV3100 [SEARCH ALL TEXT]

#	**Suchbegriffe**	**Ergebnis**
1	Enzalutamide	1
2	Xtandi	0
3	MDV3100	1
4	1 OR 2 OR 3.	2
Treffer für #4	Cochrane Reviews All: 0 Other Reviews: 0 Trials: 1 Methods Studies: 0 Technology Assessments: 1 Economic Evaluations: 0 Cochrane Groups: 0	

Update der Suche am 08.07.2013

Datenbankname Embase®
Embase® Alert

Suchoberfläche ProQuest® Dialog® DataStar®

Datum der Suche 08.07.2013

Zeitsegment Embase®: keine zeitliche Begrenzungl
Embase® Alert: keine zeitliche Begrenzung

Suchfilter Filter für randomisierte kontrollierte Studien nach Wong 2006 [18] – Strategy minimizing difference between sensitivity and specificity
Search Strategy for identifying randomized trials in Embase® and Embase® Alert entsprechend der Syntax für ProQuest® Dialog® DataStar®

#	**Suchbegriffe**	**Ergebnis**
1	Enzalutamid*	272
2	MDV3100	217
3	Xtandi	26
4	1 OR 2 OR 3.	413
5	random*	980860
6	trial*	1509775
7	control*	6933703
8	placebo	323325
9	5 AND (6 OR 7 OR 8)	7859940
10	Prostata	3087
11	prostate	223918
12	cancer	2992797
13	carcinoma	888692
14	neoplasm*	505971
15	4 AND 9 AND 10 OR 11 AND 12 OR 13 OR 14	83

Update der Suche am 08.07.2013

[18] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Datenbankname	MEDLINE®
Suchoberfläche	ProQuest® Dialog® DataStar®
Datum der Suche	08.07.2013
Zeitsegment	1950 bis aktuell
Suchfilter	Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision) entsprechend der Syntax für ProQuest® Dialog® DataStar®

#	Suchbegriffe	Ergebnis
1	Enzalutamid*	68
2	MDV3100	115
3	Xtandi	5
4	1 OR 2 OR 3.	168
5	random*	830230
6	trial*	1124307
7	control*	3693170
8	placebo	163601
9	5 AND (6 OR 7 OR 8)	4332004
10	Prostata	2071
11	prostate	129709
12	10 OR 11	130027
13	cancer	1538881
14	carcinoma	603975
15	neoplasm*	2036108
16	4 AND 9 AND 12 AND 13 OR 14 OR 15	16

Update der Suche am 08.07.2013

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

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

Studienregister	Clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	13.02.2013 + 25.06.2013
Suchstrategie	Advanced Search: Search Terms: Enzalutamide OR MDV3100 OR Xtandi Additional Criteria/Phase: Phase II OR Phase III OR Phase IV
Treffer	14

Studienregister	International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO)
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	13.02.2013 + 25.06.2013
Suchstrategie	1) Advanced Search: Title: Enzalutamide OR MDV3100 OR Xtandi Recruiting Status: All 2) Advanced Search: Intervention: Enzalutamide OR MDV3100 OR Xtandi Recruiting Status: All
Treffer	**1)** 21 Treffer für 16 Studien; **2)** 19 Treffer für 15 Studien

Update der Suche am 25.06.2013: Die Aktualisierung der Suche in den beiden Studienregistern ergab keine weiteren relevanten Studien mit der gleichen Suchstrategie.

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.

Ausschlussgrund A2 (Doppelpublikation ohne relevante Zusatzinformation) (alphabetisch geordnet):

1. Kuznar, W., 2012. Enzalutamide extends survival, improves qol in advanced prostate cancer. *American Health and Drug Benefits*, 5, (SPL.ISS. 5). [Kuznar 2012]

2. Berruti, A.; Generali, D.; Tampellini, M., 2012. Enzalutamide in prostate cancer after chemotherapy. *New England Journal of Medicine*, 367, (25), 2448. [Berruti 2012]

3. Dumas, L.; Payne, H.; Chowdhury, S., 2012. The evolution of antiandrogens: MDV3100 comes of age. *Expert review of anticancer therapy*, 12, (2), 131-133. [Dumas 2012]

Ausschlussgrund: Einschlusskriterium E2 (Intervention: Enzalutamid gemäß Zulassung) nicht erfüllt (alphabetisch geordnet):

1. Geynisman, D.M.; Szmulewitz, R.Z.; Stadler, W.M., 2012. A trial postmortem: Challenges in conducting a randomized, double-blind, phase 2 study in men with castration-resistant prostate cancer. *European Urology*, 62, (5), 864-866. [Geynisman 2012]

Ausschlussgrund: Einschlusskriterium E5 (Studientyp: randomisierte kontrollierte Studie (RCT)) nicht erfüllt (alphabetisch geordnet):

1. Aragon-Ching, J.B., 2012. Enzalutamide (formerly MDV3100) as a new therapeutic option for men with metastatic castration-resistant prostate cancer. *Asian journal of andrology*, 14, (6), 805-806. [Aragon-Ching Jeanny B. 2012]

2. Hoffman-Censits, J.; Kelly, W.K., 2013. Enzalutamide: a novel anti-androgen for patients with castrate resistant prostate cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. [Hoffman-Censits 2013]

3. Payton, S., 2010. Prostate cancer: MDV3100 has antitumor activity in castration-resistant disease. *Nature Reviews Urology*, 7, (6), 300.[Payton 2010]

Ausschlussgrund: Einschlusskriterium E6 (Vollpublikation verfügbar) nicht erfüllt (alphabetisch geordnet):

1. Fleming, M.T.; Scher, H.I.; Fizazi, K.; Taplin, M.E.; Forer, D.; Hirmand, M. et al. Long-term responders to enzalutamide (ENZA) during the phase III AFFIRM trial:

Baseline characteristics and efficacy outcomes. Journal of Clinical Oncology 2013; 31(6). [Fleming 2013]

2. Miller, K.; Scher, H.I.; Fizazi, K.; Basch, E.M.; Sternberg, C.N.; Hirmand, M. et al. Effect of enzalutamide on health-related quality of life (HRQoL) in men with metastatic castration-resistant prostate cancer (mCRPC) following docetaxel-based therapy: Results from the AFFIRM study. Journal of Clinical Oncology 2013; 31(6). [Miller 2013]

3. Mulders, P.; Scher, H.; Fizazi, K.; Saad, F.; Taplin, M.; Sternberg, C.; Miller, K.; De Wit, R.; Hirmand, M.; Selby, B.; De Bono, J., 2012. MDV3100, an androgen receptor signaling inhibitor, improves overall survival in patients with prostate cancer post docetaxel: Results from the Phase 3 AFFIRM study. Urology, 80, (3), -S30. [Mulders 2012]

4. Saad, F.; de Bono, J.S.; Shore, N.D.; Fizazi, K.; Hirmand, M.; Forer, D. et al. Efficacy outcomes by baseline prostate-specific antigen (PSA): Results from the phase 3 affirm trial. Journal of Urology 2013; 189(4): -e324. [Saad 2013]

5. Scher, H.I.; Fizazi, K.; Saad, F.; Chi, K.N.; Taplin, M.E.; Sternberg, C.N. et al. Impact of on-study corticosteroid use on efficacy and safety in the phase III AFFIRM study of enzalutamide (ENZA), an androgen receptor inhibitor. Journal of Clinical Oncology 2013; 31(6). [Scher 2013]

6. Sternberg, C.N.; De Bono, J.S.; Chi, K.N.; Fizazi, K.; Mulders, P.; Hirmand, M. et al. Outcomes in elderly patients with metastatic castration-resistant prostate cancer (mCRPC) treated with the androgen receptor inhibitor enzalutamide: Results from the phase III AFFIRM trial. Journal of Clinical Oncology 2013; 31(6). [Sternberg 2013]

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-123 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

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

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

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

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

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 4-124 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen[a]		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie[b] (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag[c] (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
9785-CL-0003	nein	ja	ja	nein	ja [Astellas Pharma Europe BV 2013]	nein

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.

d: Das Studienprotokoll liegt vor und wird im Modul 5 abgelegt.

