

Nutzenbewertung



Gemeinsamer
Bundesausschuss

**von Arzneimitteln mit neuen Wirkstoffen nach
§ 35a SGB V**

**Bewertung von Arzneimitteln für seltene Leiden
nach § 35a Absatz 1 Satz 10 i.V.m.
5. Kapitel § 12 Nr. 1 Satz 2 Verfo**

Wirkstoff: Carfilzomib

Datum der Veröffentlichung: 1. November 2016

Inhaltsverzeichnis

Inhaltsverzeichnis	2
Tabellenverzeichnis	3
Abbildungsverzeichnis	4
Abkürzungsverzeichnis	5
Hintergrund	6
1 Einführung	7
2 Nutzenbewertung	8
2.1 Fragestellung	8
2.2 Zulassungsbegründende Studien	8
2.3 Liste der verwendeten Quellen	8
2.4 Studiencharakteristika	9
2.5 Endpunkte und statistische Methoden	15
2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte	15
2.5.2 Verzerrungspotential auf Studien- und Endpunktebene	16
2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz	18
2.5.4 Statistische Methoden	24
2.6 Ergebnisse zum Zusatznutzen	27
2.6.1 Charakterisierung der Studienpopulation	27
2.6.2 Mortalität	32
2.6.3 Morbidität	33
2.6.4 Lebensqualität	35
2.6.5 Subgruppenanalysen	37
2.6.6 Sicherheit	37
3 Methodische Anmerkungen und Bewertung der Unterlagen	44
3.1 Zulassungsstatus und Zulassungspopulation von Carfilzomib	44
3.2 Design und Methodik der Studie ENDEAVOR	44
3.3 Wirksamkeit	46
3.3.1 Zusammenfassende Einschätzung zur Wirksamkeit	47
3.4 Lebensqualität	47
3.4.1 Zusammenfassende Einschätzung zur Lebensqualität	49
3.5 Sicherheit	49
3.5.1 Zusammenfassende Einschätzung zur Sicherheit	50
4 Anforderungen an eine qualitätsgesicherte Anwendung	51
5 Zusammenfassung der Bewertung	52
Referenzen	54
Anhang	57

Tabellenverzeichnis

Tabelle 1:	Charakterisierung der Studie ENDEAVOR	9
Tabelle 2:	Charakterisierung der Intervention in der Studie ENDEAVOR	13
Tabelle 3:	In der Nutzenbewertung berücksichtigte Endpunkte.....	15
Tabelle 4:	Verzerrungspotential der Studie ENDEAVOR auf Studienebene	17
Tabelle 5:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie ENDEAVOR	17
Tabelle 6:	Beschreibung der Endpunkte der Kategorie Mortalität.....	19
Tabelle 7:	Beschreibung der Endpunkte der Kategorie Morbidität.....	19
Tabelle 8:	Beschreibung der Endpunkte der Kategorie Lebensqualität.....	21
Tabelle 9:	Beschreibung der Endpunkte der Kategorie Sicherheit.....	23
Tabelle 10:	Allgemeine Angaben zur Studie ENDEAVOR, Datenschnitt 10.11.2014.....	28
Tabelle 11:	Charakterisierung der Studienpopulation der Studie ENDEAVOR zu Baseline	29
Tabelle 12:	Ergebnisse zum Gesamtüberleben der Studie ENDEAVOR (Datenschnitt: 10.11.2014).....	32
Tabelle 13:	Ergebnisse zum progressionsfreien Überleben in der Studie ENDEAVOR (Datenschnitt 10. November 2014).....	33
Tabelle 14:	Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität in der Studie ENDEAVOR, Datenschnitt: 10.11.2014.....	35
Tabelle 15:	Ergebnisse der unerwünschten Ereignisse in der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	38
Tabelle 16:	UE mit Inzidenz $\geq 10\%$ in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	39
Tabelle 17:	UE des NCI-CTCAE-Grades 3 und höher mit der Inzidenz von mind. 2 % in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 11.10.2014.....	40
Tabelle 18:	SUE mit einer Häufigkeit von mind. 1 % in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	41
Tabelle 19:	Ergebnisse zum Auftreten peripherer Neuropathien in der Studie ENDEAVOR, Sicherheitspopulation, Datenschnitt: 10.11.2014.....	43
Tabelle 20:	Zusammenfassende Darstellung der Studienergebnisse der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	52
Tabelle 21:	Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität der präspezifizierten Subskalen des EORTC QLQ-C30 in der Studie ENDEAVOR, Datenschnitt 10.11.2014	57
Tabelle 22:	Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität des EORTC QLQ-MY20 in der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	59
Tabelle 23:	Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität der Subskala „zusätzliche Beschwerden“ ⁽¹⁾ des FACT/GOG-Ntx in der Studie ENDEAVOR, Datenschnitt 10.11.2014.....	59

Abbildungsverzeichnis

Abbildung 1: Behandlungsschema mit Carfilzomib und Dexamethason allein gemäß Fachinformation [5].....	14
Abbildung 2: Flussdiagramm zur Studie ENDEAVOR	29
Abbildung 3: Kaplan-Meier-Überlebenszeitkurven für Gesamtmortalität in der Studie ENDEAVOR (Datenschnitt: 10.11.2014).....	33
Abbildung 4: Kaplan-Meier-Überlebenszeitkurven für progressionsfreies Überleben in der Studie ENDEAVOR (Datenschnitt: 10. November 2014).....	34

Abkürzungsverzeichnis

Cd	Carfilzomib und Dexamethason
CHMP	Committee for Medicinal Products for Human Use
CTCAE	Common Terminology Criteria for Adverse Events
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30	EORTC Quality of Life Questionnaire C 30
EORTC QLQ-MY20	EORTC Quality of Life Questionnaire – Multiple Myelome Module
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
HIV	Humanes Immundefizienz-Virus
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISS	International Staging System
KI	Konfidenzintervall
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference
MMRM	Gemischte Modelle für wiederholte Messungen (mixed model for repeated measures)
MW	Mittelwert
N	Anzahl
NCI-CTCAE	National Cancer Institute – Common Toxicity Criteria for Adverse Events
PT	Preferred Term
pU	Pharmazeutischer Unternehmer
OS	Gesamtüberleben (overall survival)
PFS	Progressionsfreies Überleben (progression-free survival)
SAP	Statistischer Analyseplan
SD	Standardabweichung
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
SUE	Schwerwiegende/s unerwünschte/s Ereignis/se
UE	Unerwünschte/s Ereignis/se
Vd	Bortezomib und Dexamethason
VerfO	Verfahrensordnung
ZN	Zusatznutzen

Hintergrund

Carfilzomib ist zugelassen als Arzneimittel 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. Gemäß § 35a Absatz 1 Satz 10 SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

Gemäß § 35a Absatz 2 SGB V entscheidet der Gemeinsame Bundesausschuss (G-BA), ob er die Nutzenbewertung selbst durchführt oder das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragt. Ausgehend von der gesetzlichen Vorgabe in § 35a Absatz 1 Satz 10 SGB V, dass der Zusatznutzen eines Orphan Drug durch die Zulassung als belegt gilt, hat der G-BA in seiner Sitzung vom 15. März 2012 das Verfahren der Nutzenbewertung von Orphan Drugs dahingehend modifiziert, dass bei Orphan Drugs zunächst keine eigenständige Festlegung einer zweckmäßigen Vergleichstherapie mehr durch den G-BA als Grundlage der insoweit allein rechtlich zulässigen Bewertung des Ausmaßes eines gesetzlich zu unterstellenden Zusatznutzens erfolgt. Vielmehr wird ausschließlich auf der Grundlage der Zulassungsstudien das Ausmaß des Zusatznutzens durch den G-BA bewertet.

Dementsprechend hat der G-BA in seiner Sitzung am 15. März 2012 den mit Beschluss vom 1. August 2011 erteilten Auftrag an das IQWiG zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen gemäß § 35a Absatz 2 SGB V in der Weise abgeändert, dass bei Orphan Drugs eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erfolgt, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt (vgl. § 35a Absatz 1 Satz 11 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Carfilzomib zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers heran und bewertet die darin enthaltenen Daten nach Maßgabe der im 5. Kapitel § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Carfilzomib in seiner Sitzung am 25. Oktober 2016 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 29. Juli 2016 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pharmazeutischen Unternehmers in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung
- Anzahl der GKV-Patienten in der Zielpopulation

Die Nutzenbewertung wird am 1. November 2016 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Einführung

Multiples Myelom

Das Multiple Myelom ist durch die maligne Neubildung von Plasmazellen charakterisiert, die sich im Knochenmark anreichern und infolgedessen zu Knochenabbau und Knochenmarksversagen führen [22]. Es macht etwa 10 bis 15 Prozent aller hämatologischen Neoplasien aus, mit ca. 4 bis 6 Neuerkrankungen pro 100.000 Personen pro Jahr in Deutschland [27]. Das mediane Erkrankungsalter betrug 2012 in Deutschland 72–74 Jahre [27]. Die Symptome sind unspezifisch, häufig sind Müdigkeit, Knochenschmerzen, pathologische Frakturen, Infektneigung, Gewichtsverlust. Die Erkrankung ist nicht heilbar. Bei einer Mortalität von 2 bis 3/100.000 Personen liegt das 5-Jahres-Überleben in Deutschland bei 40 % [27].

Klinischen Leitlinien zufolge [22] sollte die Behandlung einer aktiven (symptomatischen) Erkrankung erfolgen, wenn mindestens eines der sogenannten CRAB-Kriterien zur Beurteilung der Krankheitsfolgen erfüllt ist: C – erhöhtes Kalzium im Blut (hypercalcemia), R – Funktionsstörung der Niere (renal insufficiency), A – Blutarmut (anemia), B – Knochenbeteiligung (bone lesions). Durch die Behandlung kann eine Remission von unterschiedlich langer Dauer erreicht werden. Wird die Erkrankung erneut aktiv (Rezidiv), sollte wieder behandelt werden. Kann keine Remission erreicht werden, wird die Erkrankung als refraktär bezeichnet. In Abhängigkeit von den Vortherapien sowie der Ausprägung und Dauer des jeweiligen Ansprechens stehen in Deutschland Proteasom-inhibitoren (Bortezomib), immunmodulatorische Substanzen (Thalidomid, Lenalidomid, Pomalidomid), der humane monoklonale Antikörper (Elotuzumab, Daratumumab), Kortikosteroide, Alkylanzien, Histon-Deacetylase-Hemmer (Panobinostat), Hochdosischemotherapeutika und die autologe oder allogene Stammzelltransplantation für eine patientenindividuelle Therapieplanung zur Verfügung.

Carfilzomib

Carfilzomib ist ein Proteasom-inhibitor, der selektiv und irreversibel an das N-terminale aktive Zentrum des 20S-Proteasoms bindet. Der Wirkstoff zeigt wenig oder keine Aktivität gegenüber anderen Proteaseklassen. Die Hemmung des 20S-Proteasoms, das proteolytische Partikel des 26S-Proteasoms, führt zu einer Reduktion des Proteinabbaus in Zellen und ist eine der Zielstrukturen bei der Entwicklung therapeutischer Wirkstoffe gegen das Multiple Myelom. Es wird vermutet, dass Carfilzomib über diesen Mechanismus eine antiproliferative und proapoptische Wirkung bei Zellen des Multiplen Myeloms entwickelt.

Carfilzomib (Kyprolis®) ist bereits in Kombination mit Lenalidomid und Dexamethason zur Therapie vorbehandelter Erwachsener mit Multiplem Myelom zugelassen. Die Indikation wurde 2016 für die alleinige Kombination mit Dexamethason zur Therapie Erwachsener mit mindestens einer vorangegangenen Therapie des Multiplen Myeloms erweitert.

Die erstmalige Zulassung durch die Europäische Kommission wurde am 19.11.2015 erteilt. Seit dem 15. Dez. 2015 ist der Wirkstoff auf dem deutschen Markt verfügbar.

2 Nutzenbewertung

2.1 Fragestellung

Carfilzomib (Kyprolis®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 1 Satz 10 SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. In die Bewertung werden gemäß Zulassungsstatus Patienten in folgender Indikation eingeschlossen:

- in Kombination mit Dexamethason zur Behandlung von erwachsenen Patienten mit Multiplem Myelom, die mindestens eine vorangegangene Therapie erhalten haben [5].

2.2 Zulassungsbegründende Studien

Pivotal Studie

- ENDEAVOR: „A Randomized, Open-label, Phase 3 Study of Carfilzomib Plus Dexamethasone vs. Bortezomib Plus Dexamethasone in Patients with Relapsed Multiple Myeloma“. Die pivotal ENDEAVOR-Studie ist eine offene, randomisierte, multizentrische Phase-III-Studie, die die Wirksamkeit und Sicherheit von Carfilzomib in Kombination mit Dexamethason im Vergleich zu Bortezomib mit Dexamethason in Patienten mit rezidiertem und/oder progredientem Multiplen Myelom untersucht.

Supportive Studie(n)

- Zum Zeitpunkt der Nutzenbewertung lagen keine Studien vor, die supportiv zur Herleitung des Zusatznutzens herangezogen werden konnten.

2.3 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Carfilzomib wurden folgende Studien und Daten herangezogen:

- Herstellerdossier zu Carfilzomib, Modul 4 [4]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [15]
- Studienbericht [3] inkl. Studienprotokoll [2] und statistischem Analyseplan [6] der Studie ENDEAVOR (2011-003)
- Publikation zur Studie ENDEAVOR [11]

Weitere eingereichte, nicht zur Nutzenbewertung von Carfilzomib verwendete Dokumente:

- Studienprotokoll, statistischer Analyseplan und Studienreport der APIRE-Studie (PX-171-009): Eine randomisierte, offene Studie der Phase III zum Vergleich von Carfilzomib, Lenalidomid und Dexamethason vs. Lenalidomid und Dexamethason bei Personen mit rezidiertem und/oder progredientem Multiplen Myelom.
- Studienprotokoll, statistischer Analyseplan und Studienreport der FOCUS-Studie (PX-171-011): Eine multizentrische, randomisierte, offene Studie der Phase III, die die Wirksamkeit

und Sicherheit von Carfilzomib als Monotherapie im Vergleich zu „best supportive care“ untersucht.

- Studienprotokoll und statistischer Analyseplan der CLARION-Studie (2012-005): Eine randomisierte, offene Studie der Phase III zum Vergleich von Carfilzomib, Melphalan und Prednisolon vs. Bortezomib, Melphalan und Prednisolon bei Patienten mit neu diagnostiziertem Multiplen Myelom, die für eine Stammzelltherapie nicht geeignet sind.
- Studienprotokoll der ARROW-Studie (CFZ014): Eine randomisierte, offene Studie der Phase III zum Vergleich der einmaligen vs. zweimaligen wöchentlichen Gabe von Carfilzomib in Kombination mit Dexamethason bei Patienten mit rezidiviertem und refraktären Multiplen Myelom.

In den oben genannten Studien wurde Carfilzomib nicht zulassungskonform und/oder nicht entsprechend der Indikationserweiterung angewendet. Deshalb werden diese Studien nicht zur Nutzenbewertung herangezogen.

2.4 Studiencharakteristika

Die Unterlagen zum Nachweis des Ausmaßes des Zusatznutzens für Carfilzomib basieren auf der Zulassungsstudie ENDEAVOR (2011-003). Die Studie und die Intervention werden in den Tabellen 1 und 2 charakterisiert.

Tabelle 1: Charakterisierung der Studie ENDEAVOR

	Beschreibung
Design	<p>Es handelt sich um eine offene, randomisierte, multizentrische Studie der Phase III im Parallelgruppendesign (1:1) mit zwei Behandlungsarmen. Die Studienpopulation umfasst Patienten mit rezidiviertem und/oder progredientem Multiplen Myelom, die bereits mindestens eine und maximal drei Vortherapien erhalten haben.</p> <p>Patienten wurden stratifiziert randomisiert nach vorheriger Proteasomtherapie, ISS-Stadium, Anzahl vorheriger Therapien und Wahl der Applikation von Bortezomib und im Verhältnis 1:1 dem Verumarm (Cd: Carfilzomib und Dexamethason) oder dem Kontrollarm (Vd: Bortezomib und Dexamethason) zugeordnet. Die Behandlung erfolgte im Verumarm in 28-d-Zyklen und im Kontrollarm in 21-d-Zyklen bis Krankheitsprogression, Auftreten inakzeptabler Toxizität, Arztentscheidung, Rückzug der Einverständniserklärung oder Tod. Nach Beendigung der Studienmedikation gehen die Teilnehmer in die Langzeitnachbeobachtung über. Dabei werden alle 28 Tage das Auftreten eines Krankheitsprogresses bzw. nach Krankheitsprogress alle drei Monate auf Überleben überprüft.</p> <p>Der primäre Endpunkt war progressionsfreies Überleben (PFS). Weitere Endpunkte umfassten u. a. Gesamtüberleben (OS), Ansprechen, Neuropathie, Sicherheit und Lebensqualität.</p> <p>Die Studie wurde bzgl. PFS beendet. Die Teilnehmer erhalten weiterhin entsprechend der Randomisierung die Studienmedikation und die Nachbeobachtung erfolgt hinsichtlich OS und Sicherheit.</p>
Population	<p>Wesentliche Einschlusskriterien:</p> <ul style="list-style-type: none"> • Rezidiviertes oder progredientes Multiples Myelom • Messbare Erkrankung (innerhalb von 21 Tagen vor Randomisierung) definiert als Erfüllung mindestens eines der folgenden Kriterien: <ul style="list-style-type: none"> ○ M-Protein im Serum $\geq 0,5$ g/dl ○ M-Protein im Urin ≥ 200 mg/24 h

Dossierbewertung für Orphan Drugs

Carfilzomib (Kyprolis®)

	Beschreibung
	<ul style="list-style-type: none"> ○ in Patienten ohne nachweisbares M-Protein im Serum oder Urin: freie Leichtketten ≥ 100 mg/l und abnormales Kappa/Lambda-Verhältnis im Serum ○ Patienten mit Immunglobulin A Multiples Myelom: Immunglobulin im Serum (qIgA) ≥ 750 mg/dl • Dokumentiertes partielles Ansprechen auf mindestens eine Vortherapie (basierend auf Einschätzung des Studienarztes) • Mindestens eine, aber nicht mehr als drei Vortherapien (Induktionstherapie gefolgt von Stammzelltransplantation und Erhaltungstherapie werden als eine Therapielinie gewertet) • Vorbehandlung mit Bortezomib erlaubt, falls mindestens partielles Ansprechen auf Bortezomib-Vorbehandlung und das Absetzen nicht aufgrund von Toxizität erfolgte sowie ein therapiefreies Intervall von Bortezomib ≥ 6 Monate vorlag • Vorbehandlung mit Carfilzomib erlaubt, falls mindestens partielles Ansprechen auf Carfilzomib-Vorbehandlung und das Absetzen nicht aufgrund von Toxizität erfolgte sowie ein therapiefreies Intervall von Carfilzomib ≥ 6 Monate vorlag; ausgenommen sind Teilnehmer an jeglicher anderen Onyx-gesponserten Phase-III-Studie • ≥ 18 Jahre • ECOG-Performance-Status 0–2 • Linksventrikuläre Ejektionsfraktion ≥ 40 % • 21 Tage vor Randomisierung: <ul style="list-style-type: none"> ○ Adäquate Leberfunktion mit Bilirubin $< 1,5$-fache und Aspartat-Aminotransferase und Alanin-Aminotransferase < 3-fache des oberen Normwerts ○ Absolute Anzahl neutrophiler Granulozyten $\geq 1000/\text{mm}^3$ ○ Hämoglobin ≥ 8 g/dl (Erythropoese-stimulierende Faktoren und Erythrozytentransfusionen erlaubt) ○ Thrombozyten-Anzahl $\geq 50.000/\text{mm}^3$ ($\geq 30.000/\text{mm}^3$, falls Plasmazellen im Knochenmark > 50 %); mindestens eine Woche vor Messung kein Erhalt von Thrombozytentransfusionen ○ Kreatinin-Clearance ≥ 15 ml/min • Frauen im gebärfähigen Alter: negativer Schwangerschaftstest 21 Tage vor Randomisierung und Einverständnis zur Durchführung effektiver Kontrazeption bis 3 Monate nach Gabe der letzten Dosis des Studienarzneimittels • Männer: Einverständnis zur Durchführung effektiver Kontrazeption <p>Wesentliche Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Multiples Myelom des IgM-Subtyps • Gabe von Kortikosteroiden (Prednison > 30 mg/Tag oder äquivalent) innerhalb 14 Tage vor Randomisierung • POEMS-Syndrom (Polyneuropathie, Organomegalie, Endokrinopathie, Monoklonales Protein und Hautveränderungen) • Plasmazell-Leukämie oder zirkulierende Plasmazellen $\geq 2 \times 10^9/\text{l}$ • Bekannte Amyloidose • Waldenström-Makroglobulinämie • Chemotherapie oder Immunotherapie innerhalb 21 Tage vor Randomisierung • Fokale Strahlentherapie innerhalb 7 Tage bzw. großflächige, ein erhebliches Volumen an Knochenmark betreffende Strahlentherapie innerhalb von 21 Tagen vor Randomisierung • Aufwendige Operation (außer Kyphoplastik) innerhalb 28 Tage vor Randomisierung • Myokardinfarkt innerhalb 4 Monate vor Randomisierung, aktives kongestives Herzversagen (NYHA Klasse III oder IV), symptomatische Ischämie oder unkontrollierte Anomalien des Erregungsleitungssystems • Akute aktive, eine Behandlung mit systemischen Antibiotika, Virustatika oder

Dossierbewertung für Orphan Drugs

Carfilzomib (Kyprolis®)