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 4-125 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 4-125 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-125 (Anhang): Studiendesign und -methodik für die Studie AFFIRM (CRPC2)

Item[a]	Charakteristikum	Studieninformation
-	**Studienziel**	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel der Studie war die Untersuchung der Effektivität und Sicherheit von Enzalutamid im Vergleich zu Placebo bei Patienten mit einem progredienten, kastrationsresistentem Prostatakarzinom, welches auf eine oder zwei zuvor durchgeführte Chemotherapien nicht angesprochen hat, von denen mindestens eine Chemotherapie Docetaxel enthielt. Die Hypothese, die in dieser Studie untersucht werden sollte, bestand darin, dass durch die Behandlung mit Enzalutamid das Überleben bei Patienten mit einem Prostatakarzinom, welches nicht auf eine primäre Unterdrückung der Androgenproduktion oder auf eine vorausgegangene Docetaxel-haltige Chemotherapie angesprochen hat, verlängert werden kann.
-	**Methoden**	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Prospektive, multinationale, multizentrische, doppelblinde, randomisierte, placebokontrollierte, in Parallelgruppen durchgeführte Phase-III-Studie zum Vergleich der Wirksamkeit und Sicherheit der oralen Gabe von 160 mg Enzalutamid pro Tag, aufgeteilt auf vier Einzeldosen à 40 mg, mit der oralen Gabe von Placebo. Das Zuteilungsverhältnis auf die Enzalutamid- und Placebogruppe betrug 2:1.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Insgesamt wurden während der Studie drei Protokolländerungen („*protocol amendments*“) zu folgenden Zeitpunkten vorgenommen: Version — Datum Originalprotokoll Version 1.0 — 21.05.2009 Protokolländerung Nr. 1 Version 2.0 — 30.07.2009 Protokolländerung Nr. 2 Version 3.0 — 01.04.2010 Protokolländerung Nr. 3 Version 4.0 — 19.04.2011 <u>Protokolländerung Nr. 1 (30.07.2009):</u> Die Tagesdosis von Enzalutamid wurde von 240 mg auf 160 mg reduziert. Das Ziel dieser Protokolländerung war es, das Risiko-Nutzen-Profil in dieser Phase-III-Studie zu verbessern. Die Beurteilung von verfügbaren Daten aus der Phase-I-Studie (S-3100-1-01) ergab, dass die Nachhaltigkeit der Krankheitskontrolle bei Patienten, bei denen bereits eine Chemotherapie durchgeführt wurde, bei einer Enzalutamid-Tagesdosis zwischen 150 mg und 240 mg ähnlich war. Der Anteil der Patienten, bei denen bereits eine Chemotherapie durchgeführt wurde und bei denen unabhängig von der

Item[a]	Charakteristikum	Studieninformation
		Untersuchungsmethode (Prostata-spezifisches Antigen (PSA)-Wertbestimmung, radiografisch oder klinisch) keine Krankheitsprogression nachgewiesen werden konnte, lag in der 12. und 24. Therapiewoche in der Kohorte mit einer Tagesdosis von 150 mg entsprechend bei 54 % und 31 % und in der Kohorte mit einer Tagesdosis von 240 mg bei 67 % und 33 %. Aufgrund dieser vergleichbaren Wirksamkeit, der dosisabhängigen Zunahme von Müdigkeit des Grads 3 (0 % bei einer Tagesdosis von 150 mg und 9 % bei einer Tagesdosis von 240 mg), Berichten über epileptische Anfälle bei höheren Dosierungen (ein bestätigtes Ereignis bei einer Tagesdosis von 360 mg und ein weiteres bei einer Dosis von 600 mg pro Tag, ein unbestätigter epileptischer Anfall bei einer Tagesdosis von 480 mg pro Tag) und einer festen Kapselstärke von 40 mg, wurde eine Enzalutamid-Tagesdosis von 160 mg für die Untersuchung in der AFFIRM-Studie gewählt. <u>Protokolländerung Nr. 2 (01.04.2010):</u> Diese Protokolländerung ist eine Europa-spezifische – Änderung. Zur Beurteilung der Lebensqualität der Studienpatienten wurde der Fragebogen EQ-5D in die Studie integriert. Dieser Fragebogen wurde in den Wochen 1, 13, 25 und anschließend alle 12 Wochen ausgewertet. Die EQ-5D-Daten wurden deskriptiv nach Behandlungsgruppe und Studienvisite zusammengefasst. <u>Protokolländerung Nr. 3 (19.04.2011):</u> Nach Rücksprache mit dem Lenkungsausschuss und der Food and Drug Administration (FDA) wurde eine Interimanalyse hinsichtlich des Gesamtüberlebens bei ungefähr 520 Todesereignissen hinzugefügt. Die Anzahl der Todesereignisse für die Abschlussanalyse wurde von 786 auf 650 reduziert. Diese Anpassung in dem statistischen Analyseplan wurde vorgenommen, da die COU-AA-301-Studie mit Abirateronacetat plus Prednison vorzeitig beendet wurde, als in der Interimanalyse ein Überlebensvorteil gezeigt werden konnte (Hazard Ratio von 0,646). Dabei ist festzuhalten, dass die AFFIRM- und die COU-AA-301-Studie ein ähnliches Studiendesign und ähnliche Patientenpopulationen aufweisen. Diese neuen Daten legten nahe, dass die Power der AFFIRM-Studie zu hoch angesetzt wurde und daher die Ziel-Hazard-Ratio für die finale Gesamtanalyse von 0,80 auf 0,76 herabgesetzt wurde, was einer Reduktion in der Anzahl der Todesereignisse von 786 auf 650 entspricht. Die Interimanalyse erfolgte bei 520 Todesereignissen. Zudem wurden die sekundären Endpunkte vereinfacht: - Der kombinierte Endpunkt progressionsfreies Überleben, welcher sich auf den Zeitraum vom Tag der Randomisierung bis zum ersten Auftreten einer radiografischen Progression, eines skelettalen Ereignisses oder Tod bezieht, wurde in den Endpunkt radiografisch progressionsfreies Überleben geändert. Dabei bezieht sich der Endpunkt radiografisch progressionsfreies Überleben auf den Zeitraum vom Tag der Randomisierung bis zum ersten Auftreten einer radiografischen

Item[a]	Charakteristikum	Studieninformation
		Progression oder Tod. Auf den sekundären Endpunkt hinsichtlich des Auftretens einer radiografischen Progression wurde verzichtet. Der sekundäre Endpunkt radiografisch progressionsfreies Überleben wurde beibehalten.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**<u>Einschlusskriterien:</u>** Patienten, die sämtliche der folgenden Kriterien erfüllten, wurden für die Teilnahme an der Studie berücksichtigt: • Histologisch oder zytologisch bestätigtes Adenokarzinom der Prostata ohne neuroendokrine Differenzierung oder Kleinzellenhistologie; • Laufende Androgenentzugstherapie mit einem Gonadotropin-freisetzenden Hormonanologon oder einer Orchiektomie, d. h. eine chirurgische oder medikamentöse Kastration; • Bei den Patienten, bei denen keine Orchiektomie durchgeführt wurde, musste es für die Studiendauer einen Plan für die effektive Aufrechterhaltung der Behandlung mit einem Gonadotropin-freisetzenden Hormonanologon geben. • Testosteron-Serumwert von < 1,7 nmol/L (50 ng/dL) bei der Screening-Visite; • Patienten, die eine Bisphosphonat Therapie erhalten haben, mussten für mindestens vier Wochen auf einer stabilen Dosis eingestellt sein. • Vorliegen einer progredienten Erkrankung, gemessen am PSA-Wert oder basierend auf einer bildgebenden Diagnostik, nach Durchführung einer Docetaxel-haltigen Chemotherapie im Zusammenhang mit einer medikamentösen oder chirurgischen Kastration. Eine Krankheitsprogression lag vor, wenn eines oder mehr der folgenden drei Kriterien gegeben war: - Ein Anstieg des PSA-Werts, definiert als eine mindestens dreimalige Zunahme des PSA-Werts im Abstand von mindestens einer Woche zwischen den einzelnen Bestimmungen. Dabei musste der PSA-Wert bei der Screening-Visite ≥ 2 µg/L (2 ng/mL); - Progression im Weichteilgewebe gemäß der Definition des Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST, v1.1); - Progression im Knochengewebe, definiert als Nachweis von mindestens zwei neuen Läsionen in der Knochenszintigrafie; • Nicht mehr als zwei zuvor durchgeführte Chemotherapien, wobei mindestens eine dieser Therapien Docetaxel enthalten haben muss. • Eastern Cooperative Oncology Group (ECOG)-*Performance*-Status von 0 bis 2; • Geschätzte Lebenserwartung von ≥ 6 Monaten; • Fähigkeit die Studienmedikation zu schlucken und die