	Beschreibung				
	<p>Antimykotika erfordernde Infektion innerhalb von 14 Tagen vor Randomisierung</p> <ul style="list-style-type: none"> • Bekannte HIV-Infektion, Hepatitis-C- oder Hepatitis-B-Infektion (außer Patienten mit Hepatitis-B-Oberflächenantigenen und Core-Antikörper, die auf eine antivirale Hepatitis-B-Therapie ansprechen) • Bekannte Zirrhose • Zweite maligne Erkrankung innerhalb der vergangenen 3 Jahre, außer: <ul style="list-style-type: none"> ○ Adäquat behandeltes Basalzell-Karzinom, Plattenepithelzell-Karzinom der Haut ○ In situ Zervixkarzinom ○ Prostatakrebs Gleason Score ≤ 6 mit stabilen Prostata-spezifischen Antigenspiegeln über die letzten 12 Monate ○ In situ Brustkarzinom, operativ vollständig entfernt ○ Behandeltes medulläres oder papilläres Schilddrüsenkarzinom • Myelodysplastisches Syndrom • Signifikante Neuropathie (Grad 3–4 oder Grad 2 mit Schmerz) innerhalb 14 Tagen vor Randomisierung • Schwangere oder stillende Frauen • Bekannte Allergie gegenüber Captisol® • Überempfindlichkeit gegenüber Carfilzomib, Bortezomib, Boron oder Mannitol • Kontraindikation gegenüber Dexamethason • Kontraindikation gegenüber jeglicher erforderlichen gleichzeitigen Gabe von Arzneimitteln oder unterstützenden Behandlung, inkl. Überempfindlichkeit gegenüber allen Virustatika oder Unverträglichkeit gegenüber Hydrierung aufgrund bestehender eingeschränkter Lungen- oder Herzfunktion • Bestehende Graft-vs.-host-Erkrankung • Punktion erfordernder Pleuraerguss oder Aszites innerhalb 14 Tage vor Randomisierung 				
Intervention und Zahl der Patienten	<table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Carfilzomib + Dexamethason (Cd)</th> <th>Bortezomib + Dexamethason (Vd)</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">N=465</td> <td style="text-align: center;">N=464</td> </tr> </tbody> </table> <p>Es wurden 1096 Patienten für den Studieneinschluss gescreent. Davon wurden 929 randomisiert. Im Verumarm erhielten 9 und im Kontrollarm ein Patient keine Studienmedikation.</p>	Carfilzomib + Dexamethason (Cd)	Bortezomib + Dexamethason (Vd)	N=465	N=464
Carfilzomib + Dexamethason (Cd)	Bortezomib + Dexamethason (Vd)				
N=465	N=464				
Ort und Zeitraum der Durchführung	<p>Studiendauer: Studienaufnahme: 20. Juni 2012 bis 30. Juni 2015¹⁾ Datenschnitt Interimsanalyse: 10. November 2014 (ereignisgetrieben bei Erreichen von mindestens 395 PFS-Ereignissen) Die Studie ist aktiv. Es werden Daten zum OS und Sicherheit gesammelt.</p> <p>Ort der Durchführung: Die Studie wurde in 198 klinischen Zentren in 27 Ländern durchgeführt: Australien, Belgien, Brasilien, Bulgarien, Deutschland, Frankreich, Großbritannien, Griechenland, Israel, Italien, Japan, Kanada, Korea, Neuseeland, Österreich, Polen, Rumänien, Russland, Singapur, Slowakei, Spanien, Taiwan, Thailand, Tschechien, Ukraine, Ungarn, USA.</p>				

	Beschreibung
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primäres Zielkriterium:</p> <ul style="list-style-type: none"> • PFS <p>Sekundäre Zielkriterien²⁾:</p> <ul style="list-style-type: none"> • OS <ul style="list-style-type: none"> ◦ Ansprechen (ORR) • Dauer des Ansprechens (DOR) • Auftreten von Neuropathie (≥ Grad 2 und FACT/GOG-NTx-Score) • EORTC QLQ-C30: Skala: allgemeiner Gesundheitszustand • Sicherheit und Tolerierbarkeit (Veränderung der linksventrikulären Ejektionsfraktion, der rechtsventrikulären Funktion und des Lungenarteriendrucks in einer Subpopulation des Verum- und Kontrollarms) <p>Explorative Endpunkte²⁾:</p> <ul style="list-style-type: none"> • EORTC QLQ-C30: Subskalen (Fatigue, Schmerz, Physis, Rolle, Übelkeit/Erbrechen) • EORTC QLQ-MY20: Subskalen (Krankheitssymptome, Nebenwirkungen) • FACT/GOG-Ntx: Subskala „zusätzliche Beschwerden“ • TTP (Zeit bis zur Krankheitsprogression) • MRD (minimal residual disease) Negativ-Status in Patienten mit mindestens sehr gutem teilweisen Ansprechen (VGPR – very good partial response) • CBR (clinical benefit response) • DCR (disease control rate) • MRU (medical resource utilization) • Pharmakokinetik • Pharmakodynamik • Genetische Biomarker und Genexpression von Biomarkern, die prädiktiv für das Ansprechen und die Resistenz gegenüber Proteasom-inhibitoren sind
Subgruppenanalysen	<p>A priori geplant:</p> <ul style="list-style-type: none"> • Alter (18–65 vs. > 65 Jahre) • Geschlecht • Ethnie genetisch (Weiß, Schwarz, Asiate, andere) • Region (Osteuropa, Westeuropa, Nordamerika, Südamerika, Asien/Pazifik)³⁾ • Kreatininausscheidung (< 30 ml/min, 30 bis < 50 ml/min, 50 bis < 80 ml/min, ≥ 80 ml/min) • β2-Mikroglobulinspiegel (< 3,5 mg/L vs. ≥ 3,5 mg/L)³⁾ • Risikogruppe bestimmt mittels FISH (hoch, Standard): Hoch-Risiko-Gruppe besteht aus den genetischen Subtypen: t(4; 14), t(14; 16) und Deletion des 17p. Die Standard-Risikogruppe besteht aus allen anderen berichteten Subtypen. • Anzahl Vortherapien (1 vs. 2–3) • Vorherige Therapie mit Proteasom-inhibitor (ja vs. nein) • Vorherige Therapie mit Lenalidomid (ja vs. nein) • Vorherige Therapie mit Thalidomid (ja vs. nein) • Bortezomibapplikation (i.v. vs. subkutan) <p>Post hoc für das Dossier ausgewertete Subgruppen (nicht im Studienbericht)</p> <ul style="list-style-type: none"> • Keine

¹⁾ Angaben aus Modul 4. Hierzu keine Angaben im Studienprotokoll oder Studienbericht.

²⁾ Einteilung der Zielkriterien basierend auf Amendement 2 des Studienprotokolls (2. Oktober 2014) vor Datenschnitt 10. November 2014

³⁾ Definition der Subgruppen nach Datenschnitt am 10. November 2014 geändert. Ursprüngliche Planung: Region: Nordamerika, Westeuropa, restliche Welt) bzw. β2-Mikroglobulinspiegel (< 2,5 mg/L vs. ≥ 2,5 mg/L)

Abkürzungen: POEMS: Polyneuropathie, Organomegalie, Endokrinopathie, Monoklonales Protein und Hautveränderungen-Syndrom; ECOG: Eastern Cooperative Oncology Group; OS: Gesamtüberleben

Anhand der Ein- und Ausschlusskriterien der Studie ENDEAVOR wird deutlich, dass jene Patienten, die nicht auf mindestens eine von maximal drei vorangegangenen Therapien ansprachen, nicht umfasst sind. Ebenso wurden Patienten mit Multiplem Myelom des Subtyps IgM sowie Patienten mit mehr als drei Vortherapien nicht berücksichtigt. Somit entspricht die Studienpopulation nicht der Zulassungspopulation, welche alle Patienten mit mindestens einer vorangegangenen Therapie, unabhängig vom Ansprechen und Typ des Multiplen Myeloms umfasst [5].

Das Studienprotokoll wurde vier Mal geändert. Davon erfolgten die ersten beiden Änderungen vor dem Datenschnitt der Interimsanalyse und zwei Änderungen danach. Im Amendment 3 wurde die Skala des EORTC QLQ-C30 „Gesundheitszustand/Lebensqualität“ von einem sekundären zu einem explorativen Endpunkt klassifiziert. Zudem erfolgte eine Änderung der Definition des Endpunkts „periphere Neuropathien“. Ursprünglich sollten die Ergebnisse der FACT/GOG-Ntx-Skala neben MedDRA Version 15.1 herangezogen werden. Im Modul 4 wurde die Skala FACT/GOG-Ntx als separater Endpunkt in der Kategorie „gesundheitsbezogene Lebensqualität“ und „periphere Neuropathien“ gemäß MedDRA Version 15.1 in der Kategorie „Sicherheit“ betrachtet.

Tabelle 2: Charakterisierung der Intervention in der Studie ENDEAVOR

Intervention	Kontrolle
Cd: Carfilzomib, Dexamethason	Vd: Bortezomib, Dexamethason
<p>Zykluslänge: 28 Tage</p> <p>Carfilzomib: 20 mg/m² i.v. über 30 min (± 5 min) an Tag 1 und 2 des Zyklus 1, gefolgt von der Eskalation auf 56 mg/m² über 30 min (± 5 min) an den Tagen 8, 9, 15 und 16 des Zyklus 1</p> <p>Patienten, die 56 mg/m² im Zyklus 1 tolerierten, wurden fortan in jedem weiteren Zyklus an den Tagen 1, 2, 8, 9, 15 und 16 mit 56 mg/m² behandelt.</p> <p>Dexamethason: 20 mg wurden an den Tagen 1, 2, 8, 9, 15, 16, 22 und 23 mindestens 30 min, aber nicht mehr als 4 h vor Carfilzomibapplikation verabreicht (vergleiche Abbildung 1).</p>	<p>Zykluslänge: 21 Tage</p> <p>Bortezomib: 1,3 mg/m² i.v. oder subkutan an Tagen 1, 4, 8 und 11</p> <p>Dexamethason: 20 mg wurden an den Tagen 1, 2, 4, 5, 8, 9, 11 und 12 mindestens 30 min, aber nicht mehr als 4 h vor Bortezomibapplikation verabreicht</p>
<p>Darreichungsform und Darreichungsbedingungen Carfilzomib:</p> <p>Carfilzomib lag als Pulver zur Herstellung einer Infusionslösung vor und wurde mit Wasser zur Injektion mit einer Konzentration von 2,0 mg/ml aufbereitet.</p> <p>Es wurde eine intravenöse Hydratation (250–500 ml Kochsalzlösung) unmittelbar vor Verabreichung von Carfilzomib im Zyklus 1 durchgeführt und nach Ermessen des Untersuchers in Folgezyklen fortgesetzt.</p> <p>Patienten mit erhöhtem Risiko für die Entwicklung des Tumor-Lyse-Syndroms konnten bereits mit einer oralen Hydratation 48 h vor Gabe von Carfilzomib beginnen.</p>	<p>Darreichungsform und Darreichungsbedingungen Bortezomib:</p> <p>Bortezomib lag als Ampulle mit 3,5 mg Bortezomib vor und wurde mit Natriumchlorid zur Herstellung einer 1 mg/ml Injektionslösung verwendet. Zur subkutanen Applikation wurde eine 2,5 mg/ml Lösung hergestellt.</p> <p>Die Menge der verabreichten Bortezomibdosis wurde anhand der aktuellen Körperoberfläche berechnet. Eine Dosisanpassung erfolgte für Gewichtsveränderungen ≤ 10 %.</p> <p>Bortezomib wurde als 3- bis 5-Sekunden-Bolusinjektion i.v. oder subkutan entsprechend der Zulassung verabreicht. Traten lokale Reaktionen an der Einstichstelle bei subkutaner Applikation auf, so</p>

Intervention	Kontrolle																																																																																																																														
<p>Die Menge der verabreichten Carfilzomibdosis wurde anhand der Körperoberfläche berechnet. Personen mit einer Körperoberfläche > 2,2 m² erhielten eine Dosis basierend auf 2,2 m² Körperoberfläche. Die Dosis wurde für Gewichtsveränderungen ≤ 20 % angepasst.</p> <p>Carfilzomib wurde i.v. als Infusion über ca. 30 min (± 5 min) appliziert und der Patient blieb mind. für eine Stunde nach Gabe im Studienzentrum unter Beobachtung. Während der Beobachtung wurde erneut eine Hydrierung (i.v. mit 250 bis 500 ml Kochsalzlösung) im Zyklus 1 durchgeführt, welche nach Ermessen des Untersuchers in den weiteren Zyklen durchgeführt werden konnte.</p>	<p>konnte auf die geringere Konzentration von 1 mg/ml umgestellt werden.</p>																																																																																																																														
<p>Darreichungsform und Darreichungsbedingungen Dexamethason</p> <p>Dexamethason stand zur oralen Einnahme als 4 mg Tablette zur Verfügung. An Kliniktagen konnte Dexamethason i.v. verabreicht werden.</p> <p>Vergessene Dosen an Dexamethason sollten nicht zu einem späteren Zeitpunkt zusätzlich eingenommen werden. Die Einnahme von Dexamethason sollte in einem Tagebuch dokumentiert werden.</p>																																																																																																																															
<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th rowspan="3"></th> <th colspan="12">Zyklus 1</th> </tr> <tr> <th colspan="3">Woche 1</th> <th colspan="3">Woche 2</th> <th colspan="3">Woche 3</th> <th colspan="3">Woche 4</th> </tr> <tr> <th>Tag 1</th> <th>Tag 2</th> <th>Tage 3-7</th> <th>Tag 8</th> <th>Tag 9</th> <th>Tage 10-14</th> <th>Tag 15</th> <th>Tag 16</th> <th>Tage 17-21</th> <th>Tag 22</th> <th>Tag 23</th> <th>Tage 24-28</th> </tr> </thead> <tbody> <tr> <td>Kyprolis (mg/m²)</td> <td>20</td> <td>20</td> <td>–</td> <td>56</td> <td>56</td> <td>–</td> <td>56</td> <td>56</td> <td>–</td> <td>–</td> <td>–</td> <td>–</td> </tr> <tr> <td>Dexamethason (mg)</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> </tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th rowspan="3"></th> <th colspan="12">Zyklus 2 und alle nachfolgenden Zyklen</th> </tr> <tr> <th colspan="3">Woche 1</th> <th colspan="3">Woche 2</th> <th colspan="3">Woche 3</th> <th colspan="3">Woche 4</th> </tr> <tr> <th>Tag 1</th> <th>Tag 2</th> <th>Tage 3-7</th> <th>Tag 8</th> <th>Tag 9</th> <th>Tage 10-14</th> <th>Tag 15</th> <th>Tag 16</th> <th>Tage 17-21</th> <th>Tag 22</th> <th>Tag 23</th> <th>Tage 24-28</th> </tr> </thead> <tbody> <tr> <td>Kyprolis (mg/m²)</td> <td>56</td> <td>56</td> <td>–</td> <td>56</td> <td>56</td> <td>–</td> <td>56</td> <td>56</td> <td>–</td> <td>–</td> <td>–</td> <td>–</td> </tr> <tr> <td>Dexamethason (mg)</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> <td>20</td> <td>20</td> <td>–</td> </tr> </tbody> </table>			Zyklus 1												Woche 1			Woche 2			Woche 3			Woche 4			Tag 1	Tag 2	Tage 3-7	Tag 8	Tag 9	Tage 10-14	Tag 15	Tag 16	Tage 17-21	Tag 22	Tag 23	Tage 24-28	Kyprolis (mg/m ²)	20	20	–	56	56	–	56	56	–	–	–	–	Dexamethason (mg)	20	20	–	20	20	–	20	20	–	20	20	–		Zyklus 2 und alle nachfolgenden Zyklen												Woche 1			Woche 2			Woche 3			Woche 4			Tag 1	Tag 2	Tage 3-7	Tag 8	Tag 9	Tage 10-14	Tag 15	Tag 16	Tage 17-21	Tag 22	Tag 23	Tage 24-28	Kyprolis (mg/m ²)	56	56	–	56	56	–	56	56	–	–	–	–	Dexamethason (mg)	20	20	–	20	20	–	20	20	–	20	20	–
	Zyklus 1																																																																																																																														
	Woche 1			Woche 2			Woche 3			Woche 4																																																																																																																					
	Tag 1	Tag 2	Tage 3-7	Tag 8	Tag 9	Tage 10-14	Tag 15	Tag 16	Tage 17-21	Tag 22	Tag 23	Tage 24-28																																																																																																																			
Kyprolis (mg/m ²)	20	20	–	56	56	–	56	56	–	–	–	–																																																																																																																			
Dexamethason (mg)	20	20	–	20	20	–	20	20	–	20	20	–																																																																																																																			
	Zyklus 2 und alle nachfolgenden Zyklen																																																																																																																														
	Woche 1			Woche 2			Woche 3			Woche 4																																																																																																																					
	Tag 1	Tag 2	Tage 3-7	Tag 8	Tag 9	Tage 10-14	Tag 15	Tag 16	Tage 17-21	Tag 22	Tag 23	Tage 24-28																																																																																																																			
Kyprolis (mg/m ²)	56	56	–	56	56	–	56	56	–	–	–	–																																																																																																																			
Dexamethason (mg)	20	20	–	20	20	–	20	20	–	20	20	–																																																																																																																			
<p>^a Die Infusionsdauer beträgt 30 Minuten und bleibt während des Behandlungsregimes unverändert.</p>																																																																																																																															
<p><i>Abbildung 1: Behandlungsschema mit Carfilzomib und Dexamethason allein gemäß Fachinformation [5]</i></p>																																																																																																																															
<p>Dosisanpassungen Carfilzomib:</p> <p>Carfilzomib wurde von einer Dosis von 56 mg/m² im ersten Schritt auf 45 mg/m², danach auf 36 mg/m² und zuletzt auf 27 mg/m² reduziert. Falls Toxizität trotz Dosisanpassungen bestand, konnte auf einen reduzierten Dosiszyklus umgestellt werden. Dazu wurde die geringste Dosis an den Tagen 1, 2, 15 und 16 des 28-tägigen Zyklus appliziert. Erfolgte eine Umstellung auf den reduzierten Dosiszyklus, so wurde dieser für die verbleibende Studiendauer beibehalten. Wurde die Dosis von Carfilzomib nach den präspezifizierten Vorgaben reduziert, so wurde der darauffolgende Zyklus mit der reduzierten Dosis fortgeführt. Eine Reeskalation war nach Ermessen des Untersuchers möglich.</p>	<p>Dosisanpassungen Bortezomib:</p> <p>Bortezomib wurde von der Dosis von 1,3 mg/m² in einem ersten Schritt auf 1 mg/m² und danach auf 0,7 mg/m² reduziert. Bestand trotz Dosisreduktion weiterhin Toxizität, so konnte auf einen reduzierten Dosiszyklus umgestellt werden. Dabei wurde Bortezomib wöchentlich für 4 Wochen appliziert (Tage 1, 8, 15, 22), gefolgt von einer 13-tägigen Ruhephase.</p> <p><u>Einnahmedauer:</u> bis Krankheitsprogression, Auftreten inakzeptabler Toxizität, Arztentscheidung, Rückzug der Einverständniserklärung oder Tod</p>																																																																																																																														

Intervention	Kontrolle
Einnahmedauer: bis Krankheitsprogression, Auftreten inakzeptabler Toxizität, Arztentscheidung, Rückzug der Einverständniserklärung oder Tod	
Dosisanpassungen Dexamethason: Dexamethason konnte von 20 mg auf 12 und in einem weiteren Schritt auf 8 mg reduziert werden. Falls trotz Dosisreduktionen Toxizität bestand, so sollte die Einnahme von Dexamethason beendet werden. Die Einnahme von Carfilzomib und Bortezomib blieb davon unberührt. Im Falle einer Dosisreduktion wurde diese im folgenden Behandlungszyklus fortgeführt. Eine Reeskalation lag im Ermessen des Untersuchers.	
Nicht erlaubte Begleitmedikation: <ul style="list-style-type: none"> • Andere Krebsmedikamente oder Bestrahlung • Keine andere Krebstherapie vor Dokumentation einer Krankheitsprogression erlaubt 	
Erforderliche Begleitmedikation: Während der Einnahme von Dexamethason: <ul style="list-style-type: none"> • Valaciclovir (oder äquivalentes antivirales Medikament) • Lansoprazol (oder anderer Protoneninhibitor) 	
Erlaubte Begleitmedikation: <ul style="list-style-type: none"> • Allopurinol gemäß Packungsbeilage oder andere uratsenkende Wirkstoffe für Patienten mit einem erhöhten Risiko am Tumor-Lyse-Syndrom • Mycostatin oder Fluconazol • Antiemetikum und Antidiarrhoikum nach Bedarf • Myeloide Wachstumsfaktoren gegen Neutropenie • Erythrozytentransfusionen, Erythropoese-stimulierende Substanzen oder Thrombozytentransfusion, wenn klinisch indiziert • Bisphosphonate • Palliative Bestrahlung zum Schmerzmanagement mit Erlaubnis des Monitors 	

Abkürzungen: i.v.: intravenös

2.5 Endpunkte und statistische Methoden

2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte

Tabelle 3: In der Nutzenbewertung berücksichtigte Endpunkte

Endpunkt	Kategorie	Vom pU als patientenrelevant eingestuft	In der Nutzenbewertung berücksichtigt
Gesamtüberleben (OS)	Mortalität	•	•
Progressionsfreies Überleben (PFS)	Morbidität	•	o ¹⁾
Zeit bis Beginn nachfolgender Therapie		•	-
Ansprechen Gesamtansprechrates Komplettremission (sCR+CR) Sehr gute partielle Remission (sCR+CR+VGPR) Dauer der Remission		•	-

Endpunkt	Kategorie	Vom pU als patientenrelevant eingestuft	In der Nutzenbewertung berücksichtigt
Gesundheitsbezogene Lebensqualität EORTC QLQ-C30	Lebensqualität	•	•
Gesundheitsbezogene Lebensqualität EORTC QLQ-MY20		•	•
Gesundheitsbezogene Lebensqualität FACT/GOG-NTx		•	•
Periphere Neuropathie	Sicherheit	•	•
Unerwünschte Ereignisse		•	•

- Endpunkt wird als bewertungsrelevant eingestuft und in der Nutzenbewertung berücksichtigt.
- Endpunkt wird nicht als bewertungsrelevant eingestuft, aber ergänzend in der Nutzenbewertung dargestellt.
- Endpunkt wird nicht als bewertungsrelevant eingestuft und nicht in der Nutzenbewertung dargestellt.
- ¹⁾ Dieser Endpunkt wird aufgrund nicht nachvollziehbarer Operationalisierung und unklarer Patientenrelevanz nicht in der Nutzenbewertung berücksichtigt. Da es sich um den primären Endpunkt der Studie ENDEAVOR handelt, werden die Ergebnisse hierzu unter 2.6.3 ergänzend dargestellt.

Abkürzungen: sCR: stringente komplette Remission; CR: komplette Remission; VGPR: sehr gute partielle Remission

Der Einschätzung des pU, dass das Ansprechen („Gesamtansprechrate“, „Tiefe der Remission“ und „Dauer der Remission“) und die Zeit bis zum Beginn einer Folgetherapie patientenrelevante und bewertungsrelevante Endpunkte seien, konnte nicht gefolgt werden.

Ansprechen wurde gemäß IMWG-Kriterien definiert, welche entweder Laborparameter oder bildgebende Verfahren umfassen. Diese Endpunkte sind per se nicht patientenrelevant (vgl. Tabelle 7 „progressionsfreies Überleben“) und stellen auch keine validen Surrogatparameter für patientenrelevante Endpunkte dar. Ein möglicher Zusammenhang mit Lebensqualität oder Mortalität sollte sich in den Ergebnissen der entsprechenden Endpunktkategorien widerspiegeln. Endpunkte in den Endpunktkategorien Lebensqualität und Mortalität werden in der Nutzenbewertung von Carfilzomib berücksichtigt.

Die Patientenrelevanz des Endpunkts „Zeit bis Beginn einer Folgetherapie“, definiert als Zeit ab Randomisierung bis zur Aufnahme einer nachfolgenden Multiplen-Myelom-Therapie, kann nicht abschließend eingeschätzt werden. Dieser Endpunkt findet sich nicht im Studienprotokoll und wurde erst post hoc eingeführt und ausgewertet. Es liegen keine Informationen zur Operationalisierung vor. Dies betrifft insbesondere die Gründe für eine erneute Therapiebedürftigkeit als auch den Zeitpunkt, wann die Folgetherapie begonnen werden sollte. Dieser Endpunkt wird aus den o. g. Gründen nicht in die Bewertung des Zusatznutzens von Carfilzomib eingeschlossen.

2.5.2 Verzerrungspotential auf Studien- und Endpunktebene

Die Studienteilnehmer wurden den Studienarmen 1:1 randomisiert zugeordnet. Die Randomisierung erfolgte stratifiziert nach:

- Vorheriger Therapie mit Carfilzomib oder Bortezomib (ja vs. nein)
- ISS-Krankheitsstadium (Stadium 1 vs. Stadium 2 oder 3)
- Anzahl der Vortherapien (1 Vortherapie vs. 2 oder 3 Vortherapien)
- Wahl der Applikation von Bortezomib (i.v. vs. subkutan).

Die Zuteilung der Studienteilnehmer erfolgte verdeckt. Das Studienpersonal und die Studienteilnehmer waren gegenüber der Studienmedikation nicht verblindet.

Tabelle 4: Verzerrungspotential der Studie ENDEAVOR auf Studienebene

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patient	Verblindung Behandler	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
ENDEAVOR	Ja	Ja	Nein	Nein	Nein	Ja ¹⁾	Hoch

¹⁾ Ein bedeutender Teil der Studienpopulation (125 Personen, 13,5 % der Studienpopulation) wurde abweichend vom Randomisierungsschema den Studienarmen zugeteilt (Protokollverletzung). Es traten Diskrepanzen im Patientenfluss (siehe Abbildung 2 und Tabelle 10) und den Angaben zu Zensierungen beim primären Endpunkt (siehe Tabelle 13) hinsichtlich der Anzahl der Patienten, die einen Krankheitsprogress aufwiesen, lost to follow-up waren und die Einverständniserklärung zurückriefen, auf.

Das Verzerrungspotential auf Studienebene wird in der Gesamtschau als hoch eingeschätzt. Dies basiert auf dem offenen Studiendesign und dem hohen Anteil an Studienteilnehmern, die dem falschen Stratum randomisiert zugeordnet wurden (n=125, 13,5 %) sowie den Diskrepanzen zwischen Patientenfluss und Angaben zum primären Endpunkt.

Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 5 dargestellt.

Tabelle 5: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studie ENDEAVOR

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
Gesamtüberleben	Nein	Ja	Nein	Nein ¹⁾	Niedrig
Morbidität					
Progressionsfreies Überleben	Nein	Ja	Nein	Nein ¹⁾	Hoch ²⁾
Lebensqualität					
EORTC QLQ-C30	Nein	Ja	Nein	Nein ¹⁾	Hoch ³⁾
EORTC QLQ-MY20	Nein	Ja	Nein	Nein ¹⁾	Hoch ³⁾
FACT/GOG-NTx	Nein	Ja	Nein	Nein ¹⁾	Hoch ³⁾

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Sicherheit					
Periphere Neuropathie	Nein	Ja ⁴⁾	Nein	Nein ¹⁾	Hoch ⁵⁾
Unerwünschte Ereignisse	Nein	Ja ⁴⁾	Ja ⁶⁾	Nein ¹⁾	Hoch ⁵⁾

¹⁾ Trotz der hohen Zahl an Protokollverletzern durch Randomisierung in ein falsches Stratum wird dies nicht als zusätzlich verzerrender Faktor für Analysen in der Gesamtstudienpopulation gesehen, da die erfassten Baselinecharakteristika inklusive der erfassten Prognosefaktoren bzgl. der Stratifizierungsfaktoren ausgeglichen waren. Ergebnisse von Subgruppenanalysen könnten allerdings hinsichtlich der Validität eingeschränkt sein.

²⁾ Weder das Studienpersonal noch die Studienteilnehmer waren verblindet und diagnostische Messungen (Plasmozytome, Knochenläsionen, Plasmazellen im Knochenmark) erfolgten optional oder nach Vorliegen einer klinischen Indikation. Die Entscheidung, wann diagnostische Messungen vorgenommen wurden, wurde nicht eindeutig definiert und operationalisiert und wurde eventuell durch das Wissen um die Zuteilung des Patienten zum jeweiligen Studienarm beeinflusst.

³⁾ Da die Verblindung nicht gegeben war, wird das Verzerrungspotential als hoch eingestuft.

⁴⁾ Es wurde a priori festgelegt, dass dieser Endpunkt in der Sicherheitspopulation untersucht wird.

⁵⁾ Das Verzerrungspotential wird für sicherheitsrelevante Endpunkte, die vom Studienteilnehmer oder vom Studienpersonal eingeschätzt wurden, als „hoch“ angesehen, da eine Verblindung nicht vorlag. Für objektiv gemessene Laborparameter wird das Verzerrungspotential als „niedrig“ angesehen.

⁶⁾ Es bestehen Diskrepanzen zwischen geplantem und tatsächlich berichtetem Zeitraum. Laut Studienprotokoll sollen UE ab dem Tag der Randomisierung berichtet werden. Im Studienbericht werden UE ab dem Tag der ersten Studienmedikation aufgeführt.

Abkürzungen: ITT: intention-to-treat

Beim Endpunkt „Gesamtüberleben“ spielt die Verblindung eine untergeordnete Rolle, solange die Nachbeobachtung für beide Studienarme gleichermaßen operationalisiert und durchgeführt wurde. Da keine Hinweise auf ein unterschiedliches Vorgehen der Ereigniserhebung in Abhängigkeit von der Studienmedikation vorliegen, wird das Verzerrungspotential für den Endpunkt Mortalität als „niedrig“ eingestuft. Im Gegensatz dazu wird ein offenes Studiendesign für die Endpunkte PFS und Lebensqualität kritisch gesehen: wie in Tabelle 7 beschrieben, ist unklar, wann eine klinische Indikation vorlag und wann optionale diagnostische Messungen durchgeführt wurden. Definitionen hierzu wurden vom pU nicht vorgenommen. Zudem ist der Einfluss durch das Wissen des Untersuchers um die Zuteilung der Patienten auf die Diagnostik des PFS unklar, weshalb das Verzerrungspotential als hoch angesehen wird.

Bezüglich der Lebensqualität besteht die Möglichkeit einer Verzerrung auf Betroffenen-Ebene, da die Lebensqualität durch das Wissen um die Therapie bewusst oder unbewusst verzerrt wahrgenommen und angegeben werden könnte. Aus diesem Grund wird das Verzerrungspotential auch hier als „hoch“ eingestuft.

2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz

Laut SGB V § 35a Absatz 1 Satz 10 gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 sind nicht verpflichtend vorzulegen. Zur Beurteilung der Patientenrelevanz der vorgelegten Endpunkte wurden das Dossier, der zugehörige

Studienbericht und die in diesen Quellen zitierte Literatur herangezogen. Das Ergebnis der Betrachtung der Endpunkte findet sich aufgeteilt nach Endpunktkategorien in den Tabellen 6 bis 9.

Tabelle 6: Beschreibung der Endpunkte der Kategorie Mortalität

Mortalität	
Endpunkt	Gesamtüberleben
Operationalisierung	<p>Operationalisierung: Gesamtüberleben stellt die Zeit in Monaten ab Randomisierung bis Tod jeglicher Ursache dar. Ein Monat ist definiert als 30,4 Tage. Patienten, die lost to follow-up oder zum Zeitpunkt des Datenschnitts nicht verstorben waren, wurden rechts zensiert. Das Zensierungsdatum war dabei der Zeitpunkt des letzten Kontakts.</p> <p>Patienten, die die Studienmedikation abbrachen, wurden alle 28 Tage bzgl. Krankheitsprogression nachbeobachtet. Patienten, die bereits einen Krankheitsprogress erlitten hatten, wurden alle drei Monate auf Vitalstatus nachbeobachtet.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunkts ist nachvollziehbar beschrieben.</p> <p>Patientenrelevanz: Entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, <u>der Verlängerung des Überlebens</u>, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“</p> <p>Die Ergebnisse zum Endpunkt „Gesamtüberleben“ werden in der Nutzenbewertung von Carfilzomib berücksichtigt.</p>

Tabelle 7: Beschreibung der Endpunkte der Kategorie Morbidität

Morbidität	
Endpunkt	Progressionsfreies Überleben
Operationalisierung	<p>Operationalisierung: PFS stellt den Zeitraum ab Randomisierung bis zur Krankheitsprogression oder Tod jeglicher Ursache dar. Krankheitsprogression wurde anhand der IMWG-Kriterien [13,25] definiert, wobei mindestens ein Kriterium erfüllt sein musste:</p> <ul style="list-style-type: none"> • Anstieg um $\geq 25\%$ gegenüber dem niedrigsten gemessenen Wert <ul style="list-style-type: none"> ○ der M-Proteinkonzentration im Serum (absoluter Anstieg $\geq 0,5$ g/dl) und/oder ○ der M-Proteinkonzentration im Urin (absoluter Anstieg ≥ 200 mg/24 h) und/oder ○ nur bei Patienten ohne messbares M-Protein: des Unterschieds zwischen beteiligten und unbeteiligten FLC-Werten (absoluter Anstieg > 10 mg/dl) ○ Plasmazellen im Knochenmark (absoluter Anstieg $\geq 10\%$) • Entstehung neuer oder Vergrößerung bestehender Knochenläsionen oder Weichteilplasmozytome • Entwicklung einer Hyperkalzämie (korrigierter Kalziumspiegel im Serum $> 11,5$ mg/dl oder $2,65$ mmol/l), die ausschließlich der Störung in der Plasmazellproliferation zugeschrieben werden kann <p>Messmethoden und Messzeitpunkte:</p> <ul style="list-style-type: none"> • M-Proteinbestimmung: Elektrophorese zu Baseline und am Tag 28 der jeweils durchlaufenden Zyklen im Verumarm bzw. alle 4 Wochen im Kontrollarm

Morbidität	
Endpunkt	Progressionsfreies Überleben
	<ul style="list-style-type: none"> • FLC: Bestimmung im Serum zu Baseline und am Tag 28 der jeweils durchlaufenden Zyklen im Verumarm bzw. alle 4 Wochen im Kontrollarm • Plasmazellen im Knochenmark: Bestimmung mittels FISH zu Baseline und optional am Tag 28 der jeweils durchlaufenden Zyklen im Verumarm bzw. alle 4 Wochen im Kontrollarm • Plasmozytome: bidirektionale Messung zu Baseline und optional am Tag 28 der jeweils durchlaufenden Zyklen im Verumarm bzw. alle 4 Wochen im Kontrollarm • Knochenläsionen: Röntgenaufnahme des Schädels, anteroposteriore und laterale Ansichten des Beckens, der Rippen, der Femora und Humeri zu Baseline und bei Verschlimmerung der Krankheitssymptome, die auf Krankheitsprogress hindeuten oder wenn klinisch indiziert. <p>Zur Klassifizierung eines PFS-Ereignisses mussten zwei Messungen durch das Zentrallabor vorliegen. Zudem erfolgte die Einschätzung zum Vorliegen eines PFS durch ein unabhängiges Reviewkomitee.</p> <p>Nach Abbruch der Studienmedikation bevor ein Krankheitsprogress diagnostiziert wurde, trat der Patient in den Langzeit-Follow-up ein. Hierbei wurde alle 28 Tage eine Bewertung zum Vorliegen einer Krankheitsprogression vorgenommen.</p>
Bewertung	<p>Operationalisierung: Es handelt sich um einen kombinierten Endpunkt, der sich aus Mortalität und verschiedenen Endpunkten der Kategorie Morbidität zusammensetzt. Die Operationalisierung ist nicht nachvollziehbar beschrieben. Es bleibt unklar, in welchem konkreten Fall die Messung der Plasmazellen im Knochenmark und die Bestimmung der Plasmozytome im Verlauf der Studie vorgenommen wurden. Diese Messungen waren optional. Die Bestimmung der Knochenläsionen erfolgte bei Verschlimmerung der Symptome oder bei klinischer Indikation. Diese Kriterien sind nicht nachvollziehbar definiert und operationalisiert.</p> <p>Validität: Die Endpunkte der Kategorie Morbidität (M-Protein, Plasmazellzahl, Kalziumspiegel, Knochenläsionen, Plasmozytome) werden entweder biochemisch über Labormessungen oder über bildgebende Verfahren bestimmt. Der pU führt im Modul 4 aus, dass eine Krankheitsprogression mit einer verminderten Lebensqualität und einer Verkürzung der Lebenserwartung assoziiert sei. Die vom pU angeführten Quellen [16,20,21] zeigen keine kausale Verbindung, so dass dieser Argumentation nicht gefolgt werden kann.</p> <p>Darüber hinaus führt der pU aus, dass Knochenläsionen und Weichteilplasmozytome mit patientenrelevanten Endpunkten wie Schmerzen, Frakturgefahr einhergehen können. Außerdem bestünde, laut Modul 4, ein Zusammenhang zwischen der Erhöhung der Plasmazellen im Knochenmark und Symptomen wie Fatigue, Atemnot, Herzrasen und schweren Infektionen. M-Proteine könnten zudem zur Niereninsuffizienz sowie -versagen und Dialysepflicht führen. Dieser Argumentation kann nicht gefolgt werden, da keine vom pU dargelegte Quelle [7,10,26,28,30] einen kausalen Zusammenhang zwischen den gemessenen Laborparametern bzw. der Bildgebung und dem Auftreten von patientenrelevanten Endpunkten belegen kann.</p> <p>Der pU kann nicht überzeugend darlegen, dass mit einem asymptomatischen Rezidiv oder auch serologischen Progress zwingend ein symptomatischer klinischer Progress zusammenhängt. Patienten erleiden in dieser Phase nicht zwangsläufig Symptomverschlimmerungen, so dass die o. g. Endpunkte (M-Protein, Plasmazellzahl, Kalziumspiegel, Knochenläsionen, Plasmozytome) nicht als valide Surrogatparameter</p>

Morbidität	
Endpunkt	Progressionsfreies Überleben
	<p>für Morbidität angesehen werden.</p> <p>Patientenrelevanz: Die Komponente Mortalität des kombinierten Endpunkts PFS ist patientenrelevant (siehe Tabelle 6). Die Endpunkte der Kategorie Morbidität zur Bestimmung einer Krankheitsprogression (M-Protein, Plasmazellzahl, Kalziumspiegel, Knochenläsionen, Plasmozytome), entsprechend o. g. Operationalisierung, werden nicht als patientenrelevant angesehen. Da PFS der primäre Endpunkt der Studie ENDEAVOR ist, werden Ergebnisse im Abschnitt 2.6.3 ergänzend dargestellt, aber in der Nutzenbewertung von Carfilzomib nicht berücksichtigt.</p>

FLC: freie Leichtketten

Tabelle 8: Beschreibung der Endpunkte der Kategorie Lebensqualität

Lebensqualität	
Endpunkt	Gesundheitsbezogene Lebensqualität (EORTC QLQ-C30, EORTC QLQ-MY20, FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“)
Operationalisierung	<p>Operationalisierung: Folgende Instrumente wurden zur Erhebung der Lebensqualität eingesetzt: EORTC QLQ-C30, EORTC QLQ-MY20, FACT-GOG-Ntx. Die Instrumente sollten vor Verabreichung der Studienmedikation ausgefüllt werden.</p> <ul style="list-style-type: none"> • Der QLQ-C30-(Quality of Life-Core 30-Item)-Fragebogen der European Organisation for Research and Treatment of Cancer (EORTC) ist ein generisches Instrument zur Erfassung der gesundheitsbezogenen Lebensqualität bei Krebspatienten. Der Fragebogen gliedert sich in 5 Funktionsskalen (Physis, Rolle, Emotion, Kognition, soziales Leben), 3 Symptomskalen (Fatigue, Schmerz, Übelkeit und Erbrechen) und eine globale Skala zu Gesundheitszustand/ Lebensqualität auf. Darüber hinaus können einzelne Items zu Symptomen (bspw. Dyspnoe, Appetitverlust, Schlafstörungen) oder finanziellen Schwierigkeiten separat betrachtet werden. Die Fragen werden auf 4- bzw. 7-Punkte-Skalen beantwortet. Der Gesamtscore wird auf eine Skala von 0 bis 100 Punkten transformiert. Höhere Werte auf den Funktionsskalen und der globalen Gesundheits-/Lebensqualitätsskala bedeuten bessere Funktion bzw. Gesundheit/Lebensqualität. Höhere Werte auf den Symptomskalen bzw. den einzelnen Symptomitems bedeuten schwerere Symptomatik. • Der QLQ-MY20, ebenfalls entwickelt von der EORTC, ist ein Modul zur Erfassung der Lebensqualität bei Menschen mit Multiplem Myelom und wird gemeinsam mit dem QLQ-C30 angewendet. Dieser Fragebogen umfasst 20 Fragen, die auf einer 4-Punkte-Skala von den Betroffenen selbst beantwortet werden. Insgesamt umfasst der QLQ-MY20 vier Subskalen: Krankheitssymptome, Nebenwirkungen der Behandlung, Körperwahrnehmung und Zukunftsperspektive, die separat betrachtet werden. Höhere Werte auf den Symptom- und Nebenwirkungsskalen bedeuten schlechtere Symptomatik, höhere Werte auf den Skalen der Körperwahrnehmung und Zukunftsperspektive bedeuten bessere Funktion und Perspektive. • Die FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“ dient zur Erfassung spezifischer Belastungen durch eine antineoplastische Therapie. FACT/GOG-Ntx besteht aus 11 Items, welche Neurotoxizitäten im Zusammenhang mit der Therapie erfassen, aber auch andere Beschwerden wie allgemeine Abgeschlagenheit,

Dossierbewertung für Orphan Drugs

Carfilzomib (Kyprolis®)

Lebensqualität	
Endpunkt	Gesundheitsbezogene Lebensqualität (EORTC QLQ-C30, EORTC QLQ-MY20, FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“)
	<p>Hörprobleme/Tinnitus oder Gehprobleme erfragen. Der Score kann Werte zwischen 0 und 44 annehmen. Höhere Werte beschreiben eine höhere Lebensqualität.</p> <p>Der Bezugszeitraum umfasst für alle 3 Instrumente die vergangene Woche.</p> <p>Zeitpunkte:</p> <p>Alle Instrumente wurden zu Studienbeginn (Tag 1, Zyklus 1) und anschließend alle 28 Tage bis zum Auftreten einer Krankheitsprogression, Studienabbruch oder Beginn einer Folgetherapie eingesetzt. Patienten, die vor Erleiden einer Krankheitsprogression die Studienmedikation abbrechen, sollten weiterhin alle 4 Wochen bis zum Auftreten eines der o. g. Kriterien Angaben zur Lebensqualität machen.</p>
Bewertung	<p>Operationalisierung:</p> <p>Die Operationalisierung ist nachvollziehbar beschrieben.</p> <p>Validität:</p> <ul style="list-style-type: none"> • EORTC QLQ-C30: Den Ausführungen des pU zur Validität und Reliabilität des QLQ-C30 kann gefolgt werden [23]. Alle Subskalen des QLQ-C30, bis auf Subskala „Rolle“, zeichnen sich durch eine interne Konsistenz $\geq 0,7$ aus [1,18,31], die sich zwischen Personen jüngeren und höheren Alters, mit geringer und hoher Bildung und leichtem und schweren Erkrankungsstatus kaum unterscheidet. Die Konstruktvalidität des QLQ-C30 im Vergleich zum WHO-Performance-Index konnte gezeigt werden [31]. Zudem ist der QLQ-C30 veränderungssensitiv, was bei Patienten mit Multiplem Myelom, die unterschiedlich auf Therapien ansprechen, demonstriert werden konnte [18,31]. Angaben zu klinisch relevanten Responder-schwellen des QLQ-C30 variieren zwischen 8 Punkten für Verbesserung und 12 Punkten für Verschlechterung [18] bei Personen mit Multiplem Myelom und 5–10 Punkten bei Patienten mit Mamma- oder Lungenkarzinom [24]. Es kann davon ausgegangen werden, dass eine Verbesserung des Scores „Gesundheitszustand/Lebensqualität“ um 10 Punkte (Responderanalyse) eine klinisch relevante Veränderung in der jeweiligen Gruppe bedeutet. Angaben, wann Unterschiede zwischen Gruppen klinische Relevanz erreichen (minimal important difference, MID), liegen nicht vor. • QLQ-MY20: Den Ausführungen des pU zur Validität und Reliabilität des QLQ-MY20 kann gefolgt werden [23]. Die Entwicklung des krankheitsspezifischen Moduls QLQ-MY20 der EORTC ist dokumentiert [29]. Die Validität und Reliabilität wurde an 240 Personen mit Multiplem Myelom aus 7 Ländern, die an unterschiedlichsten klinischen Studien teilnahmen, untersucht [9]. Dabei zeigte sich eine hohe interne Konsistenz der vier Subskalen mit einem Cronbachs alpha $> 0,70$. Mindestens drei der vier Subskalen (Krankheitssymptome, Nebenwirkungen, Körperbild) zeigen eine gute Diskriminierung zwischen Patienten unterschiedlicher Krankheitszustände. Daneben zeigten die Subskalen Krankheitssymptome, Nebenwirkungen und Zukunftsperspektive statistisch signifikante Unterschiede in den Scores zwischen Patienten mit und ohne Frakturen. Zudem zeigte sich bei Personen, die auf die Therapie ansprechen, in den Subskalen Krankheitssymptome, Nebenwirkungen und Körperbild statistisch signifikante Unterschiede zwischen Baseline und Follow-up-Scores, sodass hier veränderungssensitive Skalen vorliegen [9]. <p>Bisher gibt es keine Untersuchungen zu MID oder Schwellenwerten für Responderanalysen. Osoba et al. [24] schlagen basierend auf empirischen Daten von Personen mit Mamma- und Bronchialkarzinom eine Responder-schwelle von</p>

Lebensqualität	
Endpunkt	Gesundheitsbezogene Lebensqualität (EORTC QLQ-C30, EORTC QLQ-MY20, FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“)
	<p>5 Punkten für den QLQ-C30 vor, die in einer anderen Arbeit für den QLQ-MY20 übernommen wurde [12].</p> <ul style="list-style-type: none"> • FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“: Den Ausführungen des pU zur Validität und Reliabilität der FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“ kann gefolgt werden [23]. Die psychometrischen Eigenschaften der FACT/GOG-Ntx-Skala „zusätzliche Beschwerden“ wurden in Patientinnen mit Ovarialkarzinomen untersucht [8]. Untersuchungen zur Kriteriums- und Veränderungssensitivität liegen vor und es konnte gezeigt werden, dass die Skala des FACT/GOG-Ntx mit dem Ansprechen auf eine Therapie mit Bortezomib korrespondiert [12]. Es liegen keine validierten Responderschwelen oder Untersuchungen zur MID zwischen Gruppen vor. <p>Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer <u>Verbesserung der Lebensqualität</u>.“ Die Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität werden in die Bewertung des Zusatznutzens von Carfilzomib eingeschlossen.</p>

Tabelle 9: Beschreibung der Endpunkte der Kategorie Sicherheit

Sicherheit	
Endpunkt	Periphere Neuropathie
Operationalisierung	<p>Operationalisierung: Periphere Neuropathien wurden mittels neurologischer ärztlicher Untersuchung diagnostiziert: Vibrationsmessung mit 128-Hertz-Stimmgabel, Sensibilitätsprüfung, Prüfung des Achillessehnenreflexes, Beurteilung der Muskelkraft, Bestimmung des Schweregrades der Neuropathie. Neuropathien wurden gemäß MedDRA Version 15.1 kodiert und der Schweregrad entsprechend der NCI-CTCAE-Kriterien Version 4.03 eingeschätzt. Es wurde der SMQN „Periphere Neuropathie“ zur Definition herangezogen. Es wurden Ereignisse ab Studienbeginn bis 30 Tage nach Ende der Einnahme der Studienmedikation bzw. bis Beginn einer nächsten Therapie gewertet¹⁾. Patienten wurden zu Screening, am Tag 1 des Zyklus 1 und vor Beginn eines jeden weiteren Therapiezyklus auf Vorliegen einer peripheren Neuropathie untersucht.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung ist nachvollziehbar beschrieben. Allerdings fehlen in den Studiendokumenten (Studienprotokoll, Studienbericht und SAP) Angaben zum Erhebungszeitraum, welche lediglich aus Modul 4 des Herstellerdossiers verfügbar waren.</p> <p>Validität: Das Auftreten von Neuropathien wird als patientenrelevanter und bewertungsrelevanter Endpunkt angesehen. Neuropathien gehen mit peripheren motorischen und sensorischen Störungen einher und können sogar lebensbedrohliche Ausmaße</p>

Sicherheit	
	<p>erreichen. Dieser Endpunkt wird daher zur Bewertung der Sicherheit von Carfilzomib herangezogen.</p> <p>Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der <u>Verringerung von Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“ Die Ergebnisse zum Endpunkt „periphere Neuropathie“ werden in die Bewertung des Zusatznutzens von Carfilzomib eingeschlossen.</p>
Endpunkt	Unerwünschte Ereignisse
Operationalisierung	<p>Operationalisierung: Ein unerwünschtes Ereignis ist jedes medizinische Ereignis, das ein Patient, der eine Form von Studienmedikation verabreicht bekommt, erfährt. Darunter zählen neu aufgetretene Ereignisse, Verschlechterungen in Schwere oder Häufigkeit eines Zustandes und Labormessungen. UE wurden gemäß MedDRA Version 15.1 erfasst und berichtet. Die Schwere der UE wurde nach NCI-CTCAE Version 4.03 eingeschätzt. Alle UE mussten fortlaufend, nachdem der Patient die Einverständniserklärung zur Studienteilnahme unterzeichnet hatte, über elektronische Fallberichtsbögen berichtet werden. Es werden alle UE bis 30 Tage nach Beendigung der Studienmedikation dokumentiert. Zusätzlich sollte der Untersucher jegliche UE, die im Zusammenhang mit der Studienmedikation stehen könnten, über den Zeitraum von 30 Tagen nach letzter Einnahme hinaus berichten.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung ist nachvollziehbar beschrieben.</p> <p>Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der <u>Verringerung von Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“ Die Ergebnisse zum Endpunkt „unerwünschte Ereignisse“ werden in die Bewertung des Zusatznutzens von Carfilzomib eingeschlossen.</p>

¹⁾ Information zum Erhebungszeitraum aus Modul 4 des Herstellerdossiers. Entsprechende Informationen waren nicht im Studienbericht, -protokoll oder SAP verfügbar.