Item[a]	Charakteristikum	Studieninformation
		Studienanforderungen zu erfüllen. • Willig und fähig, eine schriftliche Einverständniserklärung abzugeben. **<u>Ausschlusskriterien:</u>** Patienten, die eines der folgenden Kriterien erfüllten, wurden von der Studienteilnahme ausgeschlossen: • Eine schwere, gleichzeitig vorliegende Erkrankung, Infektion oder Komorbidität, die im Ermessen des Prüfarztes einen Grund für den Ausschluss des Patienten von der Studie darstellt. • Nachweis von Hirnmetastasen oder einer aktiven Epiduralerkrankung, wobei Patienten mit einer behandelten Epiduralerkrankung zur Studie zugelassen wurden. • Absolute Neutrophilenzahl von < 1.500/µL, Thrombozytenzahl von < 100.000/µL und ein Hämoglobingehalt von < 5,6 mmol/L (9 g/dL) zum Zeitpunkt der Screening-Visite. Dabei durften die Patienten keine hämatologischen Wachstumsfaktoren oder Bluttrans-fusionen innerhalb der letzten sieben Tage vor der hämatologischen Laborwertbestimmung zum Zeitpunkt der Screening-Visite erhalten haben. • Bilirubin Total, Alanin Aminotransferase (ALT), oder Aspartat Aminotransferase (AST) > 2mal höher als der Normalwert zum Zeitpunkt der Screening Visite • Serumkreatinin von > 177 µmol/L (2 mg/dL) bei der Screening-Visite; • Serumalbumin < 30 g/L (3,0 g/dL) zum Zeitpunkt der Screening-Visite; • Andere maligne Erkrankungen innerhalb der letzten fünf Jahre in der Anamnese mit Ausnahme von einem nicht-melanomatösen Hautkarzinom; • Behandlung mit einem Androgen-Rezeptor-Antagonisten (Bicalutamid, Flutamid, Nilutamid), 5α-Reduktase-Inhibitoren (Finasterid, Dutasterid), Östrogenen oder eine Chemotherapie innerhalb der letzten vier Wochen vor Studieneinschluss (Visite am Studientag 1) oder es geplant ist, während der Studienteilnahme eine Therapie mit einem dieser zuvor genannten Substanzen einzuleiten. • Behandlung mit einer therapeutischen Immunotherapie (wie z. B. PROVENGE®) zur Therapie des Prostatakarzinoms oder sofern es geplant ist, während der Studienteilnahme eine Therapie mit einem dieser zuvor genannten Substanzen einzuleiten. • Gebrauch von pflanzlichen Produkten (wie z. B. Sägepalme), durch welche der PSA-Spiegel gesenkt werden könnte, oder systemischen Kortikosteroiden in einer höheren Dosis als das Äquivalent von 10 mg Prednison/Prednisolon pro Tag innerhalb der letzten vier Wochen vor Studieneinschluss (Visite am Studientag 1) oder es geplant ist, während der Studienteilnahme eine Therapie mit einer dieser zuvor genannten Substanzen

Item[a]	Charakteristikum	Studieninformation
		einzuleiten. • Progression des Prostatakarzinoms unter einer Therapie mit Ketoconazol in der Anamnese, oder sofern es geplant ist, während der Studienteilnahme eine Therapie mit Ketoconazol einzuleiten. • Durchführung einer Radiotherapie innerhalb der letzten drei Wochen oder einer einzelnen Fraktion Radiotherapie innerhalb der letzten zwei Wochen oder die Durchführung einer Radionuklidtherapie innerhalb der letzten acht Wochen vor Studieneinschluss (Visite am Studientag 1); • Geplante palliative Maßnahmen zur Linderung von Knochenschmerzen wie z. B. Radiotherapie oder chirurgische Eingriffe; • Nachweis einer strukturell instabilen Knochenläsion, die auf eine drohende Fraktur hindeutet. • Auftreten eines epileptischen Anfalls, einschließlich eines Fieberkrampfs, Bewusstseinsverlusts oder einer transitorisch ischämischen Attacke innerhalb der letzten zwölf Monate vor Studieneinschluss (Visite am Studientag 1) oder ein Zustand, welcher für einen epileptischen Anfall prädisponiert (wie z. B. ein vorangegangener Schlaganfall, eine zerebrale arteriovenöse Malformation, ein Schädeltrauma mit Bewusstseinsverlust, welches eine Hospitalisierung erforderte.); • Vorliegen einer klinisch bedeutsamen kardiovaskulären Erkrankung, wie eine der Folgenden: - Myokardinfarkt innerhalb der letzten sechs Monate; - unkontrollierte Angina pectoris innerhalb der letzten 3 Monate; - Herzinsuffizienz New York Heart Association (NYHA)-Grad 3 oder 4 oder Patienten mit einer Herzinsuffizienz des NYHA-Grads 3 oder 4 in der Anamnese, außer wenn zum Zeitpunkt der Screening-Visite im Echokardiogramm oder in der Radionuklid-Angiografie (*multi-gated acquisition scan*) innerhalb der letzten drei Monate vor Studieneintritt eine linksventrikuläre Ejektionsfraktion von ≥ 45 % nachgewiesen werden konnte. - Diagnostiziertes oder vermutlich angeborenes Long-QT-Syndrom; - Nachweis von klinisch bedeutsamen ventrikulären Arrythmien in der Anamnese (wie z. B. eine ventrikuläre Tachkardie, Kammerflimmern, Torsades-de-pointes-Tachykardie); - Nachweis eines verlängerten korrigierten QT-Intervalls nach Anwendung der Korrekturformel nach Fridericia (QTcF) von > 470 ms im Elektrokardiogramm der Screening-Visite; - Nachweis eines AV-Blocks zweiten oder dritten Grades vom Mobitz-Typ-2 ohne einen permanenten Schrittmacher; - Hypotonie (systolischer Blutdruck < 86 mmHg oder

Item[a]	Charakteristikum	Studieninformation
		eine Bradykardie mit einer Herzfrequenz von < 50 Herzschlägen pro Minute im Elektrokardiogramm) zum Zeitpunkt der Screening-Visite oder am Studientag 1; - Unkontrollierbare Hypertonie mit einem systolischen Blutdruck > 170 mmHg oder einem diastolischen Blutdruck > 105 mmHg zum Zeitpunkt der Screening-Visite oder am Studientag 1; • Bereits vorgenommene oder geplante Einnahme der folgenden Medikamente 30 Tage vor Studieneintritt (Studientag 1) und bis zum Studienende, von denen bekannt ist, die Krampfschwelle zu senken oder das QT-Intervall zu verlängern: - Aminophyllin/Theophyllin; - Atypische Antipsychotika (wie z. B. Clozapin, Olanzapin, Risperidon, Ziprasidon); - Bupropion; - Klasse Ia- und III-Antiarrhythmika (wie z. B. Amiodaron, Bretylium, Disopyramid, Ibutilid, Procainamid, Quinidin, Sotalol); - Dolasetron; - Droperidol; - Gatafloxacin/Moxifloxacin; - Insulin; - Lithium; - Makrolidantibiotika (wie z. B. Erythromycin, Clarithromycin); - Pethidin; - Antipsychotika der Phenothiazin-Gruppe (wie z. B. Chlorpromazin, Mesoridazin, Thioridazin); - Pimozid; - Trizyklische und tetrazyklische Antidepressiva (wie z. B. Amitryptilin, Desipramin, Doxepin, Imipramin, Maprotilin, Mirtazapin); - Venlafaxin; • Eine frühere Teilnahme an einer klinischen Studie mit einer Prüfsubstanz, die die Androgensynthese (wie z. B. Abirateronacetat, TAK-700, TAK-683, TAK-448) oder den Androgenrezeptor hemmt (wie z. B: BMS-641988), außer wenn der Patient innerhalb der ersten zwölf Therapiewochen mit der Prüfsubstanz einen Anstieg des PSA-Werts gehabt hat. • Teilnahme an einer früheren klinischen Studie mit Enzalutamid; • Einnahme eines anderen Studienmedikaments innerhalb der letzten vier Wochen vor Studieneintritt (Studientag 1) oder es geplant ist, eine Behandlung mit einer anderen Prüfsubstanz während der Studienteilnahme einzuleiten. • Vorliegen einer gastrointestinalen Funktionsstörung, welche die Absorption beeinflussen könnte (wie z. B.

Item[a]	Charakteristikum	Studieninformation
		Gastrektomie, aktives Magengeschwür); • Ein größerer chirurgischer Eingriff innerhalb der letzten vier Wochen vor Studieneintritt (Studientag 1).
4b	Studienorganisation und Ort der Studiendurchführung	An der Studie nahmen 156 Studienzentren aus 15 Ländern in Nordamerika, Südamerika, Europa, Australien und Südafrika teil.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die Studienpatienten wurden auf zwei Behandlungsgruppen randomisiert: 1. Orale Gabe von 160 mg Enzalutamid pro Tag, aufgeteilt auf vier Einzeldosen à 40 mg in Kapselform 2. Orale Gabe von Placebokapseln in derselben Art und Häufigkeit wie die Gabe von Enzalutamid.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	**<u>Primäres Zielkriterium:</u>** 1. Gesamtüberleben **<u>Sekundäre Zielkriterien:</u>** 2. PSA-Progression 3. PSA-Ansprechen 4. Radiografisch progressionsfreies Überleben 5. Zeit bis zum Auftreten des ersten skelettalen Ereignisses 6. Lebensqualität 7. Schmerzlinderung 8. Radiografisches Ansprechen im Weichteilgewebe 9. Konversionsrate der im Blut zirkulierenden Tumorzellen 10. Sicherheit **<u>Definition der Zielkriterien:</u>** Zu 1.: Als primärer Wirksamkeitsendpunkt wurde das Gesamtüberleben von dem Zeitpunkt der Randomisierung bis zum Todeseintritt jeglicher Ursache gemessen. Die Gesamtüberlebensdauer wurde für die Patienten, die eines der folgenden Kriterien erfüllten, rechtszensiert: - Patienten, die seit der Randomisierung nicht mehr nachverfolgt werden konnten (*lost to follow-up*) - Patienten, von denen nicht bekannt ist, ob sie zu dem Zeitpunkt, an welchem die Studiendatenbank geschlossen wurde, bereits verstorben waren. Dies umfasste auch Patienten, von denen bekannt war, dass sie nach Schließung der Studiendatenbank verstorben sind. Zu 2.: Für die Patienten mit einem Abfall des PSA-Werts zur Studienwoche 13, war der Zeitpunkt der PSA-Progression definiert als ein Anstieg um ≥ 25 % und einem absoluten Anstieg des PSA-Werts von ≥ 2 ng/mL über dem Nadir der gemessenen PSA-Werte. Dieser PSA-Wert musste drei oder einige Wochen später durch einen zweiten nachfolgend gemessenen Wert bestätigt werden, falls die PSA-Progression

Item[a]	Charakteristikum	Studieninformation
		nicht erst bei der letzten PSA-Messung aufgetreten ist. Für die Patienten ohne Abfall des PSA-Werts zur Studienwoche 13 war der Zeitpunkt der PSA-Progression definiert als ein Anstieg um ≥ 25 % und einen absoluten Anstieg des PSA-Werts von ≥ 2 ng/mL über dem an der Baseline gemessenen PSA-Wert. Dieser PSA-Wert musste drei oder einige Wochen später durch einen zweiten nachfolgend gemessenen Wert bestätigt werden, falls die PSA-Progression nicht erst bei der letzten PSA-Messung aufgetreten ist. Das Zeitintervall bis zur PSA-Progression war definiert als die Zeit zwischen der Randomisierung und dem Zeitpunkt der eingetretenen PSA-Progression. Zu 3.: Ein bestätigtes und unbestätigtes PSA-Ansprechen war definiert als eine Verminderung des PSA-Werts um ≥ 50 % und ≥ 90 % unter dem an der Baseline gemessenen Wert. Dieser PSA-Wert bzw. das PSA-Ansprechen musste mindestens drei Wochen später durch eine zweite nachfolgende Messung bestätigt werden. Zu 4.: Das radiografisch progressionsfreie Überleben wurde von dem Zeitpunkt der Randomisierung bis zum frühesten objektiven Nachweis einer radiografischen Progression oder Todeseintritt jeglicher Ursache gemessen. Die Patienten wurden regelmäßig bei den geplanten Visiten nach einer objektiven Krankheitsprogression beurteilt. Dabei erfolgte die Beurteilung einer Krankheitsprogression nach den Konsensus Leitlinien der Prostate Cancer Clinical Trials Working Group 2 (PCWG2). [Scher 2008] Die radiografische Krankheitsprogression im Weichteilgewebe wurde definiert nach den Kriterien des RECIST v1.1 oder dem Auftreten von mindestens zwei Knochenläsionen in der Knochenszintigraphie gemäß den Leitlinien der PCWG2 [Scher 2004; Scher 2008] Eine Progression, die bei der ersten Neubewertung in der Studienwoche 13 nachgewiesen werden konnte, musste nach mindestens sechs Wochen radiografisch bestätigt werden. Jedoch wurde als Zeit bis zur Progression immer das Zeitintervall von der Randomisierung bis zur erstmals nachgewiesenen Progression gewertet, unabhängig davon, ob eine radiografische Bestätigung durchgeführt wurde. Die Patienten, die diesen Endpunkt nicht erreichten, wurden bei ihrer letzten Beurteilung unter folgenden Bedingungen rechtszensiert: - Keine Beurteilungen an der Screening/Baseline-Visite durchgeführt wurden. - Patienten, die seit der Randomisierung nicht mehr nachverfolgt werden konnten (*lost to follow-up*). - Eine radiografische Progression oder Tod nach mindestens zwei hintereinander verpassten radiografischen Untersuchungen eingetreten ist. - Patienten, von denen nicht bekannt ist, ob sie zu dem Zeitpunkt, an welchem die Studiendatenbank geschlossen wurde, eine Progression gehabt haben oder bereits verstorben waren. Dies umfasste auch