Abkürzungen: SMQN: Standardized MedDRA Query narrow scope

2.5.4 Statistische Methoden

In diesem Abschnitt werden die geplanten statistischen Methoden zur Auswertung der Studie ENDEAVOR beschrieben. Die erste Version des SAP wurde am 9. August 2012 kurz nach Aufnahme der Studie ENDEAVOR verfasst. Die zweite Version wurde nach dem Datenschnitt zur Interimsanalyse am 19. Dezember 2014 verfasst. Zur Analyse der Lebensqualitätsinstrumente wurde ein separater SAP erstellt, der am 20. Januar 2015 fertiggestellt wurde. Alle Analysen, die nach dem Datenschnitt zur Interimsanalyse geplant wurden, werden als post hoc angesehen.

Statistische Modelle und Tests:

Die Studie ENDEAVOR ist eine Überlegenheitsstudie, in der die Wirksamkeit und Sicherheit von Carfilzomib und Dexamethason im Vergleich zu Bortezomib und Dexamethason bei vorbehandelten Patienten mit rezidiviertem und/oder progredientem Multiplen Myelom untersucht wird.

Testhypothesen wurden nicht formuliert. Den Ausführungen im SAP ist zu entnehmen, dass es sich um eine Überlegenheitsstudie handelt. Die inferentiellen Tests des primären und sekundären Endpunkts wurden einseitig mit einem Signifikanzniveau von $\alpha=0,025$ durchgeführt. Das Signifikanzniveau der Interimsanalysen wurde nach O'Brien-Fleming mit einer Alpha-Funktion nach Lan und DeMets [19] bestimmt.

Die erste Interimsanalyse wurde bei Erreichen von 75 % der PFS-Ereignisse (n=395) geplant. Zeigte sich bei der Interimsanalyse zu PFS ein statistisch signifikantes Ergebnis, so sollten weitere Analysen zu den sekundären Endpunkten durchgeführt werden. Zum OS sollen insgesamt 3 Analysen stattfinden (SAP, Version 2). Die erste Analyse erfolgt bei Erreichen der geplanten Ereigniszahl zu PFS, die zweite Analyse soll bei Erreichen von 394 OS-Ereignissen und die dritte Analyse bei Erreichen von 496 OS-Ereignissen stattfinden.

PFS und OS wurden graphisch mittels Kaplan-Meier-Methode analysiert. Der Log-Rank-Test (stratifiziert nach Randomisierungsfaktoren: Vortherapie mit Proteasominhibitor, Anzahl Vortherapien, ISS-Stadium, Bortezomibapplikation) wurde zum Test auf statistisch signifikante Unterschiede in den Überlebenszeitkurven zwischen Verum- und Kontrollarm herangezogen.

Das Hazard Ratio und 95 %-Konfidenzintervall wurde mittels Cox-Proportional-Hazards-Regression berechnet, wobei auch hier eine Stratifizierung des Basalinerisikos nach o. g. Randomisierungsfaktoren zugelassen wurde.

Zur Analyse der Daten zur gesundheitsbezogenen Lebensqualität wurden gemischte Modelle für wiederholte Messungen (MMRM) herangezogen. Dabei wurden alle Messungen der Lebensqualität berücksichtigt. Außerdem wurde die Annahme getroffen, dass fehlende Werte zufällig waren („Missing at Random“). Fehlten bei einem Teilnehmer weniger als die Hälfte der Items zu einem Zeitpunkt, so wurden diese basierend auf den gemessenen Items zu diesem Zeitpunkt imputiert. Fehlten mehr als die Hälfte der Items oder lag vollständige Non-Response zu bestimmten Messzeitpunkten vor, so wurde nicht imputiert. Randomisierungsfaktoren und die Zugehörigkeit zum Behandlungsarm wurden als fixed effects in die Modelle aufgenommen.

Responderanalysen zum EORTC QLQ-C30-Score „Gesundheitszustand/Lebensqualität“ werden basierend auf der Verbesserung um mindestens 5 Punkte durchgeführt. Der Cochran-Mantel-Haenszel-Test, stratifiziert nach Randomisierungsfaktoren, wird zum Test auf Unterschiede zwischen Verum- und Kontrollgruppe herangezogen. Diese Analyse wird ebenfalls für einen Schwellenwert von 10 Punkten wiederholt. Zudem wurden mit den genannten Schwellenwerten Überlebenszeitanalysen durchgeführt. Alle statistischen Analysen zur gesundheitsbezogenen Lebensqualität wurden nach dem Datenschnitt zur Interimsanalyse geplant und werden daher als Post-hoc-Analysen angesehen.

Periphere Neuropathie: Das Auftreten peripherer Neuropathien war einer der sekundären Endpunkte und sollte nach der gruppensequenziellen Prozedur nach Holm [32] gemeinsam mit OS und ORR nach Anpassung des Signifikanzniveaus untersucht werden. Es wurden die Durchführung eines nicht-stratifizierten Mantel-Haenszel-Chi²-Tests und die Berechnung von

Odds Ratios und dazugehörige 95 %-Konfidenzintervalle geplant. Zudem wurden post hoc auch Risk Ratios für die CTC-Schweregrade ≥ 2 und ≥ 3 berechnet. Die Definition des Endpunkts „periphere Neuropathie“ wurde nach Datenschnitt für die hier vorgelegte Analyse verändert (vgl. Ausführungen zu Protokollamendements Abschnitt 2.4).

Fallzahlplanung:

Die Fallzahlplanung der Studie ENDEAVOR basierte auf dem primären Endpunkt PFS und legte folgende Annahmen zugrunde:

- Power: 90 %
- Signifikanzniveau (α -Fehler): 2,5 %
- 1:1 Randomisierung
- Hazard Ratio (Cd vs. Vd): 0,75
- Mittlere progressionsfreie Zeit: 10 Monate im Kontrollarm (Vd)

Der Stichprobenumfang betrug 888 Patienten, die in die Studie aufgenommen werden mussten, um bei Vorliegen von 526 PFS-Ereignissen die präspezifizierten Annahmen zu erfüllen. Eine Interimsanalyse sollte bei Vorliegen von 75 % (n=395) der geplanten PFS-Ereignisse erfolgen.

Für den sekundären Endpunkt OS bedeutete die Fallzahlplanung, dass bei Erreichen der geplanten Ereignisse (n=496 Todesfälle) die Studie eine Power von 69 % hat, um eine Reduktion des Hazard Ratios von 20 % zu zeigen. Basierend auf der Zwischenauswertung mit dem Datenschnitt November 2014 wird die geplante Anzahl OS-Ereignisse 2018 erreicht werden. Ursprünglich waren 631 Todesfälle für die finale Analyse geplant. Die Ereigniszahl wurde aufgrund der Interimsanalyse angepasst.

Interimsanalyse

Die vorliegenden Ergebnisse basieren auf einer Interimsanalyse mit dem Datenschnitt 11. November 2014, die vorab bei Erreichen von ca. 75 % der geplanten PFS-Ereignisse geplant wurde. Bei Ablehnung der Nullhypothese des primären Endpunkts sollten weitere statistische Tests der sekundären Endpunkte vorgenommen werden.

Studienpopulationen:

Folgende Studienpopulationen wurden definiert:

- Intention-to-Treat (ITT): Diese Population umfasst alle randomisierten Patienten entsprechend ihrer Zuteilung nach Verum- und Kontrollarm (n=464 bzw. n=465). Diese Population wurde zur Untersuchung der Wirksamkeitsendpunkte herangezogen.
- Sicherheitspopulation: Diese Population umfasst alle randomisierten Patienten, die mindestens eine Dosis einer Studienmedikation erhalten haben (Carfilzomib, Bortezomib oder Dexamethason). Die Patienten werden entsprechend der verabreichten und nicht der randomisierten Medikation analysiert. Dabei wurden n=463 im Verum- und n=456 Patienten im Kontrollarm berücksichtigt. Die Neuropathieereignisse und Sicherheitsendpunkte werden in dieser Studienpopulation untersucht.

Sensitivitätsanalysen

Zum primären Endpunkt PFS wurden folgende Sensitivitätsanalysen durchgeführt:

- PFS beurteilt durch den Prüfarzt
- PFS beurteilt durch einen Computeralgorithmus (ORCA)
- Ohne Berücksichtigung der Stratifizierungsvariablen
- Beginn einer anderen Krebstherapie gewertet als PFS-Ereignis
- Beginn einer anderen Krebstherapie weder als PFS-Ereignis noch Zensierung gewertet
- Fehlende Erhebungen werden im Vd-Arm als Zensierung und im Cd-Arm als PFS-Ereignis gewertet

Die Ergebnisse waren vergleichbar mit denen der Primäranalyse und werden im Folgenden nicht näher betrachtet.

Für die Erhebung der gesundheitsbezogenen Lebensqualität mittels EORTC QLQ-C30 wurden ebenfalls zwei a priori definierte Sensitivitätsanalysen durchgeführt. Dabei wurde einerseits zur Überprüfung der Robustheit der Ergebnisse für fehlende Werte eine Indikatorvariable ins Modell aufgenommen (Muster der fehlenden Werte). Andererseits wurden nur Messzeitpunkte berücksichtigt, zu denen Informationen von mind. 40 % der randomisierten Patienten vorlagen. Die Ergebnisse ähneln denen der Primäranalyse.

Subgruppenanalysen

Subgruppenanalysen waren vorab geplant (siehe Tabelle 1).

2.6 Ergebnisse zum Zusatznutzen

2.6.1 Charakterisierung der Studienpopulation

Insgesamt wurden 1096 Personen zur Aufnahme in die Studie ENDEAVOR gescreent, wovon 929 Patienten den beiden Studienarmen randomisiert zugeteilt wurden. Zum Zeitpunkt des Datenschnitts nahmen noch ca. 43 % der Patienten im Verumarm und ca. 22 % im Kontrollarm die Studienmedikation ein. Der andere Teil der Studienpopulation hatte bereits die Einnahme der Studienmedikation abgebrochen. Hauptgrund stellte Krankheitsprogression dar, gefolgt vom Auftreten unerwünschter Ereignisse und Entscheidung der Patienten.

Die mediane Behandlungsdauer unterschied sich zwischen den Armen und betrug im Verumarm 39,9 Wochen und im Kontrollarm 26,8 Wochen. Die mediane Anzahl begonnener Therapiezyklen betrug 10 im Verum- und 8 im Kontrollarm. Die Beobachtungszeit war im Median in beiden Studienarmen ähnlich lang. Fast alle Studienteilnehmer nahmen mindestens eine Begleitmedikation ein. Dabei stellten Protonenpumpenhemmer und Viruzide die Medikamentengruppen dar, die am häufigsten eingenommen wurden. Relevante Unterschiede in der Begleitmedikation zwischen den Studienarmen wurden nicht berichtet.

Tabelle 10: Allgemeine Angaben zur Studie ENDEAVOR, Datenschnitt 10.11.2014

ENDEAVOR	Cd (n=464)	Vd (n=465)
Randomisiert, ITT-Population, n (%)	464 (100)	465 (100)
Randomisiert, aber keine Einnahme der Studienmedikation, n (%)	1 (0,2)	9 (1,9)
Derzeitig Einnahme der Studienmedikation, n (%)	200 (43,1)	105 (22,6)
Abbruch der Einnahme der Studienmedikation, n (%)	263 (56,7)	351 (75,5)
Krankheitsprogression, n (%)	117 (25,2)	168 (36,1)
aufgrund von UE, n (%)	65 (14,0)	73 (15,7)
Patientenentscheidung, n (%)	40 (8,6)	45 (9,7)
Arztentscheidung, n (%)	18 (3,9)	35 (7,5)
Entzug der Einverständniserklärung, n (%)	6 (1,3)	19 (4,1)
Tod, n (%)	13 (2,8)	9 (1,9)
Protokollverletzung, n (%)	4 (0,9)	1 (0,2)
Lost to follow-up, n (%)	0	1 (0,2)
Patienten im LTFU ¹⁾ , n (%)	218 (47,0)	298 (64,1)
Krankheitsprogression, n (%)	47 (10,1)	71 (15,3)
Mediane Behandlungsdauer in Wochen (min; max) ²⁾	39,9 (1,0; 108,1)	26,8 (1,0; 106,1)
Mediane Anzahl begonnener Therapiezyklen (min; max) ²⁾	10,0 (1,0; 26,0)	8,0 (1,0; 35,0)
Mediane Beobachtungsdauer in Monaten (95%-KI) ³⁾	11,9 (11,2; 12,4)	11,1 (10,2; 11,4)
Patienten mit mindestens einer Begleitmedikation, n (%)	462 (99,6)	456 (98,1)

¹⁾ Patienten, die vor Eintritt einer Krankheitsprogression die Einnahme der Studienmedikation abbrechen, treten in die Langzeitnachbeobachtung ein.

²⁾ Angaben basieren auf Sicherheitspopulation

³⁾ Die mediane Beobachtungsdauer variiert je nach zeitabhängig gemessenem Endpunkt; dargestellt ist hier die Beobachtungsdauer des primären Endpunkts (PFS).

Abkürzungen: Cd: Carfilzomib + Dexamethason; Vd: Bortezomib + Dexamethason; LTFU: Langzeit-Follow-up

Der Patientenfluss der Studie ENDEAVOR ist zusätzlich in Abbildung 2 dargestellt.

Angaben zur Demographie, krankheitsbezogene Daten und Vortherapie der Teilnehmer der Studie ENDEAVOR sind in Tabelle 11 dargestellt. Die Baselinecharakteristika scheinen gleich über beide Studienarme verteilt zu sein. Leichte Imbalancen traten bei Neuropathien in der Vergangenheit (46,3 % im Cd-Arm vs. 52,5 % im Vd-Arm) und in der Vortherapie mit Thalidomid auf (45,5 % im Cd-Arm vs. 53,1 % im Vd-Arm).

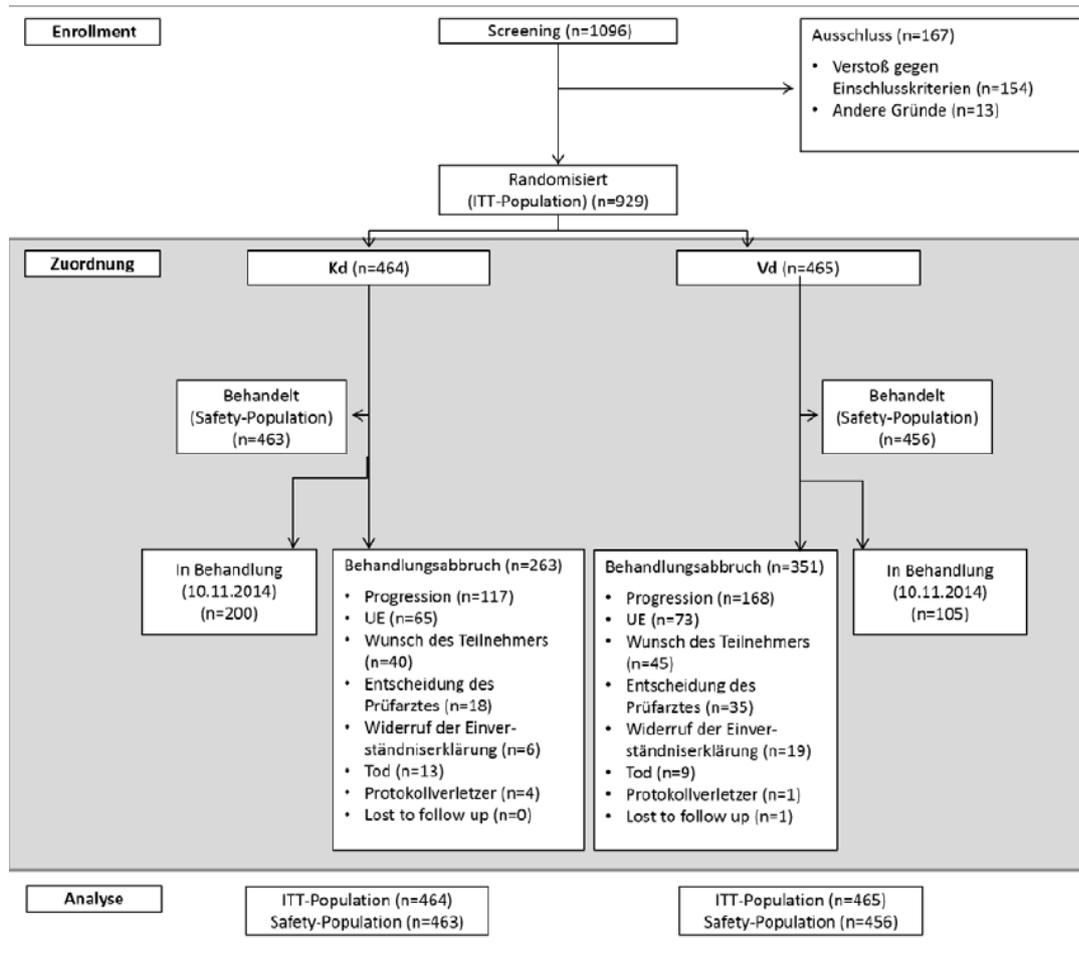


Abbildung 2: Flussdiagramm zur Studie ENDEAVOR

Tabelle 11: Charakterisierung der Studienpopulation der Studie ENDEAVOR zu Baseline

Studie ENDEAVOR	Cd N=464	Vd N=465
Alter (Jahre)		
Median (min; max)	65 (35,0; 89,0)	65 (30,0; 88,0)
Geschlecht, n (%)		
männlich	240 (51,7)	229 (49,2)
weiblich	224 (48,3)	236 (50,8)
Ethnie, n (%)		
lateinamerikanisch	19 (4,1)	21 (4,5)
nicht-lateinamerikanisch	379 (81,7)	385 (82,8)
keine Angabe	66 (14,2)	59 (12,7)

Studie ENDEAVOR	Cd N=464	Vd N=465
<i>Ethnie (genetisch), n (%)</i>		
weiß	348 (75,0)	353 (75,9)
schwarz	8 (1,7)	9 (1,9)
asiatisch	56 (12,1)	57 (12,3)
Hawaiianer/andere pazifische Inselbewohner	2 (0,4)	0
keine Angabe	50 (10,8)	45 (9,7)
mehrere	0 (0)	1 (0,2)
<i>Region, n (%)</i>		
Osteuropa	135 (29,1)	121 (26,0)
Westeuropa	182 (39,2)	169 (36,3)
Nordamerika	35 (7,5)	49 (10,5)
Südamerika	10 (2,2)	15 (3,2)
Asien/Pazifik	102 (22,0)	111 (23,9)
<i>Körperoberfläche (m²); Median (min; max)¹⁾</i>	1,8 (1,2; 3,0)	1,8 (1,2; 2,8)
<i>ECOG-Status</i>		
0	221 (47,6)	232 (49,9)
1	211 (45,5)	203 (43,7)
2	32 (6,9)	30 (6,5)
<i>Kreatinin-Clearance (ml/min); Median (min; max)</i>	73,0 (14; 185)	72 (12; 208)
<i>LVEF²⁾ in %; Median (min; max)</i>	63,0 (40; 83)	64 (40; 84)
<i>Neuropathie in der Vergangenheit, n (%)</i>		
Ja	215 (46,3)	244 (52,5)
Nein	249 (53,7)	221 (47,5)
<i>Neuropathie beim Screening</i>		
Grad 1 (mild)	133 (28,7)	159 (34,2)
Grad 2 (moderat)	10 (2,2)	10 (2,2)
Krankheitsbezogene Daten		
<i>Zeit seit Diagnose bis Randomisierung³⁾ (Monate); Median (min; max)</i>	44,3 (4,0; 246,6)	43,3 (5,4; 306,2)
<i>ISS-Krankheitsstadium, n (%)</i>		
Stadium I	212 (45,7)	205 (44,1)
Stadium II	138 (29,7)	151 (32,5)
Stadium III	114 (24,6)	109 (23,4)
<i>Plasmozytome, n (%)</i>		
Ja	29 (6,3)	24 (5,2)
Nein	435 (93,8)	441 (94,8)
<i>Knochenläsionen, n (%)</i>		
Ja	348 (75)	340 (73,1)
Nein	116 (25)	125 (26,9)
<i>Plasmazellen im Knochenmark (%)</i>		
Median (min; max)	14,0 (0; 100)	16,0 (0; 100)

Studie ENDEAVOR	Cd N=464	Vd N=465
<i>M-Protein heavy chain isotype, n (%)</i>		
IgG	286 (61,6)	284 (61,1)
IgA	90 (19,4)	105 (22,6)
IgD	6 (1,3)	4 (0,9)
Unbekannt	82 (17,7)	72 (15,5)
Vorbehandlung		
<i>Systemische Therapie, n (%)</i>		
Ja	464 (100)	465 (100)
<i>Anzahl Vortherapien, n (%)</i>		
1	232 (50)	232 (49,9)
2	157 (33,8)	145 (31,2)
3	75 (16,2)	87 (18,7)
4	0 (0)	1 (0,2)
<i>Stammzelltherapie, n (%)</i>		
Ja	266 (57,3)	272 (58,5)
Autolog	262 (56,5)	268 (57,6)
Allogen	6 (1,3)	4 (0,9)
Nein	198 (42,7)	193 (41,5)
<i>Vortherapie mit IMiDen, n (%)</i>		
Lenalidomid	177 (38,1)	177 (38,1)
Thalidomid	211 (45,5)	247 (53,1)
<i>Vortherapie mit Proteasominhibitor, n (%)</i>		
Carfilzomib	2 (0,4)	1 (0,2)
Bortezomib	250 (53,9)	252 (54,2)
Weder Carfilzomib noch Bortezomib	212 (45,7)	212 (45,6)
<i>Zeit seit letzter Proteasominhibitortherapie (Monate)</i>		
Median (min; max)	22,7 (4,4; 96,5)	21,0 (6,0; 92,6)
<i>Refraktär gegenüber vorheriger Bortezomibtherapie⁴⁾, n (%)</i>		
Ja	15 (3,2)	19 (4,1)
Nein	449 (96,8)	446 (95,9)

¹⁾ Angaben beruhen auf n=455 im Cd-Arm und n=463 im Vd-Arm

²⁾ Angaben beruhen auf n=457 im Cd-Arm und n=461 im Vd-Arm

³⁾ Angaben beruhen auf n=464 im Cd-Arm und n=464 im Vd-Arm

⁴⁾ Refraktär wurde definiert als Nichterreichen des Minimalansprechens oder besser, Krankheitsprogress während der Therapie oder innerhalb von 60 Tagen nach Beendigung der Therapie.