Item[a]	Charakteristikum	Studieninformation
		Patienten, von denen bekannt war, dass sie eine Progression gehabt haben oder verstorben sind, nachdem die Studiendatenbank geschlossen wurde. Zu 5.: Die Zeit bis zum Auftreten des ersten skelettalen Ereignisses wurde vom Zeitpunkt der Randomisierung bis zum ersten Auftreten eines skelettalen Ereignisses gemessen. Die Patienten wurden regelmäßig bei den geplanten Visiten auf skelettale Ereignisse hin untersucht. Ein skelettales Ereignis war definiert als eine Radiotherapie oder chirurgische Operation am Knochen, eine pathologische Knochenfraktur, eine Rückenmarkkompression oder eine Änderung der antineoplastischen Therapie zur Behandlung von Knochenschmerzen. Die Patienten, die diesen Endpunkt nicht erreichten, wurden bei ihrer letzten Untersuchung rechtszensiert oder für die Patienten, die nach der Baseline-Visite nicht auf ein skelettales Ereignis hin untersucht wurden, wurde das Datum der Randomisierung herangezogen. Zu 6.: Ein Therapieansprechen hinsichtlich der Lebensqualität lag vor, wenn die Patienten im Gesamtwert des Functional Assessment of Cancer Therapy-Prostate (FACT-P) an zwei aufeinanderfolgenden Messungen im Abstand von mindestens drei Wochen im Vergleich zur Baseline-Visite eine Verbesserung um zehn Punkte gehabt haben. Zudem wurde bei jeder Studienvisite die Lebensqualität von jeder Behandlungsgruppe auch anhand des European Quality of Life Five-Domain Scale (EQ-5D) deskriptiv zusammengefasst. Zu 7.: Der Anteil der Patienten mit einer Schmerzlinderung wurde für die Patienten mit einer konstanten und hohen Schmerzbelastung ausgewertet. Die Schmerzbelastung wurde anhand des *Brief Pain Inventory-short form* gemessen und im Patiententagebuch festgehalten. Diese Skala maß den Schmerz auf einer Skala von 0 bis 10, wobei 0 Punkte bei Vorliegen von Schmerzfreiheit und 10 Punkte bei Vorliegen von stärksten Schmerzen vergeben wurden. Die Schmerzlinderung wurde in der Studienwoche 13 für den Anteil der Männer bestimmt, bei denen bei der Baseline-Visite eine Knochenmetastasierung nachgewiesen werden konnte und deren Schmerz bei der Baseline-Visite auf die Knochenmetastasierung zurückzuführen war. Die Patienten wurden gebeten, ihr Patiententagebuch jeweils sieben Tage vor dem Studientag 1 und der Studienwoche 13 auszufüllen. Für jeden dieser sieben Tage gab der Patient gemäß dem *Brief Pain Inventory* questionnaire seinen schlimmsten erfahrenen Schmerz der vergangenen 24 Stunden und den Gebrauch von Analgetika an. Um für diese Analyse ausgewertet werden zu können, musste der Patient für mindestens vier der sieben Tage in der *Run-in-Phase* der Baseline Angaben gemacht haben. Zudem musste der Patient bei der Baseline-Visite einen stabilen Schmerzustand und Analgetikaverbrauch aufweisen. Dieser lag vor, wenn die Variation bei den täglichen erhobenen Schmerzscores nicht größer als zwei Punkte war und sich der Gebrauch von Analgetika nicht um mehr als 30 % unterschied. Um für die

Item[a]	Charakteristikum	Studieninformation
		Analyse der Schmerzlinderung ausgewertet werden zu können, musste der Patient außerdem während der *Run-in-Phase* der Baseline einen durchschnittlichen Schmerzscore von ≥ 4 Punkten aufweisen. Eine Schmerzlinderung war definiert als eine Verminderung des durchschnittlichen Schmerzscores um ≥ 30 % in der Studienwoche 13 im Vergleich zur Baseline ohne einen Anstieg des Analgetikaverbrauchs um ≥ 30 %. Zu 8.: Das radiografische Ansprechen im Weichteilgewebe, welches während der Studie anhand der RECIST v1.1-Kriterien beurteilt wurde, wurde unter Verwendung der durch den Prüfarzt durchgeführten Untersuchungen und der abgeleiteten Auswertungen nach Behandlungsgruppe hinsichtlich eines Therapieansprechens zusammengefasst. In diese Analyse wurden nur die Patienten aufgenommen, bei denen bei der Screening-Visite eine messbare Weichteilerkrankung vorlag. Eine messbare Weichteilerkrankung war vorhanden, wenn bei der Screening-Visite mindestens eine Zielläsion anhand der RECIST v1.1-Kriterien nachgewiesen werden konnte. Der numerische und prozentuale Anteil der Patienten aus jeder der folgenden Kategorien wurde tabellarisch dargestellt: Vollständiges Therapieansprechen, partielles Ansprechen, stabile Erkrankung, progrediente Erkrankung, nicht auswertbar und objektives Therapieansprechen. Zu 9.: Zur Untersuchung der Konversionsrate der im Blut zirkulierenden Tumorzellen wurden bei einem Teil der Studienpatienten Blutproben gesammelt. Zu 10.: Die Sicherheit wurde anhand der behandlungsbedingten unerwünschten Ereignisse, der unerwünschten Ereignisse, welche zum Tod oder zum Abbruch der Studienmedikationseinnahme geführt haben, der laborchemischen Testergebnisse, der Änderungen in den Vitalzeichen und der Elektrokardiogramme beurteilt. Alle unerwünschten Ereignisse, welche während der Behandlungsphase aufgetreten sind, wurden in die Tabellen der unerwünschten Ereignisse integriert. Dabei wurde die Behandlungsphase definiert als das Zeitintervall von der ersten bis 30 Tage nach der letzten Einnahme der Studienmedikation oder bis zur Einleitung einer anderen antineoplastischen Therapie, je nachdem welcher Zustand zuerst eintrat. Die Analyse der Todesursachen in der *safety population* schloss alle Todesereignisse ein, welche bei oder vor der Schließung der Studiendatenbank am 25.09.2011 aufgetreten sind. Dabei wurden sowohl die Todesereignisse während der Behandlungsphase als auch die im Anschluss an die Therapiephase aufgetretenen Ereignisse eingeschlossen. **Erhebungszeitpunkte:** Siehe Tabelle 9.5.1-1, Schedule of Assessment
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Änderungen der Zielkriterien im Vergleich zum Originalprotokoll vom 21.05.2009:

Item[a]	Charakteristikum	Studieninformation
		Mit der Protokolländerung vom 19.04.2011 wurden die sekundären Endpunkte vereinfacht: - Der kombinierte Endpunkt progressionsfreies Überleben, welcher sich auf den Zeitraum vom Tag der Randomisierung bis zum ersten Auftreten einer radiografischen Progression, eines skelettalen Ereignisses oder Tod bezieht, wurde in den Endpunkt radiografisch progressionsfreies Überleben geändert. Dabei bezieht sich der Endpunkt radiografisch progressionsfreies Überleben auf den Zeitraum vom Tag der Randomisierung bis zum ersten Auftreten einer radiografischen Progression oder Tod. Auf den sekundären Endpunkt hinsichtlich des Auftretens einer radiografischen Progression wurde verzichtet. Der sekundäre Endpunkt radiografisch progressionsfreies Überleben wurde beibehalten.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Bei der Berechnung der Fallzahl wurde von einer geschätzten medianen Überlebenszeit von 12 Monaten in der Placebogruppe und von geschätzten 15,7 Monaten in der Enzalutamid-Gruppe ausgegangen. Auf Grundlage dieser Annahme, wurde eine Mindestanzahl von insgesamt 650 Todesfällen berechnet, um die angestrebte Hazard Ratio von 0,76 bei einer Teststärke (*power*) von ca. 90 % und einem zweiseitigen Signifikanzniveau von 0,05 nachweisen zu können. Um die berechnete Anzahl von 650 Todesereignissen zu erreichen, mussten insgesamt ca. 1.080 Patienten (720 Patienten in der Enzalutamid-Gruppe, 360 Patienten in der Placebogruppe) randomisiert werden. Mit der dritten Protokolländerung vom 19.04.2011 wurde die Berechnung der Teststärke (*power*) geändert. Der Grund für diese Änderung lag darin, dass sich die angestrebte Hazard Ratio von 0,76 innerhalb des beobachteten 95 %-Konfidenzintervalls der Hazard Ratio für das Gesamtüberleben aus der COU-AA-301-Studie befand. Bei der COU-AA-301-Studie handelte es sich um eine Phase-III-Studie, in der die Kombination aus Abirateronacetat und Prednison mit der Kombination aus Placebo und Prednison bei Patienten mit kastrationsresistentem Prostatakarzinom und einer zuvor durchgeführten Docetaxel-basierten Chemotherapie miteinander verglichen wurden. Die angenommene mediane Überlebenszeit von 12 Monaten in der Placebogruppe aus dem Originalprotokoll wurde nicht verändert, da diese auf Präzedenzfällen mit kastrationsresistentem Prostatakarzinom und einer zuvor durchgeführten Docetaxel-basierten Chemotherapie basiert. Eine statistische Analyse hinsichtlich des Gesamtüberlebens war bei Vorliegen von insgesamt 650 Todesereignissen vorgesehen. Unter der Annahme, dass bis zu 8 % der Patienten im Laufe der Studie nicht nachverfolgt werden können (*lost to follow-up*), wurde eine Fallzahl von insgesamt 1.170 Studienpatienten bestimmt (780 Patienten in der Enzalutamid-