Abkürzungen: Cd: Carfilzomib und Dexamethason; Vd: Bortezomib und Dexamethason; LVEF: Linksventrikuläre Ejektionsfraktion; ISS: International Staging System; ECOG: Eastern Cooperative Oncology Group

2.6.2 Mortalität

Im Folgenden werden die Ergebnisse zum Endpunkt „Gesamtüberleben“ für die ITT-Population der Studie ENDEAVOR präsentiert (Tabelle 12 und Abbildung 3). Es handelt sich um Ergebnisse der ersten Interimsanalyse, in die 163 der ursprünglich 631 geplanten Ereignisse eingeflossen sind.

Tabelle 12: Ergebnisse zum Gesamtüberleben der Studie ENDEAVOR (Datenschnitt: 10.11.2014)

Studie ENDEAVOR	Cd (N=464)	Vd (N=465)
Tod, n (%)	75 (16,2)	88 (18,9)
Zensierungen, n (%)	389 (83,8)	377 (81,1)
Lebend	377 (81,3)	342 (73,5)
Lost to follow-up	5 (1,1)	5 (1,1)
Entzug der Einverständniserklärung	7 (1,5)	30 (6,5)
Überlebensdauer (Monate), Median (95%-KI) ¹⁾	n.b. (n.b.; n.b.)	24,3 (24,3; n.b.)
HR ¹⁾ (95%-KI)	0,79 (0,58; 1,08)	
p-Wert ²⁾ (1-seitig; $\alpha=0,025$)	0,065	

¹⁾ berechnet mittels Cox-Proportional-Hazards-Regression stratifiziert nach Vortherapie mit Proteasominhibitor, Anzahl Vortherapien, ISS-Stadium, Bortezomibapplikation

²⁾ einseitiger Log-Lank-Test stratifiziert nach zur Randomisierung verwendeten Stratavariablen: Vortherapie mit Proteasominhibitor, Anzahl Vortherapien, ISS-Stadium, Bortezomibapplikation

Abkürzungen: HR: Hazard Ratio; n.b.: nicht berechenbar KI: Konfidenzintervall, Vd: Bortezomib und Dexamethason; Cd: Carfilzomib und Dexamethason

Bis zum Datenschnitt sind im Verumarm 75 (16,5 %) und im Kontrollarm 88 (18,9 %) der Studienteilnehmer verstorben. Ein geringer Teil der Studienpopulation war lost to follow-up. Die Einverständniserklärung wurde von deutlich weniger Patienten im Verumarm (n=7, 1,5 %) als im Kontrollarm (n=30, 6,5 %) zurückgezogen. Die mediane Nachbeobachtungszeit im Verumarm betrug 12,5 und im Kontrollarm 11,9 Monate.

Insgesamt zeigte sich kein statistisch signifikanter Vorteil im Gesamtüberleben in der Gruppe mit Carfilzomib und Dexamethason gegenüber der Gruppe, die mit Bortezomib und Dexamethason therapiert wurde.

Eine zweite Interimsanalyse ist nach Auftreten von 394 Ereignissen geplant. Die finale Analyse soll bei Vorliegen von 496 Ereignissen durchgeführt werden. Im Modul 4 wird eine zusätzliche Analyse, die nicht a priori geplant war, präsentiert. Zum Zeitpunkt des Datenschnitts am 3. März 2016 lagen 322 Ereignisse zum Gesamtüberleben vor. Es zeigte sich kein statistisch signifikanter Vorteil für die Therapie mit Carfilzomib und Dexamethason im Vergleich zu Bortezomib und Dexamethason (HR=0,81; 95%-KI: 0,65; 1,00; einseitiger p-Wert: 0,053).

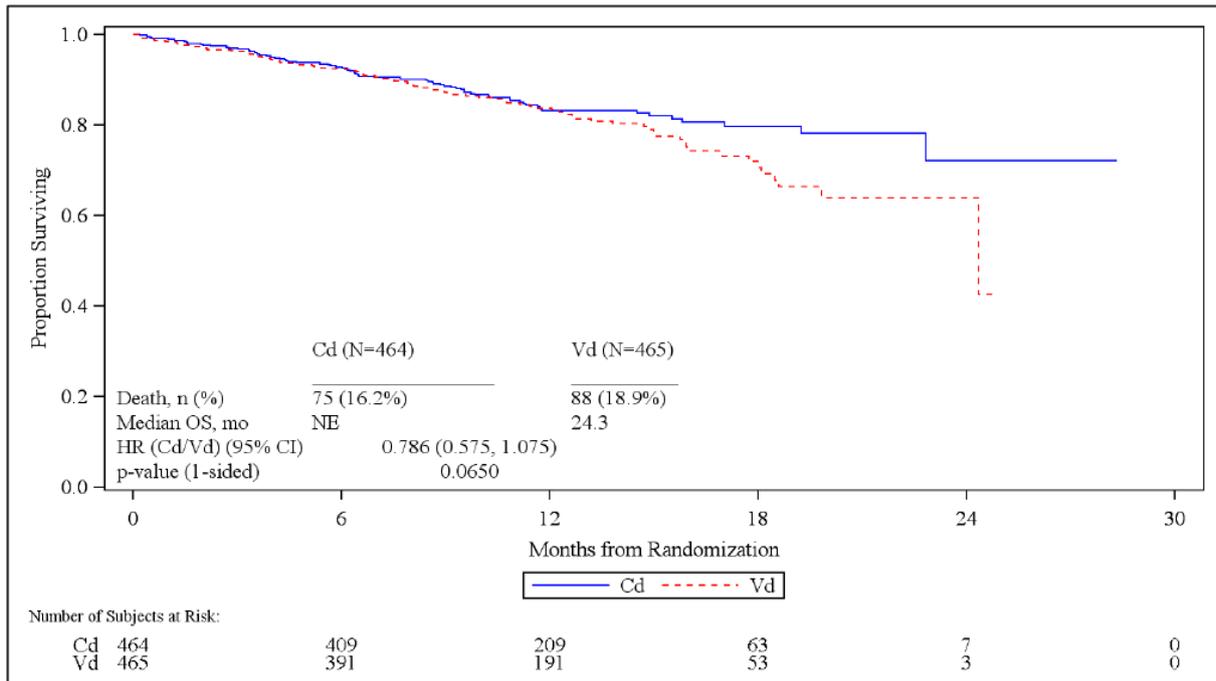


Abbildung 3: Kaplan-Meier-Überlebenszeitkurven für Gesamtmortalität in der Studie ENDEAVOR (Datenschnitt: 10.11.2014)

Die Kaplan-Meier-Überlebenszeitkurven verlaufen bis zum Studienmonat 12 nahezu gleich. Danach reduzieren sich die Patienten unter Risiko aus, so dass eine sinnvolle Interpretation und verlässliche Aussage über den Studienmonat 12 hinaus nicht möglich ist. Die mediane Überlebenszeit betrug im Kontrollarm 24 Monate und wurde im Verumarm zum Zeitpunkt der ersten Interimsanalyse nicht erreicht.

2.6.3 Morbidität

Der primäre Endpunkt „progressionsfreies Überleben“ wird im Folgenden ergänzend dargestellt, aber nicht zur Nutzenbewertung von Carfilzomib herangezogen (vgl. Tabelle 7).

Es werden Ergebnisse der ITT-Population präsentiert. Die PFS-Ereignisse wurden durch ein unabhängiges verblindetes Reviewkomitee bestätigt.

Tabelle 13: Ergebnisse zum progressionsfreien Überleben in der Studie ENDEAVOR (Datenschnitt 10. November 2014)

Studie ENDEAVOR	Cd (N=464)	Vd (N=465)
Ereignisse, n (%)		
Krankheitsprogress ¹⁾	171 (36,9)	243 (52,3)
Tod	150 (32,3)	228 (49,0)
	21 (4,5)	15 (3,2)
Zensurierungen, n (%)		
Keine Visite zu Baseline	293 (63,1)	222 (47,7)
Lebend ohne Progression	0	0
Ereignis nach mehr als einer verpassten Visite	258 (55,6)	157 (33,8)
Neue Krebstherapie	4 (0,9)	6 (1,3)
Lost to follow-up oder Entzug der Einverständniserklärung	23 (5,0)	29 (6,2)
	8 (1,7)	30 (6,5)

Studie ENDEAVOR	Cd (N=464)	Vd (N=465)
Dauer PFS (Monate), Median (95%-KI) ²⁾	18,7 (15,6; n.b.)	9,4 (8,4; 10,4)
HR ³⁾ (95%-KI)	0,53 (0,44; 0,65)	
p-Wert ⁴⁾ (1-seitig; $\alpha=0,025$)	< 0,0001	

¹⁾ Einschätzung, ob ein Ereignis vorlag, wurde von einem unabhängigen Review-Komitee getroffen, welches gegenüber der vom Studienteilnehmer erhaltenen Studienmedikation verblindet war.

²⁾ bestimmt nach der Kaplan-Meier-Methode. KI berechnet nach Klein und Moeschberger 1997 [17]

³⁾ stratifiziert nach zur Randomisierung verwendeten Stratavariablen: Vortherapie mit Proteasominhibitor, Anzahl Vortherapien, ISS-Stadium, Bortezomibapplikation

⁴⁾ Log-rank p-Wert, stratifiziert nach zur Randomisierung verwendeten Stratavariablen: Vortherapie mit Proteasominhibitor, Anzahl Vortherapien, ISS-Stadium, Bortezomibapplikation

Abkürzungen: HR: Hazard Ratio, PFS: progressionsfreies Überleben, KI: Konfidenzintervall, Vd: Bortezomib und Dexamethason; Cd: Carfilzomib und Dexamethason

Zum Datenschnitt ist bei ca. einem Drittel der Patienten im Verumarm und bei der Hälfte der Patienten im Kontrollarm ein PFS-Ereignis aufgetreten. Die Zensurierungen erfolgten größtenteils bei Patienten, die zum Datenschnitt lebten und keinen Krankheitsprogress aufwiesen. Es wurden ebenfalls Studienteilnehmer zensuriert, die eine neue Krebstherapie begannen, ein PFS-Ereignis nach einer erfolgten Studienvisite erlitten oder zur Nachbeobachtung nicht mehr zur Verfügung standen. Letzteres trat häufiger im Kontroll- als im Verumarm auf.

Das Hazard Ratio für PFS für die Intervention im Vergleich zur Kontrollmedikation war statistisch signifikant und belief sich auf HR=0,53 (95%-KI: 0,44; 0,65).

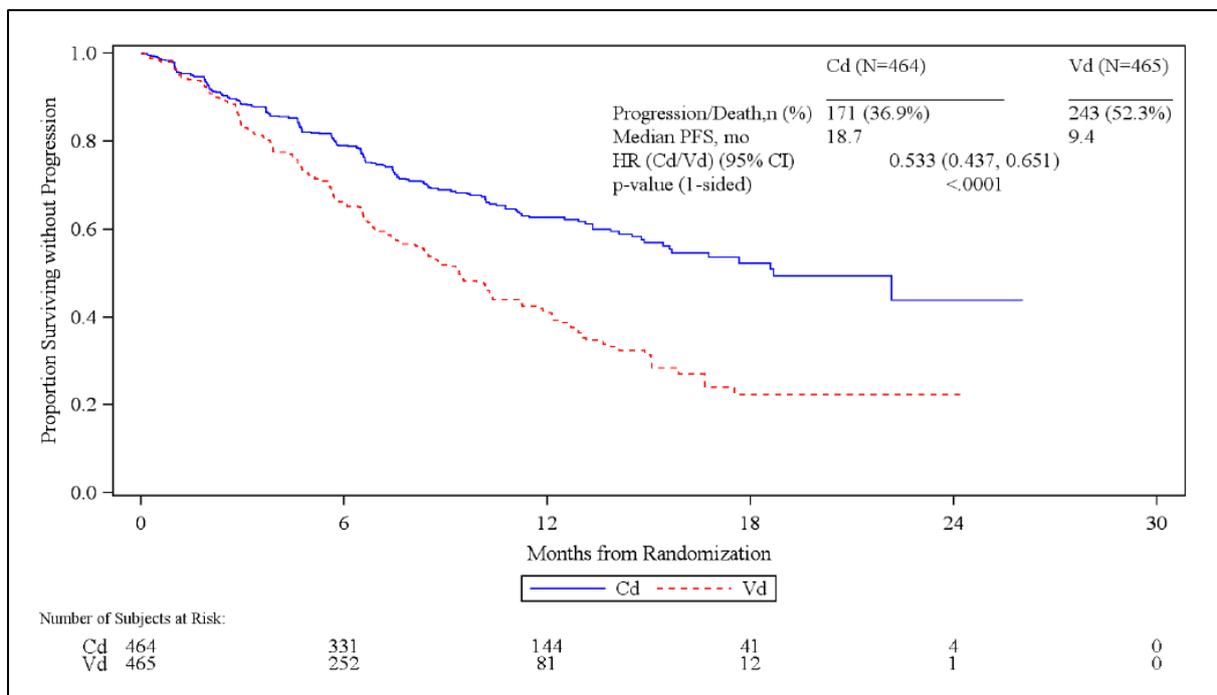


Abbildung 4: Kaplan-Meier-Überlebenszeitkurven für progressionsfreies Überleben in der Studie ENDEAVOR (Datenschnitt: 10. November 2014)

In Abbildung 4 wird ab Nachbeobachtungsmonat 3 ein Unterschied zwischen den Überlebenskurven der Kontrollmedikation und der Intervention sichtbar. Im Verlauf der Studie

wird dieser Unterschied größer. Die mediane Überlebenszeit für PFS betrug im Verumarm 18,7 Monate (95%-KI: 15,6; n.b.) und im Kontrollarm 9,4 Monate (95%-KI: 8,4; 10,4).

2.6.4 Lebensqualität

Als Lebensqualitätsendpunkte standen Ergebnisse zu Veränderungen ausgewählter Skalen der Instrumente EORTC QLQ-C30, EORTC QLQ-MY20 und FACT/GOG-Ntx zur Verfügung. Die Instrumente wurden zu Baseline und darauffolgend alle vier Wochen eingesetzt. Dies bedeutet, dass im Verumarm die gesundheitsbezogene Lebensqualität stets vor Beginn eines neuen Therapiezyklus gemessen wurde. Im Kontrollarm betrug die Zykluslänge nur 21 Tage, so dass die Erhebung häufig innerhalb des Therapiezyklus stattfand. Um eine Verzerrung durch unterschiedliche Erhebungszeitpunkte während des Therapiezyklus zu verringern, werden die Ergebnisse zu Woche 12, 24, 36, 48, 60 und 72 dargestellt. Zu diesen Zeitpunkten waren jeweils die Therapiezyklen in der Verum- und Kontrollgruppe abgeschlossen.

Die Rücklaufquoten, bezogen auf die verfügbare Anzahl Studienteilnehmer zu den jeweiligen Zeitpunkten, waren akzeptabel bei > 70 %, teils > 80 %. Legt man allerdings die Anzahl der randomisierten Teilnehmer zugrunde, dann sinkt die Rücklaufquote beträchtlich auf bis zu 2,4 % zu Studienwoche 72 (Daten nicht dargestellt). Die Rücklaufquote war für alle drei Instrumente höher im Verum- als im Kontrollarm. Teilnehmer wurden bis zum Auftreten einer Krankheitsprogression, Abbruch der Studienmedikation oder bis zur Aufnahme einer Folgetherapie bzgl. Lebensqualität nachbeobachtet.

EORTC QLQ-C30, EORTC QLQ-MY20:

Die Skalen des EORTC QLQ-C30 und QLQ-MY20 umfassen Wertebereiche zwischen 0 und 100, wobei für die Skala „Gesundheitszustand/Lebensqualität“, die Funktionsskalen „Physis“ und „Rolle“ ein hoher Score-Wert und für die Symptomskalen „Fatigue“, „Übelkeit/Erbrechen“, „Schmerz“ sowie „Krankheitssymptome“ und „Nebenwirkungen“ ein niedriger Score-Wert einen besseren Gesundheitszustand reflektieren.

Zu Baseline zeigte sich in keiner Skala des EORTC QLQ-C30 oder QLQ-MY20 ein statistisch signifikanter Unterschied in den Scorewerten zwischen Verum- und Kontrollarm (Ergebnisse nicht dargestellt).

Tabelle 14: Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität in der Studie ENDEAVOR, Datenschnitt: 10.11.2014

ENDEAVOR	Cd (N=464)		Vd (N=465)	
	n ¹ /N ² (%)	MW (SD)	n ¹ /N ² (%)	MW (SD)
EORTC QLQ-C30 Skala „Gesundheitszustand/Lebensqualität“⁽³⁾				
Baseline	407/464 (87,7)	61,5 (21,3)	392/465 (84,3)	63,7 (21,7)
Woche 12	383/408 (93,9)	61,3 (20,1)	336/388 (86,6)	59,4 (20,7)
Woche 24	298/343 (86,9)	63,6 (18,0)	222/254 (87,4)	59,9 (19,0)
Woche 36	235/258 (91,1)	63,7 (19,2)	142/162 (87,7)	63,3 (18,0)

Woche 48	137/159 (86,2)	62,2 (18,7)	73/80 (91,3)	63,9 (14,8)
Woche 60	73/82 (89,0)	66,9 (16,9)	28/37 (75,7)	68,2 (18,7)
Woche 72	41/44 (93,2)	68,9 (19,6)	11/12 (91,7)	68,2 (14,3)

¹⁾ Anzahl Studienteilnehmer, die den Fragebogen ausfüllten

²⁾ Anzahl Studienteilnehmer, die zu dem Zeitpunkt zum Ausfüllen des Fragebogens zur Verfügung standen

³⁾ Die Spannweite der Werte liegt zwischen 0 und 100. Ein höherer Score bedeutet bessere gesundheitsbezogene Lebensqualität.

Abkürzungen: MW: Mittelwert; KI: Konfidenzintervall; SD: Standardabweichung

In Tabelle 14 sind die Mittelwerte und Standardabweichungen des Scores des EORTC QLQ-C30 der Skala „Gesundheitszustand/Lebensqualität“ dargestellt. In beiden Armen verbesserte sich die gesundheitsbezogene Lebensqualität. Unterschiede zwischen den Armen waren zu Beginn der Therapie (Woche 12 und 24) zu beobachten, aber im weiteren Studienverlauf kaum vorhanden. Der größte Unterschied belief sich auf 4 Punkte zu Studienwoche 24 zugunsten der Therapie mit Carfilzomib.

Die Mittelwerte und Standardabweichungen der Subskalen ab Baseline bis Studienwoche 72 sind im Anhang (siehe Tabelle 21) dargestellt. Auf der Subskala „Fatigue“ stieg der Score in beiden Studienarmen nach Baseline an und fiel gegen Mitte bzw. Ende der Nachbeobachtungszeit auf Baseline-Niveau ab. Die größten Unterschiede beliefen sich auf 5 Punkte zu Woche 72 und auf 4,6 Punkte zu Woche 12 zugunsten von Carfilzomib.

Der Score auf der Subskala „Schmerz“ sank im Verlauf der Studie in beiden Armen ab, was mit einer besseren gesundheitsbezogenen Lebensqualität assoziiert ist. Bis Woche 48 war der Abfall des Scores deutlicher im Kontroll- als im Verumarm mit bis zu 4,2 Punkten Unterschied zuungunsten von Carfilzomib. Zu Woche 60 und 72 fielen Unterschiede zugunsten von Carfilzomib aus.

Der Score der Subskala „Physis“ stieg in beiden Studienarmen im Verlauf der Studie an. Unterschiede zwischen den Armen waren bis Woche 60 eher gering mit maximal 3 Punkten zu Woche 24 zugunsten von Carfilzomib. Zu Woche 72 wurde ein Unterschied von 8,7 Punkten zugunsten von Carfilzomib beobachtet.

In beiden Studienarmen sank der Score in der Subskala „Rolle“ im Verlauf der Studie und stieg ab Woche 60 wieder an. Dies bedeutet eine Verschlechterung der Lebensqualität während des ersten Studienjahres. Unterschiede zwischen den Armen waren gering und betragen maximal 3 Punkte zugunsten von Carfilzomib zu Woche 24.

Auf der Subskala „Übelkeit/Erbrechen“ verbesserte sich die Lebensqualität im Verlauf der Studie in beiden Armen leicht. Unterschiede zwischen den Armen betragen maximal 1,9 Punkte zu Studienwoche 24 zugunsten von Carfilzomib.

Der Score zur Skala „Krankheitssymptome“ des myelomspezifischen Instruments EORTC QLQ-MY20 (Tabelle 22) sank im Verlauf der Studie in beiden Studienarmen, was für eine Verbesserung der Krankheitssymptomatik spricht. Unterschiede zwischen den Armen waren gering und betragen nicht mehr als 2 Punkte zuungunsten von Carfilzomib.

Die Skala „Nebenwirkungen“ zeigte zu Beginn der Therapie in beiden Studienarmen eine Verschlechterung, die ab Woche 36 bzw. 48 wieder zurückging und in beiden Therapiearmen Baseline-Niveau erreichte. Unterschiede zwischen den Therapiearmen waren moderat, teils

zugunsten, teils zuungunsten von Carfilzomib und beliefen sich auf nicht mehr als 3,7 Punkte zugunsten von Carfilzomib.

Die klinische Relevanz der Unterschiede zwischen den Studienarmen ist aufgrund der fehlenden MID für den EORTC QLQ-C30 und QLQ-MY20 unklar.

Die Ergebnisse der vorab geplanten Analysen mittels MMRM wurden nicht in der Nutzenbewertung dargestellt. In diese Analysen gingen alle erhobenen Informationen zur Lebensqualität (alle 4 Wochen) ein. Eine Verzerrung durch die Erhebung der Lebensqualität zu unterschiedlichen Zeitpunkten vor und innerhalb der Therapiezyklen in Verum- und Kontrollarm erscheint wahrscheinlich. Zudem lässt ein globaler Schätzer über den gesamten Zeitraum von 72 Wochen bei sehr wenigen zur Verfügung stehenden Beobachtungen ab Studienwoche 48 kaum belastbare Aussagen über die gesamte Studienpopulation zu. Aus diesen Gründen wird auf eine Darstellung der Ergebnisse verzichtet.

Ebenso wurde in der Nutzenbewertung auf die Darstellung der Ergebnisse der Responderanalysen verzichtet. Im Rahmen dieser Analysen wurde Bezug auf die gesamte randomisierte Population genommen und Responder über die Messzeitpunkte der Studie addiert. Dabei wurde nicht berücksichtigt, dass ein bedeutender Anteil der Studienpopulation nicht bis Woche 72 nachbeobachtet wurde. Die im Modul 4 präsentierten Ergebnisse sind daher nicht sinnvoll interpretierbar.

FACT/GOG-Ntx:

Zu Baseline unterschieden sich die Mittelwerte zwischen den Studienarmen der Skala „weitere Beschwerden“ des FACT/GOG-Ntx nicht (siehe Tabelle 23). Im Verlauf der Nachbeobachtung traten kaum Veränderungen auf. Innerhalb des Verumarms beliefen sich die Veränderungen auf 0,8 Punkte und im Kontrollarm auf maximal 1,5 Punkte. Unterschiede zwischen den Armen betragen höchstens 1,5 Punkte zugunsten von Carfilzomib. Die klinische Relevanz dieser Unterschiede ist allerdings aufgrund einer fehlenden MID sowie Responseschwelle unklar.

2.6.5 Subgruppenanalysen

Für den primären Endpunkt PFS und den sekundären Endpunkt Gesamtüberleben wurde eine Reihe von Subgruppenanalysen vorgelegt. Dabei beschränkte sich die Darstellung auf Effektschätzer für jedes einzelne Stratum (24 Stratavariablen, 64 Strata). Ergebnisse zu Interaktionstests der jeweiligen Stratavariablen mit der Studienmedikation wurden nicht präsentiert.

Bezugnehmend auf die geforderten Subgruppenanalysen zu Alter, Geschlecht, Region und Krankheitsschwere unterschieden sich die Ergebnisse in den Subgruppen in den Endpunkten PFS und Gesamtüberleben wenig von den Gesamtergebnissen. Insgesamt ist die Ereigniszahl zu gering, um verlässliche Rückschlüsse auf präspezifizierte Subgruppen ziehen zu können.

2.6.6 Sicherheit

Zur Darstellung von sicherheitsrelevanten Endpunkten wurde die Sicherheitspopulation herangezogen, die alle randomisierten Personen mit mindestens einer Anwendung der Studienmedikation (Cd: n=463; Vd: n=456) umfasst. Die Behandlungsdauer unterschied sich in beiden Studienarmen deutlich mit im Median 39,9 Wochen (Spanne: 1–108) im Verumarm und

26,8 Wochen (Spanne: 1–106) im Kontrollarm. Die mediane Anzahl begonnener Zyklen betrug im Carfilzomibarm 10 (Spanne: 1–26) und im Bortezomibarm 8 (Spanne: 1–35).

In Tabelle 15 sind die Ergebnisse zu unerwünschten Ereignissen (UE) in der Studie ENDEAVOR zusammengefasst. Hinsichtlich des Erfassungszeitraumes der UE besteht eine Diskrepanz zwischen Studienprotokoll und Studienbericht. Laut Studienprotokoll sollten UE ab Tag der Randomisierung bis 30 Tage nach Abschluss der Studienmedikation erfasst werden. Im Studienbericht werden UE allerdings erst ab Tag der ersten Einnahme der Studienmedikation bis 30 Tage nach Abschluss der Studienmedikationseinnahme aufgeführt.

Insgesamt traten bei fast allen Studienteilnehmern UE auf. Numerisch traten häufiger UE des NCI-Grades ≥ 3 , SUE sowie UE, die zum Tod führten, im Verum- als im Kontrollarm auf. Dieser Unterschied war unter Beachtung der Expositionszeit nicht statistisch signifikant. UE, die zum Abbruch von mindestens einer Studienmedikation führten, traten numerisch häufiger im Kontroll- als im Verumarm auf. Unter Beachtung der Expositionszeit zeigte sich ein statistisch signifikanter Vorteil zugunsten von Carfilzomib.

Tabelle 15: Ergebnisse der unerwünschten Ereignisse in der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014

Patienten mit mindestens einem..., n (%)	Cd N=463	Vd N=456	Risk Ratio ¹⁾ (95 %-KI)
UE	455 (98,3)	447 (98,0)	n.d. ²⁾
UE des NCI-CTCAE-Grades ≥ 3	339 (73,2)	305 (66,9)	0,92 (0,78; 1,07)
SUE	224 (48,4)	162 (35,5)	1,06 (0,87; 1,30)
SUE des NCI-CTCAE-Grades ≥ 3	k.A.	k.A.	k.A.
UE, das zum Abbruch mind. einer Studienmedikation führte	92 (19,9)	95 (20,8)	0,71 (0,53; 0,94)
UE, das zum Tod führte	23 (5,0)	18 (3,9)	0,93 (0,50; 1,73)

¹⁾ Cd vs. Vd; Risk Ratio wurde unter Berücksichtigung der Expositionszeit berechnet. Es handelt sich hierbei um Ergebnisse aus Post-hoc-Analysen.

²⁾ Patientenrelevanz nicht eindeutig beurteilbar.

Abkürzungen: n.d.: nicht dargestellt; k.A. keine Angabe durch pU

Unerwünschte Ereignisse

Unerwünschte Ereignisse mit einer Inzidenz von 10 % oder höher sind in Tabelle 16 dargestellt. Unterschiede von 5 %-Punkten oder mehr zwischen den Studienarmen fanden sich in fast allen Systemorganklassen. Für die SOC: „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, „Infektionen und parasitäre Erkrankungen“, „Erkrankungen der Atemwege, des Brustraums und des Mediastinums“, „Erkrankungen der Muskel, des Skeletts und des Bindegewebes“, „Erkrankungen des Blutes und Lymphsystems“, „Gefäßerkrankungen“ und „Untersuchungen“ fielen die Unterschiede von ≥ 5 %-Punkten zwischen den Armen zuungunsten von Carfilzomib aus. Für die SOC: „Erkrankungen des Gastrointestinaltrakts“, „Erkrankungen des Nervensystems“ und „Augenerkrankungen“ fielen die Unterschiede von ≥ 5 %-Punkten zugunsten von Carfilzomib aus.

Tabelle 16: UE mit Inzidenz $\geq 10\%$ in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014

MedDRA Systemorganklasse, Preferred Term¹⁾, n (%)	Cd N=463	Vd N=456
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	342 (73,9)	287 (62,9)
Fatigue	136 (29,4)	130 (28,5)
Pyrexie	130 (28,1)	63 (13,8)
Periphere Ödeme	101 (21,8)	78 (17,1)
Asthenie	94 (20,3)	75 (16,4)
Infektionen und parasitäre Erkrankungen	341 (73,1)	294 (64,5)
Infektionen der oberen Atemwege	94 (20,3)	67 (14,7)
Bronchitis	76 (16,4)	41 (9,0)
Nasopharyngitis	66 (14,3)	51 (11,2)
Pneumonie	41 (8,9)	48 (10,5)
Erkrankungen des Gastrointestinaltraktes	264 (57,0)	291 (63,8)
Diarrhö	143 (30,9)	175 (38,4)
Nausea	90 (19,4)	82 (18,0)
Konstipation	68 (14,7)	123 (27,0)
Erbrechen	65 (14,0)	40 (8,8)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	264 (57,0)	161 (35,3)
Dyspnoe	132 (28,5)	60 (13,2)
Husten	115 (24,8)	64 (14,0)
Erkrankungen der Muskel, des Skeletts und des Bindegewebes	259 (55,9)	224 (49,1)
Rückenschmerzen	86 (18,6)	71 (15,6)
Muskelkrämpfe	86 (18,6)	27 (5,9)
Arthralgie	47 (10,2)	46 (10,1)
Knochenschmerzen	47 (10,2)	38 (8,3)
Schmerzen in den Extremitäten	47 (10,2)	49 (10,7)
Erkrankungen des Blutes und Lymphsystems	237 (51,2)	184 (40,4)
Anämie	182 (39,3)	123 (27,0)
Thrombozytopenie	95 (20,5)	78 (17,1)
Erkrankungen des Nervensystems	217 (46,9)	328 (71,9)
Kopfschmerzen	79 (17,1)	46 (10,1)
Periphere Neuropathie	43 (9,3)	121 (26,5)
Schwindel	37 (8,0)	67 (14,7)
Parästhesien	36 (7,8)	74 (16,2)
Periphere sensorische Neuropathie	27 (5,8)	67 (14,7)
Neuralgie	9 (1,9)	70 (15,4)

MedDRA Systemorganklasse, Preferred Term¹⁾, n (%)	Cd N=463	Vd N=456
Gefäßerkrankungen	209 (45,1)	113 (24,8)
Hypertonus	115 (24,8)	40 (8,8)
Stoffwechsel- und Ernährungsstörungen	205 (44,3)	187 (41,0)
Hypokaliämie	50 (10,8)	45 (9,9)
Hyperglykämie	49 (10,6)	41 (9,0)
Verminderter Appetit	40 (8,6)	57 (12,5)
Untersuchungen	196 (42,3)	150 (32,9)
Verminderte Thrombozytenzahl	55 (11,9)	39 (8,6)
Erhöhter Kreatininspiegel im Blut	48 (10,4)	26 (5,7)
Psychische Erkrankungen	168 (36,3)	177 (38,8)
Insomnie	117 (25,3)	119 (26,1)
Erkrankungen der Haut und des Unterhautzellgewebes²⁾	123 (26,6)	128 (28,1)
Augenerkrankungen²⁾	86 (18,6)	112 (24,6)
Kardiale Erkrankungen²⁾	99 (21,4)	43 (9,4)
Erkrankungen der Nieren und Harnwege²⁾	94 (20,3)	61 (13,4)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen²⁾	73 (15,8)	88 (19,3)

¹⁾ Patienten wurden nur einmal in jeder Systemorganklasse und Preferred Term gezählt.

²⁾ Kein Preferred Term wurde mit $\geq 10\%$ der Patienten beobachtet.

UE des CTCAE-Grades 3 oder höher

Es zeigten sich leichte Unterschiede zuungunsten von Carfilzomib. Diese Unterschiede betragen zwischen den Studienarmen nicht mehr als 5 %-Punkte bis auf den PT Hypertonus, der bei 8,9 % der Patienten im Verumarm und bei 2,6 % der Patienten im Kontrollarm beobachtet wurde (siehe Tabelle 17).

Tabelle 17: UE des NCI-CTCAE-Grades 3 und höher mit der Inzidenz von mind. 2 % in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 11.10.2014

UE \geq Grad 3 nach NCI-CTCAE, Preferred Term, n (%)	Cd N=463	Vd N=456
Patienten mit mindestens einem Ereignis	339 (73,2)	305 (66,9)
Anämie	67 (14,5)	45 (9,9)
Hypertonus	41 (8,9)	12 (2,6)
Thrombozytopenie	39 (8,4)	43 (9,4)
Pneumonie	32 (6,9)	36 (7,9)
Verminderte Lymphozytenzahl	26 (5,6)	8 (1,8)
Dyspnoe	25 (5,4)	10 (2,2)
Fatigue	25 (5,4)	32 (7,0)
Lymphopenie	20 (4,3)	12 (2,6)

UE ≥ Grad 3 nach NCI-CTCAE, Preferred Term, n (%)	Cd N=463	Vd N=456
Hyperglykämie	18 (3,9)	16 (3,5)
Verminderte Thrombozytenzahl	17 (3,7)	24 (5,3)
Asthenie	16 (3,5)	14 (3,1)
Diarrhö	16 (3,5)	34 (7,5)
Hypophosphatämie	11 (2,4)	5 (1,1)
Pyrexie	11 (2,4)	3 (0,7)
Knochenschmerz	10 (2,2)	6 (1,3)
Bronchitis	10 (2,2)	4 (0,9)
Herzversagen	10 (2,2)	3 (0,7)
Hyponatriämie	10 (2,2)	6 (1,3)
Neutropenie	10 (2,2)	10 (2,2)
Rückenschmerzen	8 (1,7)	12 (2,6)
Hypokaliämie	7 (1,5)	13 (2,9)
Insomnie	7 (1,5)	11 (2,4)
Periphere Neuropathie	6 (1,3)	24 (5,3)
Konstipation	2 (0,4)	9 (2,0)
Synkope	0	12 (2,6)

Schwerwiegende UE

SUE sind in der Tabelle 18 dargestellt. Insgesamt traten deutlich häufiger SUE im Verum- als im Kontrollarm auf: Pneumonien wurden in beiden Armen am häufigsten berichtet. Die verbleibenden PT der SUE traten deutlich seltener auf und Unterschiede in den PT zwischen den Studienarmen beliefen sich auf nicht mehr als 2,8 %-Punkte.

Tabelle 18: SUE mit einer Häufigkeit von mind. 1 % in einem der Studienarme der Sicherheitspopulation der Studie ENDEAVOR, Datenschnitt 10.11.2014

SUE, Preferred Term, n (%)	Cd N=463	Vd N=456
Patienten mit mindestens einem Ereignis	224 (48,4)	162 (35,5)
Pneumonie	28 (6,0)	39 (8,6)
Pyrexie	15 (3,2)	3 (0,7)
Dyspnoe	14 (3,0)	1 (0,2)
Lungenembolie	10 (2,2)	3 (0,7)
Herzversagen	8 (1,7)	3 (0,7)
Akutes Nierenversagen	8 (1,7)	5 (1,1)
Krankheitsprogression	7 (1,5)	5 (1,1)
Infektion der oberen Atemwege	7 (1,5)	3 (0,7)

SUE, Preferred Term, n (%)	Cd N=463	Vd N=456
Bronchopneumonie	6 (1,3)	0 (0)
Sepsis	6 (1,3)	3 (0,7)
Vorhofflimmern	5 (1,1)	4 (0,9)
Rückenschmerz	5 (1,1)	2 (0,4)
Diarrhö	5 (1,1)	9 (2,0)
Multiples Myelom	5 (1,1)	1 (0,2)
Plasmozytome	5 (1,1)	0 (0)
Infektion der Atemwege	5 (1,1)	5 (1,1)
Harnwegsinfekt	5 (1,1)	4 (0,9)
Erbrechen	5 (1,1)	2 (0,4)
Hyperkalzämie	0 (0)	5 (1,1)
Thrombozytopenie	4 (0,9)	6 (1,3)

Zu einem Abbruch der Behandlung mit der Studienmedikation führende UE

Im Verumarm traten 92 (19,2 %) und im Kontrollarm 95 (20,8 %) UE auf, die zum Abbruch einer Studienmedikation führten. Am häufigsten brachen Patienten aufgrund peripherer Neuropathien mit 19 Ereignissen (4,2 %) im Kontrollarm und 1 Ereignis (0,2 %) im Verumarm eine der Studienmedikationen ab. Dies war gefolgt von Fatigue mit 8 (1,8 %) Ereignissen vs. 1 (0,2 %) Ereignis zugunsten von Carfilzomib und Herzversagen mit 0 vs. 6 (1,3 %) Ereignissen zuungunsten von Carfilzomib. Die anderen Gründe traten seltener auf und waren über die Studienarme annähernd gleich verteilt mit weniger als 1 %-Punkt Unterschied zwischen den Armen.

Vom pU beschriebene UE von besonderem Interesse

Es wurden 15 UE von besonderem Interesse präspezifiziert. Diese Auswahl beruht auf Erfahrungen mit Patienten mit Multiplem Myelom und bekannten UE der Proteasominhibitoren Bortezomib und Carfilzomib aus früheren Studien. Folgende UE wurden dabei berücksichtigt: Herzerkrankungen, hämatologische Erkrankungen, Blutungen, Lebererkrankungen, Hypersensitivitäten, Infektionen, Tumorerkrankungen, Neuropathien, nicht-infektiöse Enzephalopathien, Lungenerkrankungen, Nierenerkrankungen, thromboembolische Ereignisse, thrombotische Mikroangiopathien, Tumor-Lyse-Syndrom und vaskuläre Ereignisse.

Unterschiede von ≥ 5 %-Punkten zwischen den Studienarmen wurden für: Herzversagen (Cd: n=38 (8,2 %) vs. Vd: n=13 (2,9 %)), kardiale Arrhythmien (Cd: n=75 (16,2 %) vs. Vd: n=43 (9,4 %)), Infektionen und parasitäre Erkrankungen (Cd: n=341 (73,7 %) vs. Vd: n=294 (64,5 %)), Dyspnoe (Cd: n=143 (30,9 %) vs. Vd: n=78 (17,1 %)), embolische und thrombotische Ereignisse (Cd: n=66 (14,3 %) vs. Vd: n=26 (5,7 %)), Hypertonus (Cd: n=120 (25,9 %) vs. Vd: n=44 (9,6 %)) zuungunsten von Carfilzomib und hämatopoetische Erythropenie (Cd: n=185 (40,0 %) vs. Vd: n=125 (27,4 %)), hämatopoetische Thrombozytopenie (Cd: n=55 (11,9 %) vs. Vd: n=116 (25,4 %)), periphere Neuropathien (Cd: n=87 (18,8 %) vs. Vd: n=235 (51,5 %)), vaskuläre

hypotensive Erkrankungen (Cd: n=27 (5,8 %) vs. Vd: n=50 (11,0 %)) zugunsten von Carfilzomib beobachtet und berichtet.

Weitere sicherheitsrelevante Aspekte:

Das Auftreten peripherer Neuropathien ist ein sekundärer Endpunkt der Studie ENDEAVOR. Häufigkeiten und relative Risikoschätzer sind in Tabelle 19 dargestellt.

Tabelle 19: Ergebnisse zum Auftreten peripherer Neuropathien in der Studie ENDEAVOR, Sicherheitspopulation, Datenschnitt: 10.11.2014

ENDEAVOR	Cd N=463	Vd N=456	Cd vs. Vd
	Ereignisse (%)	Ereignisse (%)	Risk Ratio (95% σ -KI) p-Wert
Alle Schweregrade	87 (18,8)	235 (51,5)	k.A.
CTC-Grad ≥ 2	28 (6,0)	146 (32,0)	0,19 (0,13; 0,28) < 0,0001
CTC-Grad ≥ 3	10 (2,2)	37 (8,1)	0,27 (0,13; 0,53) < 0,0001
CTC-Grad ≥ 4	0 (0)	1 (0,2)	k.A.

Abkürzungen: k.A. keine Angaben im Modul 4 oder Studienbericht

Es entwickelten im Verlauf der Studie deutlich mehr Patienten unter Bortezomibtherapie periphere Neuropathien als unter Therapie mit Carfilzomib. Dieser Unterschied war für die Schweregrade ≥ 2 bzw. ≥ 3 statistisch signifikant.

3 Methodische Anmerkungen und Bewertung der Unterlagen

3.1 Zulassungsstatus und Zulassungspopulation von Carfilzomib

Das hier betrachtete Anwendungsgebiet umfasst Carfilzomib in Kombination mit Dexamethason zur Behandlung von erwachsenen Patienten mit Multiplem Myelom, die mindestens eine vorangegangene Therapie erhalten haben. Die zulassungsbegründende Studie ENDEAVOR schloss Patienten mit rezidiviertem oder progredientem Multiplem Myelom ein, die mindestens eine, aber nicht mehr als drei Vortherapien erhalten haben. Zudem wurden nur Personen berücksichtigt, die ein partielles Ansprechen auf mindestens eine vorangegangene Therapie zeigten und nicht den Subtyp IgM des Multiplen Myeloms aufwiesen. Personen, die nicht auf mindestens eine der Vortherapien ansprachen und den Subtyp IgM des Multiplen Myeloms aufwiesen, kämen laut Zulassung auch für eine Behandlung mit Carfilzomib infrage. Eine Aussage zu diesen Teilpopulationen ist auf Basis der vorgelegten Daten nicht möglich.

3.2 Design und Methodik der Studie ENDEAVOR

Die Studie ENDEAVOR ist eine offene, randomisierte, multizentrische und multinationale Studie der Phase III mit zwei Behandlungsarmen zur Untersuchung der Wirksamkeit und Sicherheit von Carfilzomib in Kombination mit Dexamethason im Vergleich zu Bortezomib in Kombination mit Dexamethason bei erwachsenen vortherapierten Patienten mit Multiplem Myelom. Die Studie ist weiterhin aktiv, allerdings beschränkt sich die Datenerhebung nur noch auf die Endpunkte „Gesamtüberleben“ und Sicherheit.

Die Randomisierung der Studienteilnehmer erfolgte stratifiziert mit verdeckter Zuteilung. Aus der Liste der Protokollverletzungen wird deutlich, dass ein bedeutender Teil der Population (125 Teilnehmer, 13,5 % der Studienpopulation) in das falsche Stratum randomisiert wurde. Weitere Details zu diesen Protokollverletzungen sind weder im Dossier noch im Studienprotokoll oder -bericht aufgeführt. Es ist unklar, inwiefern dies zu Imbalancen insbesondere in Subgruppen geführt haben könnte. Subgruppenanalysen könnten möglicherweise durch diese Protokollverletzungen verzerrt sein und keine validen Aussagen zulassen. Zwischen den Studienarmen konnte zu Baseline auf keine ungleiche Verteilung der Studienteilnehmer bezüglich der Stratifizierungsfaktoren und der berichteten Baselinecharakteristika geschlossen werden, so dass von keiner hohen Verzerrung durch die Protokollverletzung ausgegangen wurde. Insgesamt scheint die Population bzgl. der dargelegten Charakteristika zwischen den Studienarmen ausgeglichen.