Item[a]	Charakteristikum	Studieninformation
		Gruppe, 390 Patienten in der Placebogruppe).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Eine Interimanalyse für das Gesamtüberleben wurde durchgeführt bei einer Anzahl von 520 Sterbefällen. Zur Zeit des „Data-cutoff“, gab es 308 Todesfälle in der Enzalutamid Gruppe und 212 Todesfälle von Patienten im Placeboarm. Die Analyse der stratifizierten Daten des Gesamtüberlebens, zeigte eine statistisch signifikante Verlängerung der Überlebensdauer der Enzalutamid behandelten Patienten im Vergleich zur Placebogruppe (p<0,0001). Die mittlere Überlebensdauer betrug 18,4 Monate für Enzalutamid im Vergleich zu 13,6 Monaten im Placeboarm.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Methode zur Generererierung der zufälligen Zuteilung basierte auf einem computergenerierten Randomisierungsschema in einem Verhältnis 2:1 Enzalutamid vs. Placebo.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Es wurde eine zentral regulierte, randomisierte, permutierte Block Methode verwendet, stratifiziert nach den folgenden Faktoren: - Studienbeginn ECOG-*Performance*-Status 0 oder 1 gegenüber 2 - Studienbeginn BPI Frage 3 Auswertung, stratifiziert durch Wert < 4 gegenüber ≥ 4 auf einer Skala von 0 – 10.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Nachdem die schriftliche Einverständniserklärung eingeholt wurde, rief das jeweilige Prüfzentrum das interaktive Sprachdialogsystem (*Interactive Voice Response System* = IVRS) oder das interaktive Webdialogsystem (*Interactive Web Response System* = IWRS) an, um eine dem jeweiligen Teilnehmer zugewiesene Randomisierungsnummer und eine Nummer für das Behandlungs-Kit zu erhalten. Diese basierten auf einem computergenerierten Randomisierungsschema. Das korrespondierende Behandlungs-Kit wurde zusammen mit den Anweisungen zur Verabreichung der Studienmedikation ausgeteilt. Die Liste mit den Randomisierungscodes und den zugewiesenen Behandlungen wurden im Studienprotokoll festgehalten
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	United BioSource Corporation (San Franzisco), kürzlich umbenannt in Bracket, führte die Randomisierung durch mittels IVRS/IWRS. Catalent (Somerset) verschiffte das Medikament zu den Studienzentren mittels IVRS Management.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht	Alle Patienten, Prüfärzte, Personal, Mitarbeiter des Sponsors und die Personen, die bei der Durchführung der Studie beteiligt waren, waren verblindet. Die Verteilung des Studienmedikaments erfolgte über IVRS/IWRS.

Item[a]	Charakteristikum	Studieninformation
	verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Enzalutamid- und Placebokapseln waren in Aussehen und Verpackung identisch. Das Produkt wurde als weiße Gelatine Kapsel gegeben. Jeweils 180 Kapseln wurden in Polyethylen Flaschen mit Kindersicherung verpackt, die tägliche Dosis betrug 160 mg, d. h. 4 Kapseln a 40 mg.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Analyse des Zielkriteriums Gesamtüberleben war in zwei Meilensteinen geplant. Eine Interimanalyse sollte nach 520 berichteten Todesfällen stattfinden. Sollte die Interimanalyse nicht zur Beendigung der Studie, aufgrund Erreichung des Endpunktes führen, war eine weitere Anaylse nach 650 Todesfällen geplant. Da der Endpunkt bereits in der Interimanalyse erreicht wurde, wird sich im Folgenden auf die Interimanalyse bezogen. Die Enzalutamid- und Placebo Gruppe wurden mit einem lok rank Test, stratifiziert nach ECOG-*Performance*-Status und dem Schmerzfragebogen (BIP), verglichen. Ein zweiseitiger Test mit einem Signifikanzniveau von 0,05 wurde festgelegt. Zur statistischen Beschreibung wurde der Kaplan-Meier-Plot auf Grundlage der mittleren Überlebenszeit und derer 95 % Konfidenzintervalle gewählt. Zusätzlich wurde der Nutzen von Enzalutamid gegenüber Placebo anhand des einzelnen Hazard Ratios einschließlich 95 % Konfidenzintervall basierend auf einem stratifizertem COX Regression Modell ausgewertet. Die drei sekundären Endpunkte „PSA Progression“, „Radiographisches- Progressionsfreies Überleben“ und „Zeit bis zum Auftreten des ersten skelettbezogenen Ereignisses“, wurden anhand eines log- rank Testes stratifiziert nach ECOG-*Performance*-Status und BPI ausgewertet. Der Vergleich der Enzalutamid- und Placebogruppe unterlag einem zweiseitigen Test mit einem Signifikanzniveau von 0,05. Zusätzlich wurde der Vorteil Enzalutamids gegenüber Placebo mittels einzelner Hazard Ratios und zugehörigem 95 % Konfidenzintervall basierend auf dem COX Regression Modell ausgewertet. Dargestellt wurden die jeweiligen Ergebnisse in Kaplan-Meier-Plots unter Angabe der 95 % Konfidenzintervalle. Das 50. Perzentil wurde festgelegt, zur Abschätzung der mittleren Zeit zur PSA Progression, der mittleren radiographisch progressionsfreien Überlebenszeit und dem mittleren Zeitpunkt bis zum Auftreten des ersten skelettalen Ereignisses.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Vordefinierte **<u>Subgruppenanalysen</u>** über das Gesamtüberleben wurden durchgeführt, um zu bestimmen, ob der Behandlungseffekt unter den Subgruppen konform war. Dafür wurden die folgenden Subgruppen-Analysen prädefiniert: Für den Endpunkt Gesamtüberleben wurden folgende Subgrupppen definiert: • ECOG-*Performance*-Status zu Studienbeginn (0 oder 1 vs. 2); • Mittlerer Schmerzscore basierend auf Frage 3 des *Brief Pain Inventory – short form* zu Studienbeginn (< 4 vs. ≥ 4); • Alter (< 65 vs. ≥ 65); • Geographische Region (Nord America vs. Europa vs. den