Auf eine Verblindung der Studienteilnehmer und des Studienpersonals wurde verzichtet. Das offene Design der Studie ENDEAVOR birgt die Gefahr der Verzerrung: zum einen auf Teilnehmerseite, wenn beispielsweise Lebensqualität verzerrt, teils auch unbewusst, wahrgenommen und berichtet wird; zum anderen durch ungleiche Erfassung, Bewertung und Behandlung von Symptomen und unerwünschten Ereignissen auf Untersucherseite. Letzteres trifft insbesondere für Einzelkomponenten des Endpunkts PFS zu, bei welchen „optional“ eine weiterführende Diagnostik zur Bestimmung der Plasmazellen und Plasmazytome im Verumarm durchgeführt wurde. Im Kontrollarm erfolgte die Diagnostik regulär alle 4 Wochen. Wann diese „optionale“ Bestimmung im Verumarm vorgenommen wurde, ist nicht nachvollziehbar operationalisiert und unterliegt vermutlich der Entscheidung des Studienpersonals, welche

Kenntnis über die Zuteilung des Patienten zum jeweiligen Studienarm hatten. Eine Verzerrung der Ergebnisse kann nicht ausgeschlossen werden. Darüber hinaus erfolgte die Messung der Knochenläsionen, wenn eine „klinische Indikation“ vorlag. Vorab definierte Kriterien für eine klinische Indikation waren nicht vorhanden. Möglicherweise erfolgte die Entscheidung, wann eine Messung der Knochenläsionen vorgenommen wurde, nicht unabhängig vom Wissen über die Zuteilung zum Verum- oder Kontrollarm des Patienten. In diesem Fall ist auch hier eine Verzerrung durch das offene Studiendesign nicht auszuschließen.

In beiden Studienarmen wurde in jedem Therapiezyklus acht Mal 20 mg Dexamethason eingenommen. Allerdings unterschied sich die Länge der Therapiezyklen zwischen den Studienarmen, sodass im Verlauf der Studie die eingenommene Menge an Dexamethason im Kontrollarm höher als im Verumarm war. Darüber hinaus erfolgte die Therapie im Kontrollarm ab dem 9. Therapiezyklus nicht fachinformationskonform. Laut Zulassung ist Bortezomib bis zu maximal acht Therapiezyklen anzuwenden. In der Studie ENDEAVOR wurden Patienten hingegen über acht Zyklen hinaus mit Bortezomib therapiert.

Der Patientenfluss der Studie ENDEAVOR, dargestellt in Tabelle 10 und Abbildung 2, ist nicht konsistent mit Angaben zum primären Endpunkt PFS (siehe Tabelle 13). Diskrepanzen treten bezüglich der Anzahl der Patienten, die einen Progress erlitten, und Patienten, die lost to follow-up waren und die die Einverständniserklärung zurückzogen, auf. Diese Abweichungen sind nicht durch unterschiedliche Definitionen von PFS erklärbar (beispielsweise Festlegung durch Untersucher vs. Festlegung durch unabhängiges Reviewkomitee). Die Diskrepanzen im Patientenfluss können mit den vorliegenden Unterlagen nicht abschließend geklärt werden.

Die Anzahl der Studienteilnehmer, die die Studienmedikation abbrachen, unterschied sich zwischen den Studienarmen. Unter Therapie mit Carfilzomib entzogen im Vergleich zum Kontrollarm weniger Patienten die Einverständniserklärung und weniger Patienten brachen aufgrund der Entscheidung des Studienarztes die Studienmedikation ab als unter Bortezomibtherapie. Diese Abbruchgründe könnten durchaus im Zusammenhang mit einer bevorstehenden Krankheitsprogression stehen. Das offene Studiendesign könnte zudem zu einem häufigeren Abbruch der Studienmedikation auf Anraten des Arztes beigetragen haben. Gründe, wann Studienärzte zu einem Abbruch der Studienmedikation rieten, wurden nicht aufgeführt. Eine Verzerrung durch ungleiche Medikationsabbrecher kann in der vorliegenden Studie nicht ausgeschlossen werden.

Die Fallzahlplanung und die statistische Analyse sind im SAP dargelegt. Der pU stellt nicht dar, wie er zu den Annahmen einer Reduktion des Hazard Ratios für PFS um 25 % und mittleren progressionsfreien Zeit von 10 Monaten im Kontrollarm gelangt, welche zur Fallzahlplanung herangezogen werden. Die Fallzahlplanung ist somit nicht eindeutig nachvollziehbar. Änderungen im SAP erfolgten teils nach dem hier betrachteten Datenschnitt. Diese Änderungen betrafen u. a. Testprozeduren der sekundären Endpunkte, Definition des Endpunkts „periphere Neuropathie“, Analysemethoden zur Untersuchung von UE und die Klassifizierung von Endpunkten. Es ist unklar, warum Änderungen erst nach einem Interimsdatenschnitt erfolgten und wie diese Änderungen, welche als post hoc betrachtet werden müssen, zu einer ergebnisgesteuerten Berichterstattung geführt haben könnten.

3.3 Wirksamkeit

Mortalität

Für den patientenrelevanten Endpunkt „Gesamtüberleben“ liegen Ergebnisse aus der ersten Interimsanalyse vor, in die 163 der 631 geplanten Ereignisse eingegangen sind. Es zeigte sich kein statistisch signifikantes Ergebnis im Endpunkt „Gesamtüberleben“. Anhand der Kaplan-Meier-Kurven zeigte sich bis zum Studienmonat 12 kein Unterschied im Überleben zwischen den beiden Studienarmen. Eine Interpretation über den Studienmonat hinaus ist nicht sinnvoll, da sich die Personen unter Risiko ab diesem Zeitpunkt stark reduzierten.

Die Ergebnisse der Subgruppenanalysen lassen durch die geringen Ereigniszahlen keine belastbaren Aussagen zu. Eine Einschätzung zum Vorliegen qualitativer oder quantitativer Interaktionen ist nicht möglich.

Aus den Studienunterlagen gingen keine Informationen zur Postprogressionstherapie oder zur Therapie nach Abbruch der Studienmedikation hervor. Beim Endpunkt „Gesamtüberleben“ besteht die Gefahr einer Verzerrung durch Aufnahme einer Folgetherapie nach Progress oder Abbruch der Studienmedikation, insbesondere, wenn Personen aus dem Kontrollarm auf wirksamere Therapien als im Verumarm oder auf die Studienmedikation selbst umstellen. Ein Einfluss eines solchen Effekts kann aufgrund der fehlenden Informationen zu Art und Zeitpunkt der Folgetherapie nicht eingeschätzt werden.

Progressionsfreies Überleben (primärer Endpunkt)

Eine Bewertung der Patientenrelevanz des Endpunkts „progressionsfreies Überleben“ ist aufgrund fehlender Belege zur Validität als Surrogatparameter für Morbidität, Mortalität und/oder gesundheitsbezogene Lebensqualität nicht möglich. Die Teilkomponente „Tod“, die patientenrelevant ist, wird bereits durch den Endpunkt „Gesamtüberleben“ abgebildet. Insgesamt ergab sich ein statistisch signifikanter Vorteil zugunsten von Carfilzomib für das „progressionsfreie Überleben“, was allerdings aufgrund der fraglichen Patientenrelevanz und Validität nicht zur Herleitung des Zusatznutzens herangezogen werden kann. Diese Ergebnisse basieren auf der ersten Interimsanalyse, in welche 414 von geplanten 526 PFS-Ereignissen eingegangen sind.

Die Operationalisierung des Endpunkts war bzgl. der Kriterien, wann eine Bestimmung von Knochenläsionen, Plasmazytomen und Plasmazellen im Knochenmark vorgenommen werden sollte, nicht nachvollziehbar beschrieben. Diese unzureichende Definition von Kriterien zur Bestimmung von diagnostischen Messungen wurde auch durch den Einsatz eines unabhängigen Reviewkomitees nicht ausgeglichen. Zwar schätzte das Reviewkomitee die erhobenen Daten zum PFS unabhängig und verblindet ein. Jedoch lagen von Personen, die keine klinische Indikation zur Messung von Knochenläsionen aufwiesen, solche diagnostischen Informationen nicht vor. Bestimmungen der Plasmazytome und der Plasmazellen im Knochenmark waren optional im Verumarm und es ist unklar, wer im Verlauf der Studie die Entscheidung traf, dass diese Messungen vorgenommen werden sollten. Durch das offene Studiendesign hatte das Studienpersonal Kenntnis über die Zuordnung des jeweiligen Patienten zum Verum- oder Kontrollarm. Es kann nicht ausgeschlossen werden, dass eine Einschätzung über die weiterführende Diagnostik zur Bestimmung eines Krankheitsprogresses durch dieses Wissen,

wenn auch unterbewusst, beeinflusst war. Eine Verzerrung der Ergebnisse zum Endpunkt PFS ist daher nicht auszuschließen.

Die Ergebnisse der Subgruppenanalysen lassen durch die Vielzahl der präspezifizierten Strata, die relativ geringen Ereigniszahlen und durch die Inflation des Fehlers 1. Art keine belastbaren Aussagen zu. Eine Einschätzung zum Vorliegen qualitativer oder quantitativer Interaktionen ist nicht möglich.

3.3.1 Zusammenfassende Einschätzung zur Wirksamkeit

Basierend auf der ersten Interimsanalyse konnte kein Unterschied zwischen den Studienarmen im Endpunkt „Gesamtüberleben“ beobachtet werden. Allerdings beruht dieses Ergebnis auf lediglich 33 % der geplanten Ereignisse, so dass eine abschließende Bewertung zu diesem Zeitpunkt nicht möglich ist.

Eine zweite Interimsanalyse ist bei Erreichen von ca. 394 Ereignissen (79 % der geplanten Ereigniszahl) vorgesehen. Die abschließende Analyse soll bei Vorliegen von 496 Ereignissen durchgeführt werden.

3.4 Lebensqualität

Die gesundheitsbezogene Lebensqualität wurde in der Studie ENDEAVOR mittels ausgewählter Skalen der Instrumente EORTC QLQ-C30, EORTC QLQ-MY20 und FACT/GOG-Ntx erhoben. Zu Baseline lagen keine statistisch signifikanten Unterschiede auf den betrachteten Skalen zwischen den Studienarmen vor. Im Verlauf der Studie veränderte sich die gesundheitsbezogene Lebensqualität gemessen mittels EORTC QLQ-C30 leicht. Die mittleren Unterschiede zwischen den Armen betragen weniger als 5 Punkte auf allen betrachteten Skalen. Die gesundheitsbezogene Lebensqualität erhoben mit dem krankheitsspezifischen Instrument EORTC QLQ-MY20 verbesserte sich auf der Skala „Krankheitssymptome“ in beiden Studienarmen deutlich, aber zeigte kaum Unterschiede zwischen den Studienarmen. Auf der Skala „Nebenwirkungen“ zeigten sich ebenso kaum Unterschiede zwischen den Armen (weniger als 4 Punkte Unterschied). Dieses Ergebnis ist konsistent mit den Beobachtungen der Subskala „zusätzliche Beschwerden“ des FACT/GOG-Ntx, auf der ebenfalls kaum Veränderungen innerhalb und zwischen den Studienarmen beobachtet wurden.

Da keine validierten Schwellenwerte für klinisch relevante Veränderungen zwischen den Gruppen (MID) vorliegen, können die beobachteten mittleren Unterschiede zwischen dem Verum- und Kontrollarm nicht eingeschätzt werden.

Die Rücklaufquoten der Instrumente bezogen auf die zur Verfügung stehenden Personen war im Kontrollarm etwas geringer als im Verumarm, aber lag fast zu allen Zeitpunkten bei über 80 %. Bezieht man den Rücklauf auf die Anzahl der randomisierten Personen, so war bereits sehr früh im Studienverlauf der Rücklauf bei unter 70 % bis hin zu < 3 % zu Studienwoche 72.

Die Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität wurden alle 4 Wochen eingesetzt. Zur Bewertung der gesundheitsbezogenen Lebensqualität wurden ausschließlich die Messzeitpunkte zu Baseline und zur Woche 12, 24, 36, 48, 60 und 72 herangezogen. Dieses Vorgehen begründet sich darin, dass die Therapiezyklen im Kontroll- und Verumarm unterschiedlich lang sind und so eine Vergleichbarkeit der Studienarme ermöglicht wird. Im Kontrollarm fanden drei von vier Erhebungen während eines laufenden Therapiezyklus statt.

Während einer Chemotherapie kann die Lebensqualität deutlich geringer sein als nach einer Pause, die nach Abschluss der Chemotherapie in beiden Armen erfolgte. Vor diesem Hintergrund wurden Informationen zur Lebensqualität nur zu den Messzeitpunkten betrachtet, als in beiden Studienarmen der jeweilige Therapiezyklus beendet war, um die Vergleichbarkeit der Daten zu gewährleisten.

Es liegen Ergebnisse aus Analysen mittels gemischter Modelle für wiederholte Messungen (MRMM) vor. Die Ergebnisse wurden in der vorliegenden Nutzenbewertung von Carfilzomib nicht berücksichtigt. Folgende Gründe sind hierfür zu nennen: In die MMRM-Modelle wurden alle Messzeitpunkte (alle 4 Wochen) berücksichtigt, was grundsätzlich zu begrüßen ist. Allerdings ergibt sich durch die unterschiedliche Länge der Therapiezyklen in Verum- und Kontrollarm eine Verzerrung der Lebensqualität, da im Kontrollarm die Lebensqualität häufig während des laufenden Therapiezyklus erhoben wurde (siehe Ausführungen im vorherigen Absatz). Dies könnte zu einer Verzerrung zugunsten von Carfilzomib führen. Weiterhin werden das Auftreten von Missings und die getroffene Annahme des zufälligen Auftretens von Missings als kritisch gesehen. So zeigte der pU, dass das Auftreten selbst und der Zeitpunkt des Auftretens der Missings von der zuvor gemessenen Lebensqualität abhängig waren. Die Lebensqualität ist wiederum von der Schwere der Erkrankung und dem Therapieregime abhängig, die zwar zu Baseline gleich verteilt waren, aber im Verlauf der Studie durch unterschiedliche Wirksamkeit der Therapieregime nicht mehr gleich verteilt sein müssen. Die Annahme des zufälligen Auftretens der Missings (MAR) zu Baseline wäre zwar erfüllt. Allerdings könnte im Verlauf der Studie diese Annahme zunehmend verletzt sein, was wiederum zu einer erheblichen Verzerrung der Ergebnisse führen könnte, deren Ausmaß und Richtung nicht einschätzbar ist [14]. Daher werden die Ergebnisse MMRM-Modelle über den gesamten Studienzeitraum als wenig valide und aussagekräftig angesehen.

Es wurden ebenfalls Responderanalysen zu allen untersuchten Skalen vorgelegt. Diese Responderanalysen nahmen zu allen dargestellten Zeitpunkten Bezug auf die gesamte randomisierte Studienpopulation, obwohl nur für einen Bruchteil der Population Informationen zur gesundheitsbezogenen Lebensqualität verfügbar waren. Darüber hinaus wurden die Responder über den Verlauf der Studie addiert. Somit wurden Personen, die zu einem Zeitpunkt eine Response zeigten, auch zu allen anderen nachfolgenden Messzeitpunkten als Responder gewertet, obwohl deren Nachbeobachtungszeit früher endete, z. B. aufgrund von Tod, Krankheitsprogress oder kürzerer Nachbeobachtungszeit. Die Ergebnisse der Responderanalysen sind in der Form nicht interpretierbar und wurden deshalb nicht in die Nutzenbewertung eingeschlossen.

Zusätzlich wurden Überlebenszeitanalysen zu Respondern vorgelegt. Hierbei wurde der Zeitpunkt berücksichtigt, wann eine klinisch relevante Veränderung der gesundheitsbezogenen Lebensqualität eintrat. Dies ist grundsätzlich zu begrüßen. Allerdings sprachen mehrere Gründe gegen die Berücksichtigung dieser Ergebnisse. Zum einen wurden, wie schon bei den MMRM-Modellen, alle verfügbaren Messzeitpunkte berücksichtigt, was aufgrund der Messung während eines laufenden Therapiezyklus im Kontrollarm zu einer Verzerrung des Ergebnisses führen könnte. Zum anderen handelt es sich um Post-hoc-Analysen, zu denen keine weiteren methodischen Informationen zur Verfügung standen, so dass eine Beurteilung der methodischen Qualität der Datenanalyse nicht möglich war. Darüber hinaus könnte eine informierte Zensurierung durch PFS vorliegen, da bei Eintritt einer Krankheitsprogression die Lebensqualität nicht länger

erhoben wurde und ein Fortschreiten der Erkrankung möglicherweise mit einer Minderung der Lebensqualität verbunden sein könnte.

Insgesamt wäre es wünschenswert gewesen, die Veränderung der Lebensqualität auf der Basis methodisch fundierter Responderanalysen beurteilen zu können. Solche Analysen sollten insbesondere die Verzerrung durch unterschiedlich lange Therapiezyklen berücksichtigen und nur die vergleichbaren Messpunkte heranziehen sowie konkurrierende Ereignisse, wie Tod und Krankheitsprogression, entsprechend berücksichtigen. Diese Kriterien wurden bei keiner der vom pU präsentierten Analysen erfüllt.

3.4.1 Zusammenfassende Einschätzung zur Lebensqualität

Es wird begrüßt, dass die Lebensqualität mittels unterschiedlichster, teils generischer, teils krankheitsspezifischer Instrumente erfasst wurde. Jedoch unterliegen die Ergebnisse Einschränkungen bezüglich methodischer Qualität der statistischen Analyse und der Interpretierbarkeit durch das offene Studiendesign. Vor diesem Hintergrund werden die Ergebnisse zur Lebensqualität als wenig belastbar eingeschätzt. Die beobachteten mittleren Veränderungen der Lebensqualität auf den präspezifizierten Skalen waren gering und können bezüglich der klinischen Relevanz nicht eingeschätzt werden, da zu allen Skalen keine validierten klinischen Relevanzschwellen vorliegen.

3.5 Sicherheit

Zur Beurteilung von sicherheitsrelevanten Endpunkten diente die Sicherheitspopulation (Cd: n=463; Vd: n=456). Die mediane Behandlungsdauer war im Verumarm deutlich länger als im Kontrollarm (39,9 vs. 26,8 Wochen), die mediane Anzahl begonnener Zyklen unterschied sich weniger stark mit 10 Zyklen im Carfilzomibarm und 8 Zyklen im Bortezomibarm. Bezüglich des Erhebungszeitraums besteht eine Diskrepanz zwischen Studienprotokoll und -bericht. Entgegen der Planungen im Studienprotokoll wurden im Studienbericht keine UE zwischen Randomisierung und erster Einnahme der Studienmedikation berichtet. Falls der Zeitraum zwischen Randomisierung und erster Einnahme der Studienmedikation zwischen den Studienarmen unterschiedliche Längen aufwies, könnte dies zu einer Verzerrung der Ergebnisse geführt haben.

Unerwünschte Ereignisse

Es traten bei fast allen Studienteilnehmern UE auf. Folgende UE mit einer Inzidenz $\geq 10\%$ unterschieden sich um mind. 5 %-Punkte zuungunsten von Carfilzomib: Pyrexie, Infektionen der oberen Atemwege, Bronchitis, Erbrechen, Dyspnoe, Husten, Muskelkrämpfe, Anämie, Kopfschmerzen und Hypertonus. Hingegen traten insbesondere in der SOC Erkrankungen des Nervensystems mit dem PT periphere Neuropathie, Schwindel, Parästhesie, periphere sensorische Neuropathie und Neuralgien UE deutlich häufiger im Kontroll- als im Verumarm auf.

Unerwünschte Ereignisse des NCI-CTCAE-Grades 3 und höher

UE des Grades 3 und höher traten numerisch häufiger im Verum- als im Kontrollarm auf. Anämie war dabei der PT, der am häufigsten in beiden Armen beobachtet wurde. Im PT Hypertonus unterschied sich die Häufigkeit um mehr als 5 %-Punkte zuungunsten von Carfilzomib.

Schwerwiegende unerwünschte Ereignisse

SUE traten deutlich häufiger im Verum- als im Kontrollarm auf. Pneumonie war hierbei der PT, der in beiden Studienarmen am häufigsten zu beobachten war, wobei hier ein Unterschied von 2,6 %-Punkten zugunsten von Carfilzomib zu verzeichnen war. Im PT Lungenembolie zeigte sich ein Unterschied von 2,5 %-Punkten zuungunsten von Carfilzomib. Für SUE des Grades 3 und höher lagen keine Informationen in den Studienunterlagen vor.

Zu einem Abbruch der Behandlung einer der Studienmedikationen führende UE

UE, die zum Abbruch einer Studienmedikation führten, waren in beiden Studienarmen annähernd gleich. Wurde die Expositionszeit beachtet, so zeigte sich ein statistisch signifikanter Unterschied zugunsten von Carfilzomib. Unter Bortezomib brachen deutlich häufiger Patienten aufgrund von peripherer Neuropathie die Therapie ab. Im Gegensatz waren im Carfilzomibarm Fatigue und Herzversagen häufiger Ursachen für Therapieabbrüche als im Kontrollarm.

Vom pU beschriebene UE von besonderem Interesse

Unterschiede von ≥ 5 %-Punkten zwischen den Studienarmen zuungunsten von Carfilzomib wurde insbesondere für Herzerkrankungen, Infektionen und Gefäßerkrankungen beobachtet. Hingegen fiel auf, dass zugunsten von Carfilzomib deutlich weniger periphere Neuropathien auftraten als im Kontrollarm.

Periphere Neuropathie

Periphere Neuropathie ist ein bekanntes UE der Proteasominhibitoren [15]. Unter der Therapie mit Carfilzomib traten im Vergleich zu Bortezomib deutlich weniger periphere Neuropathien des Grades 2 und 3 auf. Dieser Unterschied war statistisch signifikant zugunsten von Carfilzomib. Allerdings wurden bei der Berechnung des Effektmaßes nicht die Expositionszeiten, wie bei UE, die zum Abbruch einer Studienmedikation führten, berücksichtigt.

3.5.1 Zusammenfassende Einschätzung zur Sicherheit

In der Gesamtschau traten bei fast allen Studienteilnehmern UE im Verlauf der Studie auf. Hingegen überwogen numerisch UE des Grades 3 und höher als auch SUE und UE, die zum Tod führten, bei Carfilzomib. Diese numerischen Unterschiede waren jedoch statistisch nicht signifikant. UE, die zum Abbruch einer Studienmedikation führten, traten unter Berücksichtigung der Expositionszeit statistisch signifikant seltener unter der Therapie mit Carfilzomib als unter Bortezomib auf. Vorteile wurden auch im Endpunkt „periphere Neuropathie“ zugunsten von Carfilzomib beobachtet. Allerdings ist anzumerken, dass keine Informationen zu SUE des Grades 3 und höher vorlagen und unter Carfilzomib häufiger UE von besonderem Interesse, insbesondere Herz- und Gefäßerkrankungen und Infektionen, berichtet wurden.

Zusammenfassend scheint das Sicherheitsprofil von Carfilzomib keinen deutlichen Nachteil gegenüber Bortezomib zu haben. In Bezug auf periphere Neuropathien zeigte Carfilzomib gegenüber Bortezomib einen deutlichen signifikanten Vorteil.

4 Anforderungen an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen.

Die Patienten sind über die bekannten Risiken und Nebenwirkungen zu informieren und sorgfältig auf entsprechende Anzeichen hin zu beobachten, um diese unerwünschten Ereignisse zu vermeiden, frühzeitig zu erkennen und/oder einen schwerwiegenden Verlauf zu verhindern.

Wenn es bei Auftreten von Neutropenien, Thrombozytopenien oder renalen Toxizitäten trotz einer Dosisanpassung zu keiner Besserung kommt, sollte ein Abbruch der Behandlung mit Carfilzomib erwogen werden.

Die Behandlung muss durch Ärzte eingeleitet werden, die Erfahrung in der Anwendung von Krebstherapien haben.

5 Zusammenfassung der Bewertung

Carfilzomib ist zugelassen in Kombination mit Dexamethason für erwachsene Patienten mit Multiplem Myelom, die mindestens eine vorangegangene Therapie erhalten haben. Die Nutzenbewertung von Carfilzomib basiert auf der zulassungsbegründenden Studie ENDEAVOR (2011-003), einer offenen, kontrollierten, randomisierten, multizentrischen und multinationalen Studie der Phase III mit progressionsfreiem Überleben als primären Endpunkt. Die Studie umfasst zwei Therapiearme: Carfilzomib in Kombination mit Dexamethason (n=465) und Bortezomib in Kombination mit Dexamethason (n=464).

Die Ergebnisse der Studie ENDEAVOR werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Carfilzomib ein statistisch signifikant positiver Effekt bei hohem oder unklarem (\uparrow) oder niedrigem ($\uparrow\uparrow$) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem (\downarrow) oder niedrigem ($\downarrow\downarrow$) Verzerrungspotential bzw. kein Unterschied (\leftrightarrow) gezeigt werden konnte.

Tabelle 20: Zusammenfassende Darstellung der Studienergebnisse der Studie ENDEAVOR, Datenschnitt 10.11.2014

Endpunkt	Ergebnis	Effekt
Mortalität	Cd vs. Vd, n (%); HR (95%-KI), p-Wert¹⁾	
Gesamtüberleben ²⁾ (Interimsanalyse basierend auf 33 % der geplanten Ereignisse)	75 (16,2) vs. 88 (18,9) 0,79 (0,58; 1,08) 0,065	\leftrightarrow
Gesundheitsbezogene Lebensqualität	Cd vs. Vd, MW (SD)	
EORTC QLQ-C30 ³⁾ : Skala Gesundheitszustand/Lebensqualität	<u>Baseline:</u> 61,5 (21,3) vs. 63,7 (21,7) <u>Wo. 36:</u> 63,7 (19,2) vs. 63,3 (18,0) <u>Wo. 72:</u> 68,9 (19,6) vs. 68,2 (14,3)	*
EORTC QLQ-MY20 ⁴⁾ : Skala „Krankheitssymptome“	<u>Baseline:</u> 27,8 (20,5) vs. 25,1 (20,6) <u>Wo. 36:</u> 16,4 (16,2) vs. 16,2 (16,2) <u>Wo. 72:</u> 14,2 (17,3) vs. 16,2 (16,9)	*
FACT/GOG-Ntx ⁵⁾ : Skala „zusätzliche Beschwerden“	<u>Baseline:</u> 37,0 (6,0) vs. 37,0 (6,3) <u>Wo. 36:</u> 37,0 (5,8) vs. 36,2 (6,5) <u>Wo. 72:</u> 37,1 (6,1) vs. 36,8 (5,8)	*
Sicherheit	Cd vs. Vd, n (%); RR (95% KI)⁶⁾	
UE des NCI-CTCAE-Grades \geq 3	339 (73,2) vs. 305 (66,9) 0,92 (0,78; 1,07)	\leftrightarrow
SUE	224 (48,4) vs. 162 (35,5) 1,06 (0,87; 1,03)	\leftrightarrow
SUE des NCI-CTCAE-Grades \geq 3	k.A.	**

Endpunkt	Ergebnis	Effekt
UE, das zum Abbruch einer Studienmedikation führte	92 (19,9) vs. 95 (20,8) 0,71 (0,53; 0,94)	↑
UE, das zum Tod führte	23 (5,0) vs. 18 (3,9) 0,93 (0,50; 1,73)	↔
Periphere Neuropathie CTC-Grad ≥ 2	28 (6,0) vs. 146 (32,0) 0,19 (0,13; 0,28)	↑

¹⁾ einseitiger Test, Signifikanzniveau: $\alpha=0,025$

²⁾ Eine weitere Interimsanalyse bei Erreichen von ca. 394 Ereignissen ist geplant. Finale Analyse bei Vorliegen von 496 Ereignissen.

³⁾ Die Spanne des Scores beträgt 0 bis 100. Höhere Werte bedeuten eine bessere Lebensqualität.

⁴⁾ Die Spanne des Scores beträgt 0 bis 100. Niedrigere Werte bedeuten eine bessere Lebensqualität.

⁵⁾ Die Spanne des Scores beträgt 0 bis 44. Höhere Werte bedeuten eine bessere Lebensqualität.

⁶⁾ Risk Ratio wurde, ausgenommen für den Endpunkt „periphere Neuropathie“, unter Berücksichtigung der Expositionszeit berechnet. Post-hoc-Analyse.

* Nicht einschätzbar, da keine validierten MID für Gruppenvergleiche vorliegen und die angewendeten Methoden zur statistischen Analyse keine validen oder interpretierbaren Ergebnisse lieferten.

** Nicht einschätzbar, da keine Angaben im Modul 4 oder Studienbericht vorgelegt wurden.

Abkürzungen: MW: Mittelwert; SD: Standardabweichung; k.A. keine Angaben im Modul 4 oder Studienbericht; Wo.: Studienwoche

Referenzen

1. **Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al.** The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85(5):365-376.
2. **Amgen.** Clinical Study Protocol (2011-003): Amendment 4: A Randomized, Open-label, Phase 3 Study of Carfilzomib Plus Dexamethasone vs. Bortezomib Plus Dexamethasone in Patients with Relapsed Multiple Myeloma (ENDEAVOR). 2015.
3. **Amgen.** Clinical Study Report (2011-003): A Randomized, Open-label, Phase 3 Study of Carfilzomib Plus Dexamethasone vs. Bortezomib Plus Dexamethasone in Patients with Relapsed Multiple Myeloma (ENDEAVOR). 2015.
4. **Amgen.** Dossier zur Nutzenbewertung gemäß § 35a SGB V; Wirkstoff Carfilzomib (neues Anwendungsgebiet - Kombination mit Dexamethason allein). Modul 1-4 [unveröffentlicht]. 2016.
5. **Amgen.** Kyprolis 60 mg. Pulver zur Herstellung einer Infusionslösung [online]. 06.2016. [Zugriff: 01.09.2016]. URL: <http://www.fachinfo.de/>.
6. **Amgen.** Statistical Analysis Plan (2011-003): A Randomized, Open-label, Phase 3 Study of Carfilzomib Plus Dexamethasone vs. Bortezomib Plus Dexamethasone in Patients with Relapsed Multiple Myeloma (ENDEAVOR). 2014.
7. **Augustson BM, Begum G, Dunn JA, Barth NJ, Davies F, Morgan G, et al.** Early mortality after diagnosis of multiple myeloma: analysis of patients entered onto the United Kingdom Medical Research Council trials between 1980 and 2002--Medical Research Council Adult Leukaemia Working Party. *J Clin Oncol* 2005;23(36):9219-9226.
8. **Calhoun EA, Welshman EE, Chang CH, Lurain JR, Fishman DA, Hunt TL, et al.** Psychometric evaluation of the Functional Assessment of Cancer Therapy/Gynecologic Oncology Group-Neurotoxicity (Fact/GOG-Ntx) questionnaire for patients receiving systemic chemotherapy. *Int J Gynecol Cancer* 2003;13(6):741-748.
9. **Cocks K, Cohen D, Wisloff F, Sezer O, Lee S, Hippe E, et al.** An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-MY20) in assessing the quality of life of patients with multiple myeloma. *Eur J Cancer* 2007;43(11):1670-1678.
10. **Dimopoulos MA, Kastritis E, Rosinol L, Blade J, Ludwig H.** Pathogenesis and treatment of renal failure in multiple myeloma. *Leukemia* 2008;22(8):1485-1493.
11. **Dimopoulos MA, Moreau P, Palumbo A, Joshua D, Pour L, Hajek R, et al.** Carfilzomib and dexamethasone versus bortezomib and dexamethasone for patients with relapsed or refractory multiple myeloma (ENDEAVOR): a randomised, phase 3, open-label, multicentre study. *Lancet Oncol* 2016;17(1):27-38.
12. **Dubois D, Dhawan R, van de Velde H, Esseltine D, Gupta S, Viala M, et al.** Descriptive and prognostic value of patient-reported outcomes: the bortezomib experience in relapsed and refractory multiple myeloma. *J Clin Oncol* 2006;24(6):976-982.

13. **Durie BG, Harousseau JL, Miguel JS, Blade J, Barlogie B, Anderson K, et al.** International uniform response criteria for multiple myeloma. *Leukemia* 2006;20(9):1467-1473.
14. **European Medicines Agency (EMA).** Guideline on Missing Data in Confirmatory Clinical Trials EMA/CPMP/EWP/1776/99 Rev. 1 [online]. 02.07.2010. London (GBR): EMA; 2010. [Zugriff: 01.10.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf.
15. **European Medicines Agency (EMA).** Kyprolis/Carfilzomib: Group of variations including an extension of indication assessment report EMEA/H/C/003790/II/0001/G [online]. 26.05.2016. London (GBR): EMA; 2016. [Zugriff: 01.09.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/003790/WC500211642.pdf.
16. **Jordan K, Proskorovsky I, Lewis P, Ishak J, Payne K, Lordan N, et al.** Effect of general symptom level, specific adverse events, treatment patterns, and patient characteristics on health-related quality of life in patients with multiple myeloma: results of a European, multicenter cohort study. *Support Care Cancer* 2014;22(2):417-426.
17. **Klein JP, Moeschberger ML.** Survival Analysis: Techniques for Censored and Truncated Data. New York; Springer; 2003.
18. **Kvam AK, Fayers PM, Wisloff F.** Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. *Eur J Haematol* 2011;87(4):330-337.
19. **Lan KKG, Demets DL.** Discrete sequential boundaries for clinical trials. *Biometrika* 1983;70(3):659-663.
20. **Ludwig H, Miguel JS, Dimopoulos MA, Palumbo A, Garcia Sanz R, Powles R, et al.** International Myeloma Working Group recommendations for global myeloma care. *Leukemia* 2014;28(5):981-992.
21. **Mols F, Oerlemans S, Vos AH, Koster A, Verelst S, Sonneveld P, et al.** Health-related quality of life and disease-specific complaints among multiple myeloma patients up to 10 yr after diagnosis: results from a population-based study using the PROFILES registry. *Eur J Haematol* 2012;89(4):311-319.
22. **National Comprehensive Cancer Network (NCCN).** Multiple Myeloma. NCCN Clinical Practice Guidelines in Oncology, Version 2.2016 [online]. Fort Washington (USA): NCCN; 2016. [Zugriff. URL: http://www.nccn.org/professionals/physician_gls/pdf/myeloma.pdf.
23. **Onyx Pharmaceuticals.** PROs in the ENDEAVOR Clinical Trial. An Expert Report (Final) [unveröffentlicht]. 2015.
24. **Osoba D, Rodrigues G, Myles J, Zee B, Pater J.** Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998;16(1):139-144.
25. **Rajkumar SV.** Multiple myeloma: 2011 update on diagnosis, risk-stratification, and management. *Am J Hematol* 2011;86(1):57-65.

26. **Ralston SH, Gallacher SJ, Patel U, Campbell J, Boyle IT.** Cancer-associated hypercalcemia: morbidity and mortality. Clinical experience in 126 treated patients. *Ann Intern Med* 1990;112(7):499-504.
27. **Robert Koch Institut (RKI), Gesellschaft der epidemiologischen Krebsregister in Deutschland (GeKiD).** Krebs in Deutschland 2011/2012; 10. Ausgabe [online]. Berlin (GER): Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland; 2015. [Zugriff: 01.09.2016]. (Gesundheitsberichterstattung des Bundes). URL: http://www.gekid.de/Doc/krebs_in_deutschland_2015.pdf.
28. **Saad F, Lipton A, Cook R, Chen YM, Smith M, Coleman R.** Pathologic fractures correlate with reduced survival in patients with malignant bone disease. *Cancer* 2007;110(8):1860-1867.
29. **Stead ML, Brown JM, Velikova G, Kaasa S, Wisloff F, Child JA, et al.** Development of an EORTC questionnaire module to be used in health-related quality-of-life assessment for patients with multiple myeloma. European Organization for Research and Treatment of Cancer Study Group on Quality of Life. *Br J Haematol* 1999;104(3):605-611.
30. **Stringer S, Basnayake K, Hutchison C, Cockwell P.** Recent advances in the pathogenesis and management of cast nephropathy (myeloma kidney). *Bone Marrow Res* 2011;2011:493697.
31. **Wisloff F, Eika S, Hippe E, Hjorth M, Holmberg E, Kaasa S, et al.** Measurement of health-related quality of life in multiple myeloma. Nordic Myeloma Study Group. *Br J Haematol* 1996;92(3):604-613.
32. **Ye Y, Li A, Liu L, Yao B.** A group sequential Holm procedure with multiple primary endpoints. *Stat Med* 2013;32(7):1112-1124.

Anhang

Tabelle 21: Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität der präspezifizierten Subskalen des EORTC QLQ-C30 in der Studie ENDEAVOR, Datenschnitt 10.11.2014

Studie ENDEAVOR	Cd (N=464)		Vd (N=465)	
Zeitpunkt	n ¹ /N ² (%)	MW (SD)	n ¹ /N ² (%)	MW (SD)
„Gesundheitszustand/Lebensqualität“⁽³⁾				
Baseline	407/464 (87,7)	61,5 (21,3)	392/465 (84,3)	63,7 (21,7)
Woche 12	383/408 (93,9)	61,3 (20,1)	336/388 (86,6)	59,4 (20,7)
Woche 24	298/343 (86,9)	63,6 (18,0)	222/254 (87,4)	59,9 (19,0)
Woche 36	235/258 (91,1)	63,7 (19,2)	142/162 (87,7)	63,3 (18,0)
Woche 48	137/159 (86,2)	62,2 (18,7)	73/80 (91,3)	63,9 (14,8)
Woche 60	73/82 (89,0)	66,9 (16,9)	28/37 (75,7)	68,2 (18,7)
Woche 72	41/44 (93,2)	68,9 (19,6)	11/12 (91,7)	68,2 (14,3)
Subskala „Fatigue“⁽⁴⁾				
Baseline	407/464 (87,7)	33,6 (23,3)	392/465 (84,3)	34,5 (24,0)
Woche 12	383/408 (93,9)	35,0 (23,3)	336/388 (86,6)	39,6 (24,0)
Woche 24	298/343 (86,9)	34,8 (20,6)	222/254 (87,4)	37,6 (22,8)
Woche 36	235/258 (91,1)	34,4 (21,7)	142/162 (87,7)	32,9 (19,1)
Woche 48	137/159 (86,2)	35,3 (21,2)	73/80 (91,3)	32,9 (22,2)
Woche 60	73/82 (89,0)	31,4 (19,7)	28/37 (75,7)	31,7 (25,8)
Woche 72	41/44 (93,2)	29,3 (19,8)	11/12 (91,7)	34,3 (18,2)
Subskala „Schmerz“⁽⁴⁾				
Baseline	407/464 (87,7)	33,1 (28,5)	392/465 (84,3)	31,2 (27,4)
Woche 12	383/408 (93,9)	28,2 (25,9)	336/388 (86,6)	24,3 (24,4)
Woche 24	298/343 (86,9)	24,6 (24,0)	222/254 (87,4)	20,4 (20,2)
Woche 36	235/258 (91,1)	23,0 (22,2)	142/162 (87,7)	21,8 (22,7)
Woche 48	137/159 (86,2)	22,6 (24,4)	73/80 (91,3)	22,1 (21,5)
Woche 60	73/82 (89,0)	17,3 (22,4)	28/37 (75,7)	21,5 (21,2)
Woche 72	41/44 (93,2)	24,2 (29,2)	11/12 (91,7)	17,1 (20,6)

Studie ENDEAVOR	Cd (N=464)		Vd (N=465)	
Zeitpunkt	n ¹ /N ² (%)	MW (SD)	n ¹ /N ² (%)	MW (SD)
„Gesundheitszustand/Lebensqualität“³				
Subskala „Physis“³				
Baseline	407/464 (87,7)	73,8 (22,0)	392/465 (84,3)	74,0 (22,0)
Woche 12	383/408 (93,9)	74,0 (20,2)	336/388 (86,6)	72,4 (21,8)
Woche 24	298/343 (86,9)	76,5 (17,3)	222/254 (87,4)	73,5 (20,6)
Woche 36	235/258 (91,1)	76,6 (18,1)	142/162 (87,7)	76,3 (17,7)
Woche 48	137/159 (86,2)	76,4 (17,8)	73/80 (91,3)	75,9 (18,7)
Woche 60	73/82 (89,0)	78,1 (16,6)	28/37 (75,7)	78,8 (22,1)
Woche 72	41/44 (93,2)	79,0 (17,4)	11/12 (91,7)	70,3 (23,7)
Subskala „Rolle“³				
Baseline	407/464 (87,7)	73,7 (28,5)	392/465 (84,3)	73,6 (28,6)
Woche 12	383/408 (93,9)	72,5 (25,3)	336/388 (86,6)	69,6 (28,9)
Woche 24	298/343 (86,9)	74,2 (22,0)	222/254 (87,4)	71,2 (25,9)
Woche 36	235/258 (91,1)	75,4 (24,5)	142/162 (87,7)	74,8 (22,9)
Woche 48	137/159 (86,2)	73,6 (24,3)	73/80 (91,3)	76,0 (21,9)
Woche 60	73/82 (89,0)	77,4 (19,7)	28/37 (75,7)	76,8 (28,8)
Woche 72	41/44 (93,2)	77,6 (23,2)	11/12 (91,7)	75,8 (29,2)
Subskala „Übelkeit/Erbrechen“⁴				
Baseline	407/464 (87,7)	4,9 (11,3)	392/465 (84,3)	5,2 (12,7)
Woche 12	383/408 (93,9)	4,7 (11,2)	336/388 (86,6)	3,8 (10,9)
Woche 24	298/343 (86,9)	3,0 (9,2)	222/254 (87,4)	4,9 (11,9)
Woche 36	235/258 (91,1)	3,9 (8,4)	142/162 (87,7)	3,9 (9,0)
Woche 48	137/159 (86,2)	3,0 (8,6)	73/80 (91,3)	3,0 (7,0)
Woche 60	73/82 (89,0)	3,0 (6,5)	28/37 (75,7)	3,7 (8,0)
Woche 72	41/44 (93,2)	1,5 (5,0)	11/12 (91,7)	2,4 (6,0)

¹) Anzahl Studienteilnehmer, die den Fragebogen ausfüllten

²) Anzahl Studienteilnehmer, die zu dem Zeitpunkt zum Ausfüllen des Fragebogens zur Verfügung standen

³) Die Spanne des Scores beträgt 0 bis 100. Höhere Werte bedeuten eine bessere Lebensqualität.

⁴) Die Spanne des Scores beträgt 0 bis 100. Niedrigere Werte bedeuten eine bessere Lebensqualität.

Abkürzungen: SD: Standardabweichung; MW: Mittelwert; KI: Konfidenzintervall

Tabelle 22: Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität des EORTC QLQ-MY20 in der Studie ENDEAVOR, Datenschnitt 10.11.2014

Studie ENDEAVOR	Cd (N=464)		Vd (N=465)	
Zeitpunkt	n ¹ /N ² (%)	MW (SD)	n ¹ /N ² (%)	MW (SD)
Subskala „Krankheitssymptome“³⁾				
Baseline	407/464 (87,7)	27,8 (20,5)	392/465 (84,3)	25,1 (20,6)
Woche 12	383/408 (93,9)	18,6 (17,0)	336/388 (86,6)	18,2 (17,2)
Woche 24	298/343 (86,9)	16,9 (16,2)	222/254 (87,4)	16,5 (15,8)
Woche 36	235/258 (91,1)	16,4 (16,2)	141/162 (87,0)	16,2 (16,2)
Woche 48	137/159 (86,2)	16,7 (15,9)	73/80 (91,3)	15,3 (16,6)
Woche 60	73/82 (89,0)	15,0 (16,1)	28/37 (75,7)	13,5 (14,8)
Woche 72	41/44 (93,2)	14,2 (17,3)	11/12 (91,7)	16,2 (16,9)
Subskala „Nebenwirkungen“³⁾				
Baseline	407/464 (87,7)	16,8 (13,7)	392/465 (84,3)	16,8 (13,3)
Woche 12	383/408 (93,9)	17,5 (13,6)	336/388 (86,6)	20,4 (13,9)
Woche 24	298/343 (86,9)	17,3 (13,5)	222/254 (87,4)	21,0 (15,2)
Woche 36	235/258 (91,1)	20,0 (12,9)	141/162 (87,0)	17,7 (13,5)
Woche 48	137/159 (86,2)	17,9 (11,9)	73/80 (91,3)	18,4 (14,0)
Woche 60	73/82 (89,0)	16,7 (12,5)	28/37 (75,7)	18,4 (14,1)
Woche 72	41/44 (93,2)	15,3 (12,3)	11/12 (91,7)	16,5 (14,6)

¹⁾ Anzahl Studienteilnehmer, die den Fragebogen ausfüllten

²⁾ Anzahl Studienteilnehmer, die zu dem Zeitpunkt zum Ausfüllen des Fragebogens zur Verfügung standen

³⁾ Die Spanne des Scores beträgt 0 bis 100. Niedrigere Werte bedeuten eine bessere Lebensqualität.

Abkürzungen: SD: Standardabweichung; MW: Mittelwert; KI: Konfidenzintervall

Tabelle 23: Ergebnisse zu Rücklaufquoten und gesundheitsbezogener Lebensqualität der Subskala „zusätzliche Beschwerden“¹⁾ des FACT/GOG-Ntx in der Studie ENDEAVOR, Datenschnitt 10.11.2014

Studie ENDEAVOR	Cd (N=464)		Vd (N=465)	
Zeitpunkt	N ² /N ³ (%)	MW (SD)	N ² /N ³ (%)	MW (SD)
Baseline	407/464 (87,7)	37,0 (6,0)	392/465 (84,3)	37,0 (6,3)
Woche 12	383/408 (93,9)	36,8 (6,5)	336/388 (86,6)	35,5 (6,9)
Woche 24	298/343 (86,9)	37,0 (6,5)	222/254 (87,4)	35,5 (6,9)
Woche 36	235/258 (91,1)	37,0 (5,8)	141/162 (87,0)	36,2 (6,5)
Woche 48	137/159 (86,2)	36,2 (6,5)	73/80 (91,3)	36,0 (7,8)
Woche 60	73/82 (89,0)	37,2 (5,3)	28/37 (75,7)	36,9 (7,1)
Woche 72	41/44 (93,2)	37,1 (6,1)	11/12 (91,7)	36,8 (5,8)

¹⁾ Die Spanne des Scores beträgt 0 bis 44. Höhere Werte bedeuten eine bessere Lebensqualität.

²⁾ Anzahl Studienteilnehmer, die den Fragebogen ausfüllten

³⁾ Anzahl Studienteilnehmer, die zu dem Zeitpunkt zum Ausfüllen des Fragebogens zur Verfügung standen

Abkürzungen: SD: Standardabweichung; MW: Mittelwert; KI: Konfidenzintervall

Dossierbewertung für Orphan Drugs

Carfilzomib (Kyprolis®)