Item[a]	Charakteristikum	Studieninformation
		Rest der Welt); • Gleason Score bei Diagnose (≤ 7 vs. ≥ 8); • Anzahl der vorherigen Chemotherapie Regimes (1 vs. ≥ 2); • Typ der Progression zu Studienbeginn (nur PSA Progression vs. radiographische Progression mit oder ohne PSA Progression); • Viszerale Erkrankung zu Studienbeginn (ja vs. nein); • PSA Wert zu Studienbeginn (≤ Median vs. > Median); • Laktat Dehydrogenase Wert zu Studienbeginn (≤ Median vs. > Median); und • Hämoglobin Wert zu Studienbeginn (≤ Median vs. > Median). Diese Faktoren wurden gewählt, da sie entweder prognostische Faktoren für Prostatakarzinom, demographische Merkmale von Interesse oder Faktoren, die regionale Unterschiede der Behandlungsmuster repräsentieren, darstellen. Unerwünschte Ereignisse wurden in folgenden Subgruppen ausgewertet: • Alter (< 65, ≥ 65, ≥ 75) • Anamnese einer Hypertonie (ja vs. nein), und • Anamnese kardiovaskulärer Erkrankungen (ja vs. nein). Für die Endpunkte Gesamtüberleben, radiographisches progressionsfreies Überleben, Zeit bis zur ersten skelettbezogenen Komplikation, Schmerz, Lebensqualität sowie die sicherheitsbezogenen Ereignisse wurden folgende Subgruppen analysiert (soweit sie nicht bereits durch die prädefinierten Analysen für das Gesamtüberleben und die sicherheitsbezogenen Ereignisse abgedeckt waren): • Gabe von Bisphosphonaten während der Studie (ja vs. nein) • Gabe von Steroiden während der Studie (ja vs. nein) • ECOG-*Performance*-Status bei Studienbeginn (0 oder 1 vs. 2) • Durchschnittliche Schmerzpunktzahl basierend auf Frage 3 des BPI - *short form* bei Studienbeginn (< 4 vs.≥ 4) • Alter (< 65 vs. ≥ 65) • Anzahl vorheriger Chemotherapien (1 vs. ≥ 2) • Viszerale Erkrankung (Lunge und/oder Leber) zum Zeitpunkt des Screenings (ja vs. nein) Der Vorteil von Enzalutamid verglichen zu Placebo wird ausgewertet anhand von Hazard Ratio und 95 % Konfidenzintervall basierend auf einem unstratifiziertem COX Regressions Modell für jede Subgruppe. Hazard Ratios und 95 % Konfidenzintervalle werden in Forrest Plots dargestellt. Folgende Angaben zu geplanten **Sensitivitätsanalysen** werden im Studienbericht gemacht: **Gesamtüberleben:** Es wurde das Hazard Ratio (Enzalutamid/Placebo) sowie das 95 % Konfidenzintervall auf der Basis eines stratifizierten Cox Regressions Modells berechnet. Im Rahmen der Sensitivitätsanalyse wurde ebenfalls der unstratifizierte log-

Item[a]	Charakteristikum	Studieninformation
		rank Test sowie ein unstratifiziertes Cox Regressions Modell angewendet (Hazard Ratio für Enzalutamid/Placebo und 95 % KI). [Studienbericht Kapitel 9.7.4] **Radiographisches Progressionsfreies Überleben:** Folgende Sensitivitätsanalysen wurden für das Investigator-beurteilte sowie die daraus abgeleitete Progression definiert: 1. Im Fall einer Krankheitsprogression, die zwischen zwei geplanten radiographischen Beurteilungen dokumentiert wurde, wurde das erste Datum einer radiographischen Beurteilung, an dem Evidenz für eine Krankheitsprogression vorlag, als Datum der Krankheitsprogression verwendet. 2. Um den Einfluss der Compliance für die radiographische Untersuchung zu beurteilen, wurden Patienten mit Krankheitsprogression oder im Todesfall nach zwei oder mehr aufeinanderfolgenden radiographischen Beurteilungen, so beurteilt, dass sie ein „Ereignis“ gehabt hatten. [Studienbericht Kapitel 9.7.5.2] **Zeit bis zur ersten skelettbezogenen Komplikation:** Im Rahmen einer Sensitivitätsanalse wurde – im Gegensatz zur sonst verwendeten Definition für die erste skelettbezogene Komplikation - eine Änderung der anti-neoplastischen Therapie, die aufgrund der Notwendigkeit der Behandlung von Knochenschmerzen aufgrund von Knochenmetastasen erfolgte, nicht als Teil der Definition für das „erste skelettbezogene Ereignis“ gewertet. [Studienbericht Kapitel 9.7.5.3]
-	**Resultate**	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flussdiagramm Patientenfluss in der Studie AFFIRM
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Zwischen 22.September 2009 und 15.November 2010 wurden 1199 Patienten in diese Studie randomisiert: 800 in die Enzalutamid-Gruppe und 399 in die Placebogruppe. Alle randomisierten Patienten haben die geplante Behandlung erhalten und gingen auch später in die Analyse des Zielkriteriums mit ein. a) Randomisiert: Enzalutamid: n = 800 Placebo: n = 399 b) Erhielten tatsächlich die geplante Behandlung/Intervention: Enzalutamid: n = 800 Placebo: n = 399 c) In der Analyse des primären Zielkriteriums berücksichtigt: Enzalutamid: n = 800 Placebo: n = 399
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen	

Item[a]	Charakteristikum	Studieninformation
	Patienten nach Randomisierung mit Angabe von Gründen	Siehe Abbildung 4
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Zwischen 22 September 2009 und 15 November 2010 wurden 1199 Patienten in diese Studie randomisiert: 800 in die Enzalutamid-Gruppe und 399 in die Placebogruppe. Duration of follow up: alle 12 Wochen nach *data cutoff date*. Das *data cutoff* Datum war September 2011.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde nach der Interimanalyse geplant beendet.
a: nach CONSORT 2010.		

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

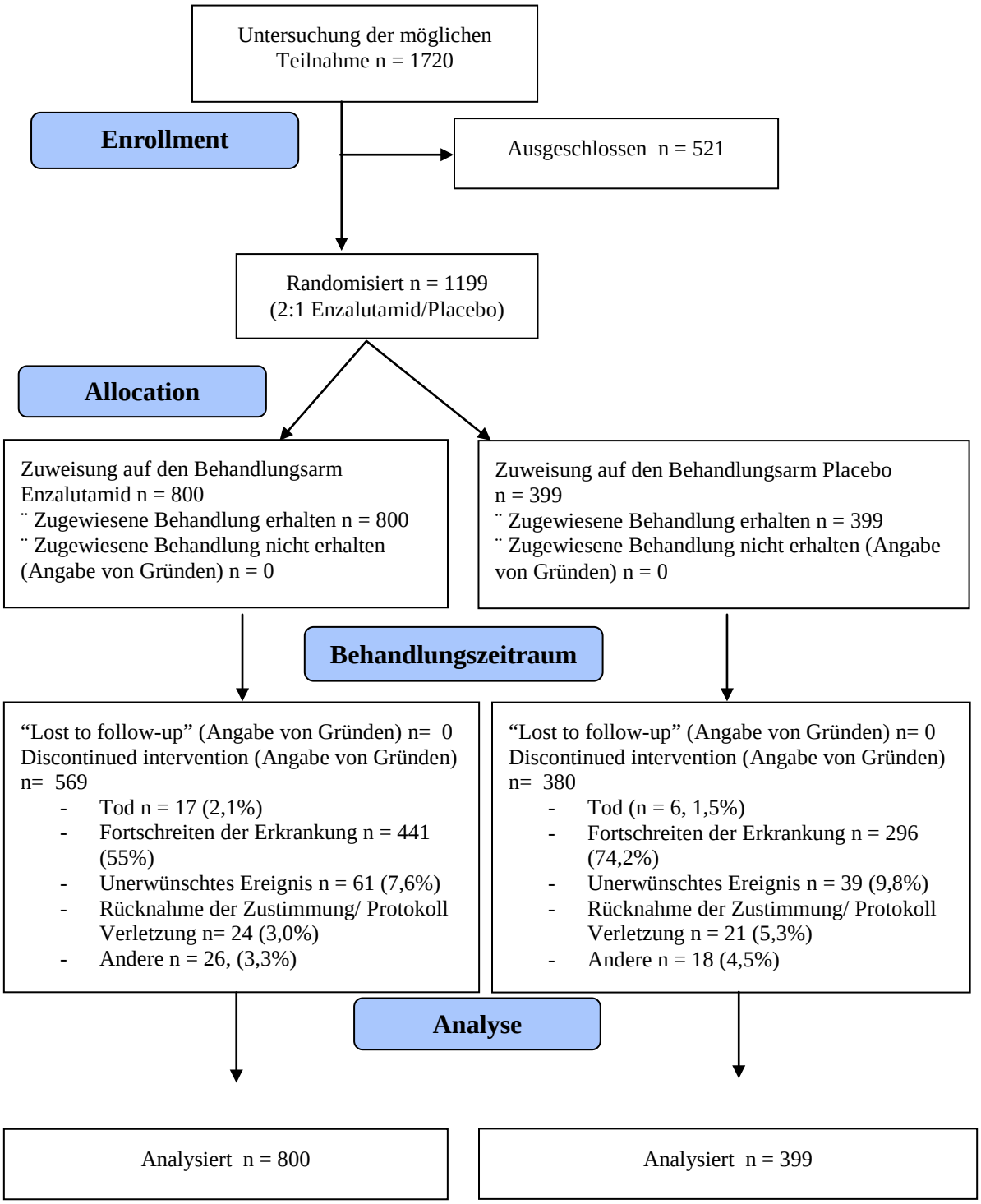


Abbildung 4: Flussdiagramm Patientenfluss in der Studie AFFIRM

Tabelle 4-126: Übersicht über den zeitlichen Ablauf der Erhebungen und Untersuchungen innerhalb des Studienprogramms

Tag der Studie	**Screening Besuch**	**1**	**8**	**29**	**57**	**85**	**113**	**141**	**169**	**Sicherheit F/U**	**Ungeplanter Besuch**	**Langzeit F/U**
Woche	-4 to -1 (28 Tage)	1	2	5	9	13	17	21	25 und alle nachfolgenden 12 Wochen	30 Tage nach der letzten Dosis	n/a	Alle 12 Wochen
Intervall (Tage)			±2	±3	±3	±7	±3	±3	±7	±7	n/a	±7
Einwilligung	X											
Anamnese	X											
Einschluss/Ausschlusskriterien	X	X										
Randomisierung (IVRS)		X										
Vitalparameter	X	X	X	X	X	X	X	X	X	X	X	
Ärztliche Untersuchung, Gewicht	x[f]	X	X	X	X	X	X	X	X	X	X	
12 Kanal EKG	X	X	X	X	X	X	X	X	X	X	X	
MUGA/Echokardiogramm	X											
Klinische Laboruntersuchungen	X	X	X	X	X	X	X	X	X	X	X	
PSA	X	X				X	X	X	X	X		
PK		X	X	X	X	X			X		X	
CTCs, molekulares Profiling, Marker zur Messung des Knochenaufbaus	X	X			X	X			X			
CT/MRT und Knochenscan	X					X			X			
Röntgenthoraxaufnahme oder CT Thorax	X											
Eastern Cooperative Oncology Group	X	X	X	X	X	X	X	X	X	X	X	
Vergabe Schmerztagebuch	x[p]				X							
Abgabe Schmerztagebuch und *Brief Pain Inventory - short form*		X				X						

Brief Fatigue Inventory und *Fatigue Severity Assessment*		X										
FACT-P		X				X	X	X	X			
EQ-5D		X				X			X			
Unerwünschte Ereignisse		X	X	X	X	X	X	X	X	X	X	
Begleitmedikationen	X	X	X	X	X	X	X	X	X	X	X	
Studienmedikament Vergabe		X		X		X	X		X			
Langzeituntersuchungen												X
Studienmedikamenttherapie		X	X	X	X	X	X	X	X			

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 4-127 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie AFFIRM

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: AFFIRM

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht der Studie A Multinational Phase 3, Randomized, Double-Blind, Placebo-Controlled Efficacy and Safety Study of Oral MDV3100 in Patients with Progressive Castration-Resistant Prostate Cancer Previously Treated with Docetaxel-Based Chemotherapy	CRPC2
Publikation zur Studie Scher, H.I.; Fizazi, K.; Saad, F.; Taplin, M.E.; Sternberg, C.N.; Miller, K. et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. N.Engl.J.Med. 27.9.2012. 367, (13), 1187-1197. [Scher 2012]	[Scher 2012]

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 handelt sich um eine multinationale, randomisierte, doppelblinde, Placebo kontrollierte, Phase 3 Studie

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

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

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

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

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

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja**: Eines der folgenden Merkmale trifft zu:

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

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

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

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

---

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja**: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine 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:

Die Studie wurde doppelt verblindet durchgeführt. Weder Patient noch der Prüfarzt waren über die zugewiesene Behandlungsgruppe informiert.

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 doppelt verblindet durchgeführt. Weder der Patient noch der Prüfarzt waren über die zugewiesene Behandlungsgruppe informiert

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: Gesamtüberleben

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet.

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: Radiographisches progressionsfreies Überleben

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: Zeit bis zur ersten skelettbezogenen Komplikation

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: Schmerzlinderung

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

☐ **ja:** Eines der folgenden Merkmale trifft zu:

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

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

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

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

Die Auswertungspopulation *für die Schmerzlinderung* umfasste 64 Patienten, und somit nur 5 % der ITT-Population (1.199 Patienten). Damit wurde das ITT-Prinzip nicht adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

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

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

☐ **nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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

☒ **ja**

☐ **nein**

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

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:

Da das ITT-Prinzip nicht adäquat umgesetzt wurde, wird das Verzerrungspotenzial der Ergebnisse des Endpunkts mit hoch bewertet.

Endpunkt: Schmerzprogressionsrate

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

☐ **ja:** Eines der folgenden Merkmale trifft zu:

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

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

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

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

Die Auswertungspopulation für die Schmerzprogressionsrate umfasste 884 Patienten. Somit befanden sich lediglich 74 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

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

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

☐ **nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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

☒ **ja**

☐ **nein**

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

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:

Da das ITT-Prinzip nicht adäquat umgesetzt wurde, wird das Verzerrungspotenzial der Ergebnisse des Endpunkts mit hoch bewertet.

Endpunkt: Zeit bis zur Schmerzprogression

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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 nach FACT-P

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

☐ **ja:** Eines der folgenden Merkmale trifft zu:

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

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

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

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

Die Auswertungspopulation für FACT-P umfasste 871 Patienten. Somit befanden sich, unabhängig von der Auswertung des Endpunkts als Responder-Analyse oder als Veränderung gegenüber Studienbeginn maximal 73 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

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

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

☐ **nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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

☒ **ja**

☐ **nein**

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

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:

Da das ITT-Prinzip nicht adäquat umgesetzt wurde, wird das Verzerrungspotenzial der Ergebnisse des Endpunkts mit hoch bewertet.

Endpunkt: Gesundheitsbezogene Lebensqualität nach EQ-5D

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

☐ **ja:** Eines der folgenden Merkmale trifft zu:

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

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

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

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

Die Auswertungspopulation für den EQ-5D umfasste 214 Patienten. Somit befanden sich, unabhängig von der Art und dem Zeitpunkt der Auswertung maximal 18 % der ITT-Population (1.199 Patienten) in der Auswertung. Damit wurde das ITT-Prinzip nicht adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

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

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

☐ **nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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

☒ **ja**

☐ **nein**

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

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:

Da das ITT-Prinzip nicht adäquat umgesetzt wurde, wird das Verzerrungspotenzial der Ergebnisse des Endpunkts mit 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

Die Studie wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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 nach CTCAE Grad ≥ 3

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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: Schwerwiegende 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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: Studienabbruch aufgrund unerwünschter 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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: Tod aufgrund unerwünschter 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 wurde doppelt verblindet durchgeführt. Sowohl die Patienten als auch das behandelnde Personal und/oder der Prüfarzt waren während der gesamten Studie verblindet

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: